Armed Groups: Competition and Political Violence*

Martin Gassebner † Paul Schaudt‡ Melvin H. L. Wong §

This version: December 2022

Abstract

We show that the proliferation of armed groups increases the amount of organized political violence. The natural death of a tribal leader provides quasi-experimental variation in the number of armed groups across districts in Pakistan. Employing event study designs and IV-regressions allows us to isolate the effect of the number of armed groups on political violence from locational fundamentals of conflict, e.g., local financing and recruiting opportunities or government capacity. In line with the idea that armed groups compete for resources and supporters, we estimate semi-elasticities of an additional armed group on political violence ranging from 50 to 60%. Introducing a novel proxy for government counter-insurgency efforts enables us to show that this increase is driven by insurgency groups and not the state. Moreover, we show that groups splitting-up compensate for their capacity loss by switching to non-capital intensive attacks.

Keywords: Political violence, competition, armed groups, conflict, terrorism, double-counting

JEL Classification: D74, F52, H56

*We thank Baloch journalist Malik Siraj Akbar who gave us deep insights into the conflict in Balochistan and the different armed groups. We also thank Nicola Limodio for sharing some of his data with us. We are grateful for helpful suggestions by Roland Hodler, Krisztina Kis-Katos, Todd Sandler, and comments from seminar participants at St.Gallen, the RITES seminar series, and conference participants at EPCS, and HiCN. Finally, we like to thank Lukas Bertram, Tom Görges, Linyun Yu for excellent research assistance. All errors are ours. Note, that this paper has been previously circulated as "Armed groups in conflict: Competition and political violence in Pakistan". The findings, interpretations, and conclusions expressed in this article are entirely those of the authors. They do not necessarily represent the views of the KfW Development Bank and its affiliated organizations.

†Leibniz University Hannover, Institute of Macroeconomics; KOF Swiss Economic Institute; CESifo Munich; e-mail: *gassebner@mak.uni-hannover.de*

‡University of Bern, SIAW University of St.Gallen, e-mail: *paul.schaudt@unibe.ch*

§KfW Development Bank; e-mail: *melvin.wong@kfw.de*

1. Introduction

The proliferation of armed groups is often associated with a rise of organized political violence[1] and failing states. Prominent examples include Libya and Syria since 2011, and the Democratic Republic of Congo during the Great War of Africa.[2] An additional armed group can destabilize the status quo by threatening the influence of incumbent groups and the government. The additional group may amplify the threat if it claims to fight for the same cause as an established group. In such cases, the additional group not only challenges the monopoly of violence from other actors but threatens their distinct support base, e.g., financial supporters and recruits. The local capacity of a group is a key driver of political violence (Limodio, 2022; Sviatschi, 2022), and a new entrant can directly reduce this capacity. A prominent example is the appearance of Hamas in the Gaza Strip and the West Bank challenging the Palestinian Liberation Organization (PLO) as the sole agent of the Palestinians. In summary, the potentially opposing competition and capacity effects for the groups in a location do not allow for a clear ex-ante expectation on how an additional armed group affects organized violence.

An additional armed group is likely to increase violence if competition for resources induces groups to commit more attacks than they would find optimal otherwise (much like oligopolistic firms) in order to signal relevance. The capacity effect is likely to be negative if groups exhibit increasing returns to scale in generating attacks[3] or become resource-constraint.

Empirical evidence is limited. Currently, the literature only reports positive correlations between the number of armed groups and the frequency and severity of political violence (Findley and Young, 2012; Nemeth, 2014; Conrad and Greene, 2015). The main problem in estimating the causal effect of an increase in the number of armed groups on political violence (apart from potentially opposing mechanisms) is that the number of armed groups within a given geographic area is endogenous. First, groups most likely select themselves into given areas (Gaibulloev, 2015). The selection, in turn, reasonably depends on the strength of incumbent actors as well as attributes inert to the area in question. Prominent examples are weak state capacity (Fearon and Laitin, 2003), local financing opportunities (Berman et al., 2017; Limodio, 2022), and the attitude of the local population (Berman et al., 2011). Second, groups have varying goals and strategies, respond to different incentives, and might have diverse support groups (see Stanton, 2013; Kis-Katos et al., 2014; Toft and Zhukov, 2015). Hence, new groups may form to cater to previously neglected interests and grievances. Finally, political violence itself affects the

[1]We use the term organized political violence as a general term for politically motivated violence, such as civil war, terrorism, and counter-insurgencies.

[2]Taken to the extreme, the proliferation of armed actors means a war of everyone against everyone, famously making life "solitary, poor, nasty, brutish, and short" (Hobbes, 1969).

[3]Adaptation of cheap technological innovations makes this a very likely scenario (Faria, 2014).

number of armed groups, as some groups bleed out during a conflict or are attracted by the fighting itself, e.g., hunting their opponents across locations.

This paper provides quasi-experimental evidence on the effect of group proliferation on the frequency and intensity of organized political violence. We exploit a unique setting in which the number of armed groups increases through a split of a separatist group that is plausibly exogenous to the conflict dynamics. Specifically, we exploit the split of the Baloch Liberation Army (BLA) into the BLA and United Baloch Army (UBA), operating primarily within districts of the Balochistan province in Pakistan. The split between the BLA and UBA goes back to a leadership dispute between two brothers who, in short, could not agree on the organization's leadership. While disagreement between the brothers could be related to some unobserved conflict dynamics, the split of the BLA has the additional feature that the groups only effectively split after the father of the two brothers, who suppressed open conflict between them, died of natural causes following a relatively short and severe illness.[4]

The exogenous timing of the father's death and the groups' overlapping area of operations allows us to specify event studies and generalized difference-in-difference (DiD) specifications. We test if districts in which the BLA has traditionally been more active experience more violence following the split. Moreover, we use the DiD setup as a shift share instrument for the number of active armed groups operating within districts.

We estimate that an additional active armed group increases the quantity of political violence between 50% and 80% and the severity of violence (the sum of individuals wounded or killed) between 50% and 100%. The results suggest that the competition effect (for publicity, recruits, and/or financing) between armed groups dominates on average in our setting. Concerns about unobserved confounders explaining the UBA formation are relatively small since the general goals, target audience, primary opponent, and tactics of the BLA and UBA are similar.[5] Moreover, we do not find evidence that our results are driven by infighting between armed groups, increased counter-insurgency efforts by the government, changes in politically disenfranchised populations, or local financing opportunities.

Taking our analysis to the group level, we leverage the UBA split from the BLA to i) test for the capacity effect experienced by the BLA and ii) investigate how the BLA allocates its attacks in response to increased competition. We show that the BLA primarily conducts additional attacks in districts in which other groups, as well as the UBA, are active. Hence, we can rule out that increased violence is driven by competition

[4]Khair Baksh Marri died within five days after being admitted to the hospital (Khan, 2014; News International, 2014).

[5]Looking at raw data shows that on the district-year level, 21 % of BLA attacks do not cause bodily harm, while this number is 26% for the UBA. Both groups conduct a singular severe attack in 52 % of the district-years in which they are active. Regarding targets, both groups target private citizens one third of the time and businesses about 20% (BLA) and 23% (UBA), respectively.

between the BLA and UBA alone. Moreover, we document that the BLA conducts more non-capital intensive attacks following its split, which provides suggestive evidence that the split is indeed a negative capacity shock. However, the fatalities inflicted by the BLA, both in absolute terms and relative to other groups operating within the same district, do not decrease. Hence, the BLA seems to be able to compensate for the negative capacity effect by switching strategies, which is in line with theoretical predictions of Bueno de Mesquita (2013).

Our empirical analysis combines data from multiple publicly available data sources on political violence committed by the various armed groups within Pakistan. To measure the number of armed groups correctly, we systematically document all mergers and splits of armed groups in Pakistan between 1990 and 2018. Thus, we provide a unified analysis of organized political violence, including terrorism, guerilla warfare, as well as more symmetric forms of political violence. This allows us to test if armed groups change their strategies in response to increased competition. Recent theoretical and empirical work highlights that groups alter their strategies in response to changing constraints, of which increased competition could be an important factor.[6]

Combining data on terrorism from the Global Terrorism Database (GTD) (START, 2019) and political violence more broadly from the UCDP Georeferenced Event Dataset (GED) (Sundberg and Melander, 2013) allows us to increase coverage and proxy for government counter-insurgency efforts. We differentiate insurgency from counter-insurgency by exploiting the different inclusion criteria of events for each database. Accounting for counter-insurgency activity highlights that the violence is primarily driven by armed groups and not by the government's reaction to the split of the BLA.

We contribute to various strands of the literature. Our results show that the proliferation of armed groups increases organized political violence, adding additional insights to the literature on the determinants of political violence (see Blattman and Miguel, 2010; Gaibulloev and Sandler, 2019, for excellent overviews). Conceptually, we highlight that the proliferation of actors has an independent effect on political violence, even if local determinants of conflict, such as opportunity cost (Dube and Vargas, 2013) or state capacity (Fearon and Laitin, 2003; Dube and Naidu, 2015) remain constant. We also provide evidence that group proliferation seems not to affect infighting between groups in settings where group finances do not depend on the extraction of natural resources (as in Morelli and Rohner, 2015; Adhvaryu et al., 2018; Gehring et al., 2019), but mostly on local contributions (Limodio, 2022). On the econometric side, we show that group proliferation is a potential omitted variable in many studies and cannot be captured by fixed effects in monadic settings. Moreover, the issue cannot be resolved by focusing on

[6] For a theoretical model see Bueno de Mesquita (2013). For empirical evidence showing how different groups use different strategies, see Stanton (2013). For the varying impact of shocks and support groups on different groups, see Dube and Naidu (2015); Toft and Zhukov (2015); Limodio (2022).

smaller units such as grid-cells.[7]

Methodologically, we provide a novel approach to proxy for counter-insurgency activity by exploiting different coding criteria across databases. In doing so, we also provide a transparent way to account for potential double-counting, which can result from combining multiple databases on political violence. We tackle the issue with a data-driven approach. Conceptually, we implement an uncertainty-based measure applying spatial and temporal buffers surrounding each incident from one dataset and flag incidents in the second dataset that fall within the joint buffer. In essence, the approach provides a transparent way to trade off potential false-positive vs. false-negative assignments of double-counts.

Finally, we provide new time-variant data on the armed group level itself. Specifically, we collected the universe of mergers and splits for armed groups in Pakistan since 1990. Most current group level variables are time-invariant ideology and support group characteristics (Kis-Katos et al., 2011; Polo and Gleditsch, 2016). Two notable exceptions are the contributions by König et al. (2017) and Trebbi and Weese (2019) that document observed and unobserved coalition structures over time. We complement the latter two by de facto group changes.

The remainder of the paper is structured as follows: Section 2 introduces our setting in detail. Section 3 presents our data and the definition of our core variables. Section 4 discusses our empirical strategy. Section 5 reports our main results. Section 6 explores alternative mechanisms and extends our baseline analysis to the group level. Finally, Section 7 provides a brief overview of the robustness tests, and Section 8 concludes.

2. Setting and background

The Balochistan conflict is an ethnic dispute concentrated in the Balochistan province[8] of Pakistan[9]. It started in 1948 when newly independent Pakistan annexed the autonomous Baloch state of Kalat. Since the start, there have been several violent periods between Pakistan and Balochi insurgents: 1958-59, 1962-63, 1973-77, and ongoing since the early 2000s (Times of India, 2016). One of the most important figures that emerged during the 1970s insurgency was Kahir Bakhsh Marri (KBM), who led the Balochistan People's Liberation Front (BPLF). After concessions from the government, the conflict de-escalated, although it smoldered beneath the surface until it flared up again in the early 2000s. Most current insurgent groups (the BPLF no longer exists) call for an

[7]See Buhaug and Rød (2006); Tollefsen et al. (2012); Besley and Reynal-Querol (2014); Condra et al. (2018) for prominent examples.

[8]One of the four provinces in Pakistan which form the first sub-national layer together with two autonomous territories and the Federal Territory of Islamabad.

[9]Traditional Balochistan has been divided between Iran, Afghanistan, and Pakistan following the colonial period.

independent Balochistan. Among the many reasons for the insurgency are systemic repression and marginalization of Baloch people and the exploitation of natural resources without improvements in local living conditions, an issue that has continuously been raised since the 1960s.[10] As Dashti (2017, chapter 1) puts it: "[t]he Baloch are considered the poorest people while their land is amongst the richest in the world." The recent development follows a vicious cycle of violence: Pakistan follows a "pick up and dump strategy" whereby the Baloch opposition is rounded up and subsequently tortured and killed (Rashid, 2014). The insurgents initially attacked the military, but they have also turned against non-Baloch natives recently.

The BLA is one of the key players in the insurgency movement led by the Marri tribe. It was founded around 2000 by the eldest son of KBM. Other Baloch insurgency groups exist, such as the Baloch Liberation Front (BLF), Baloch Republican Army (BRA), Balochistan Liberation United Front (BLUF), or United Baloch Army (UBA).[11] The groups' area of operations is concentrated mainly across districts within Balochistan. All of the Baloch insurgency groups are considered terrorist organizations by the Pakistani government (NACTA, 2020).

Despite the similarity of the groups, Baloch insurgency groups are distinct entities that compete against each other. Groups primarily compete for attention, financial backers, and recruits within the Balochistan province but rarely fight each other. Hence, visibility is key for each group. Jetter (2017) highlights that a reduction in media attention decreases the attention pay-offs for a group, which in turn reduces the group's capabilities. Attacks on protected government institutions and incidences with high casualties demonstrate the capability of a group and will generate more attention. This logic seems especially crucial in this setting since the established insurgency groups of Balochistan have similar platforms. Furthermore, Baloch insurgency groups rely heavily on financing from other governments, wealthy individuals, and the local middle-class (Economist, 2012).

How did the UBA enter the conflict, and is it plausible that its appearance is exogenous with respect to the local conflict dynamics? Baloch groups usually do not openly communicate who their leaders are. In the case of the BLA, KBM seems to be the person who has been calling the shots. In 2007, the previous leader of the BLA, Balach Marri, was killed in action (Dawn.com, 2014). Balach Marri is one of six sons of KBM and BLA leadership passed to the next-born brother, Hyrbyair Marri. His younger brother Mehran Marri was in dispute with Hyrbyair regarding leadership and strategy. Personal

[10] The Baloch region is abundant, among other things, in natural gas, copper, and gold (Shah, 2017). It also provides access to the Straits of Hormuz. De Luca et al. (2018) document that while most of Pakistan's gas is produced in Balochistan, the central government charges lower prices for it and pays fewer royalties compared to gas from other regions.

[11] The set of Baloch insurgency groups, apart from the appearance of the UBA, has remained constant since 2005. Note that other groups, such as the Taliban, also have a large presence within the Balochistan province.

correspondence with Baloch journalist Malik Siraj Akbar revealed that the BLA recruited from non-Marri tribes starting from 2006 onward. Some members did not agree with recruiting people that are outside their tribe. Mehran Marri supposedly stole weapons and money to form his group—the UBA. KBM, however, asked the BLA leader to pardon his younger brother's theft and uprising. Thus the UBA initially operated as a faction within the BLA starting in 2011 (Ali, 2015; Nabeel, 2017; Balochistan Post, 2018).

The actual split of the BLA into two distinct groups occurred after the death of the brothers' father in June of 2014 due to a brain hemorrhage (Khan, 2014). Such cerebral bleeding occurs suddenly, and the most frequent reason for such bleeding types is high blood pressure. He was admitted to the hospital, and physical damage to his head is unlikely to go unnoticed and under-reported, given his popularity. This is not to say that alienation between the two factions could not have already been progressing before his death. However, the first recorded clash between the two factions/groups occurred five months after the death of KBM (see START, 2019; Sundberg and Melander, 2013). What is more, individual UBA incidents started being recorded around that time and concentrate heavily in the former area of operations of the BLA. We discuss the geographical overlap in more detail and how we leverage it for identification in Section 4 below.

In summary, the timing of the actual split between the BLA and UBA is not likely to be driven by the competition of the already established groups nor by some external factors influencing political violence within Balochistan. As such, we are confident that the group split provides exogenous variation in the number of armed groups operating within Balochistan.

3. Data

The units of observation are the districts of Pakistan between 1995 and 2018.[12] Pakistan's districts correspond to the third administrative layer (first-tier of local government). The main variables of interest are the level of organized political violence, and the number of active armed groups correcting for group mergers, group splits, and naming conventions (e.g., "Al-Qa'ida" vs. "Al-Qaida").

[12]We require a balanced panel for most of our estimations which prohibits using the GTD prior to 1993 as this year is missing in the dataset (see https://www.start.umd.edu/gtd/about/). Moreover, 1994 is lost due to the differencing of some variables. Our approach needs an uninterrupted time-series. 2018 is the final year in our sample because the extensive data work was conducted in the spring and fall of 2019 using a team of several RAs.

3.1. Dependent variable: Organized political violence

Our dependent variable is organized political violence. We take the number of incidents committed by armed groups to measure the frequency of organized political violence and the number of casualties (sum of people wounded and killed) to proxy for the severity of political violence. Note that we do not explicitly focus on the extensive margin of violence because the detection of any group requires at least one incident in a location.

The main data source is the "Global Terrorism Database" (GTD) (START, 2019), complemented by information from the "UCDP Georeferenced Event Dataset" (GED) (Sundberg and Melander, 2013). The GTD, officially tracking terrorism, is our preferred source due to two reasons.[13] First, since our armed groups of interest are classified as terrorist organizations, the coverage of incidents in which they have been involved turns out to be most comprehensively tracked by the GTD. The GTD codes more than 500 incidents committed by either the BLA or UBA, while alternative open source databases such as the GED or the "Armed Conflict Location & Event Data Project" (ACLED) (Raleigh et al., 2010), contain far fewer incidents (333 and 90, respectively) in which one of the two groups is involved.[14] Second, the GTD does not have a fatality threshold to include incidents – as is the case for the GED – or has known geographic biases in the recording of incidents – as has been shown for ACLED (Eck, 2012).[15] Note that we can only use incidents from the GTD and GED, which contain information on the district where they occur. This results in a loss of 95 incidents in the GTD and 180 incidents in the GED, leaving us with 14,063 and 5,611 incidents in the respective database.

Counting casualties deserves some special consideration. First, casualties in the databases are recorded with considerable uncertainty. Incidences are always reported if there is newspaper coverage. On the contrary, fatalities and people wounded may not be stated if the source is too vague or may not state how many people died during an incident. Most notably, the most recent source is used for the fatality and wounded estimate. If several newspapers report fatalities and wounded for an incident, the modal figure will be included in the database. Second, the number of fatalities and wounded is subject to a larger degree of randomness. While armed groups may conduct their attacks

[13]The GTD defines a terrorist attack as: "the threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation."(START, 2019)

[14]Since most events are purely domestic, the ITERATE database is not applicable.

[15]The GTD is, however, likely to suffer from general reporting biases as is common to all open source databases relying on news reports to track organized political violence (Van der Windt and Humphreys, 2016). While this reporting biases could be related to some district level characteristics that also attract more groups (cities vs. isolated rural areas), our setting is unlikely to be affected by them. We employ district- and time fixed effects as well as district-group and district-year fixed effects in our group-district level analysis. These fixed effects should already purge much of the potential bias. Moreover, our identifying variation comes from relative changes in the amount of political violence committed in treated vs. untreated districts over time. To the best of our knowledge, there is no evidence that differential reporting changes between the treatment and control group over time and is thus unlikely to bias our results.

with certain expectations with regard to how "big" an attack should be, there are a couple of factors that contribute to the actual number of deaths. In the case of a specific assassination, collateral damage may be acceptable depending on how reliant the group is on public support by the affected civilians (as in Toft and Zhukov, 2015). Moreover, the perpetrators are included in the death toll. For example, a suicide attack resulting solely in the perpetrators' death is coded as a fatal attack. Even though casualty rates are difficult to predict, they are informative of the group's intention and capabilities.

A downside of the GTD database is its' focus on terror attacks. Although the applied definition of terrorism is rather broad, it is not clear if a "proper" battle between an armed group and the Pakistani government on a "clearly defined" battlefield would be coded. It should not—as this constitutes symmetric warfare. Furthermore, the GTD does not code counter-insurgency operations by the government. An example would be an airstrike in northwest Pakistan, killing 20 militants by the Pakistani government reported on the 28th of June 2015, which is included in the GED but not the GTD. To answer our research question, we need to capture these types of events as well. Thus we supplement the GTD data with data from the GED. Specifically, we complement it with GED data on internal armed conflict and one-sided violence against civilians.[16] Using both databases also allows us to test if our results are driven by database-specific coding criteria.

Employing two databases that track organized political violence comes at a cost. The risk of double-counting incidents introduces potential measurement errors. Double-counting arises if both the GED and GTD code the same incidents for the same groups. We propose to address this issue by assigning an uncertainty measure for double-counting to each incident in the GED dataset. Specifically, we introduce several temporal and spatial buffers around each incident in the GTD database and flag GED incidents that fall within the buffer. Thus, the reader may decide with which buffer she is comfortable. The only assumption necessary for this approach to work is that double-counting is only an issue between databases but not within them.

3.2. Independent variable: Number of armed groups

Our primary independent variable of interest is the number of active armed groups. We consider all actors in the GTD and GED as armed groups if they have an individual name. That means we exclude actors such as "gunmen" or "tribesmen".[17] After independently cleaning the data, we compare our groups with the groups reported in Hou et al. (2020)

[16]The GED defines an event as: "an incident where armed force was by an organized actor against another organized actor, or against civilians, resulting in at least 1 direct death at a specific location and a specific date" (Stina, 2019, p.4).

[17]We also exclude so-called "one-hit wonders" (Blomberg et al., 2010), which are groups that only commit a single attack. We test for the sensitivity of our results to including them in the robustness section. A complete list of all armed groups is provided in Table D.

and find no omissions. We define any group as "active" within a district if it commits at least one attack during the year in that district. The number of active armed groups is then just the count of those groups.

On average, there are roughly 0.4 groups active within a district in a year during our study period. Only 15% of district-year observations host a positive number of active groups. That is not to say that most districts never experience group activity. Only 25% of 141 districts in our sample do not experience any activity during the sample period.[18]

Counting groups only as active in a district if they commit an attack during a year is by no means the only way how to think about group presence. For one, it ignores the strategic choice of locality (Marineau et al., 2020). Hence, we employ alternative measures of the number of active groups, such as the potential number of active groups. That is, we set existing groups as potentially active in all districts in which they have ever been active in any year if they are active somewhere in Pakistan in a given year. Groups that cease to exist cannot be potentially active in a district. The idea behind the potential active group measure is that a group reveals the set of districts in which it competes to us only over time while other groups are already aware of them. Furthermore, we are ambivalent about the exact locational choice in a specific year that might be driven by operational or strategic concerns that we cannot observe.

Other issues when counting the number of independent armed groups are splits and mergers of armed groups and related measurement errors within our source databases. The GTD and GED do not track the split and mergers of different armed groups but assign the perpetrator or conflict party of a given incident based on who claimed involvement in an incident or a third party that attests to the identity of the included actors. Hence there is the potential to attribute an incident to a group called "X-A", which is simply a faction of "A", but might later become an independent group. Much like in the case of the BLA and UBA. Note that both the GTD and GED change past entries in their databases if they receive new information, and it is not clear if they also backward correct specific names. However, given that our estimation sample only runs until 2018, this specific problem should be minimized, assuming that most corrections occur within the first two years rather than later on.

To address the issue of potential splits and mergers, we conduct an in-depth analysis of all armed groups within Pakistan and track if they split from or merged with other groups during our sample period.[19] The analysis is based on full-text online searches of major media outlets.

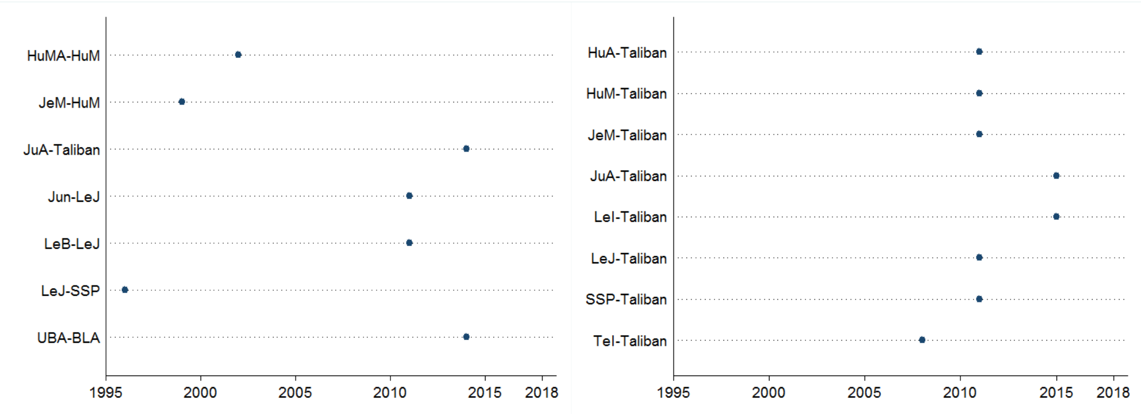
Figure I provides an overview of the timing of all splits and mergers occurring

[18] Figure A-1 reports the active group distribution for districts, as well as the distribution for districts in which the BLA has been active (or not active) prior to treatment separately. The distribution of the number of armed groups is skewed slightly more to the right for districts in which the BLA has been active prior to treatment compared to those in which it has not been active.

[19] Conducted during the first three quarters of 2019.

Figure I
Armed groups splits and mergers

(A) Group splits (B) Group mergers



Notes: Reports the year in which groups split (panel A) or merge (panel B): Baloch Liberation Army (BLA), Harakat ul-Mujahidin Al-Almi (HuMA), Harakat ul-Mujahidin (HuM), Jaish-e-Mohammad (JeM), Jamaat-ul-Ahrar(JuA), Jundullah (Jun), Lashkar-e-Balochistan (LeB), Lashkar-e-Islam (LeI), Lashkar-e-Jhangvi (LeJ), Sipah-e-Sahaba/Pakistan (SSP), Tehrik-e-Islami (TeI), United Baloch Army (UBA)

in our sample. Apart from several splits outside of Balochistan, we observe a major consolidation of the Taliban which absorb several groups between 2011 and 2015,[20] Using the information in Figure I we can reassign incidents and casualties to the corresponding pre-merger or post-split groups and adjust the number of groups for each district, to reflect splits and mergers correctly. Note that we will not use the other splits or mergers to identify the competition effect since we cannot rule out that the timing of the mergers and splits are endogenous to the conflict dynamics within Pakistan. However, neglecting the other group splits and mergers would result in the measurement error of our independent variable. Full descriptive statistics for our variables of interest are reported in Table A-1.

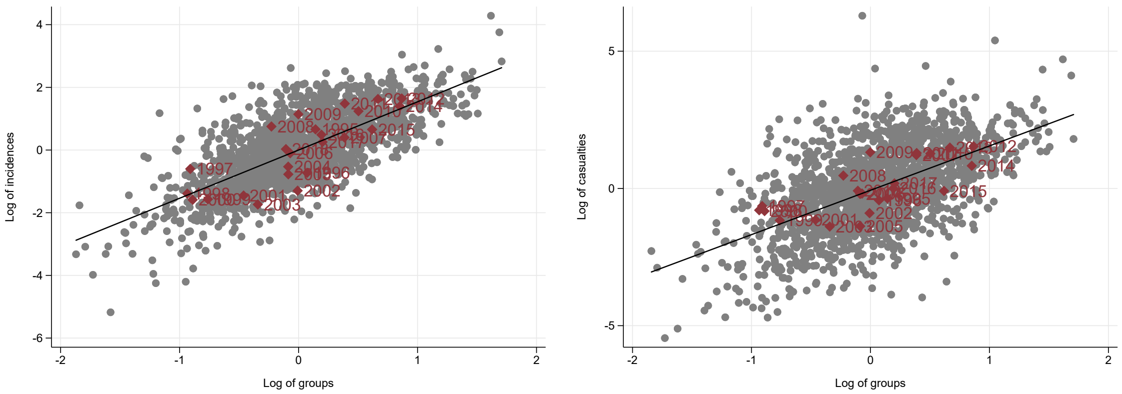
How unique is Pakistan as a case study for our proposed mechanism? To get an initial idea, we plot the elasticity between aggregated incidents and casualties on the number of active armed groups at the country-year level for all countries included in the GTD between 1995 and 2018. Figure II shows the results, highlighting Pakistan-Year observations in dark red. All observations are demeaned by country and year.

Figure II points to a positive net effect, i.e., a dominance of the competition effect. First, there is an apparent correlation between the number of armed groups active within a country and the number of organized political violence perpetrated. Second, Pakistan is no outlier but fits the linear prediction quite well. Of course, this is only suggestive evidence on the country level, but it is supportive of the notion that the proliferation of armed groups leads to more political violence.

[20]Table D-2 and Table D-3 in Appendix D provides detailed documentation of each case.

Figure II
Armed groups and political violence

(A) Incidents-Group-Elasticity (B) Casualties-Group-Elasticity



Notes: Depicts a scatter plot of the (log of) groups vs. (log of) incidents & casualties created by these groups, demeaned by country and year. The unit of observation is country-year. Pakistan is represented in dark red. The black line illustrates the best linear fit using the global GTD sample between 1995 and 2018.

4. Empirical Strategy

In the spirit of Draca et al. (2011), we will use two complementary identification strategies to test if group proliferation increases organized political violence. First, we run event study estimations in which we regress political violence on a set of binned treatment indicators. The main goal of the event studies is to understand the reduced form effect of the BLA split on political violence within Pakistan. Second, we use the DiD version of the reduced form as a shift-share instrument for the number of armed groups within districts in 2SLS regressions. The goal of the 2SLS specifications is to estimate the semi-elasticity of an additional armed group on political violence, which is the causal effect we are after.

The reduced form specification is a standard event study with an effect window running from $\underline{s}$ to $\overline{s}$ for all $t = \underline{t}, \ldots, \overline{t}$

$$Y_{it} = \sum_{s=\underline{s}}^{\overline{s}} \beta_s b_{it}^s + OG_{it} + \mathbf{X}_{\mathbf{it}}^{'} \psi + \eta_i + \gamma_t + \xi_{it} + \epsilon_{it} \tag{1}$$

where Y_{it} is the log of political violence (either incidents+1 or casualties+1) perpetrated in district i during year t. b_{it}^s are treatment change indicators binned at the endpoints $\underline{s} = -4 \; \forall \; t \leq -4$ and $\overline{s} = 3 \; \forall \; t \geq 3$, with $s = 0$ representing the treatment year 2014.[21] Specifically, each b_{it}^s corresponds to the interaction $BLA\ share_i \times BLA\ split_t$. $BLA\ share_i$ is the share of years in which the BLA has been active in the district prior to treatment. $BLA\ split_t$ is a variable taking on the value one for the years 2014 and

[21] Corresponding to years 2010 and before or 2017 and later, respectively.

later and zero otherwise. OG_{it} is the number of other active groups present within the district, which we discuss momentarily. $\mathbf{X}'$ is a vector of control variables we use to control for potentially unobserved confounders between the control and treatment districts over time. We include the log of the population to normalize the count of incidents relative to the local population and proxies for conflict suitability that are plausibly exogenous, such as the difference in average rainfall and temperature within districts (Fearon and Laitin, 2003; Buhaug et al., 2009).[22] η_i and γ_t are district and year fixed effects, ξ_{it} are division-specific-linear-trends that capture trends in the upper layer administrative units.[23] As is common in event studies, we omit b_{it}^{-1}. Hence, all effects have to be interpreted relative to this baseline.

The intuition behind our reduced form event study is that districts in which the BLA has been more active are more affected by the split, i.e., the districts are more likely to have a UBA presence compared to those districts with less BLA activity. Limodio (2022), for example, provides empirical evidence in line with the idea that terrorist groups in Pakistan face internal frictions in their capital and labor markets, i.e., groups are more active in locations in which they have more personnel and capital. After the BLA split, it is reasonable that districts with a larger presence prior to treatment are more likely to host both groups post-treatment, all else being equal.

Figure III illustrates the point. Panel (A) shows that there is a large overlap between areas in which the BLA and UBA operate, primarily within the Baloch province, which is highlighted by the green border. 95% of all incidents of the BLA and UBA are committed within the Baloch province. Panel (B) of Figure III highlights further that the districts in which the BLA and UBA overlap are those in which the BLA has already been more active prior to treatment. Panel (C) shows that the number of active groups in the post-BLA split period rises more often in districts with a high BLA presence allowing us to run instrumental variable specifications. Moreover, Figure A-8 in Section A-1 highlights that the UBA reduces its area of operation over time to the areas in which the BLA has the highest pre-treatment presence. The second stage 2SLS specification is defined as:

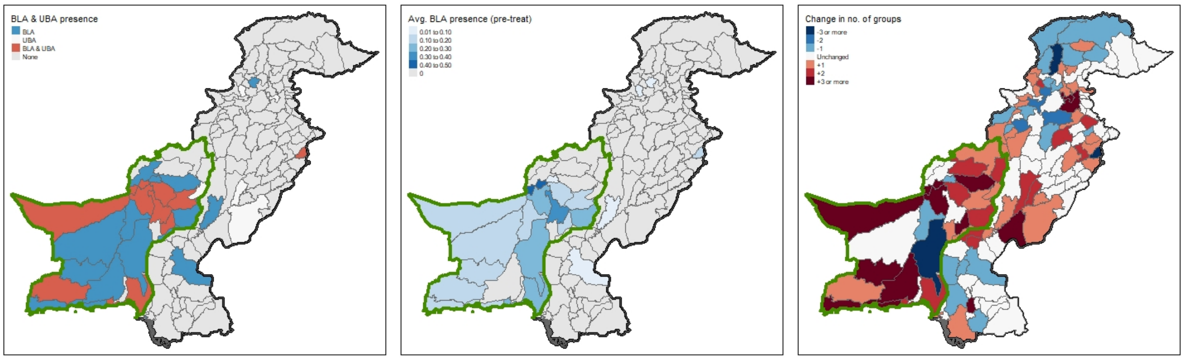
$$Y_{it} = \delta \widehat{AG_{it}} + \mathbf{X}'_{\mathbf{it}}\psi + \eta_i + \gamma_t + \xi_{it} + \epsilon_{it} \tag{2}$$

[22]The log of population density is calculated based on the GWP (CIESIN, 2018). Note that the GWP is only provided every five years and only provides detailed spatial population estimates for the reference years 1990, 1995, 2000, 2010, and 2015. We linearly interpolate and extrapolate the population data between those reference years and 2018, the last year of our sample. The rainfall and temperature differences are calculated using information from temperature and rainfall rasters provided by Hersbach et al. (2018). We aggregate the 0.25-degree raster information to the district level, take the yearly means and then take the difference. The rainfall measure is scaled by a factor of 1,000.

[23]Divisions are the second subnational administrative layer of Pakistan hosting on average about 4.5 districts. We cannot use district-specific trends due to degrees of freedom constraints.

Figure III
Identifying variation

(A) Presence BLA & UBA (B) Avg. BLA presence (C) Δ no. active groups



Notes: Panel (A) plots the districts in which the BLA & UBA have both been present at any point in our sample in red, those in which only the BLA has been present in blue, and those in which only the UBA has been present in light grey. Panel (B) plots the avg. presence of the BLA prior to treatment, i.e., the fraction of years in which the BLA has committed at least one attack in a district prior to treatment. The highest presence is observed in Quetta, the capital district of the Balochistan province. Panel (C) plots the change in the number of armed groups operating on average in a district in the pre- vs. post-treatment period. The Balochistan province is highlighted by the bold green borders.

with the corresponding first stage:

$$AG_{it} = \beta IV_{it} + \mathbf{X'_{it}}\psi + \eta_i + \gamma_t + \xi_{it} + \epsilon_{it} \tag{3}$$

where AG_{it} is the number of active groups within a district-year (including BLA and UBA) and the instrument IV_{it} is $BLA\ share_i \times BLA\ split_t$, all else is defined as before.

Before we turn to our core results, let us briefly discuss the identifying assumptions of our two approaches. Our event-study design relies on the standard assumption that unobserved time-varying confounders affect districts that are more or less treated similarly, i.e., with respect to $BLA\ share_i \times BLA\ split_t$. This is the standard parallel trends assumption in the presence of heterogeneous treatment effects. Stated differently, there should be no anticipation effect of KBMs death (and the BLA split) depending on the level of pre-treatment BLA activity within districts. As outlined above, KBM died in a hospital from a brain hemorrhage. Hence it seems implausible that districts with a higher BLA presence anticipate his death more precisely compared to districts with a lower average presence.[24] The treatment heterogeneity caused by the variation in $BLA\ share_i$ is another matter. Potentially the locational fundamentals with respect to political violence, such as state capacity or the demand for armed groups, change differently in districts in which the BLA has traditionally been active over time. KBM himself could have had some impact on the locations, apart from mitigating tensions

[24]His old age was public knowledge and not limited to members of the BLA.

between his two sons and keeping the BLA together. We tackle these issues below by investigating the pre-treatment coefficients in the event study and explicitly testing for potential confounders of the kind just mentioned.

In the 2SLS case, we require the usual instrumental variable assumptions of excludability and relevance. Relevance (or power) is not a concern, as we show below. Excludability, in turn, needs to be argued for. There are many potential ways in which KBM's death (and the BLA split) could have affected political violence differently in the respective treatment and control groups apart from an increase in the number of armed groups, i.e., by altering the demand for armed groups. We use our reduced-form specification and further extensions to the 2SLS models to alleviate concerns with respect to obvious violations of exclusion restriction in Section 6.

5. Results

5.1. Reduced from evidence

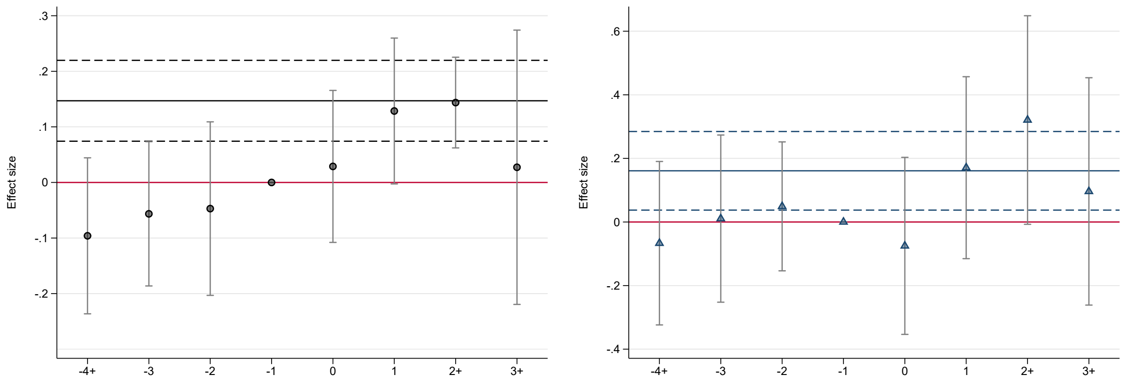
The main results of our event study are depicted in Figure IV. Panel (A) plots the results for the incidents specification, panel (B) for the casualties specification. Neither specification exhibits any pre-trends, as can be observed from the statistically insignificant coefficients prior to treatment. This is also the case if we extend the pre-treatment event sequence (see Figure A-2). Following the split of the BLA, we observe an increase in the incidents of political violence in districts with a comparably higher pre-treatment BLA presence, stating in $t+1$ which rises in $t+2$ and reverts to the baseline for all periods $t+3$ or later. The casualty effect, in turn, is limited to t+2. Note that the $BLA\ share_i$ variable is standardized. Hence all coefficients can be read as the differential effect following the BLA split for districts with a one standard deviation higher pre-treatment BLA presence compared to others.

The obtained reduced form effects are sizeable. We estimate that violence increases by roughly 15% in the first two years following treatment. Comparing a district with an average BLA presence prior to treatment compared to districts in which the BLA has not been active results in a 30% increase of incidents in the first two years following treatment. The casualty estimate is larger, although much less precisely estimated. We estimate an increase in casualties of about 35% in the second year following treatment. Comparing a district with an average pre-treatment BLA presence to a district with no prior BLA activity results in an estimated increase of casualties of 72%.

In both panels, we also report the generalized DiD estimates for our reduced form specification in which we predict the differential change across the entire post-treatment period ($BLA\ share_i \times BLA\ split_t$). For incidents, the point estimate of the generalized DiD is close to identical to the event study estimates in $t+1$ and $t+2$. For casualties,

Figure IV
Reduced form evidence: Main results

(A) Ln incidents (B) Ln casualties



Notes: The figure reports our event study point estimates and their 95% confidence intervals regressing the log of incidents + 1 (panel A) or the log of casualties + 1 (panel B). Coefficients are calculated based on standardized variables. The CIs are based on standard errors clustered at the district level. The horizontal lines report the corresponding DiD estimate (solid) and its CI (dashed).

the effect is about 50% smaller compared to the $t+2$ event study estimates. However, in neither case are we able to reject that the separate post-treatment estimates are identical to the DiD estimate.

Our incident results remain qualitatively similar if we refrain from the log transformation of our incidents and estimate the event studies using Poisson pseudo maximum likelihood estimator, employ an inverse hyperbolic sine transformation of the dependent variables, or add a smaller constant before taking the logs (see panels (A) to (C) in Figure A-3).[25] The casualties estimates are slightly off, but those are in general more volatile. We can also refrain from using control variables or limit our sample to districts within the Balochistan province, whose independence is the official goal of the BLA (see panel (D) and (E) in Figure A-3). The estimated effect sizes are nearly identical, despite reducing the sample to 20% of its original size. We can also calculate our pre-treatment BLA presence only based on years prior to 2011 when the UBA faction formed within the BLA. Again, results remain qualitatively the same and do not suggest that the UBA faction already influenced the area of operations of the BLA (see panel (F) in Figure A-3).[26] Summing up we observe qualitatively similar results in all of the different specifications. Crucially, the absence of observable pre-treatment trends makes us confident that we can proceed under the common trends assumption and employ the DiD version of the reduced-form as an instrument for the number of active armed groups in our 2SLS specification.

[25] As in Dube and Vargas (2013) and Limodio (2022)

[26] We do not have information on which faction within the BLA carried out an attack before the official split in 2014.

5.2. Group competition and political violence

We now turn to estimating the relationship between the number of armed groups and political violence. Before turning to the 2SLS specification, we run a simple OLS regression of the log of political violence on the number of active groups within districts.

TABLE I
Competition and political evidence

	OLS		*2SLS*			
	Dependent variable:					
	Ln incidents	Ln casualties	Ln incident	Ln casualties	Ln incident	Ln casualties
	(1)	(2)	(3)	(4)	(5)	(6)
No. active groups	0.4490	0.7945	0.8431	1.1767	0.5356	0.6997
	(0.0444)	(0.1024)	(0.1712)	(0.2767)	(0.1277)	(0.1777)
			1st stages			
			Dependent variable: No. active groups			
$BLA\ share_i \times BLA\ split_{post}$	–		0.4559		–	
	–		(0.0868)		–	
$BLA\ share_i \times BLA\ split_{2014}$	–		–		0.7526	
	–		–		(0.1429)	
$BLA\ share_i \times BLA\ split_{2015}$	–		–		0.4662	
	–		–		(0.0991)	
$BLA\ share_i \times BLA\ split_{2016}$	–		–		0.3105	
	–		–		(0.0857)	
$BLA\ share_i \times BLA\ split_{2017+}$	–		–		-0.0276	
	–		–		(0.1110)	
No. act grps in **all** districts:			*Mean: 0.4078*			
			SD: 1.0480			
No. act grps in **act** districts:			*Mean: 1.7468*			
			SD: 1.5384			
Controls	✓	✓	✓	✓	✓	✓
District-FE	✓	✓	✓	✓	✓	✓
Year-FE	✓	✓	✓	✓	✓	✓
Division-trend	✓	✓	✓	✓	✓	✓
Adj. R^2	0.730	0.634	0.263	0.332	0.495	0.387
Obs	3384	3384	3384	3384	3384	3384
F-stat IV (1st stage)	–	–	11.44	11.44	16.53	16.53

Notes: The table reports the results of regressing the log of incidents + 1 and the log of casualties on the number of active groups. Columns 1 and 2 use OLS estimates. Columns 3 to 6 report the first and second stage results based on 2SLS specification stated in eq. 2. Columns 3 and 4 use the interaction $BLA\ share_i \times BLA\ split_{post}$ as the instrument for the number of active groups operating within a district. Columns 5 and 6 use a dynamic version of $BLA\ share_i \times BLA\ split_t$ in which each post-treatment period (2014-2017/18) is allowed to have a different effect on the number of groups. Standard errors are clustered at the district level in parenthesis.

Columns 1 and 2 of Table I report the results. The estimated semi-elasticity of the number of active groups on incidents of political violence is 0.449, meaning that an additional group is expected to increase the frequency of political violence by about 58%. The corresponding casualty semi-elasticity is 81%. On average, a district hosts about 0.4 groups in a year. Hence, an increase by one means a percentage increase of just above 150%, implying an elasticity for the average district of political violence of between 0.39 and 0.54 (columns 1 and 2, respectively). Taken at face value, the severity of violence increases more compared to the incidents. This is in line with the idea that groups compete with one another for public attention to garner recruits and financial contributions (Jetter, 2017). However, the number of groups active in a location is most likely endogenous to the local conflict dynamics. Columns 3 to 6 present the second stage results from our 2SLS specification. In columns 3 and 4 we use the generalized DiD estimate of $BLA\ share_i \times BLA\ split_t$ as our instrument. In columns 5 and 6, we use the set of post-treatment indicators from the event study $\sum_{s=0}^{s=3} \beta_s b_{it}^s$, i.e., we allow the interaction $BLA\ share_i \times BLA\ split_t$ to have a dynamic effect on the number of groups within districts over time. Regardless of the IV choice, the first stage F-stat suggests that our IVs have enough power.

On average, we observe an increase in the number of active armed groups in treatment compared to control districts of about 0.5. Yet, the estimated initial increase is higher (about 0.75 in columns 5 and 6 of Table I) and then falls over time before the estimate turns insignificant for years three or more after treatment. Note that we again standardize the pre-treatment BLA share. Comparing a district in which the BLA has not been active prior to treatment with the average presence of the BLA results in a DiD estimate of about 0.98. This corresponds, e.g., to the appearance of the independent UBA. Moreover, the general spatial distribution of group activity remains rather stable.[27]

The observed pattern in the first stages of Table I is consistent with our interpretation of the reduced form effects shown in Figure IV: They primarily capture the increase in the number of armed groups due to the appearance of the UBA. If the number of groups in treatment compared to control districts becomes indistinguishable in $t + 3$, we would not expect that treatment districts experience more political violence compared to the control districts.

The magnitude of 2SLS estimates for incidents are in general larger compared to the OLS estimate, although much less so for the dynamic IV. This is not the case for the more imprecise causality estimates. Focusing on the incidents, we estimate semi-elasticities of 85% and 64%.[28] If our argument has merit, then it makes sense that the 2SLS estimates are larger because the endogenous selection of groups into districts no longer biases the results. In other words, smaller OLS estimates are consistent with groups not selecting into districts where they cannot compete.

[27] Figure A-7 shows that the cross-sectional distribution of active groups across districts is relatively stable (pre- to post-treatment), at least with respect to the ordering,

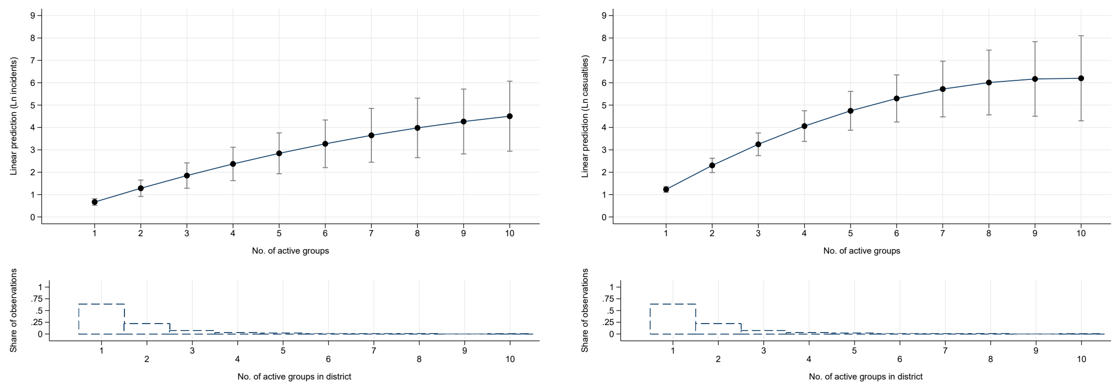
[28] We obtain similar patterns if we use the inverse hyperbolic sine transformation of the dependent variables or focus exclusively on districts within Balochistan (see Table A-3 and Table A-4).

So far, we have focused on the average effect of additional groups on political violence. Yet, if competition is the primary driver we identify, we would expect that the effect of additional groups depends on the number of other groups already competing within an area. This also relates to potential selection effects. Groups could either be deterred from mighty incumbents that do not tolerate competition or avoid areas in which competition is so high that they are unlikely to garner any support (the demand for armed groups is already saturated).

We test for the nonlinear effect of the number of armed groups on political violence and investigate the direction of selection using a control function approach. The control function approach has two advantages over 2SLS specifications in this setting. First, the first-stage residual shows if selection is likely to be significantly different from zero, as well as the direction of potential selection. Second, we only have to use the residual of the number of armed groups to control for the endogeneity of the baseline as well as the squared term, which makes the estimation more efficient (Wooldridge, 2010, 2015).

Figure V
Predicted political violence over number of active groups

(A) Incidents (B) Casualties



Notes: Panel (A) plots the predicted amount of political violence (Log incidents+1) by the number of active groups, based on column 5 of Table A-5 and the corresponding frequency distribution of these groups across districts with group activity. Panel (B) plots the predicted amount of political violence (Log casualties+1) over the number of armed groups, based on column 6 of Table A-5. Confidence intervals are 95% CI based on standard errors clustered at the district level.

Figure V plots the predicted amount of political violence over the number of armed groups, based on our preferred nonlinear control function specifications (columns 5 & 6 of Table A-5).[29] Both panels (A) and (B) suggest that the increase in violence, an additional group causes diminishes in the number of armed groups. This result is consistent with the idea that the political benefit of a successful attack diminishes in the number of attacks

[29]Note that we bin the number of active armed groups at 10 or more because empirical support is missing for some of the higher numbers going up to 15. Table A-5 replicates Table I with control function methods. In addition, it includes the squared term for the number of active armed groups.

conducted by other groups, which at some point will be below the costs of conducting an attack. Relatedly, the negative point coefficients of the first-stage residuals highlight that OLS specifications underestimate the effect of an additional armed group on political violence. This suggests that the marginal group selects itself into areas in which many other groups already operate.

Indeed we can confirm a similar pattern with the UBA. Figure A-8 in the Appendix documents that while the UBA initially operates within several districts in which the BLA has traditionally been active, it concentrates its activity over time in the districts around the provincial capital Quetta. Quetta and its surrounding districts, in turn, are among those districts with the highest number of active groups within our sample (see Figure A-7). We do not observe a similar trend for the much larger BLA, which keeps a relatively constant area of operations in the years following treatment.[30]

6. Alternative channels and extensions

What drives this increase in political violence? We argue that our reduced-form estimates capture the plausibly exogenous increase in the number of active armed groups with respect to local conflict dynamics, which increases organized political violence. Given that we control for the number of other armed groups present in districts, this seems plausible.[31] It is also in line with the 2SLS results which we reported above. However, our reduced-form estimates could also capture potential other differential changes in the conflict dynamic between the treated and control districts over time, which would violate the exclusion restriction in the 2SLS models.

In this section, we further scrutinize how our treatment affects competition between armed groups. We explore if the type of organized political violence changes through the treatment, specifically if our results are driven by increased violence primarily between groups (Section 6.1). Furthermore, we show that local determinants of political violence at the district level – such as government capacity, the politically excluded population, and financing possibilities for armed groups – do not change differently between the control and treatment districts over time (Section 6.2). Thus, they are unlikely to explain our effects. Finally, we provide evidence that the BLA indeed conducts more attacks in districts in which other groups are active as well. Moreover, we show that the BLA split did not cause the BLA to lose its relevance in the local conflict dynamics (Section 6.3). In fact, the BLA compensates for the negative capacity shock of the split by switching to non-capital intensive attacks, which is in line with the theoretical predictions of Bueno de

[30] The UBA commits roughly 60 incidents in the post-treatment period while the BLA conducts more than 200.

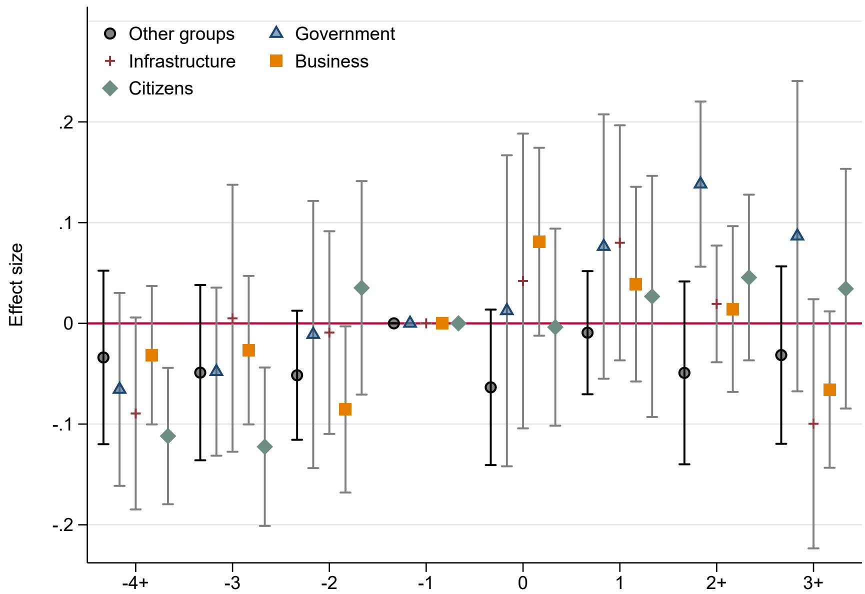
[31] The general size of our effects is not sensitive to dropping the other active group control or all controls (see Figure A-4). The effects become only smaller if we start to include district times decade fixed effects on top of our current fixed effects, trends, and controls.

Mesquita (2013).

6.1. Targets of armed groups within districts

Does increased infighting drive our results, i.e., are armed groups attacking each other? If KBM was a unifying figure, he might have stopped different groups from attacking each other (such as his sons). Our data allows us to test this alternative explanation directly. The GTD list the target type of incidents, e.g., "Terrorists/Non-State Militia" or "Violent Political Party" among others. We create an alternative incident count using only incidents that target either of those categories and rerun our reduced form event study. In addition, we test whether groups changed their target selection post-treatment. The specific incident categories we employ as dependent variables are attacks against; *i*) other armed groups, *ii*) the government, *iii*) public infrastructure, *iv*) private business, and *v*) private citizens.[32]

FIGURE VI
Group targets: Ln incidents



Notes: Reports the coefficient and their accompanying 95% CIs of our event study specification as stated in eq. 1 for different incidents counts (see Table A-2). We add 1 to all incident counts and take the log of them in all specifications. The CIs are based on standard errors clustered at the district level.

Figure VI reports the event study estimates for the different incident measures. It shows that our treatment does not affect infighting in the treatment vs. the control group differently (black dots). Hence, KBM's death is unlikely to have caused increased infighting between groups. The remaining point coefficients in Figure VI suggest that

[32] Table A-2 provides the specific definitions for each of the measures.

violence primarily increases against government targets. However, using only a subset of the incidents reduces the precision of our estimates.

6.2. District determinants of political violence

State capacity increases or decreases could change differently between the treatment and control districts, which could explain why we observe more violence and more groups in treatment compared to control districts (Fearon and Laitin, 2003). We proxy for state capacity using the counter-insurgency effort of the government. Note that the effect of government counter-insurgency on political violence is theoretically ambiguous, as it can increase as well as decrease mobilization (Bueno De Mesquita, 2005).[33]

Obtaining a suitable proxy for counter-insurgency operations is not without problems. Recall that the GTD only codes terrorist events and hence misses counter-insurgency operations, such as the airstrike mentioned in Section 3.1. The GED, on the other hand, codes event dyads, but those are not directional. That is, there is no indicator variable indicating whether the government or an armed group initiated an incident. We circumvent the issue and classify incidents between the government and armed groups as counter-insurgency incidents if the incident is reported in the GED but not in the GTD. The assumption is that if we subtract the incidents between the government and any armed group included in the GTD and thus identified as a terrorist activity by the GTD, the events left can be used as reasonable proxies of operations instigated by the government. The main operational obstacle is dealing with measurement uncertainty between the two databases. We tackle this issue with our proposed double-counting procedures, which we explain in detail in Appendix B. In short, we draw a buffer of 25km around each GED event and flag it as a potential double count if the GTD codes an event of the same armed group during the same day.[34] Events that are flagged as potential double counts are excluded from the analysis, which leaves us with a set of incidents that will use as our counter-insurgency proxy (roughly 47% of all incidents in the GED in which the government is involved).

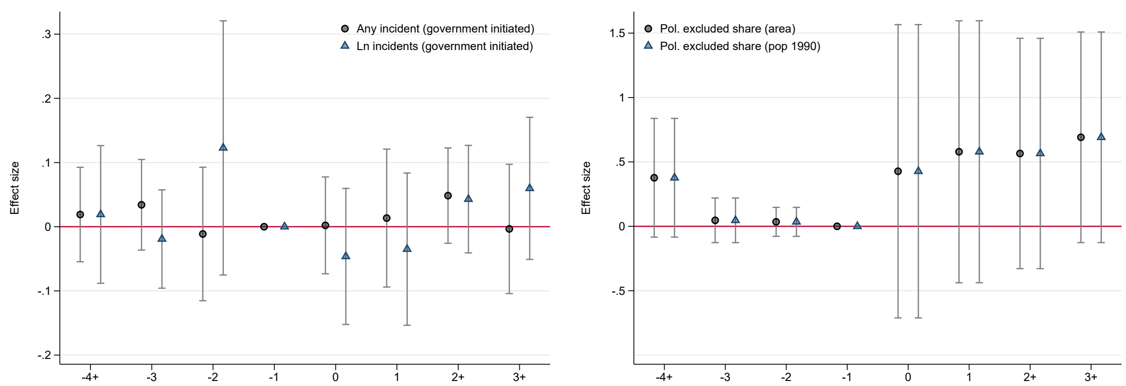
Panel (A) of Figure VII reports our event study estimates for counter-insurgency efforts by the government. We do not observe any significant effect on the likelihood that the government initiates any counter-insurgency effort, nor is the intensity of counter-insurgency, proxied by the log of counter-insurgency incidents + 1 (blue triangles), affected. Hence, we cannot reject the null hypothesis that state capacity has evolved similarly between treatment and control districts over time.

[33]At least for intermediate values of state capacity, for which groups are not deterred from forming in the first place.

[34]We use only events for which the geographic precision provided by the GED is 1 to 25km for this exercise.

Figure VII
District determinants of political violence

(A) Counter-insurgency (B) Political excluded population



Notes: Panel (A) of the figure reports our event study (as specified in eq. 1) coefficients of interest and the accompanying 95% CI based on event studies regressing the probability of a counter-insurgency incident and the log counter-insurgency incidents + 1. Panel (B) uses the share of the politically excluded population (based on areas or the 1990 population) as the dependent variable in our event study. The CI are based on standard errors clustered at the district level.

Demand for armed groups is another explanation for our findings. If demand for armed groups that challenge the government increases in an area (or district), it is more likely to observe more groups operating in this area. Relatedly, demand should be correlated with the willingness of people to either join an armed group in the area or support it otherwise. We proxy for the local demand of armed groups by calculating the share of the politically excluded population within districts over time. Hence, we assume that when the share of politically excluded people in a location increases, demand and hence potential support for armed groups is likely to go up (Bormann et al., 2019).

Our "demand" proxy is based on the geocoded version of the Ethnic Power Relations (geoEPR) data (Wucherpfennig et al., 2011; Vogt et al., 2015). The geoEPR dataset provides polygons and time-varying political power status information for politically relevant ethnic groups worldwide. Figure D-1 plots the respective groups for Pakistan. To reassign the political power of different ethnic groups to districts, we weigh the political status of groups either by their homeland area share in the district or by a proxy for their 1990 population share.[35] The politically excluded population (the "demand" proxy) is the share of people classified as "discriminated against" or "powerless". Details and descriptive statistics are provided in Appendix D.

Panel (B) of Figure VII plots the results of using either the area or population-based measures for the locally politically excluded population as dependent variables in our event study. Once more, we do not find evidence in favor of a diverging trend between

[35] The 1990 population share is the share of the population within the districts that reside in the EPR homeland, based on the GHSL population grid.

our treatment and control group.

Group financing opportunities can also vary across locations and time, thus potentially explaining our results. In his seminal paper, Limodio (2022) provides evidence in line with the idea that terrorist groups in Pakistan face frictions both in their internal capital and labor markets. Specifically, he shows that increases in local financing opportunities increase local attacks. To proxy for local financing opportunities, we use an annual district level equivalent of the identification strategy employed by Limodio (2022). In short, we exploit that the threshold for the mandatory levy (Zakat donations) for Sunni before Ramadan is dependent on the silver price, which leads to a differential impact in donations between majority Sunni and other districts. This, in turn, affects the financing opportunities for armed groups primarily composed of Sunni in majority Sunni districts more than other groups within those districts and elsewhere. Note that our setting only exploits changes in the average global silver price across years and not the price variation just before Ramadan (requiring within-year variation due to the moving dates of Ramadan over the years) which is exploited for causal identification in Limodios analysis. For details we refer to Limodio (2022).

TABLE II
Group financing and district determinants

	Dependent variable:			
	Ln incidents	Ln casualties	Ln incidents	Ln casualties
	(1)	(2)	(3)	(4)
No. active groups	0.5067	0.6376	0.5468	0.7029
	(0.1438)	(0.1925)	(0.1406)	(0.1983)
$Sunni\ share \times silverprice_t$	0.0049	0.0246	-0.0906	-0.1348
	(0.0707)	(0.1734)	(0.0641)	(0.1379)
Share politically excluded (pop)			0.0157	-0.0190
			(0.0273)	(0.0604)
Ln counter-insurgency			0.2323	0.5085
			(0.0515)	(0.0668)
Controls	✓	✓	✓	✓
District-FE	✓	✓	✓	✓
Year-FE	✓	✓	✓	✓
Division-trend	✓	✓	✓	✓
Adj. R^2	0.510	0.364	0.560	0.430
F-stat IV	14.34	14.34	10.13	10.13
Obs	2882	2882	2667	2667

Notes: The table replicates columns 5 and 6 of Table I controlling for the districts determinants of political violence. Standard errors are clustered at the district level in parenthesis.

Columns 1 & 2 of Table II replicate our preferred 2SLS specifications (columns 5

& 6 of Table I) controlling for the interaction between the Sunni share of a districts population and the log of the annual global silver price ($Sunni\ share_i \times silverprice_t$).[36] Thus, we capture some of the potentially different local financing opportunities that vary across districts over time. Our results remain virtually unchanged. Columns 3 and 4 add the other two proxies for the district determinants of political violence, although it is unclear if they are bad controls. Regardless, our coefficients are within a standard error distance from our baseline results. Note further that our results do not depend on the inclusion of any specific district (see Figure A-6).[37] In fact, our results are somewhat stronger if we drop the Quetta district, which suggests again that there are diminishing returns to competition in terms of violence. In summary, it seems unlikely that changes in local government capacity, the demand and potential support for armed groups, or local differences in financing opportunities explain our results.

6.3. Within group evidence:

We now turn our attention to how the BLA split has affected the BLA itself. Moreover, we want to understand if the relative increase of political violence in "BLA districts" is driven by the BLA itself, competition between the BLA and UBA, or by other groups operating within those districts. The two issues are interrelated. If the BLA experiences a negative capacity shock, e.g., due to a loss of manpower or equipment, other groups might try to challenge the BLA. In such a case, we would expect the BLA to commit less violence than other groups following treatment. On the flip side, the BLA might engage in even more violence to signal its continued importance to potential recruits and financial backers. Hence, the net effect is unclear, at least ex-ante. Moreover, the BLA could simply change the type of violence it commits, i.e., hitting softer targets (Bueno de Mesquita, 2013) or using less capital intensive attacks.[38] We investigate those scenarios, running a triple-difference specification on a BLA-within-district panel, to test how the BLA responds to other groups (Section 6.3.1). In addition, we specify event study specifications on the group-district-year level that allows us to test how the BLA and UBA behave compared to other groups (Section 6.3.2). In conjunction, the two sets of results are consistent with the idea that the BLA keeps its relevance and is most active in districts in which it faces competition.

[36] Data on the Sunni share has been provided to us by Limodio (2022). The global silver price is taken from https://www.metalary.com/.

[37] Furthermore, we obtain similar results using Conley standard errors with a spatial cutoff of up to 400km.

[38] Where capital can be either human or physical capital. Empirical evidence highlights the importance of both (e.g. Benmelech and Berrebi, 2007; Limodio, 2022).

6.3.1. Within BLA evidence

To test if our results are solely driven by competition between the BLA and UBA we run a within BLA triple-diff specification;

$$
\begin{aligned}
Y_{it} = & \beta_1(UBA_{it} \times Post_t) + \beta_2(Post_t \times OG_{it}) + \beta_3(Post_t \times OG_{it} \times UBA_{it}) \\
& + \beta_4 OG_{it} + \mathbf{X}_{\mathbf{ijt}}^{'}\psi + \eta_{\mathbf{i}} + \gamma_{\mathbf{t}} + \epsilon_{\mathbf{it}}
\end{aligned}
\tag{4}
$$

where Y_{it} are the log of BLA incidents (+1) in districts-years, UBA_{it} is an indicator that is unity if the UBA is active within a district in a year, OG_{it} is the count of groups (excluding the BLA and UBA) active within a district in a year. The coefficients of interest are β_1 to β_4. The idea of the specification is that we test if the BLA commits more attacks in districts in which if faces competition by the UBA (something that only occurs after the BLA split), compared to districts in which it faces other groups or is by itself. If the BLA only competes with the UBA, we would expect that only β_1 matters.

TABLE III
Within BLA evidence

	Dependent variables:					
	Ln incidents (all)	Ln casualties (all)	Ln incident (civilians)	Ln incident (capital intensive)	Ln incident (non-capital intensive)	
	(1)	(2)	(3)	(4)	(5)	
POST × UBA	-0.1181	0.0332	-0.1281	-0.0372	-0.1134	
	(0.1288)	(0.2290)	(0.0741)	(0.0765)	(0.0962)	
POST × OTHER GROUPS	0.0233	0.0371	-0.0135	0.0126	0.0100	3
	(0.0088)	(0.0151)	(0.0198)	(0.0088)	(0.0087)	
POST × OTHER GROUPS × UBA	0.2757	0.3193	0.1179	0.0657	0.3008	
	(0.1159)	(0.0993)	(0.0647)	(0.0688)	(0.0989)	
OTHER GROUPS	0.0513	0.0830	0.0532	0.0346	0.0235	
	(0.0244)	(0.0491)	(0.0339)	(0.0157)	(0.0138)	
District-FE	✓	✓	✓	✓	✓	
Year-FE	✓	✓	✓	✓	✓	
Adj. R^2	0.310	0.252	0.271	0.227	0.200	
Obs	3384	3384	3384	3384	3384	

Notes: The table reports a triple-diff analysis for the BLA only. We regress the log of incidents, casualties, and specific incident types +1 on a UBA presence indicator interacted with the post-treatment period, an interaction of the number of other groups (not including the BLA and UBA) with the post-treatment period, and the interaction of the two interactions. All columns include district and year fixed effects. Standard errors are clustered at the district level.

Table III provides the results of the specification across incidents, causalities, and different types of attacks (against civilians, capital intensive, and non-capital intensive).[39] Counter to the idea that competition is only driven by the BLA and UBA we observe that the coefficient of $(UBA_{it} \times Post_t)$ is negative and mostly statistically insignificant. In turn, the presence of other groups consistently predicts increased BLA activity (both before and after the split). Note that the marginal increase of UBA presence in districts in which also other groups are present has a substantial increase, which supports the idea that the BLA might be particularly sensitive to UBA presence in contested districts. However, it could also point to the fact that the UBA and BLA have higher capabilities in those districts and can react to the presence of other groups more strongly. In the next subsection, we address this issue by leveraging within district-group fixed effects.

6.3.2. Within group-district diff-in-diff

To test how the relative activity of the BLA compared to that of other groups after its negative capacity shock, we run event study specifications on the group-district-year level,

$$Y_{ijt} = \sum_{s=\underline{s}}^{\overline{s}} \beta_s b_{jt}^s + \mathbf{X}_{\mathbf{ijt}}^{'} \psi + \eta_{ij} + \gamma_{it} + \xi_{jt} + \epsilon_{ijt} \tag{5}$$

where Y_{ijt} is political violence committed by group j within district i at time t (e.g., incidents committed by the BLA in the district Quetta in 2015), b_{ijt}^s is a set of binned treatment dummies of the interaction term $(BLA_j \times BLA\ split_t)$. Hence, we only treat the BLA as a group and not districts in which the BLA has been present. $\mathbf{X}_{\mathbf{ijt}}^{'}$ includes the triple interaction of Sunni groups with the Sunni share of districts and the global silver price (as in Limodio, 2022),[40] and a set of time-invariant group ideology indicators (taken from Kis-Katos et al. (2014)) interacted with year fixed effects. The goal of the first interaction is to control for local differences in financing opportunities, while the second set of interactions captures global shocks for different types of groups, i.e., increased counter-insurgency against particular types of groups. η_{ij} are district-group fixed effects controlling for the time-invariant capacity a group has within a district, as well as group-specific selection into districts at the extensive margin. γ_{it} are district-year fixed effects controlling for competition between groups, state capacity, and local demand within districts over time. Finally, ξ_{jt} are group-specific linear time-trends and ϵ_{ijt} is the error term.[41]

The within-group event study represents by far the most restrictive specification that

[39] Attacks against civilians are the sum of incidents defined as attacks against civilians in the GTD and GED, capital intensive attacks are defined following Limodio (2022), see Appendix D for details. Note that capital intensive and non-capital intensive attacks do not sum to total attacks due to missing information on the attack type.

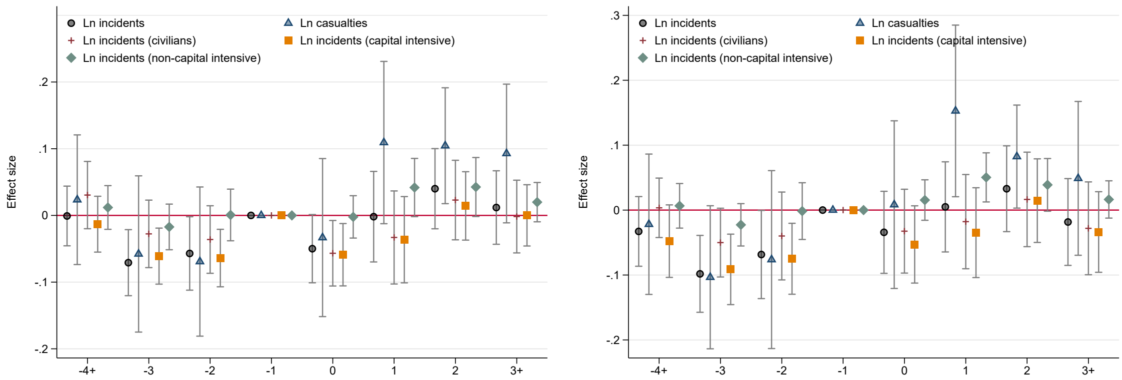
[40] We classify groups as "Sunni" following Table-D3 in Limodio (2022).

[41] Note that we cannot include group-year fixed effects because they would absorb our treatment.

we employ. Identifying variation is now restricted to differences between the BLA and other armed groups within districts over time (the UBA is excluded for now but included below). Note that this limits the set of armed groups to those which operate at least once in the pre-and post-treatment period.

Figure VIII
Within Baloch separatist groups evidence

(A) Pol. violence BLA (B) Pol. violence BLA & UBA



Notes: Reports the event study coefficient and their accompanying 95% CIs for within-group event study specification as stated in eq. 5. 95% CI are based on standard errors clustered at the district level.

The event study specification coefficients for the log of incidents+1 (black dots) and casualties+1 (blue triangles) are depicted in panel (A) of Figure VIII. The results suggest that the BLA engaged in fewer incidents during the year KBM died (although the point estimate is only marginally significant) but then returned to business as usual in $t+1$. However, this seems not to have been the case for casualties inflicted by the BLA, which rise compared to other groups following treatment. The temporary drop in $t = 0$ is similar in size and more precisely estimated if we focus on attacks against civilians and those which are comparably capital intensive, represented by the red crosses and orange squares in panel (A) of Figure VIII. Attacks that are not capital intensive (green diamonds) do not fall compared to other groups and increase in $t+1$ and $t+2$. The respective magnitudes correspond to a 5% decrease for capital intensive attacks and attacks against civilians compared to other groups and an increase of about 5% in non-capital intensive attacks in $t+1$ and $t+2$. The fatality estimates imply an increase of about 10%.

Panel (B) replicates panel (A) but treats the BLA and UBA as a single group. This tests if the aggregate amount of political violence committed by the two splinter groups jointly has changed. Treating the two groups as one negates the temporary drop in $t = 0$ and reaffirms the increase in fatalities and non-capital intensive attacks.

In summary, the two sets of results suggest that there has been some negative shock to the BLA's capacity. Still, this shock did not change the relative importance with respect to the political violence occurring within districts. In fact, the BLA seems to keep its relevance by compensating for the capacity shock by switching strategies as they commit more non-capital intensive attacks.[42] Moreover, the BLA members that split away to form the UBA seem to follow a similar strategy. Both results are in line with our general argument. In the presence of increased competition (particularly by a similar actor), both the BLA and UBA become more lethal compared to other groups. However, the magnitudes of the effect highlight that the additional violence is not driven by the BLA or UBA alone.

7. Robustness tests

We perform several additional tests to understand the sensitivity of our findings, which we report briefly here and in greater detail in the Appendix.

We start by testing how our results are affected if we create our dependent variables from two separate datasets in Appendix B. We show that using both incidents from GTD and GED does not affect our baseline results in a meaningful way and take this as suggestive evidence that our results are not driven by a change in strategy of the groups operating in the treatment districts towards events more likely to be covered by the GTD. Moreover, we can show that our results remain qualitatively and quantitatively the same if we only focus on incidents officially claimed by a group. Thus, uncertainty about the perpetrator seems not to increase with more competition. The results are also inconsistent with the idea that groups try to claim more events if competition is more fierce. We also find little evidence that "potential" double-counting affects our combined results using both the GTD and GED. However, our probability-based approach to assess the likelihood of potential double-counts suggests that double counting is an issue for around 10% of GED events for realistic scenarios in our case.

We also further probe our concept of active armed groups (see Appendix C). The skeptical reader might be worried that our measure of active armed groups increases violence by construction because groups are only counted as active within districts if they commit at least one attack. To avoid potential selection on the extensive margin, we introduce the concept of potentially active groups. They are defined as groups that are active anywhere in the country and have been active at least once in a specific district and year. This approach acknowledges uncertainty about the spatial choices of the armed groups that we do not observe. Again, our results remain remarkably robust. Note that the measure of potential active groups and active armed groups are highly correlated (0.77). The overlap highlights another property of our setting. Specifically, the armed groups in our sample seem to have well-defined areas of operation. We also extended

[42]This empirical result is in line with theoretical work by Bueno de Mesquita (2013).

the potential active armed groups measure to cover all districts falling within the convex hull of a group's incidents (similar to König et al., 2017). Again our results remain stable. Finally, we find no evidence that the inclusion of "one-hit wonders" (Blomberg et al., 2010) in our measure of active armed groups affects our results. Note that "one-hit wonders" are counted identically in both potential and realized armed group counts since they commit only a single incident.

8. Conclusion

This paper studies the effect of the proliferation of armed groups on organized political violence. While the arguments in favor of such a mechanism have long been present in the literature, we are the first paper to provide quasi-experimental evidence on the matter. We exploit a unique setting in Pakistan where the unexpected death of a pivotal figure leads to the split of a major armed group, allowing us to provide quasi-experimental evidence on the net effect of group proliferation and differentiate between opposing competition and capacity effects.

Our estimates predict that one additional active armed group increases the incidents of organized political violence by about 60% and causes casualties to rise by roughly 75%. These sizeable effects and dynamics that we document are consistent with the idea of competition between armed groups for local dominance. In a communication to the Indian newspaper *The Hindu* (Bhattacherjee, 2019), the BLA indicated that "they are planning to intensify the struggle against Pakistan as they remain 'the most popular' militant organization in Balochistan."

Moreover, our 2SLS results suggest that groups seem to endogenously select into locations in which other groups already operate. Hence, for given locational fundamentals (e.g., resources, state capacity), the effect of an additional armed group is likely to be underestimated because we find a diminishing effect with respect to the number of existing groups. This also has some implications for the generalizability of our results. If the presence of additional groups is mostly occurring due to more available resources or less state capacity, the group effect itself will be smaller, while total violence could increase even more. However, comparing cross-country correlations, we do not find that Pakistan in general, is a very particular case. In fact, it seems to be a rather regular one.

Exploring the determinants and consequences of group appearance, mergers, and splits is a promising avenue for future research. Currently, there is little evidence on how local determinants of conflict, such as state capacity, the demand for armed groups, and financing opportunities, affect armed groups and are affected by them. Future research needs to trace why new groups form or split up and encroach upon the territories of other groups. Understanding within-group dynamics is largely absent from the literature so far. We believe this to be a major obstacle when it comes to policy recommendations. Consider the evaluation of counter-insurgency efforts against a specific group, for example. It is impossible to evaluate whether the policy can reduce political violence if we ignore how other groups are indirectly affected. Our study offers a toolkit to engage in those kinds of studies by providing a method to calculate proxies for counter-insurgency efforts by combining the GTD and GED databases. What is more, matching of incidences between the GTD and GED datasets enables researchers to analyze political violence of armed groups and increase coverage holistically.

Finally, our results suggest that politicians and military leaders should be careful if they employ targeted killing strategies against the leaders of armed groups to incapacitate large groups. Splitting up a larger group into competing splinter groups can actually increase violence in the short term.

References

Adhvaryu, A., J. E. Fenske, G. Khanna, and A. Nyshadham (2018). Resources, conflict, and economic development in Africa. NBER Working Paper No. 24309.

Ali, N. S. (2015). Situationer: Who's who of Baloch insurgency. https://www.dawn.com/news/1185401. Accessed 2020-03-13.

Balochistan Post (2018). Warring armed organisations BLA and UBA call truce. https://thebalochistanpost.net/2018/03/warring-armed-organisations-bla-uba-call-truce/. Accessed 2020-03-13.

Benmelech, E. and C. Berrebi (2007). Human capital and the productivity of suicide bombers. *Journal of Economic Perspectives 21*(3), 223–238.

Berman, E., J. N. Shapiro, and J. H. Felter (2011). Can hearts and minds be bought? The economics of counterinsurgency in Iraq. *Journal of Political Economy 119*(4), 766–819.

Berman, N., M. Couttenier, D. Rohner, and M. Thoenig (2017). This mine is mine! How minerals fuel conflicts in Africa. *American Economic Review 107*(6), 1564–1610.

Besley, T. and M. Reynal-Querol (2014). The legacy of historical conflict: Evidence from Africa. *American Political Science Review 108*(2), 319–336.

Bhattacherjee, K. (2019). Explained: The Baloch Liberation Army. https://www.thehindu.com/news/international/explained-the-baloch-liberation-army/article28273960.ece. Accessed 2020-03-13.

Blattman, C. and E. Miguel (2010). Civil war. *Journal of Economic Literature 48*(1), 3–57.

Blomberg, S. B., R. C. Engel, and R. Sawyer (2010). On the duration and sustainability of transnational terrorist organizations. *Journal of Conflict Resolution 54*(2), 303–330.

Bormann, N.-C., L.-E. Cederman, S. Gates, B. A. T. Graham, S. Hug, K. W. Strøm, and J. Wucherpfennig (2019). Power sharing: Institutions, behavior, and peace. *American Journal of Political Science 63*(1), 84–100.

Bueno De Mesquita, E. (2005). The quality of terror. *American Journal of Political Science 49*(3), 515–530.

Bueno de Mesquita, E. (2013). Rebel tactics. *Journal of Political Economy 121*(2), 323–357.

Buhaug, H., S. Gates, and P. Lujala (2009). Geography, rebel capability, and the duration of civil conflict. *Journal of Conflict Resolution 53*(4), 544–569.

Buhaug, H. and J. K. Rød (2006). Local determinants of African civil wars, 1970–2001. *Political Geography 25*(3), 315–335.

CIESIN - Center for International Earth Science Information Network - Columbia University (2018). Gridded population of the World, version 4. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).

Condra, L. N., J. D. Long, A. C. Shaver, and A. L. Wright (2018). The logic of insurgent electoral violence. *American Economic Review 108*(11), 3199–3231.

Conrad, J. and K. Greene (2015). Competition, differentiation, and the severity of terrorist attacks. *Journal of Politics 77*(2), 546–561.

Dashti, N. (2017). *The Baloch Conflict with Iran and Pakistan.* Trafford Publishing.

Dawn.com (2014). Baloch nationalist leader Khair Bakhsh Marri passes away. https://www.dawn.com/news/1111835. Accessed 2020-03-13.

De Luca, G., R. Hodler, P. A. Raschky, and M. Valsecchi (2018). Ethnic favoritism: An axiom of politics? *Journal of Development Economics 132*, 115–129.

Draca, M., S. Machin, and R. Witt (2011). Panic on the streets of London: Police, crime, and the July 2005 terror attacks. *American Economic Review 101*(5), 2157–81.

Dube, O. and S. Naidu (2015). Bases, bullets, and ballots: The effect of US military aid on political conflict in Colombia. *Journal of Politics 77*(1), 249–267.

Dube, O. and J. F. Vargas (2013). Commodity price shocks and civil conflict: Evidence from Colombia. *Review of Economic Studies 80*(4), 1384–1421.

Eck, K. (2012). In data we trust? A comparison of UCDP GED and ACLED conflict events datasets. *Cooperation and Conflict 47*(1), 124–141.

Economist (2012). We only receive back the bodies. https://www.economist.com/asia/2012/04/07/we-only-receive-back-the-bodies. Accessed 2020-03-13.

Faria, J. R. (2014). The economics of technology in terrorist organizations. *Brown Journal of World Affairs 20*(2), 285–296.

Fearon, J. D. and D. D. Laitin (2003). Ethnicity, insurgency, and civil war. *American Political Science Review 97*(1), 75–90.

Findley, M. G. and J. K. Young (2012). Terrorism and civil war: A spatial and temporal approach to a conceptual problem. *Perspectives on Politics 10*(2), 285–305.

Gaibulloev, K. (2015). Terrorist group location decision: An empirical investigation. *Oxford Economic Papers 67*(1), 21–41.

Gaibulloev, K. and T. Sandler (2019). What we have learned about terrorism since 9/11. *Journal of Economic Literature 57*(2), 275–328.

Gehring, K., S. Langlotz, and K. Stefan (2019). Stimulant or depressant? Resource-related income shocks and conflict. CESifo Working Paper No. 7887.

Hersbach, H., B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee, and J.-N. Thépaut (2018). Era5 hourly data on single levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview.

Hobbes, T. (1969). *Leviathan, 1651.* Menston: Scholar P.

Hou, D., K. Gaibulloev, and T. Sandler (2020). Introducing extended data on terrorist groups (edtg), 1970 to 2016. *Journal of Conflict Resolution 64*(1), 199–225.

Jaeger, D. A. and M. D. Paserman (2008). The cycle of violence? An empirical analysis of fatalities in the Palestinian-Israeli conflict. *American Economic Review 98*(4), 1591–1604.

Jetter, M. (2017). The effect of media attention on terrorism. *Journal of Public Economics 153*, 32–48.

Khan, M. I. (2014). Baloch nationalist leader Nawab Khair Bakhsh Marri dies. https://www.bbc.com/news/world-asia-27801253. Accessed 2020-03-13.

Kis-Katos, K., H. Liebert, and G. G. Schulze (2011). On the origin of domestic and international terrorism. *European Journal of Political Economy 27*, 17–36.

Kis-Katos, K., H. Liebert, and G. G. Schulze (2014). On the heterogeneity of terror. *European Economic Review 68*, 116–136.

König, M. D., D. Rohner, M. Thoenig, and F. Zilibotti (2017). Networks in conflict: Theory and evidence from the Great War of Africa. *Econometrica 85*(4), 1093–1132.

Limodio, N. (2022). Terrorism financing, recruitment, and attacks. *Econometrica 90*(4), 1711–1742.

Marineau, J., H. Pascoe, A. Braithwaite, M. Findley, and J. Young (2020). The local geography of transnational terrorism. *Conflict Management and Peace Science 37*(3), 350–381.

Morelli, M. and D. Rohner (2015). Resource concentration and civil wars. *Journal of Development Economics 117*, 32–47.

Nabeel, F. (2017). Factionalism in the Balochistan insurgency – An overview. https://stratagem.pk/armed-dangerous/factionalism-balochistan-insurgency-overview/. Accessed 2020-03-13.

NACTA (2020). Proscribed organizations. National Counter Terrorism Authority – Pakistan. https://nacta.gov.pk/proscribed-organizations/. Accessed 2020-03-13.

Nemeth, S. (2014). The effect of competition on terrorist group operations. *Journal of Conflict Resolution 58*(2), 336–362.

News International (2014). Baloch nationalist leader Khair Bakhsh Marri passes away. https://www.thenews.com.pk/archive/print/638441-baloch-nationalist-leader-khair-bakhsh-marri-passes-away. Accessed 2020-03-13.

Polo, S. M. and K. S. Gleditsch (2016). Twisting arms and sending messages: Terrorist tactics in civil war. *Journal of Peace Research 53*(6), 815–829.

Raleigh, C., A. Linke, H. Hegre, and J. Karlsen (2010). Introducing ACLED: An armed conflict location and event dataset: Special data feature. *Journal of Peace Research 47*(5), 651–660.

Rashid, A. (2014). Balochistan: The untold story of Pakistan's other war. BBC News. https://www.bbc.com/news/world-asia-26272897. Accessed 2020-03-13.

Shah, A. Z. (2017). Geopolitical significance of Balochistan: Interplay of foreign actors. *Strategic Studies 37*(3), 126–144.

Stanton, J. A. (2013). Terrorism in the context of civil war. *Journal of Politics 75*(4), 1009–1022.

START (2019). Global Terrorism Database codebook: Inclusion criteria and variables. http://www.start.umd.edu/gtd/downloads/Codebook.pdf.

Stina, H. (2019). UCDP GED, Codebook version 19.1. Department of Peace and Conflict Research, Uppsala University.

Sundberg, R. and E. Melander (2013). Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research 50*(4), 523–532.

Sviatschi, M. M. (2022). Spreading gangs: Exporting US criminal capital to El Salvador. *American Economic Review 112*(6), 1985–2024.

Times of India (2016). The Balochistan conflict: 10 key points. https://timesofindia.indiatimes.com/the-balochistan-conflict-10-key-points/listshow/53688031.cms. Accessed 2020-03-13.

Toft, M. D. and Y. M. Zhukov (2015). Islamists and nationalists: Rebel motivation and counterinsurgency in Russia's North Caucasus. *American Political Science Review 109*(2), 222–238.

Tollefsen, A. F., H. Strand, and H. Buhaug (2012). PRIO-GRID: A unified spatial data structure. *Journal of Peace Research 49*(2), 363–374.

Trebbi, F. and E. Weese (2019). Insurgency and small wars: Estimation of unobserved coalition structures. *Econometrica 87*(2), 463–496.

Van der Windt, P. and M. Humphreys (2016). Crowdseeding in eastern Congo: Using cell phones to collect conflict events data in real time. *Journal of Conflict Resolution 60*(4), 748–781.

Vogt, M., N.-C. Bormann, S. Rüegger, L.-E. Cederman, P. Hunziker, and L. Girardin (2015). Integrating data on ethnicity, geography, and conflict: The ethnic power relations data set family. *Journal of Conflict Resolution 59*(7), 1327–1342.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

Wooldridge, J. M. (2015). Control Function Methods in Applied Econometrics. *Journal of Human Resources 50*(2), 420–445.

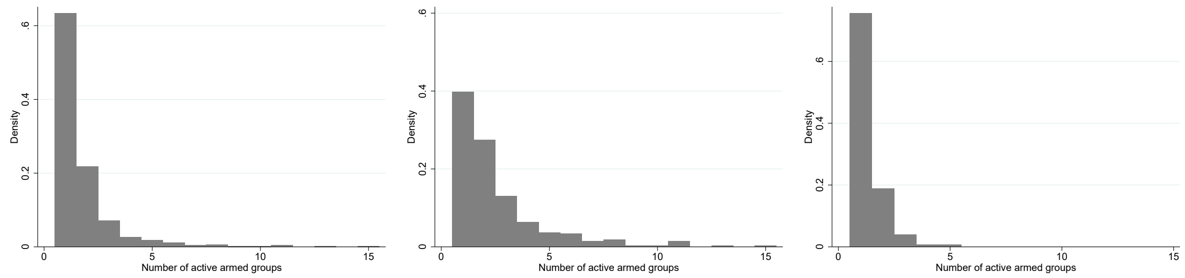
Wucherpfennig, J., N. B. Weidmann, L. Girardin, L.-E. Cederman, and A. Wimmer (2011). Politically relevant ethnic groups across space and time: Introducing the GeoEPR dataset. *Conflict Management and Peace Science 28*(5), 423–437.

A. Additional Figures and Tables

A-1. Figures

Figure A-1
Distribution: Number of armed groups

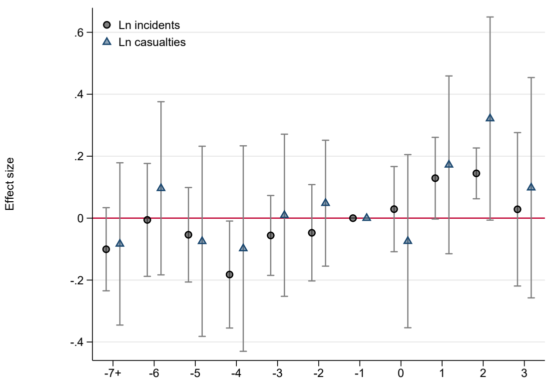
(A) All districts (B) $BLA\ share_i > 0$ (C) $BLA\ share_i = 0$



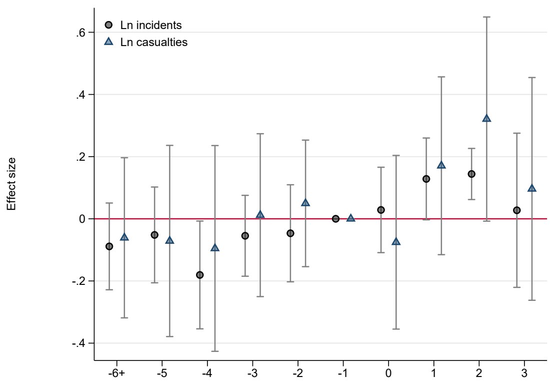
Notes: Panel A depicts the distribution of the number of *active* armed groups in district-years with at least one active group for all districts. Panel (B) and (C) plot the distribution of the number of *active* armed groups in district-years with at least one active group for all districts within our treatment and control group separately.

Figure A-2
Reduced form: Pretrends

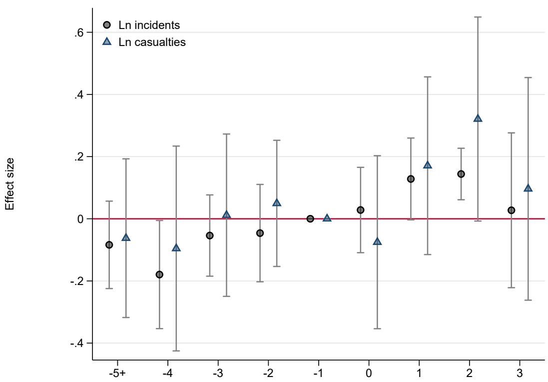
(a) -7+



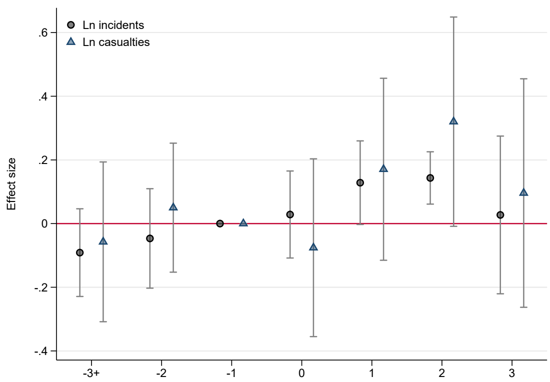
(b) -6+



(c) -5+



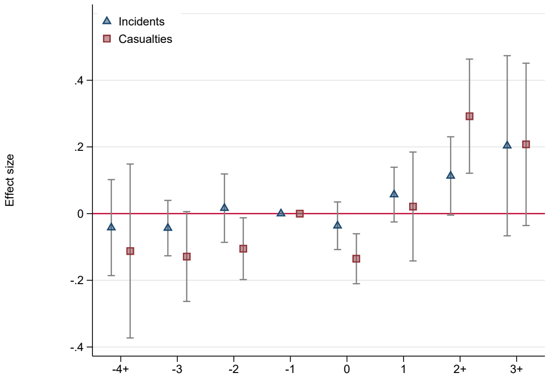
(d) -3+



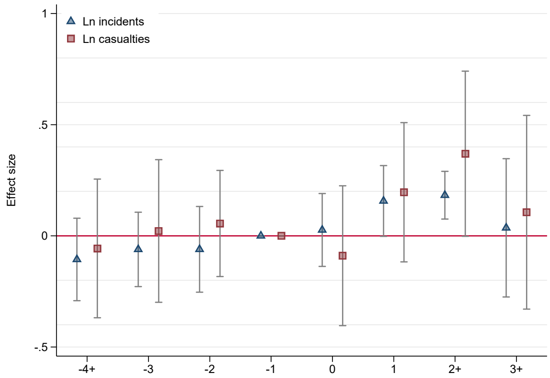
Notes: Reports the coefficients and their accompanying 95% CIs for an event study as specified in eq. 1 in which we use more single estimates prior to treatment to probe the presence of pretrends. CIs are based on standard errors clustered at the district level.

Figure A-3
Reduced form: Alternative specifications

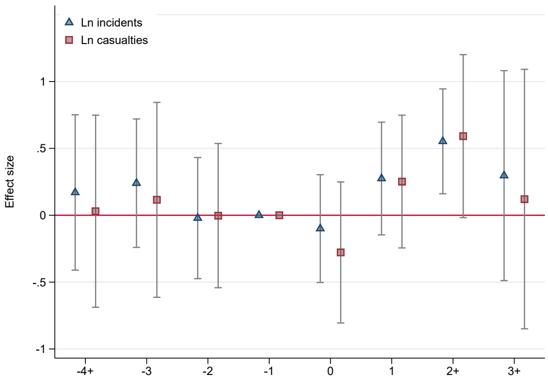
(a) (PPML)



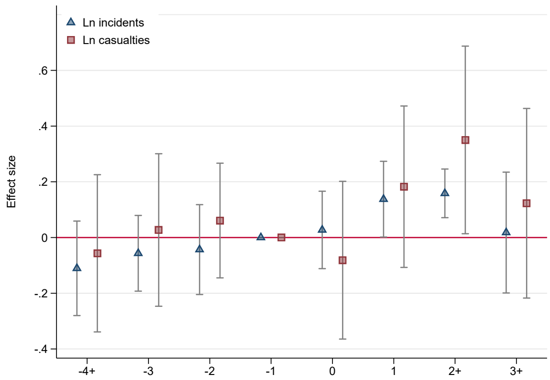
(b) (Inverse hyperbolic sine transformation DV)



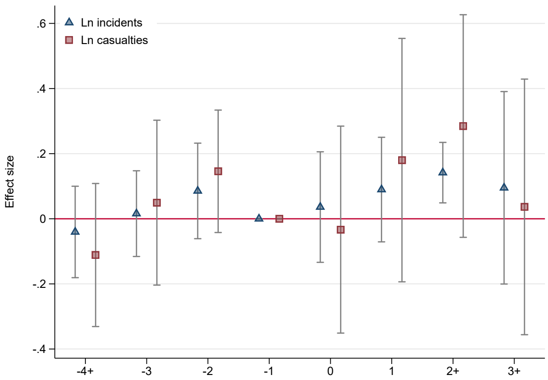
(c) (Log DV + 0.01)



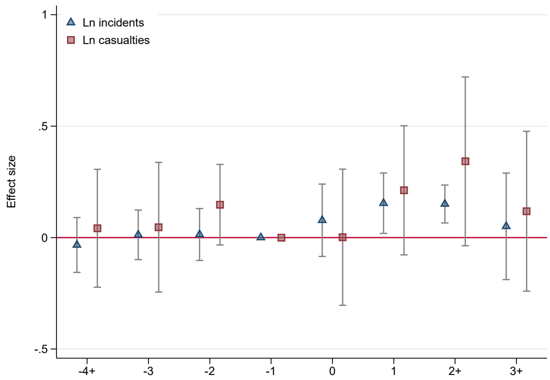
(d) (Excluding controls)



(e) (Within Balochistan)



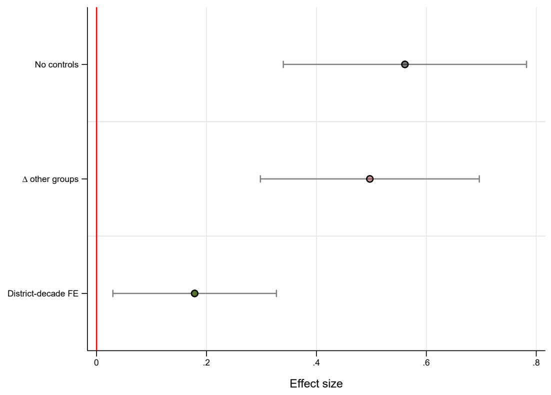
(f) (*BLA share* before 2011)



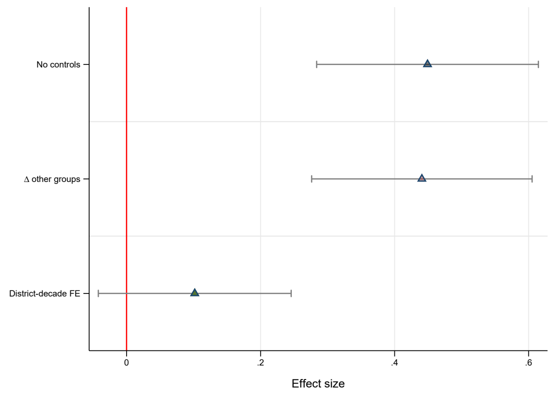
Notes: Reports the coefficients and their accompanying 95% CIs for an event study as specified in eq. 1. CIs are based on standard errors clustered at the district level.

FIGURE A-4
Reduced form: robustness

(A) Ln incidents



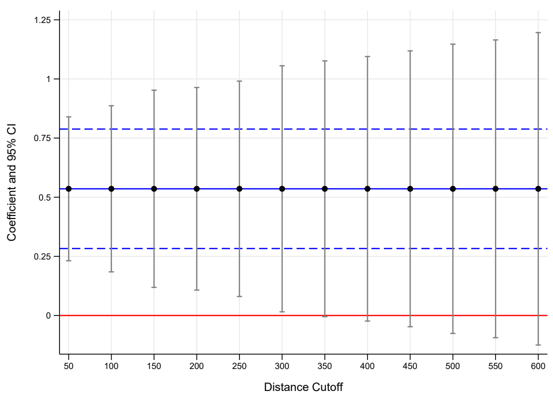
(B) Ln casualties



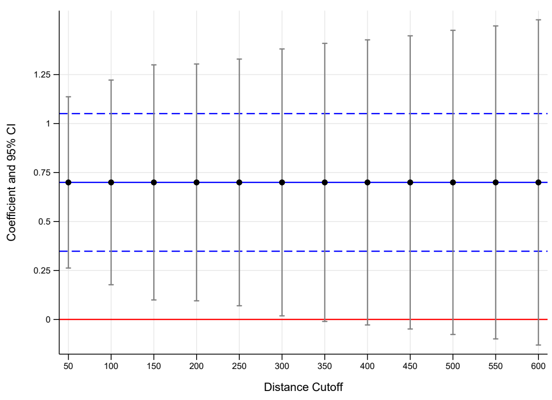
Notes: Reports the coefficients and their accompanying 95% CIs for different replications of our generalized DiD reduced form specification (corresponding to the horizontal lines in Figure IV). We first drop all controls, then include the change in other groups, and finally add decade-district fixed effects. The CIs are based on standard errors clustered at the district level.

FIGURE A-5
Baseline estimate: Arbitrary spatial clustering

(A) Ln incidents



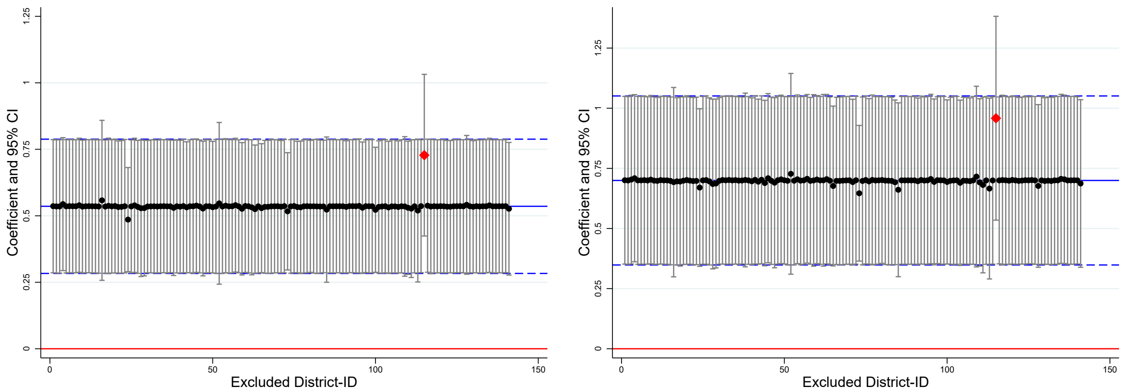
(B) Ln casualties



Notes: Reports the coefficients and their accompanying 95% CIs for replications of our generalized DiD reduced form specification (corresponding to the horizontal lines in Figure IV), in which we calculate the CIs based on Conley standard errors with varying distance cutoffs. The blue dashed lines depict the CI based on standard errors clustered at the district level.

Figure A-6
Leave one out test: Districts

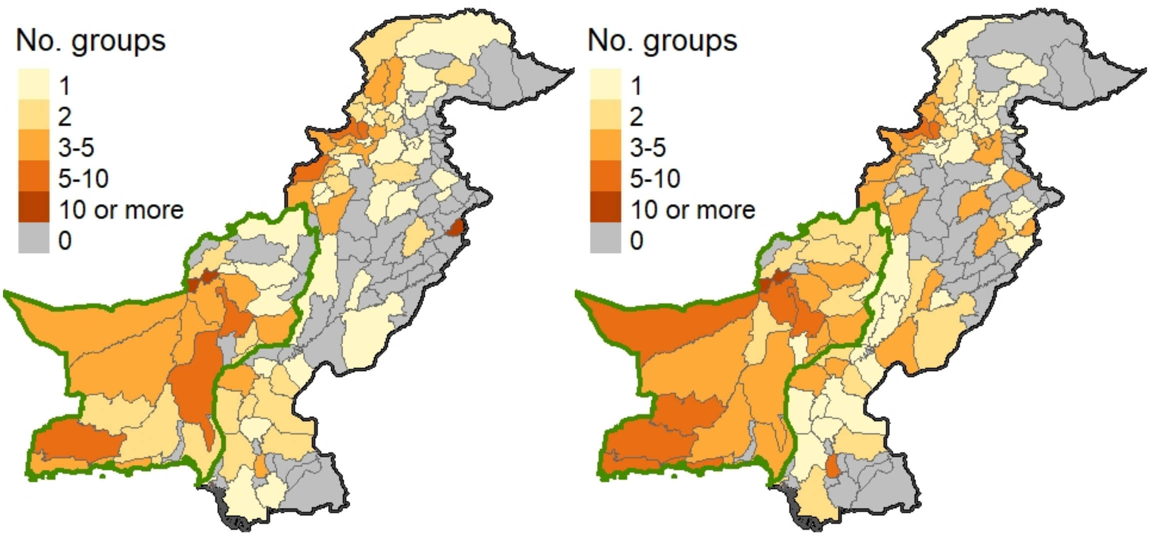
(A) Ln incidents (B) Ln casualties



Notes: Reports the coefficients and their accompanying 95% CIs for replications of our generalized DiD reduced form specification (corresponding to the horizontal lines in Figure IV), in which we drop one district at the time. The CIs are based on standard errors clustered at the district level. The red diamond highlights the exclusion of the Baloch capital district Quetta.

Figure A-7
Number of armed groups across districts

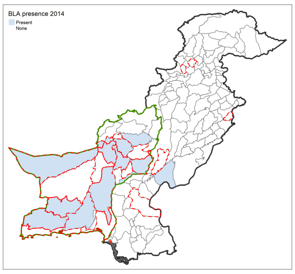
(A) Number of groups within districts (pre-treatment) (B) Number of groups within districts (post-treatment)



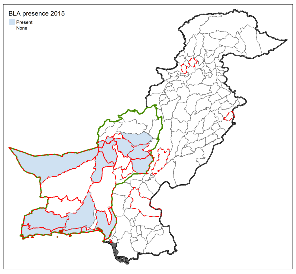
Notes: Panel A reports the number of groups active in districts in the pre-treatment period. Panel B plots the number of groups active within a district in the post-treatment group. The bold green line represents the borders of the Baloch province.

Figure A-8
BLA & UBA presence post treatment

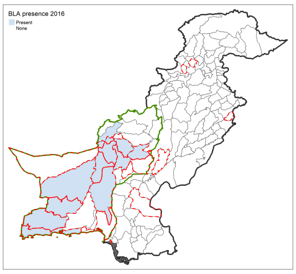
(a) BLA 2014



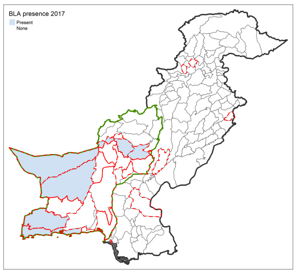
(b) BLA 2015



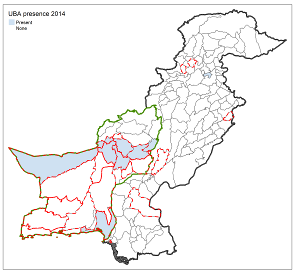
(c) BLA 2016



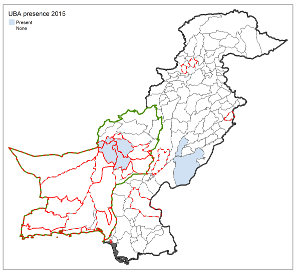
(d) BLA 2017



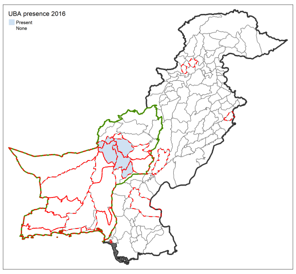
(e) UBA 2014



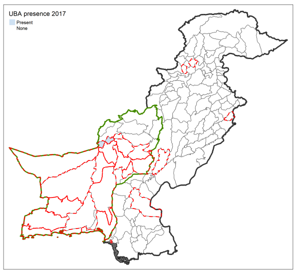
(f) UBA 2015



(g) UBA 2016



(h) UBA 2017



Notes: Panels A to D plot the presence of the BLA across districts for the years 2014 to 2017. Panels E to H plot the UBA presence across districts for the years 2014 to 2017. Presence is defined as committing at least one attack within a year. Districts with a positive BLA presence pre-treatment are highlighted by dashed red borders. The border of the Balochistan province is highlighted in green border.

A-2. Tables

Table A-1
Summary statistics

Variable	Mean	SD	Min	Max	N
Panel (A) District sample					
Dependent variables					
Incidents	0.96	4.23	0.00	96	3,384
Casualties	8.96	50.90	0.00	1,244	3,384
Ln incidents+1	0.26	0.64	0.00	4.57	3,384
Ln casualties+1	0.46	1.22	0.00	7.13	3,384
Ln incidents+1 (vs. other groups)	0.03	0.17	0.00	2.30	3,384
Ln incidents+1 (vs. government)	0.14	0.45	0.00	3.78	3,384
Ln incidents+1 (vs. infrastructure)	0.05	0.24	0.00	3.47	3,384
Ln incidents+1 (vs. businesses)	0.04	0.20	0.00	2.30	3,384
Ln incidents+1 (vs. civilians)	0.10	0.37	0.00	4.11	3,384
Counter-insurgency dummy	0.12	0.32	0.00	1.00	3,384
Ln counter-insurgency+1	0.17	0.57	0.00	4.96	3,384
Ln incidents+1 (GTD+GED)	0.42	0.90	0.00	5.58	3,384
Ln casualties+1 (GTD+GED)	0.69	1.50	0.00	7.57	3,384
Ln incidents+1 (GTD+GED) incl. non groups	0.71	1.15	0.00	6.42	3,384
Ln casualties+1 (GTD+GED) incl. non groups	0.89	1.59	0.00	7.60	3,384
Ln incidents+1 (claimed)	0.19	0.54	0.00	3.87	3,384
Ln casualties+1 (claimed)	0.35	1.08	0.00	7.12	3,384
Ln incidents+1 (incl. one-hit wonders)	0.26	0.66	0.00	4.57	3,384
Ln casualties+1 (incl. one-hit wonders)	0.47	1.24	0.00	7.19	3,384
Treatment variables					
No. armed groups	0.41	1.05	0.00	15.00	3,384
No. armed groups (potential)	1.21	1.79	0.00	21.00	3,384
No. armed groups (potential - convex hull)	3.48	2.71	0.00	13.00	3,384
Avg. BLA presence (pre-treatment)	0.02	0.06	0.00	0.45	3,384
Avg. BLA presence (pre 2011)	0.01	0.05	0.00	0.35	3,384
Control variables					
Ln population	13.43	1.20	7.15	16.07	3,384
Δ avg. rainfall	-0.04	0.49	-1.81	1.97	3,384
Δ avg. temperature	0.06	0.67	-2.18	1.96	3,384
No. other armed groups	0.34	0.93	0.00	14.00	3,384
Δ No. other armed groups	0.02	0.47	-4.00	5.00	3,384
Share politically excluded (area based)	0.41	0.46	0.00	1.00	3,151
Share politically excluded (pop based)	0.39	0.46	0.00	1.00	3,151
Ln silver price per ton in USD	13.00	0.53	12.21	14.06	3,102
Panel B: Group sample					
Ln incidents+1	0.01	0.13	0.00	4.04	77,814
Ln casualties+1	0.02	0.26	0.00	6.54	77,814
Ln incidents+1 (capital intensive)	0.01	0.10	0.00	3.78	77,814
Ln incidents+1 (non-capital intensive)	0.01	0.08	0.00	3.18	77,814
Ln incidents+1 (vs. civilians)	0.01	0.09	0.00	3.50	77,814

Continued on next page

Table A-1 – *Continued from previous page*

Variable	Mean	SD	Min	Max	N
BLA dummy	0.04	0.19	0.00	1.00	77,814
Sunni group dummy	0.07	0.26	0.00	1.00	77,814
Left wing group dummy	0.04	0.19	0.00	1.00	77,814
Right wing group dummy	0.11	0.31	0.00	1.00	77,814
Ethnic separatist group dummy	0.04	0.19	0.00	1.00	77,814
Anti ethnic separatist group dummy	0.81	0.39	0.00	1.00	77,814
Islamist group dummy	0.22	0.42	0.00	1.00	77,814

Notes: The table reports the summary statistics of our variables of interests across samples.

Table A-2

Definition of the dependent variables in Figure VI

Incidence against...	GTD targtype coding rule
Other groups	Terrorists/Non-State Militia
	Violent Political Party
Government	Military
	Police
	Government (Diplomatic)
	Government (General)
Public infrastructure	Airports & Aircraft
	Food or Water Supply
	Telecommunication
	Transportation
	Utilities
Business	Business
	Tourists
Citizens	Private Citizens & Property

Table A-3
Inverse hyperbolic sine transformation (DV): 2SLS

	OLS		*2SLS*			
	Dependent variable:					
	Ln incidents (1)	Ln casualties (2)	Ln incident (3)	Ln casualties (4)	Ln incident (5)	Ln casualties (6)
No. active groups	0.7616	0.9240	1.0143	1.3080	0.6446	0.7933
	(0.0512)	(0.1259)	(0.2054)	(0.3094)	(0.1614)	(0.2127)
No. act grps in **all** districts:	*Mean: 0.4078*					
	SD: 1.0480					
No. act grps in **act** districts:	*Mean: 1.7468*					
	SD: 1.5384					
Controls	✓	✓	✓	✓	✓	✓
District-FE	✓	✓	✓	✓	✓	✓
Year-FE	✓	✓	✓	✓	✓	✓
Division-trend	✓	✓	✓	✓	✓	✓
Adj. R^2	0.730	0.634	0.263	0.332	0.495	0.387
Obs	3384	3384	3384	3384	3384	3384
F-stat IV (1st stage)	–	–	11.44	11.44	16.53	16.53

Notes: The table reports the results of regression the inverse hyperbolic sine transformation of incidents and casualties on the number of active groups. Columns 1 and 2 use OLS estimates. Columns 3 to 6 report the first and second stage results based on the 2SLS specification stated in eq. 2. Columns 3 and 4 use the interaction $BLA\ share_i \times BLA\ split_t$ as the instrument for the number of active groups operating within a district. Columns 5 and 6 use a dynamic version of $BLA\ share_i \times BLA\ split_t$ in which each post treatment period (t to t+4) is allowed to have a different effect on the number of groups. Standard errors clustered at the district level in parenthesis.

TABLE A-4
2SLS: Within Baloch province

	OLS		*2SLS*			
	Dependent variable:					
	Ln incidents	Ln casualties	Ln incident	Ln casualties	Ln incident	Ln casualties
	(1)	(2)	(3)	(4)	(5)	(6)
No. active groups	0.4319	0.7346	0.7276	1.1164	0.4434	0.7469
	(0.0362)	(0.0666)	(0.1299)	(0.1723)	(0.1385)	(0.2306)
No. act grps in **all** districts:				*Mean: 0.6964* *SD: 1.3149*		
No. act grps in **act** districts:				*Mean: 2.0617* *SD: 1.5185*		
Controls	✓	✓	✓	✓	✓	✓
District-FE	✓	✓	✓	✓	✓	✓
Year-FE	✓	✓	✓	✓	✓	✓
Division-trend	✓	✓	✓	✓	✓	✓
Adj. R^2	0.793	0.693	0.316	0.314	0.521	0.418
Obs	672	672	672	672	672	672
F-stat IV (1st stage)			11.38	11.38	10.08	10.08

Notes: The table replicates Table I using only districts within the Balochistan province. Standard errors clustered at the district level in parenthesis.

TABLE A-5

Number of groups: Diminishing returns

	OLS		*Control function approach*			
	Dependent variable:					
	Ln incidents (1)	Ln casualties (2)	Ln incident (3)	Ln casualties (4)	Ln incident (5)	Ln casualties (6)
No. active groups	0.5926	1.1782	1.0743	1.7700	0.7622	1.2742
	(0.0415)	(0.0718)	(0.1423)	(0.2253)	(0.1021)	(0.1164)
No. active groups^2	-0.0209	-0.0557	-0.0225	-0.0578	-0.0223	-0.0565
	(0.0044)	(0.0061)	(0.0052)	(0.0068)	(0.0053)	(0.0062)
First-stage residual			-0.4882	-0.5998	-0.1758	-0.0995
			(0.1379)	(0.2245)	(0.0966)	(0.1169)
No. act grps in **all** districts:	*Mean: 0.4078* *SD: 1.0480*					
No. act grps in **act** districts:	*Mean: 1.7468* *SD: 1.5384*					
Controls	✓	✓	✓	✓	✓	✓
District-FE	✓	✓	✓	✓	✓	✓
Year-FE	✓	✓	✓	✓	✓	✓
Division-trend	✓	✓	✓	✓	✓	✓
Adj. R^2	0.744	0.661	0.751	0.664	0.746	0.436
Obs	3384	3384	3384	3384	3384	3384

Notes: The table reports the results of replicating Table I adding the squared number of armed groups as an independent variable while using control functions instead of 2SLS models in columns 3 to 6. Standard errors are clustered at the district level in parenthesis.

B. Counting incidents and casualties

In this appendix, we take a closer look at how the incidents and casualties reported by the GTD and GED enter our specifications. We start by testing the robustness of our results regarding different issues in the data, such as if an armed group has officially claimed an event. We then proceed and discuss further what problems arise if one combines events of organized political violence across databases and how we deal with those issues.

B-1. Counting all forms of organized political violence

The main specifications primarily rely on incidents and casualties provided by the GTD. The reason is that the GTD has the best coverage for most of the actors we are interested in (see Section 3). However, having the best coverage for our actors of interest among the available databases has the potential to bias our results. The GTD, by its definition, focuses on organized political violence that fits its definition of terrorism. A case could arise in which all groups would switch to committing more violence fitting the GTD definition following our treatment. GTD would still provide the best coverage but our estimated increase in violence would not correspond to the actual level of overall violence. In such a world the actors could be involved in fewer events that fulfill the criteria of internal armed conflict (i.e., the primary focus of the GED) post-treatment.

We test for the aggregated effect across databases columns 1 and 2 of Table B-1, where we replicate columns 5 & 6 of Table I including all incidents committed, and casualties inflicted by armed groups from the GTD and GED.[43] Note that the combined casualties are the sum of fatalities and wounded (GTD) and the best estimate of battle-related deaths (GED).[44] Columns 1 and 2 of Table B-1 show that the size and statistical significance of our point estimates remain roughly constant. Thus, the competition effect, which we obtained in our main specifications, does not seem to be driven by a switch in a strategy that is over-reported by a specific database.

A related issue is our classification of an armed group. Recall that we only count actors as armed groups if they have a unique name that identifies them; hence, we exclude events from actors such as "Tribesmen" or "Gunmen". However, excluding the events committed by those actors might also bias our results in unexpected ways. Furthermore, given that both the GTD and GED rely on publicly available data, those names could also reflect uncertainty about the actual perpetrator of the event. Fortunately, the GTD codes if an armed group has officially claimed an event, which allows us to test if our results are driven by uncertainty about events or our group definition. Columns 2 and 3 of Table B-1 again replicate our preferred specifications using the log of claimed incidents + 1, as well as the log of claimed casualties + 1. Columns 5 and 6 leverage all events included in the

[43] See Section 3 for our definition of an armed group.

[44] The GED does not report estimates on those wounded in an event.

GTD and GED.

TABLE B-1
Alternative dependent variables

	Dependent variable:					
	GED & GTD events		*Claimed events*		*All events*	
	Ln incidents	Ln casualties	Ln incident	Ln casualties	Ln incidents	Ln casualties
	(1)	(2)	(3)	(4)	(5)	(6)
No. active groups	0.4498	0.6652	0.5636	0.7784	0.2201	0.4347
	(0.1279)	(0.1990)	(0.1288)	(0.1980)	(0.0866)	(0.1525)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	3384	3384	3384	3384	3384	3384
F-stat IV	16.53	16.53	16.53	16.53	16.53	16.53

Notes: The table reports the results of regression, the log of incidents + 1, and the log of casualties on the number of active groups. Columns 1 and 2 use all the incidents and casualties reported in the GTD and GED for known armed groups to construct the measures. Columns 3 and 4 use only those incidents and the resulting casualties that have been publicly claimed by a group. Columns 5 and 6 use all events included in GTD and GED (adds the events not committed by armed groups). The first stage instrument is the dynamic version of $BLA\ share_i \times BLA\ splitt$ in which each post-treatment period is allowed to have a different effect on the number of groups. Standard errors are clustered at the district level in parenthesis.

Using all events included in the GTD and GED reduces the economic size of our estimated effects somewhat. This is not surprising because the number of groups is now less well correlated with the number of attacks (e.g., attacks committed by "Gunmen" are now included, but "Gunmen" do not count as a group). However, the results show that even 'unorganized' violence perpetrated by gunmen or rioting mobs increases in the number of active groups. The estimates for the claimed incidents and casualties are again indistinguishable from our main result. This is reassuring because it suggests that more groups do not automatically lead to more uncertainty about the perpetrators of incidents, nor do groups seem to over-claim in the presence of competition.

B-2. Double counting organized political violence

Using two data sources on organized political violence can result in double-counting. Double counting arises if both the GED and GTD code the same incidents for the same groups. Even though they have different definitions of organized political violence, this is not implausible, at least for a subset of events that might fit both definitions.

Testing for double counting is not straightforward due to two reasons. First, GTD and GED have slightly different group names and different levels of aggregation. Generally, the level of aggregation is usually higher in the GED compared to the GTD. For example, GED will code a group "X", and the GTD will code the same group "X" as "$X-1$" referring to some faction within "X" and "$X-2$", referring to another faction within "X". Second, each event is coded based on publicly available source material subject to human interpretation. Thus, GTD and GED may attribute events to different actors due to conflicting and or different source material. This may also lead to alternative coding decisions concerning the day or exact location of a specific event.

The first issue is easily solved by harmonizing the group names. In practice, we aggregate the GTD group names up to the GED group name for all matches.[45] The second issue has no clear solution. Hence, we allow for varying time and location precision levels in both datasets.

With these limitations in mind, we propose a solely uncertainty-based approach to address the issue. Specifically, we quantify the likelihood that a GED event is a potential double count of a GTD event, depending on the distance both in space and time from one GED event committed by group i to all GTD events also committed by group i. The advantage of our approach compared to other possibilities, i.e., checking the source material of each incident, is the scalability for samples containing more than 10,000 events, as in our setting. Furthermore, it is transparent and allows the reader to decide with which buffer size she is comfortable instead of crosschecking our individual coding decisions.

The procedure of incidence-matching is simple. We classify a GED event committed by group i as being a potential double-count if the temporal distance to any GTD event committed by the same group i is within a bandwidth t starting from $+/-$ one day and simultaneously below a certain distance threshold d starting from one kilometer. To allow for multiple different events within close proximity at the same time period, we require the number of people killed reported by the GTD to fall within the bounds (low and high estimate) of the fatality count provided by GED.[46] Note that the matching of fatalities becomes more important as soon as the temporal and distance buffer increase, as the likelihood of false positives increases with wider buffers. Note that for this approach to be valid, we need to assume that double counting is only an issue between databases but not within. We do not find this assumption problematic since both databases conduct internal quality control, and it seems unlikely that they systematically misinterpret the set of sources they judged to be meaningful. Finally, the approach may also be performed without matching groups and fatalities from both databases. The only effect this has is that one is more likely to classify GED events as potential double counts, which in truth are not. In general, the approach has a clear trade-off between the likelihood of committing type 1 vs. type 2 errors, depending on the thresholds and the same-group

[45] The specific matching of group names between the GTD and GED dataset is documented in Table D

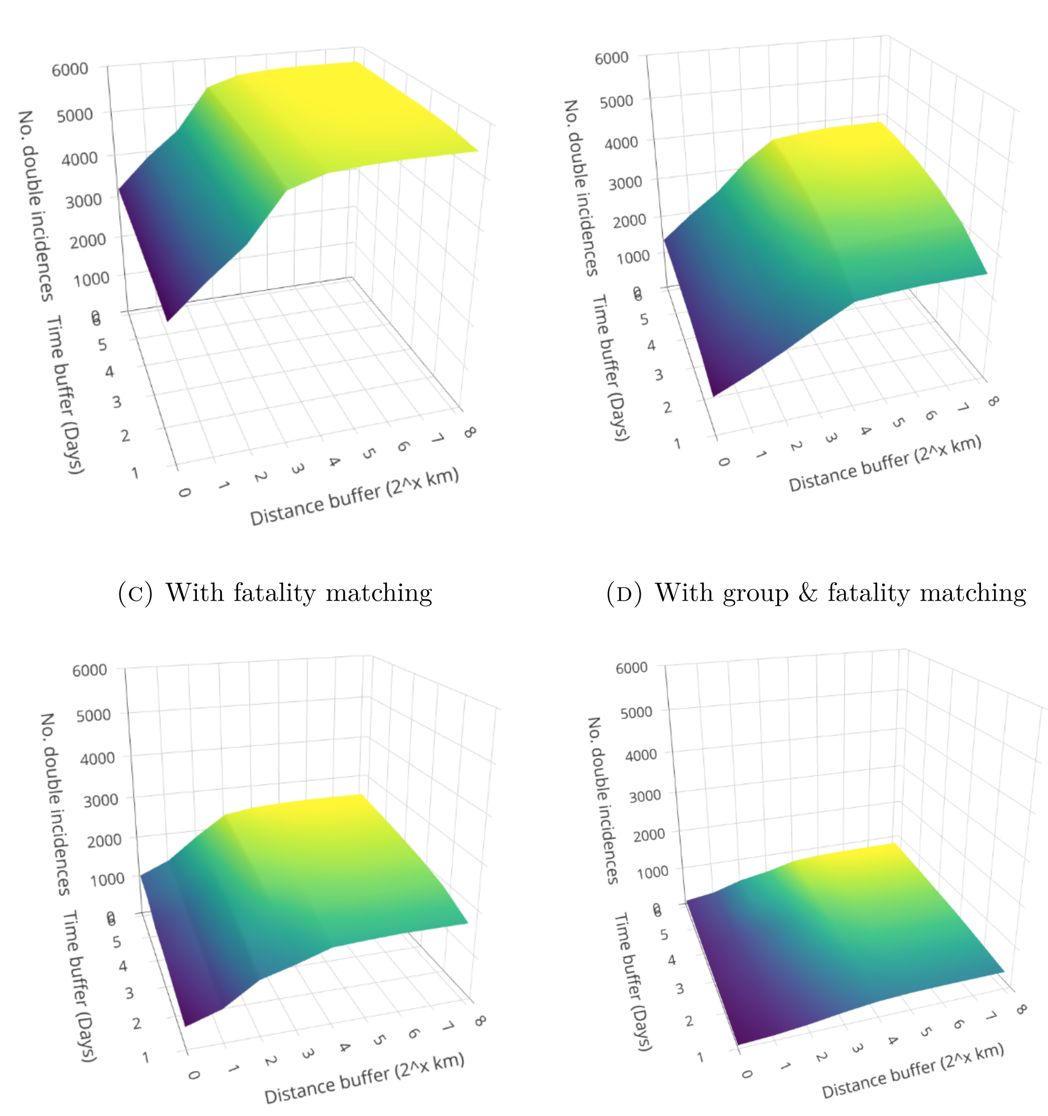
[46] Both the GED and GTD count all fatalities related to a specific incident. If our approach is applied to sources using different methods for counting fatalities, this criteria should be dropped from the procedure.

criterion.

Figure B-1
No. of incidents of GTD and GED double coding

(A) Without group matching (B) With group matching

(C) With fatality matching (D) With group & fatality matching



Notes: The figure plots the amount of GED incidents flagged as potential double counts across the distance and temporal thresholds across the different approaches: Panel (A) naive approach, panel (B) group matching approach, panel (C) fatality matching approach, panel (D) group, and fatality match approach.

We illustrate the impact of the double counts for each criterion. Panel A of Figure B-1 shows the naive approach (excluding the same group name and fatality criteria). The number of GED events that are considered potential double counts of GTD events for parameter constellations of daily temporal bandwidth $t = (1, 2, 3, 4, 5, 6)$, and a distance threshold ranging from $d = (2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8)$, which corresponds to 1, 2, 4, 8, 16, 32, 64, 128, and 256 kilometers, respectively. Note that the temporal threshold affects the potential double count status of a GED event much less compared to the distance threshold. As soon as we use a distance threshold of 32km, every GED event is a potential double count. In other words, whenever a GED event is coded in any day during our sample period, there is at least 1 GTD event coded within 32km distance of that GED event.

Panel B of Figure B-1 replicates the approach but applies the same-group name criterion. The picture becomes much more nuanced. Raising the distance threshold affects the number of assigned double counts beyond 32km. Furthermore, the interaction between the temporal and distance buffer is more pronounced. Lastly, the maximum amount of GED incidents flagged as potential double counts are 65% of those in the naive approach.

Panel C and D show that the inclusion of the fatality-match decreases the number of assigned double counts dramatically. The maximum amount of assigned double counts falls to just above 2000 without name matching and below 1000 if one includes the name match. The general effect of increasing the distance and temporal thresholds remains similar to the respective approach without fatality matching.

We conclude that the naive approach may involve too many type 1 errors. Type 1 errors are likely to be high if one ignores the same name criterion, given a high number of armed groups active within Pakistan. Hence, name-matching seems a necessary condition for a meaningful application of the approach in our setting. The fatality criterion seems more optional if one already includes the same-name requirement and keeps the distance thresholds moderate. Still, it is unclear when the trade-off between type 1 and type 2 errors is minimized for the temporal and distance buffers.

Next, we assess the stability of our core results concerning double-counting. Since we have no clear guidance regarding the optimal thresholds for t and d we test the stability of our results for all combinations of t and d introduced above. Note that there is no upper limit on the combination one could test. Nevertheless, we restrict ourselves to the introduced set for brevity. We focus here on the approach using the same name criterion only (results are basically the same using the other approaches).[47]

Panels A to F of Figure B-2 plot the point coefficients of our preferred incidents specification (column 5 Table I) for increasing distance thresholds across different time buffers, starting with +/− one day in Panel A, in the upper part of each panel and the fraction of GED events in use (not flagged as a potential double-count) in the lower panel. Furthermore, each panel plots the baseline estimate of column 5 Table I with its 95% CIs (black solid and dashed lines), as well as the estimate using all events attributed to armed groups in GTD and GED (corresponding to column 1 Table B-1) with its 95% CIs (blue solid and dashed lines).

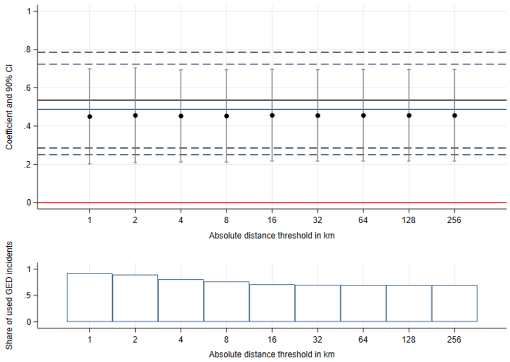
Figure B-2 shows that our core results seem remarkably robust to double-counting

[47]Results not reported but available upon request

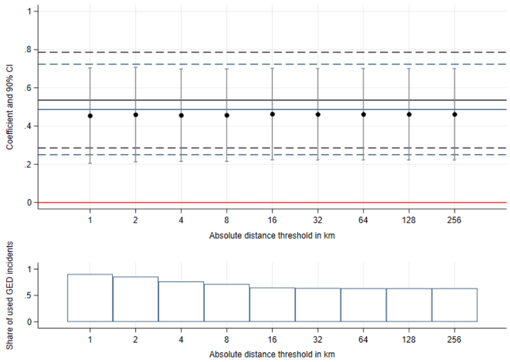
since we cannot reject that they are identical to either of the reported baseline effects. Furthermore, we see that the share of incidents marked as doubly counts falls rather quickly, regardless of the timing and distance thresholds. Note that those thresholds have more of an impact if the same name criterion is dropped. In summary, the results make us confident that our combined results using the full set of available events of organized political violence are not biased by some systematic measurement error due to double counting.

Figure B-2
Potential double coding (Same Names): GTD & GED

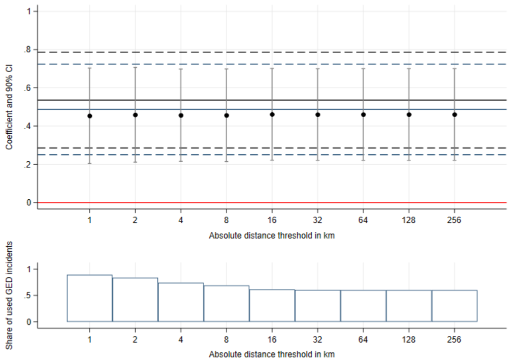
(a) A: +/- 1 days



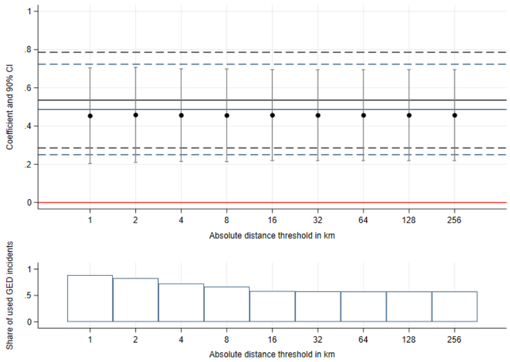
(b) B: +/- 2 days



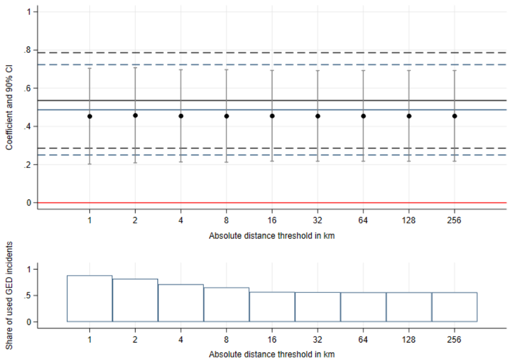
(c) C: +/- 3 days



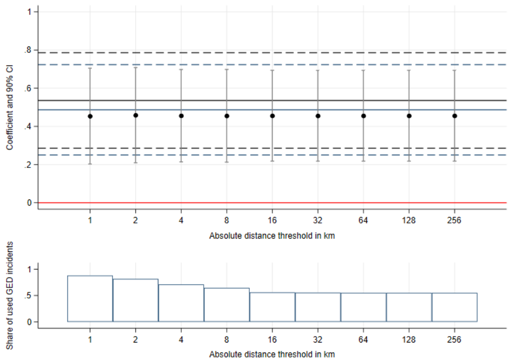
(d) D: +/- 4 days



(e) E: +/- 5 days



(f) F: +/- 6 days



Notes: Reports the coefficients and their accompanying 95% CIs for different potential double-counting thresholds both in time and space. The horizontal dark-grey solid line is the baseline coefficient using only GTD incidents. The two horizontal dark-grey dashed lines the accompanying 95% CIs. The horizontal blue solid line is the coefficient using all GED and GTD incidents ignoring potential double-counting. The two horizontal blue dashed lines the accompanying 95% CIs.

C. Counting independent groups

How many groups are competing within a district at any point in time? Given that our competition argument is based primarily on the number of active armed groups within a locality, this question is of paramount importance. However, as in the case of measuring organized political violence, measurement choices are abundant, theoretical guidance is limited, and empirical best practices are absent.

Recall, the number of active groups in our main specifications is a simple count of the armed groups that commit at least one incident within a district in a given year. Yet, one might argue that the actual incident committed is a strategic choice that maximizes utility over several dimensions for the group, e.g., ease of committing the attack vs. potential payoff (Marineau et al., 2020). If strategic considerations drive actual attacks, counting groups only as active within a district if they commit an attack during a year is an imperfect measure of their presence.[48] In turn, this will lead to an imperfect count of active groups and thus an imperfect proxy for group competition. It seems plausible that group effort is a result of the actual competition and not the imperfect perceived one. In general, we assume that local groups, as well as the government, have better information about the currently active groups within a district. Another issue closely related is the treatment of "one-hit wonders" (Blomberg et al., 2010), defined as armed groups that only commit one attack during our sample. In our main analysis, we exclude those armed groups entirely. Nevertheless, it seems prudent to test if our coding of active groups is sensitive to them.

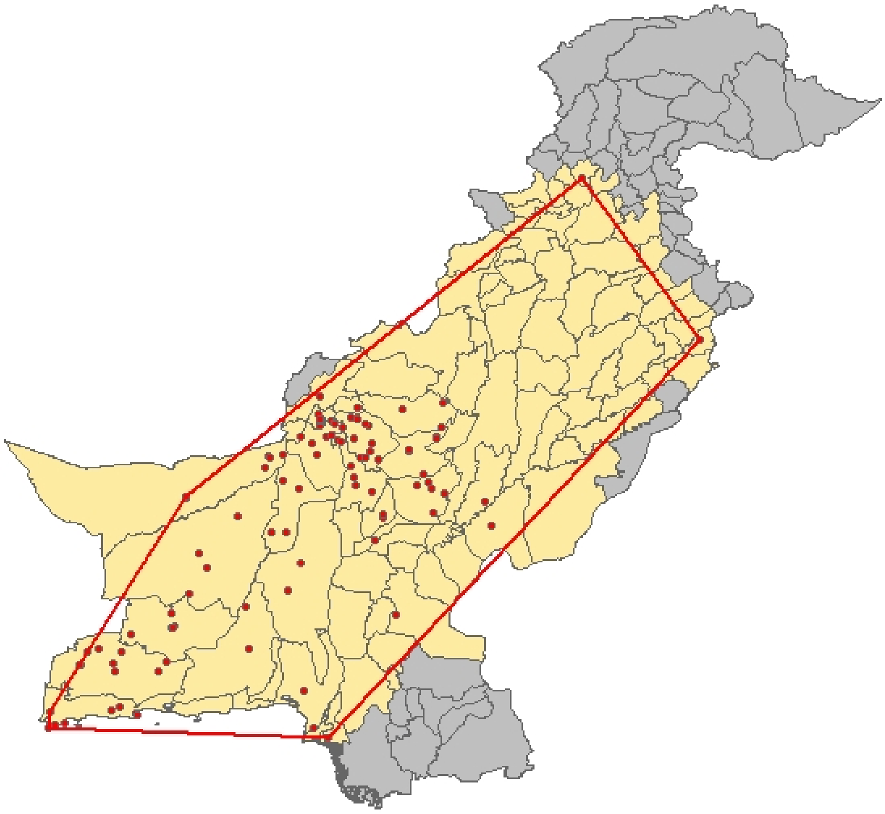
'Potential' active armed groups & one-hit wonders: We propose an alternative measure of active groups, which we call "potential active groups". We define a group as potentially active in all districts in which it has ever been active (during our sample) if it commits at least one incident in any district in the current year. The number of potentially active groups is then simply the sum of all potential groups within a district. Note that this measure will, by definition, always be greater or equal to the number of active groups within a district, since being active in one district automatically assumes potential group activity in any other district where the group has ever been active during our sample period.

Defining a group's area of potential presence as the set of districts in which a group has committed at least one incident is not our only option. We can also define an area of operations in which we count a group as active whenever it is active somewhere. We follow König et al. (2017) and define a group's area of operations as the convex hull drawn around its incidents over the sample period. Specifically, we treat all districts as belonging to the group's area of operations if the convex hull of incidents intersects

[48] Groups could also have the goal to be unpredictable (Jaeger and Paserman, 2008).

with them. Figure C-1 below illustrates the area of operations for the BLA based on this definition. Note that the convex hull is at the opposite end of our standard measure of group presence. In the case of the BLA three single attacks outside Baluchistan are enough to flag many districts outside of Balochistan with a potential BLA presence. Note that "one-hit wonders" are not affected by either of the potential group measures, since they commit by definition only one incident. Nonetheless, we can include them in the main count of active armed groups and test if our results are affected.

Figure C-1
Convex hull of BLA incidents



Notes: Depicts the convex hull of all BLA incidents (red line) and the incident locations of the BLA (red dots). Districts intersected by the convex hull are colored and counted as the area of operations for the BLA following this approach.

We replicate columns 5 and 6 of Table I using the different measures of active armed groups and report the second stage results in Table C-1. Columns 1 to 4 show that using either the set of ever-active districts or the convex hull makes little difference. The point estimates are roughly 20% smaller, although we cannot reject that they are not statistically different from our baseline specification. Interestingly, the first-stage F-stats on our instruments show that we actually have more power when using the potential number of active armed groups. Including one-hit wonders has a negligible effect. The point coefficients as well as the standard errors remain virtually the same.

In summary, our results seem not to be driven by our specific choice of group count but are robust to various sensible perturbations.

TABLE C-1
Potential active armed groups

	Potential groups		*Potential groups (hull)*		*Incl. one-hit-wonders*	
	Dependent variable:					
	Ln incidents	Ln casualties	Ln incidents	Ln casualties	Ln incidents	Ln casualties
	(1)	(2)	(3)	(4)	(5)	(6)
No. active groups	0.4225	0.5670	0.4506	0.5871	0.5488	0.7362
	(0.0874)	(0.1395)	(0.0890)	(0.1295)	(0.1290)	(0.1837)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	3384	3384	3384	3384	3384	3384
F-stat IV	28.98	28.98	23.10	23.10	16.53	16.53

Notes: The table reports the results of regressing the log of incidents + 1 and the log of casualties + 1 on the number of active groups. Columns 1 and 2 use the measure of potential active groups. Columns 3 and 4 use the measure of potential active groups based on convex hulls of group activity. Columns 5 and 6 include one-hit wonders in the incident and casualty counts. The first stage instrument is the dynamic version of $BLA\ share_i \times BLA\ split_t$ in which each post-treatment period is allowed to have a different effect on the number of groups. Standard errors are clustered at the district level in parenthesis.

D. Data appendix

Table D-1
Group matches GTD & GED

No.	GTD Group Name	GED Group Name	Matched Group Name
1		Abdullah Azzam Brigades	Abdullah Azzam Brigades
2		Abu Hafs Katibatul al-Ghurba al-Mujahideen	Abu Hafs Katibatul al-Ghurba al-Mujahideen
3		al-Intiqami al-Pakistani	Al-Intiqami al-Pakistani
4		Al-Intiqami al-Pakistani	Al-Intiqami al-Pakistani
5		Al-Jihad	Al-Jihad
6		Al-Mansoorian	Al-Mansoorian
7		Al-Qa‘ida	Al-Qaida
8		Al-Qaida	Al-Qaida
9	Al-Qaida	Al-Qaida	Al-Qaida
10		Al-Qaida in the Indian Subcontinent	Al-Qaida
11		Al-Qa’ida in the Indian Subcontinent	Al-Qaida
12		Amr Bil Maroof Wa Nahi Anil Munkir	Amr Bil Maroof Wa Nahi Anil Munkir
13	Ansaar ul-Islam		Ansaar ul-Islam
14		Ansarul Islam (Pakistan)	Ansaar ul-Islam
15		Ansar Al-Mujahideen	Ansar Al-Mujahideen
16		Ansar al-Sharia	Ansar al-Sharia
17			Ansar Wa Mohajir
18		Ansar Wa Mohajir (Pakistan)	Ansar Wa Mohajir
19		Ahle Sunnat Wal Jamaat (ASWJ-Pakistan)	ASWJ
20		Sipah-e-Sahaba/Pakistan (SSP)	ASWJ
21		Baba Ladla Gang	Baba Ladla Gang
22	Baloch Ittehad		Baloch Ittehad
23		Baloch Militant Defense Army	Baloch Militant Defense Army
24		Baloch Mussalah Diffah Tanzim (BMDT)	Baloch Mussalah Diffah Tanzim (BMDT)
25		Baloch National Liberation Front	Baloch National Liberation Front
26		Baloch Liberation Army (BLA)	BLA
27	BLA		BLA

Continued on next page

Appendix D-1 – continued from previous page

No.	GTD Group Name	GED Group Name	Matched Group Name
28		Baloch Liberation Front (BLF)	BLF
29	BLF		BLF
30		Baloch Liberation Tigers (BLT)	BLT
31		Balochistan Liberation United Front (BLUF)	BLUF
32		Baloch Republican Army (BRA)	BRA
33		Baloch Republican Party	BRA
34	BRA		BRA
35		Baloch Republican Guards (BRG)	BRG
36		Baloch Waja Liberation Army (BWLA)	BWLA
37		Baloch Young Tigers (BYT)	BYT
38	Fedayeen Islam		Fedayeen Islam
39	Forces of Momin Afridi		Forces of Momin Afridi
40	Forces of Shah Sahib		Forces of Shah Sahib
41	Forces of Turkestan Bhittani		Forces of Turkestan Bhittani
42		Free Balochistan Army (FBA)	Free Balochistan Army (FBA)
43	Government of Afghanistan		Government of Afghanistan
44	Government of India		Government of India
45	Government of Iran		Government of Iran
46	Government of Iraq		Government of Iraq
47	Government of Pakistan		Government of Pakistan
48	Government of United States of America		Government of United States of America
49		Hafeez Brohi Group	Hafeez Brohi Group
50		Hafiz Gul Bahadur Group	Hafiz Gul Bahadur Group
51		Halqa-e-Mehsud	Halqa-e-Mehsud
52		Haqqani Network	Haqqani Network
53		Harakat ul-Mujahidin (HuM)	Harakat ul-Mujahidin (HuM)
54		Harakat ul-Mujahidin Al-Almi	Harakat ul-Mujahidin Al-Alami
55		Harkatul Jihad-e-Islami	Harkatul Jihad-e-Islami
56	Hizb-i Islami-yi Afghanistan		Hizb-i Islami-yi Afghanistan
57		Islami Jamiat-e-Talaba (IJT)	IJT
58	IMU		IMU

Continued on next page

Appendix D-1 – continued from previous page

No.	GTD Group Name	GED Group Name	Matched Group Name
59		Islamic Movement of Uzbekistan (IMU)	IMU
60	IS		IS
61		Khorasan Chapter of the Islamic State	IS
62		Jaish al-Umar (JaU)	Jaish al-Umar (JaU)
63		Jaish as-Saiyouf (Army of Swords)	Jaish as-Saiyouf (Army of Swords)
64		Jaish Usama	Jaish Usama
65		Jaish-e-Islam	Jaish-e-Islam
66	Jaish-ul-Islam		Jaish-ul-Islam
67		Jamaat-E-Islami	Jamaat-E-Islami
68	Jamaat-ul-Ahrar		Jamaat-ul-Ahrar
69		Jamaat-ul-Ahrar	Jamaat-ul-Ahrar
70		Jeay Sindh Qaumi Mahaz (JSQM)	Jeay Sindh Qaumi Mahaz (JSQM)
71		Jaish-e-Khorasan (JeK)	JeK
72		Jaish-e-Mohammad (JeM)	JeM
73	Jondullah		Jondullah
74		Jundallah (Pakistan)	Jondullah
75		Khatm-e-Nabuwat (KeN)	Khatm-e-Nabuwat (KeN)
76		Lashkar-e-Balochistan	Lashkar-e-Balochistan
77	Lashkar-e-Islam		Lashkar-e-Islam
78		Lashkar-e-Islam (Pakistan)	Lashkar-e-Islam
79		Lashkar-e-Jarrar	Lashkar-e-Jarrar
80		Lashkar-e-Omar	Lashkar-e-Omar
81		Lashkar-e-Taiba (LeT)	Lashkar-e-Taiba (LeT)
82		Lashkar-e-Jhangvi	LeJ
83	LeJ		LeJ
84		Majlis-e-Askari	Majlis-e-Askari
85		Majlis-e-Lashkari	Majlis-e-Lashkari
86		Mohajir National Movement	MQM
87	MQM		MQM
88		Muttahida Qami Movement (MQM)	MQM
89	MQM-H		MQM-H
90		Mujahideen Ansar	Mujahideen Ansar
91		Mutahida Majlis-e-Amal	Mutahida Majlis-e-Amal
92		Afghans	NA
93		Baloch Nationalists	NA
94		Bandits	NA
95	Bangesh		NA
96		Bhittani tribe	NA
97		Brelvi Muslims	NA

Continued on next page

Appendix D-1 – continued from previous page

No.	GTD Group Name	GED Group Name	Matched Group Name
98	Civilians		NA
99		Gunmen	NA
100		Individual	NA
101	Kachai sub-tribe of Bangesh		NA
102		Kalpar Tribesmen	NA
103	Kashmir insurgents		NA
104	Lashkar of Akakhel tribe		NA
105	Lashkar of Akakhel tribe		NA
106	Lashkar of Kukikhel clan		NA
107	Lashkar of Mohmand tribe		NA
108	Lashkar of Orakzai tribe		NA
109	Lashkar of Salarzai tribe		NA
110	Lashkar of Zakakhel tribe		NA
111	Lashkhar of Wazir tribe		NA
112	Laskhar of Masozai Qaumi tribe		NA
113		Mahsud Tribe	NA
114	Mangal		NA
115		Militants	NA
116		Miscreants	NA
117	Mishti		NA
118	Mohajir		NA
119		Muslim Extremists	NA
120		Muslim extremists	NA
121		Muslim Fundamentalists	NA
122		Muslim Militants	NA
123		New People's Army (NPA)	NA
124		Orakzai Freedom Movement	NA
125		Other	NA
126	Pashtun		NA
127		Separatists	NA
128		Shia Muslim extremists	NA
129		Shiite Muslims	NA
130	Sindhi		NA
131		Sunni Muslim extremists	NA
132		Sunni Muslims	NA
133	Supporters of MQM		NA

Continued on next page

Appendix D-1 – continued from previous page

No.	GTD Group Name	GED Group Name	Matched Group Name
134	Supporters of PPP		NA
135	Supporters of Yousaf Ali Khan Magsi		NA
136	Supporters of Zulfikar Ali Khan Magsi		NA
137		Tribal Group	NA
138		Tribesmen	NA
139	Turi		NA
140		Youths	NA
141		Pakistani People's Party (PPP)	Pakistani People's Party (PPP)
142		People's Amn Committee	People's Aman Committee
143		Qari Kamran Group	Qari Kamran Group
144		Sindh Liberation Front	Sindh Liberation Front
145		Sindh Revolutionary Army	Sindh Revolutionary Army
146		Sindhu Desh Liberation Army (SDLA)	Sindhu Desh Liberation Army (SDLA)
147		Sindhudesh Revolutionary Army (SRA)	Sindhudesh Revolutionary Army (SRA)
148		Sipah-I-Mohammed	Sipah-I-Mohammed
149		Punjabi Taliban	Taliban
150	Taleban		Taliban
151		Taliban	Taliban
152		Tehrik-i-Taliban Pakistan (TTP)	Taliban
153	TTP		Taliban
154		Tanzeem al-Islami al-Furqan	Tanzeem al-Islami al-Furqan
155	Tawheed ul-Islam		Tawheed ul-Islam
156	Tawheed ul-Islam		Tawheed ul-Islam
157		Tawheedul Islam	Tawheed ul-Islam
158		Tehrik-e-Khilafat	Tehrik-e-Khilafat
159		Tehrik-e-Taliban Islami (TTI)	Tehrik-e-Taliban Islami (TTI)
160		Tehrik-e-Tuhafaz	Tehrik-e-Tuhafaz
161		Tehrik-e-Tuhafaz (Pakistan)	Tehrik-e-Tuhafaz
162		Tehrik-e-Nifaz-e-Aman Balochistan-Jhalawan Brigade (TNAB-Jhalawan Brigade)	TNAB-Jhalawan Brigade
163		Tehreek-e-Nafaz-e-Shariat-e-Mohammadi (TNSM)	TNSM

Continued on next page

Appendix D-1 – continued from previous page

No.	GTD Group Name	GED Group Name	Matched Group Name
164		Tehrik-e-Nafaz-e-Shariat-e-Mohammadi (TNSM)	TNSM
165	TTP-Islahi		TTP-Islahi
166	TTP-KM		TTP-KM
167	TTP - MR		TTP-MR
168	TTP - MT		TTP-MT
169	TTP-SM		TTP-SM
170	TTP - TA		TTP-TA
171	UBA		UBA
172		United Baloch Army (UBA)	UBA
173		Uzair Baloch Gang	Uzair Baloch Gang
174		Zehri Youth Force (ZYF)	Zehri Youth Force (ZYF)

Table D-2
Observed groups splits and reasons to split

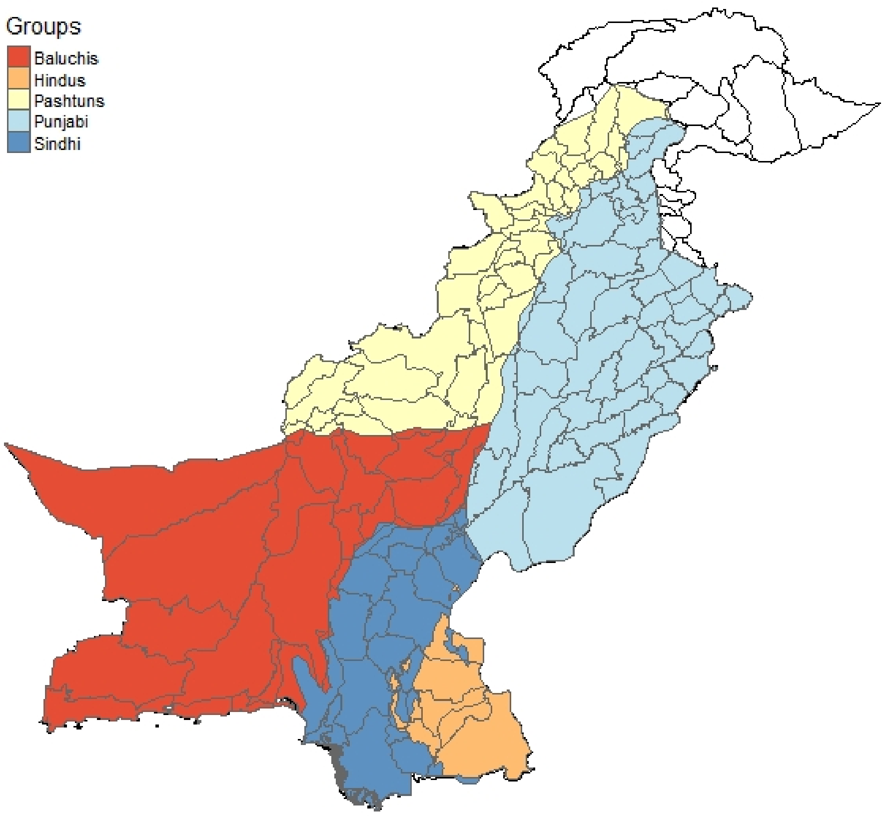
Child group	Split year	Mother group	Reason to split	Details	Source
United Baloch Army (UBA)	2014	Baloch Liberation Army (BLA)	Natural death of leader's father	The UBA split from the BLA after the father of both group leaders died. The group leaders are brothers and could not agree on who should lead the BLA.	Stanford-Mapping Militants
Jaish-e-Mohammad (JeM)	2000	Harakat ul-Mujahidin (HuM)	Lost funding by ISI	"Some sources claim that ISI [Pakistan's Inter-Services Intelligence] lost interest in funding HuM after Khalil's [founder of HuM] 1998 decision to join hands with Bin Laden. ISI may have offered Azhar assistance and funding to establish JeM following his release from prison."	Stanford-Mapping Militants
Harakat ul-Mujahidin Al-Almi	2002	Harakat ul-Mujahidin (HuM)	Dispute over organizational affairs.	"There was reportedly some pressure on the HuM after its proscription in Pakistan in 2001 to merge with the Jamiat-ul-Mujahideen. This plan met with stiff resistance from within the HuM and reportedly, the dissent led to a group breaking away from the parent outfit and calling itself the Harkat-ul-Mujahideen Al-alami."	SATP
Jundullah	2011	Lashkar-e-Jhangvi (LeJ)	No reason found		TRAC
Lashkar-e-Balochistan	2011	Lashkar-e-Jhangvi (LeJ)	No reason found		TRAC
Lashkar-e-Jhangvi (LeJ)	1996	Sipah-e-Sahaba/Pakistan (SSP)	Ideology conflicts	"Former SiS [Sipah-i-Sahaba] militants Riaz Basra, Malik Ishaq, and Akram Lahori founded LeJ in 1996 after breaking away from SiS, claiming that SiS had deviated from its founder's teachings."	Stanford-Mapping Militants
Jamaat-ul-Ahrar	2014	Tehrik-i-Taliban Pakistan (TTP)	Leadership dispute after the killing of the former leader.	"JA split from the TTP under the leadership of former TTP commander Omar Khalid Khorasani. This separation was a result of growing tensions between Khorasani and the then leader of the TTP, Maulana Fazlullah." "The broadness of the TTP's coalition also presents challenges, however, and has threatened the group's cohesion. For instance, after the death of its former amir Hakimullah Mehsud in a U.S. drone strike in November 2013, the jihadist leaders of several key TTP factions failed to reach a consensus over who should head the group."	Stanford-Mapping Militants, UNHCR-Refworld
Jundallah (Pakistan)	2014	Tehrik-i-Taliban Pakistan (TTP)	"Self-reinvigoration through ISIS" and anti-Shi'a ideals.	"When ISIS captured Mosul in July 2014, Jundullah was one of the first organizations that pledged allegiance to Abu Bakr al-Baghdadi. Due to Jundullah's strong ties to al-Qaeda, their decision to shift alliances was probably a difficult one. However, because the group lost most of its core leadership due to severe actions taken by Pakistani law enforcement, this move indicates a policy of self-reinvigoration through ISIS." "Jundallah is likely to be partly comprised of cadres from banned sectarian Deobandi tafkiri groups like LeJ or Ahle-Sunnat-Wal-Jamat (ASWJ), which consider Shi'a Muslims to be kafirs[meaning: 'infidel' or disbelievers], underlining that the group already had strong sectarian leanings even before the advent of the Islamic State."	Washington Institute, UNHCR-Refworld

TABLE D-3
Observed groups mergers and reasons to merge

Lead group	**Merge year**	**Merging group**	**Reasons to merge**	**Details**	**Source**
Tehrik-i-Taliban Pakistan (TTP)	2011	Harakat ul-Mujahidin (HuM)	Common aims and enemies and concentration of power	"Media reports on January 5, 2011, indicated that five terrorist groups had joined the TTP and were working under its umbrella TTP. With common aims and enemies, LeJ, SSP, JeM, HuM and Harkat-ul-Ansar (HuA) had 'merged' with TTP. TTP spokesman Azam Tariq declared, 'We have not forced anyone to join TTP, and the leaders and activists of the banned religious organisations have united themselves under the umbrella of the TTP on their own choice.'" "The sole objective of the Shura meeting was to unite the small militant fractions under the leadership of TTP against NATO forces in Afghanistan and to wage a defensive jihad against Pakistani forces."	SATP
Tehrik-i-Taliban Pakistan (TTP)	2011	Jaish-e-Muhammad (JeM)			
Tehrik-i-Taliban Pakistan (TTP)	2011	Lashkar-e-Jhangvi			
Tehrik-i-Taliban Pakistan (TTP)	2011	Sipah-e-Sahaba/Pakistan (SSP)			
Tehrik-i-Taliban Pakistan (TTP)	2015	Jamaat-ul-Ahrar (JuA)	Reconciliation through government operations and leadership dispute resolution	The government's commencement of the Zarb-e-Azb operation in North Waziristan district, and supplementary operations in other districts of tribal areas, served to soften the TTP's differences over the leadership and to bind these groups together against a common enemy. In addition, Fazlullah carrying out the December 16, 2014 attack on the Army Public School in Peshawar, in which 141 people (a large number of them children) were killed, outclassed all other competing jihadist groups, and Fazlullah thereby proved his mettle to rule TTP.	UNHCR-Refworld
Tehrik-i-Taliban Pakistan (TTP)	2015	Lashkar-e-Islam (LeI)	Re-organisation as a result of significant gains by security forces	Militant organization Lashkar-i-Islam (LI) has merged with the TTP as part of a re-organization plan. The decision to unify the militant groups was taken at a meeting attended by TTP leaders Mulla Fazlullah, Omar Khalid Khurasani, and LI leader Mangal Bagh. The militants announced the unification at a time when security forces are making significant gains against them in military operations in North Waziristan and Khyber Agency (government's Operation Khyber I), which were once considered their bastions.	

Politically excluded population: Our basis for constructing the politically excluded population across districts over time is the Ethnic-Power-Relations (EPR) data (Wucherpfennig et al., 2011). Specifically the geocoded version thereof (geoEPR) (Vogt et al., 2015). The EPR tracks the political access of politically relevant groups to the central state. It codes the political power of groups with some exceptions on an ordinal scale. The set of group power status are: "Monopoly", "Dominance", "Senior Partner", "Junior Partner", "Powerless", "Discriminated", "Self-exclusion" and "Irrelevant" (see Vogt et al., 2015). The first two categories imply that a group has more or less exclusive access to power, while the "partner" categories are assigned to groups that share power in government. The remaining groups are (apart from the Self-exclusion category) self-explanatory. In our sample, we only observe groups being either a junior or a senior partner (the Punjabi for the entire period), discriminated against (the Baluchi for most of the time), or powerless. Figure D-1 reports the ethnic groups across our districts.

FIGURE D-1
Politically relevant ethnic groups in Pakistan (GeoEPR)



Notes: Reports the intersection between our district shape and the GeoEPR shape.

Two issues arise when using the GeoEPR in our setting. First, the GeoEPR measures access to political power for ethnic groups and not for districts. Hence, we need to aggregate the representation of the groups living within a district to the district level. Second, the GeoEPR's time coverage only extends to 2017 and has no data for several districts.[49] Thus, we do not include the GeoEPR variables in our baseline specifications. We aggregate group representation to the district level by creating a set of dummies representing each power category in our sample and weight each group's power with its area share (or 1990 population share) in the respective districts. In the case of the population share, we simply use the population estimated to reside on the intersection of a specific geoEPR group polygon with a district. The population estimate is based on Global Human Settlement Layer (GHSL). The data is constructed by the Joint Research Centre and the Directorate General for Regional and Urban Policy of the European Commission. It can be accessed at https://ghsl.jrc.ec.europa.eu. Both procedures provide us with 4 dummy variables for the specific power statuses observed in our sample. To obtain a suitable measure of the politically excluded population, we simply combine the shares of the "Powerless" and "Discriminated" populations within districts.

Capital intensive & non-capital intensive incidents are defined following Limodio (2022). The GTD lists attack types in which it describes the primary weaponry used in an incident. Capital intensive incidents are all incidents which are either labeled as; *i*) "Bombing/Explosions", *ii*) "Unarmed Assault", or *iii*) "Assassination". For further details see Limodio (2022). Non-capital intensive incidents in turn are all incidents that are not labeled capital intensive and for which the attack type is not missing. Since there are attacks that are missing the attack type, capital intensive and non-capital intensive attacks do not sum up to total attacks.

[49]The white areas in Figure D-1.