Gassebner, Martin; Schaudt, Paul; Wong, Melvin H. L.

Working Paper

Armed Groups in Conflict: Competition and Political Violence in Pakistan

CESifo Working Paper, No. 8372

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Gassebner, Martin; Schaudt, Paul; Wong, Melvin H. L. (2020) : Armed Groups in Conflict: Competition and Political Violence in Pakistan, CESifo Working Paper, No. 8372, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/223444

CESifo WORKING PAPERS

8372
2020

June 2020

Armed Groups in Conflict: Competition and Political Violence in Pakistan

Martin Gassebner, Paul Schaudt, Melvin H. L. Wong

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8372*

Armed Groups in Conflict: Competition and Political Violence in Pakistan

Abstract

This paper studies how an increase in the number of armed groups operating within an area affects the amount of organized political violence. We use plausible exogenous variation in the number of armed groups in Pakistan, by exploiting the split of a major group due to the natural death of its leader. Employing difference-in-difference and instrumental variable regressions on geocoded incident and fatality data allows us to derive a causal effect: more groups lead to more political violence. By combining different data sources and implementing a new approach to deal with potential double-counting, we provide a proxy for counter-insurgency efforts by the government. We show that the increase in violence is primarily driven by the armed groups and not by responses of the government.

JEL-Codes: D740, F520, H560.

Keywords: political violence, conflict, terrorism, armed groups, double-counting.

*Martin Gassebner**
Institute of Macroeconomics
Leibniz University Hannover / Germany
gassebner@mak.uni-hannover.de

Paul Schaudt
Department of Economics
University of St. Gallen / Switzerland
paul.schaudt@unisg.ch

Melvin H. L. Wong
Institute of Macroeconomics
Leibniz University Hannover / Germany
wong@mak.uni-hannover.de

*corresponding author

June 2020
We wish to thank Roland Hodler, Tobias Korn, and Todd Sandler for excellent comments. Special thanks to Baloch journalist Malik Siraj Akbar who gave us deep insights into the conflict in Balochistan and the different armed groups. Finally, we like to thank Lukas Bertram, Tom Görges, Linyun Yu for excellent research assistance.

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, or 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 threat can be amplified if the additional group 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 existing players but threatens their distinct support base, e.g., financial supporters and recruits. 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.[3]

Empirical evidence is still limited, despite strong priors that more armed groups intensify political violence. Currently, the literature can only report positive correlations between the number of armed groups and the frequency and severity of political violence.[4] The main problem in estimating the causal effect of an increase in the number of armed groups on political violence is that the number of armed groups within a given geographic area is endogenous. First and foremost, groups most likely actively select themselves into given areas Gaibulloev2015. The selection in turn reasonably depends on the strength of incumbent actors as well as attributes inert to the area in question.[5] Second, groups have different goals and strategies, respond to incentives differently, and might have different support groups (see Kis-Katos et al., 2014; Toft and Zhukov, 2015). Hence, new groups might be formed to cater to previously neglected interests and grievances. Finally, political violence itself affects the number of armed groups, as some groups bleed out during a conflict, or are attracted by the fighting itself, e.g., hunting their enemies across locations.

This paper provides causal 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

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

[3] Outbidding theory predicts that the amount and severity of violence will increase as the competing groups aim to demonstrate that they are the most effective agent for the political goals of a particular population (Kydd and Walter, 2006). The logic behind outbidding theory is that groups would like to commit less costly attacks as long as they can obtain their political goals.

[4] See Findley and Young (2012); Nemeth (2014); Conrad and Greene (2015).

[5] Such as weak state capacity (Fearon and Laitin, 2003) resulting from a higher distance to the capital (Campante et al., 2019), the presence of lootable resources (Berman and Couttenier, 2015), and the hostility or acceptance of the local population (Berman et al., 2011).

primarily within 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 died of natural causes following a relatively short and severe illness.[6]

The exogenous timing of the father's death and the geographically concentrated area of operations of both groups allow us to specify difference-in-difference (diff-in-diff) specifications. We estimate the causal effect of an additional armed group on the quantity and lethality of political violence within the districts of Pakistan.[7] We argue that the main channel underlying this effect is competition (for publicity, recruits, and/or financing) between armed groups. Furthermore, we estimate semi-elasticities between the number of armed groups and political violence using the geographically differential impact of the father's death as an instrument for the number of armed groups. Additionally, concerns about unobserved confounders explaining group formation are alleviated since the general goals, target audience, primary opponent, and tactics of the BLA and UBA are similar.[8]

The 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.[9]

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 to proxy for government counter-insurgency efforts. We differentiate insurgency from counter-insurgency by exploiting the different inclusion criteria of events for each database.

Our results show that, on average, the BLA split increases the number of incidents within districts in Balochistan compared to the districts outside of Balochistan by 3.5.

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

[7]We introduce the setting and the involved actors in detail in Section 2.

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

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

This corresponds to an increase of roughly 130%, both in the reduced form specifications and in the 2SLS models in which we estimate the effect of an additional armed group. To put this effect into perspective, an increase in the number of armed groups by one is, on average, an increase in the number of armed groups operating within a given district by 37 %. Accounting for counter-insurgency measures highlights that the violence is driven by armed groups and not a reaction of the government to the split of the BLA.

We contribute to various strands of the literature. Our results confirm that the proliferation of armed groups increases organized political violence, adding additional insight to a vast literature on the determinants of political violence (see Hegre and Sambanis, 2006; Blattman and Miguel, 2010; Gaibulloev and Sandler, 2019, for excellent overviews). Note that group proliferation is a potential omitted variable in many studies and cannot be captured by fixed effects in monadic settings. Furthermore, the issue cannot be resolved by focusing on smaller units such as grid-cells.[10]

Methodologically, we provide a novel approach of incidence matching to combine two widely used datasets on armed groups. The main problem in combining different data sources, such as the GTD and GED, is the possibility of double-counting incidents and thus the introduction of measurement error. As we highlight in Section 3, the definitions of political violence and inclusion criteria used by the different datasets overlap. Hence, multiple data sources potentially include the same incidences of some groups in some regions. Such double-counting would distort the relative activity of a specific group. We provide a data-driven solution for addressing the threat of double-counting when using multiple data sources of political violence. Specifically, we implement an uncertainty-based measure applying spatial and temporal buffers surrounding each incident within the GTD database and flag all incidents included in the GED dataset involving the same group that fall within the joint buffer. Adjusting the buffer sizes allows trading off false-positive vs. false-negative assignments of double-counts. Double-counting in Pakistan applies to 5% to 10% of the GED incidents for reasonable parameters, but our main results are robust to the issue over a wide range of thresholds.[11] The cleaned match further allows to identify incidents with government involvement which are not treated as terrorist incidents by the GTD. Hence they are plausibly initiated by the government and can be used to proxy for counter-insurgency efforts. Our approach is general and allows other authors to analyze directional incidents of political violence, previously limited to specific settings (Jaeger and Paserman, 2008; Abrahams et al., 2019).

We also provide evidence of interactions between groups in a unique setting, as there are several armed groups in Balochistan fighting the government, attacking civilians and businesses, but only sporadically each other. What is more, there are up to five established

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

[11]The range between 5% and 10% is based on the procedure including name matches and distance thresholds no bigger than 32km (see Appendix B, which we find most reasonable).

separatist groups operating in Balochistan. These groups are supported by local financiers to some degree and strife for a more autonomous Balochistan. Thus, our setting is markedly different from related studies in which the conflict parties, however defined, are engaged in a battle royal, try to loot local resources in areas that they capture (as in Morelli and Rohner, 2015; Adhvaryu et al., 2018; Gehring et al., 2019), and engage in shifting alliances (König et al., 2017).

Finally, we provide new time-variant data on the armed group level, including mergers and splits for armed groups in Pakistan. The field has made tremendous strides identifying the effects of local conditions and temporal shocks, and static group characteristics affect political violence.[12] Still, the literature is mostly silent on changes within the actors that organize, political violence on a broader scale; the armed groups themselves. 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.

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 and discusses threats to identification. Section 6 explores alternative mechanisms, and estimates semi-elasticities for an additional armed group. Finally, Section 7 provides a brief overview of the robustness tests and Section 8 concludes.

2. Background: Setting

The Balochistan conflict is an ethnic dispute concentrated in the Balochistan province[13] of Pakistan[14]. It started in 1948 when newly independent Pakistan annexed the autonomous Baloch state of Kalat. Since the start, there have been several violent waves 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 burned out, 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 independent Balochistan. Among the many reasons for the insurgency is systemic repression and marginalization

[12] See Berman et al. (2011); Dube and Vargas (2013); Berman and Couttenier (2015) for examples on conditions and shocks, while Kis-Katos et al. (2014); Fortna (2015); Polo and Gleditsch (2016) examine characteristics.

[13] 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.

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

of Baloch people as well as the exploitation of natural resources without improvements in local living conditions, an issue that has continuously been raised since the 1960s.[15] 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). The groups' area of operation is concentrated on locations 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 Baloch Province, but only rarely fight each other. Hence, visibility is key for each group. A decrease in media attention decreases the attention pay-offs for a group, which may reduce the group's capabilities (Jetter, 2017). Attacks on protected government institutions and incidences with high fatalities 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 middle-class Baloch (Economist, 2012).

The set of Baloch insurgency groups, apart from the appearance of the UBA, has remained constant since 2005. Let us now discuss how the UBA came to enter the conflict, and if it is 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 correspondence with Baloch journalist Malik Siraj Akbar revealed that the BLA recruited from non-Marri tribes starting from 2006 onwards. Some members did not agree with recruiting people that are outside their tribe.

[15]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.

Mehran Marri supposedly stole weapons and money as well to form his own group, the UBA. The UBA seems to be part of the BLA as a faction, given that KBM asked the BLA leader to pardon his younger brother's theft and uprising (Ali, 2015; Nabeel, 2017). A joint truce statement of the BLA and UBA in 2018 reveals that leadership disputes lead to the formation of the UBA in 2011(Balochistan Post, 2018).

The real split and alienation occurred after the death of the brothers' father. Only five months after the death of KBM, both groups started to attack each other (see START, 2019; Sundberg and Melander, 2013). We later address anticipation effects of KBM's death and find no evidence that groups acted independently beforehand. The critical question is whether KBM died of a natural cause. All accounts support this: he died at age 86 due to a brain hemorrhage in June 2014 (Khan, 2014). Such cerebral bleed 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. Thus, we conclude that the splitting of the BLA and UBA is driven by the natural death of the leader's father.

In summary, the timing of the actual split between the BLA and UBA is not driven by the competition of the already established groups, nor by some external factors influencing the 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 1994 and 2018.[16] Districts in Pakistan correspond to the the third administrative layer. The main results are mostly based on a balanced panel for 141 districts between 1995 and 2018 (we lose 1994 due to differencing of some variables). The primary variables of interest are the level of organized political violence and the number of armed groups accounting for group mergers, group splits, and naming conventions (e.g., "Al-Qa'ida" vs. "Al-Qaida").

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 fatalities to measure the severity of political violence.

Data is primarily taken from the "Global Terrorism Database" (GTD) (START, 2019), and complemented by information from the "UCDP Georeferenced Event Dataset"

[16] We cannot start our sample earlier as GTD does not provide data for the year 1993, due to a loss of data (see https://www.start.umd.edu/gtd/about/). Our approach needs an uninterrupted time-series.

(GED) (Sundberg and Melander, 2013). The GTD, officially tracking terrorism, is our preferred source due to two reasons.[17] 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) in which one of the two groups is involved.[18] 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). The GTD is, however, likely to suffer from under-reporting as is common to all open source database relying on news reports to track organized political violence (Van der Windt and Humphreys, 2016). This last issue is less of a problem in our setting since our identifying variation will come from relative changes in the amount of political violence committed in treated vs. untreated districts. To the best of our knowledge, the under-reporting bias does not change differently between the treatment and control group over time and is thus unlikely to bias our results. Note that we can only use incidents from the GTD and GED which contain information on the district where they occur. This leads to 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 fatalities deserves some special consideration. First, fatalities in the databases are recorded with considerable uncertainty. Incidences are always reported if there is newspaper coverage. On the contrary, fatalities 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 estimate. If several newspapers report fatalities for an incident, the modal figure will be included in the database. Second, the number of fatalities is subject to a larger degree of randomness. While armed groups may conduct their attacks 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 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 fatality rates are difficult to predict, they are informative on the group's intention and capabilities.

A downside of the GTD database is its' focus on terror attacks. Although the applied

[17]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)

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

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 28th of June 2015, which is included in the GED but not the GTD. To answer our research question, we need to ensure 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.[19] Using both databases also allows us to test if groups switch from more symmetric political violence, i.e., actions against government troops (coded predominantly by the GED), to more asymmetric types of violence, i.e., against civilians or employing hit and run tactics (the bulk of incidents included in the GTD).

Employing two databases that track organized political violence comes at a cost, however. 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 sole assumption necessary for this approach to work is that double-counting is only an issue between databases but not within them. Appendix B discusses our approach and how double-counting affects our results in detail.

3.2. Independent variable: Number of armed groups

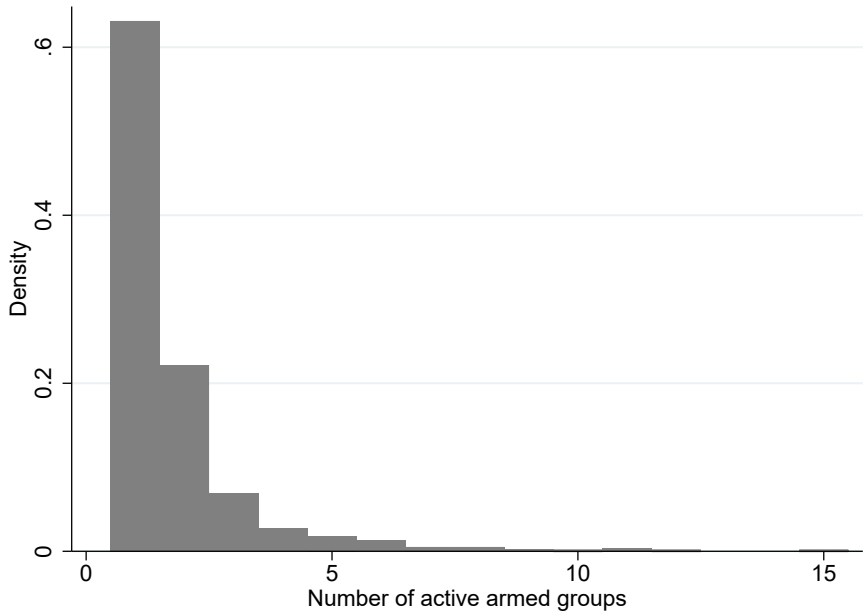
Our primary independent variable is the number of armed groups. We use these to measure the degree of competition. We consider all actors contained in the GTD and GED as armed groups if they have an individual name. That means we exclude actors such as "gunmen" or "tribesmen".[20] After independently cleaning the data, we compare our groups with the groups reported in (Hou et al., 2020) and find no omissions. In our baseline specifications, we use the active number of armed groups. 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.

Figure I depicts the distribution of the number of armed groups over district years

[19]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).

[20]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 full list of all armed groups is provided in Appendix E.

FIGURE I
Distribution: Number of armed groups



Notes: Depicts the distribution of the number of *active* groups in district-years with at least one active group. *Active* means conducting at least one attack in a given year and district.

that have a positive amount of armed groups operating within them, which are about 15% of the district-years. That is not to say that most districts never experience group activity. Of the 141 districts included in our sample, only 25% do not experience any activity during the sample period.[21]

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 ever have 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 which we are not able to observe.[22]

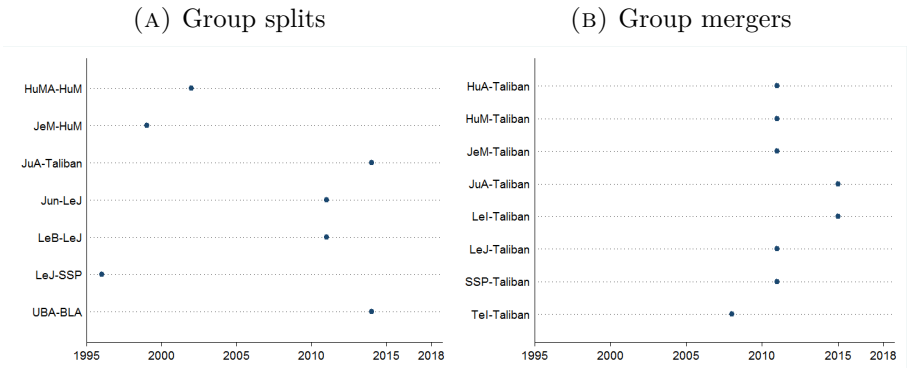
Other issues when counting the number of independent armed groups are splits and mergers of armed groups and related measurement error within our source databases. The

[21] Figure A-1 reports the active group distribution for districts in and outside of the Baloch province separately. The distribution of the number of armed groups is skewed slightly more to the right for districts within Balochistan compared to those outside of Balochistan.

[22] Appendix C shows how this changes the landscape of active groups and its effect on our core results.

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 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. To minimize the problem without reducing our sample too much, we use only data until 2018, assuming that most corrections occur within the first year, rather than later on.

FIGURE II
Armed groups splits and 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)

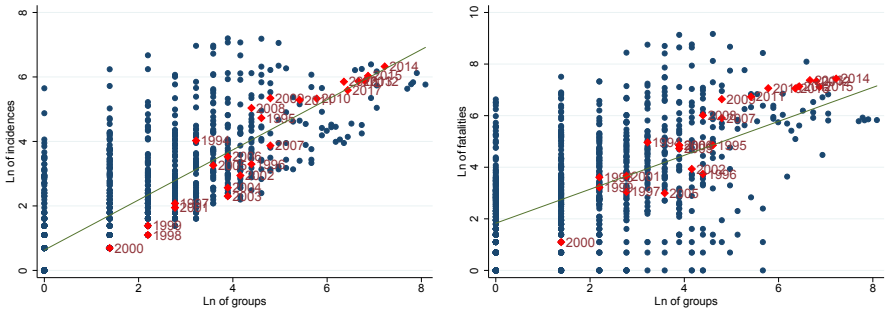
To address the issue of potential splits and mergers, we conduct an in-depth analysis of all armed groups within Pakistan and track if they split or merge with other groups during our sample period. The analysis is based on full-text online searches of major media outlets. Figure II provides an overview of the timing of all splits and mergers occurring in our sample, while Appendix F provides detailed documentation of each case. We can then reassign incidents 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, controlling for the number of groups allows us to proxy for changes in the degree of competition within our treatment and control groups unrelated to our treatment. Table A-1 shows that most group splits would neglect the increase in groups within our control group, i.e., districts outside of Balochistan. Full

descriptive statistics for our variables of interest are reported in Table A-2.

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 fatalities on the number of active armed groups at the country level for all countries included in the GTD. Figure III shows the results, highlighting Pakistan-Year observations in red.

FIGURE III
Armed groups and political violence

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



Notes: Depicts a scatter plot of unfiltered (log of) groups vs (log of) incidents & fatalities created by these groups. The unit of observation is country-year. Pakistan is represented in red. The line illustrates the best linear fit using the global GTD sample between 1994 and 2018.

The first impression is quite stark. First, there is an apparent correlation between the number of armed groups active within a country and the amount of organized political violence perpetrated. Second, Pakistan is not located in the extremes but seems to fit the linear fitted line quite well. Of course, this is only suggestive evidence on the country level, but it is supportive of notion that more armed groups lead to more political violence.[23]

4. Empirical Strategy

4.1. Estimation Framework

We are interested in the causal effect of an increase in the number of active groups on the amount of political violence within Pakistani districts. As already mentioned in the preceding sections, we exploit the fact that Khair Bakhsh Marri (KBM) died of a natural cause, which leads to the break up of the BLA into the BLA and UBA, to obtain exogenous variation in the number of competing armed groups within districts.

In the spirit of Draca et al. (2011), we will use two complementary identification strategies to answer our question. We will run panel difference-in-difference estimations in which we regress political violence on the interaction of districts in Balochistan with the

[23]The pattern is similar if we demean the measures by country and year (see Figure A-2).

post KBM death period, as well as 2SLS regressions where we use the same interaction to predict the number of armed groups active within a district. The idea behind the 2SLS approach is that the split in the BLA due to KBM's death leads to plausible exogenous variation in the number of armed groups within Balochistan compared to districts outside of Balochistan. This second strategy allows us to estimate the semi-elasticity of an additional armed group on political violence. We return to the 2SLS specification in the mechanism section of the paper and start with the reduced form specification (diff-in-diff) here.

The intuition behind the diff-in-diff specification is that the death of KBM splits the BLA. This affects the 28 districts in Balochistan more compared to districts outside of Balochistan since both the BLA and UBA almost exclusively operate in Balochistan (close to 95% of the incidents).[24] Our baseline specification is defined as follows:

$$Y_{it} = \beta(Baloch_i \times PostKBM_t) + \mathbf{X_{it}}'\psi + \eta_i + \gamma_t + \epsilon_{it} \tag{1}$$

where Y_{it} is the amount of political violence (either incidents or fatalities) perpetrated in district i during year t. $Baloch_i \times PostKBM_t$ is the interaction between districts located in the Baloch province with the *KBM post-mortem* period, $\mathbf{X}'$ 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. We also employ a set of geographic characteristics interacted with time to proxy for changes in state capacity within districts over time (Fearon and Laitin, 2003; Buhaug et al., 2009).[25] η_i and γ_t are district and year fixed effects, and ϵ_{it} is an error term assumed to be well behaved.

We also specify an event study to test for possible violations of the common trend assumption. We define:

$$Y_{it} = \sum_{s}^{S} \beta_s b_{it}^s + a_{it} + b_{it} + \mathbf{X_{it}}'\psi + \eta_i + \gamma_t + \epsilon_{it} \tag{2}$$

where b_{it}^s is our time-varying treatment indicator ($Baloch_i \times PostKBM_t$) across the sequence s.[26] s is a sequence $s = (-3, -2, 0, 1, 2, 3)$ centered at the treatment year 2014

[24] See Table A-1 for the incidences of BLA and UBA within and outside of Balochistan.

[25] 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.

We use several geographic characteristics to proxy for the government's capability to project power within a given district. We include the log of ruggedness, elevation and road density, as well as the count of airports located within a district. Note that we interact all of those time-invariant variables with year fixed effects to allow for a distinct impact each year.

[26] The treatment indicator is simply the previous introduced interaction.

for each district. We bin the endpoints of our treatment effect window with two dummies labeled a and b. a is a dummy that is unity for each district in the Baloch province following the post-treatment sequence under consideration (in our case 2018), while b is a dummy that is unity for all years before the sequence starts, i.e., 2010 or earlier. $\mathbf{X}'$ is a vector of control variables containing the same variables as in equation 1. The coefficients of interest are β_s, which report the difference-in-difference estimate separately for each year within the sequence (apart from 2013 or $t-1$, which is our omitted category). This stacked or event study approach has several advantages compared to standard difference-in-difference methods.

First and foremost, event studies provide the best insight into potential pre-treatment trends since we obtain direct estimates for them. Second, event-studies can depict the dynamics of the treatment effect and provide insight into how long it lasts.[27] However, there are also downsides. Ideally, one would have a clear idea of the timing of the effect ex-ante, i.e., for how long it should last and design the post-treatment period accordingly, but we lack such a clear prior. Note that we are also constrained since we only have four post-treatment years of data available to us. Furthermore, the estimation method is also more demanding. Borusyak and Jaravel (2017) recommend, for example, to drop the pre-treatment coefficients if the null of no pre-trends cannot be rejected in the initial event study.

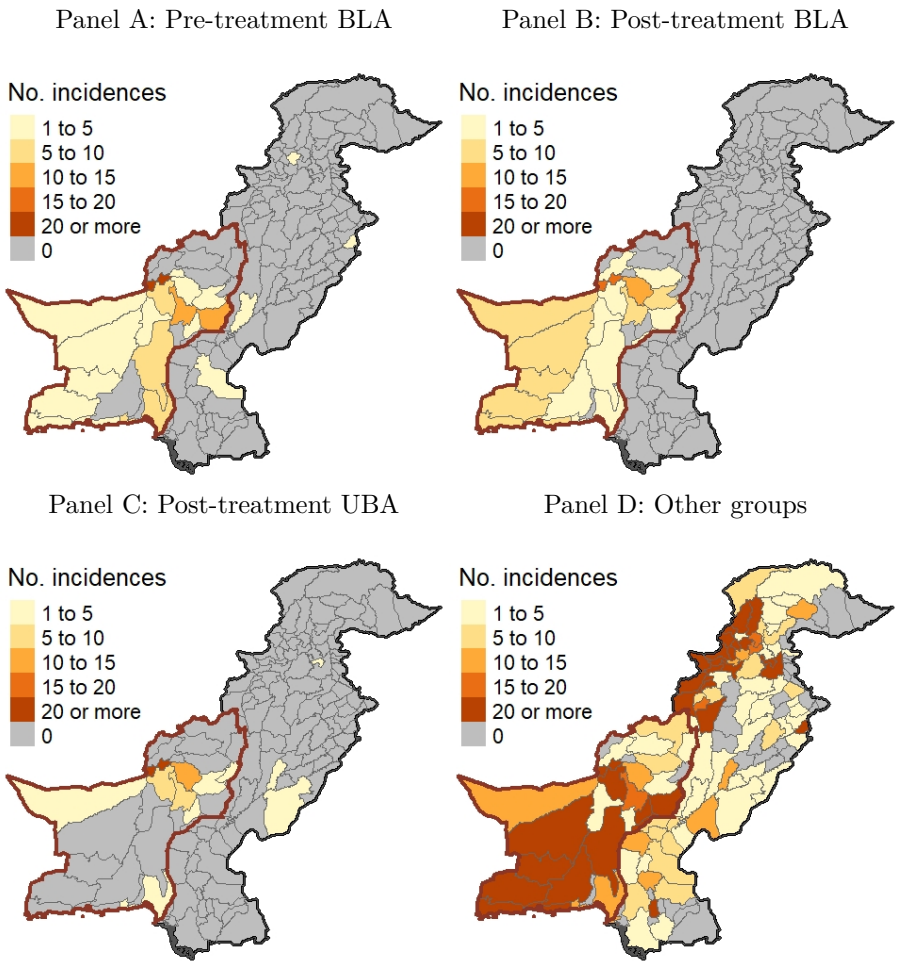
4.2. Identifying Assumptions

For standard DiD and event-study estimates to be valid, they need to satisfy two central assumptions. First, we need to assume common trends, or bias stability, which means that differences in the expected potential non-treatment outcomes over time are not related to the subsequent treatment groups. Second, we need to assume exogeneity of our control variables and pre-treatment outcomes. This implies that the timing of the treatment is exogenous. If both assumptions are satisfied, we can recover the average treatment effect on the treated (ATT).

The exogenous timing of our treatment is the main advantage of our unique setting. As outlined above, KBM died in a hospital from a brain hemorrhage. Thus, as long as we agree to assume that the brain hemorrhage was not the result of more intensive competition between the groups within the Baloch province, we have little problems in that regard. The anticipation of his death is, however, another issue. We acknowledge that KBM was already comparably old, which could have led to some of his sons anticipating his death. In such a case, one could argue that the UBA and BLA already compete before his death, i.e., when the UBA faction formed. Our analysis below will

[27] Additionally, recent studies have shown that symmetry in the pre- and post periods bring DiD estimates closer to experimental benchmarks (Chabé-Ferret, 2015).

FIGURE IV
Terrorism distribution across Pakistan

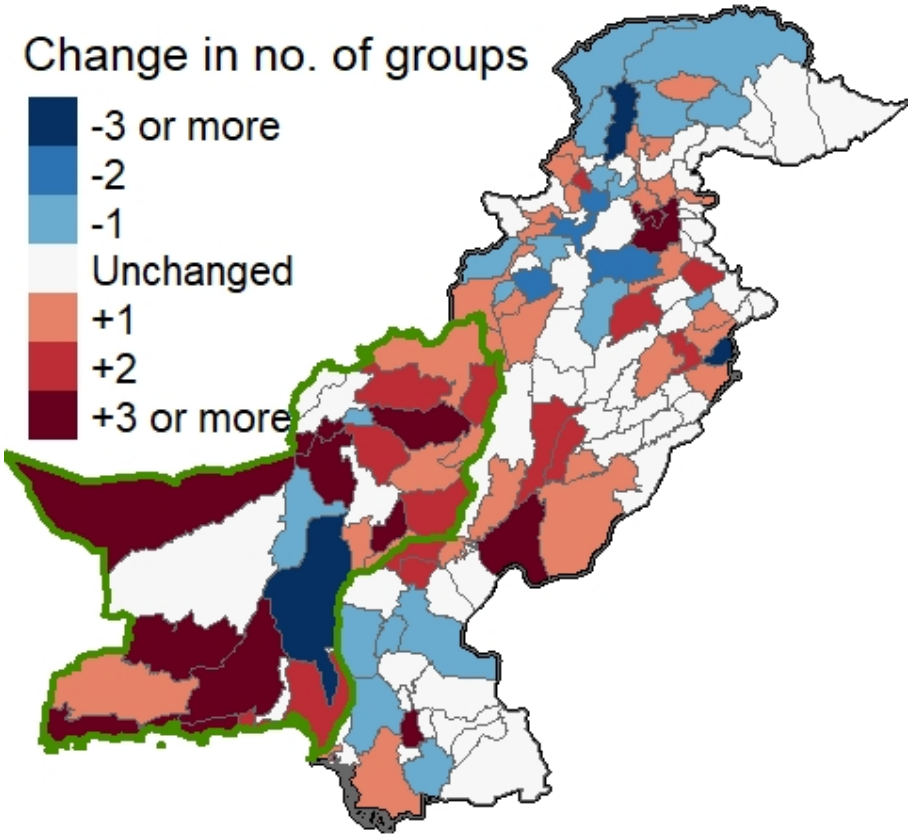


Notes: Illustrates the average number of terror events. The bold red line represents the borders of the Baloch province. The darkest shade within Balochistan in Panels A-C is Quetta, the capital.

show, however, that this seems not to be the case. The composition of the treatment groups is another crucial issue. In our main specification, we define districts in Balochistan as the treatment group, since it is the region in which the BLA has traditionally operated. If, for example, we would find that the UBA and BLA never overlap after their split, our mechanism of an exogenous increase in the number of groups could be jeopardized. Figure IV suggests that we do not face this problem, since the main area of operation remains within Balochistan. Additionally, there is a high degree of operational overlap, since almost all districts with UBA activities see BLA activities as well.

Panel A of Figure IV shows the spatial distribution of BLA attacks in the pre-treatment period, while panel B depicts the post-treatment period distribution. The activity within Balochistan has been relatively stable, while a few districts outside of Balochistan have not experienced any further violence by the BLA. In panel C, we plot the spread of UBA incident in the post-treatment period. Comparing the spatial spread

FIGURE V
Change in number of groups



Notes: Reports the change in group numbers between 2009-2013 (pre-treatment) and 2014-2017 (post-treatment). The bold green line represents the borders of the Baloch province.

of the UBA to the BLA shows that the UBA also engages in some minor activity outside of the Baloch province, but both groups share their most active districts. Finally, the last panel shows the overall distribution of incidents by other groups. It should be apparent that while Balochistan is a major hotspot for terrorism, there are other districts with a similar level of violence, e.g., the districts within the Federal Tribal Areas in the north-west (see panel D).

We provide additional suggestive evidence that our treatment operates via an increase in the number of armed groups in Figure V. Figure V plots the difference in the number of groups for the five years before treatment to the five years after treatment. As expected, after KBM's death the number of active armed groups increases most within districts where the BLA has been active. Furthermore, districts that already host more groups in the first place see a higher group count after KBM's death (Figure A-3).[28]

Finally, let us briefly discuss the parallel trends assumption and potential anticipation

[28]We directly test for the mechanism of increases in the group count after the group split using 2SLS in Table IX.

effects. Given that the UBA faction established itself already in 2011, and KBM was already relatively old, one could argue that the UBA faction was already affecting competition between groups before KBM's death. We address the issue in the next section. To preview our findings, neither parallel trends, the faction creation, nor anticipation effects seem to be a problem in our setting.

5. Results

5.1. Baseline results

Table I reports our baseline results for the difference-in-difference specifications. The coefficient in column 1 is highly statistically significant, regardless of clustering on the district or the division level, and when allowing for spatial dependence in the error terms[29]. The estimate suggests that districts within Balochistan experience an increase of around 3.5 incidents in the post-treatment period compared to districts outside Balochistan. The results for fatalities are sizeable as well, although less statistically significant. The treatment districts experience roughly 5.5 fatalities more compared to the non-treatment districts annually during the post-treatment period.

TABLE I
Baseline results

	Incidents (1)	Fatalities (2)	Ln Incidents (3)	Ln Fatalities (4)
$Baloch_i \times Post_t$	3.5904	5.5364	2.1067	1.5800
	(0.9554)***	(2.8711)*	(0.3787)***	(0.4420)***
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144
Adj. Within-R^2	0.235	0.164	0.136	0.101

Notes: The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In columns 3 and 4, we take logs of our dependent variables to address the skewness of our dependent variables. Note that we add a small constant of 0.01 to both incidents

[29]Divisions are the second administrative layer of Pakistan (below the provinces). There are 30 divisions in total, six of which lie within the Baloch Province. We also allow for arbitrary spatial dependence following Colella et al. (2019). Figure A-4 reports the results using the "acreg" command to estimate our baseline specification using Conley HAC standard errors (Conley, 1999) across several distance cutoffs for the spatial dependence. We find that our estimates remain highly statistically significant and stable even above a cutoff of 1,500 km. Table A-4 reports the alternative standard errors for Table I.

and fatalities in order to keep district-year observation in the sample that experiences no activity. Column 3 suggests that our treatment effect leads to an increase of roughly 210% in incidents, while column 4 suggests an increase in fatalities of around 160%. Since those estimates are potentially biased due to the constant that we add we also replicate the estimations using Poisson and negative binomial regressions, which produce somewhat lower effects from around 100% for incidents and between 50% to 90% for fatalities (see Table A-5 and Table A-6). The statistical significance is relatively stable throughout the specification.[30]

The qualitative effects remain stable if we employ more conservative specifications by adding district-decade fixed effects and linear time trends for each district (see Table A-8).[31] What is more, our incident results are not driven by nonlethal events, but are comparable if we focus on incidents in which at least one person was wounded or killed (Table A-10).

TABLE II
Controlling for other groups

	Incidents (1)	Fatalities (2)	Ln Incidents (3)	Ln Fatalities (4)
Panel A: Other groups in changes				
$Baloch_i \times Post_t$	3.6534	5.8153	2.1867	1.6526
	(0.9790)***	(2.9540)*	(0.3857)***	(0.4454)***
Δ No. other groups	0.9081	4.0179	1.1531	1.0458
	(0.1420)***	(0.8920)***	(0.1298)***	(0.1449)***
Adj. Within-R^2	0.249	0.180	0.239	0.166
Panel B: Other groups in levels				
$Baloch_i \times Post_t$	2.2596	0.1187	1.2159	0.6821
	(0.7291)***	(2.1189)	(0.2817)***	(0.3137)**
No. other groups	2.5476	10.3715	1.7054	1.7189
	(0.2179)***	(1.4648)***	(0.2407)***	(0.2683)***
Adj. Within-R^2	0.381	0.300	0.426	0.328
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144

Notes: The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[30] We also run log specifications without adding the constant, thus focusing on the intensive margin. Results are similar to the Poisson and negative binomial results (see column 1 Table A-7).

[31] Note that those additional fixed effects and trends can partly capture discoveries of natural resources for which we do not have data for our treatment period. Results are also similar if we exclude all controls (see Table A-9). Furthermore, our results are not driven by any specific district. In Figure A-5 we report the baseline coefficients of column 1 dropping one district at the time.

One potential unobserved factor is the number of other armed groups active within districts. The number of other groups active within treatment and control districts could bias our estimates in either direction. If, for some reason, competition declines within the control group over time, we would get an upward bias in our estimates, while our estimates would suffer from a downward bias if group activity increase for some reason within our control group. The same holds true if either development occurs within our treatment districts, or within treatment and control districts.

To address these concerns, we include the change in the number of other active groups in each district from $t-1$ to t, and alternatively, the number of other active groups as additional control variables.[32] Table II replicates the specifications of Table I including the change of other active groups in Panel A, and the amount of other active groups within the district in Panel B, as additional control variables. Columns 1 to 4 in Panel A show that our results are virtually unchanged if we control for the change in the number of active groups within districts. When controlling for the level of other active groups our coefficients are roughly halved, except for the fatality specification in which the coefficient drops to 0.1. Note that the level specification of other groups performs significantly better in terms of model fit, which is why we slightly prefer this specification. What is more, our log estimations move much closer to the nonlinear estimates (see Table A-5 and Table A-6).

5.2. Threats to identification

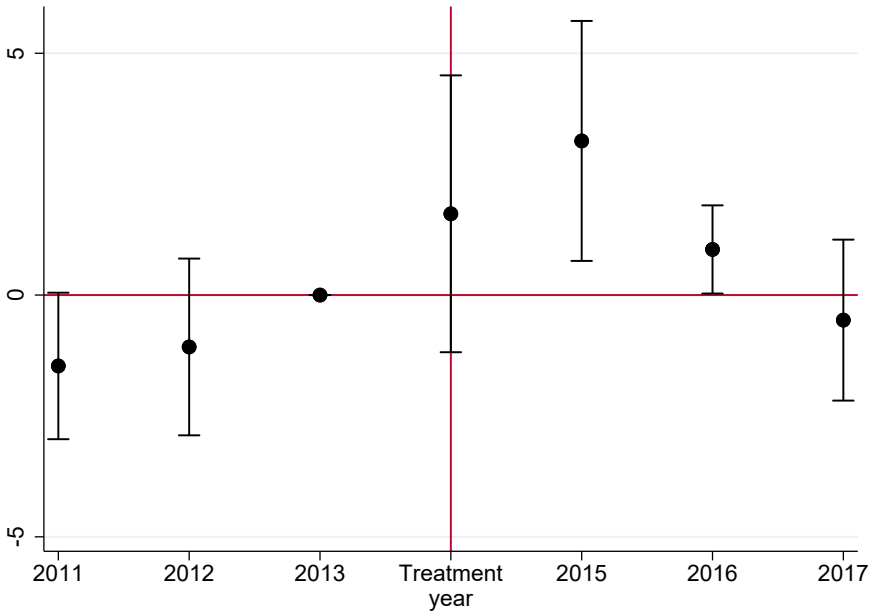
As outlined in Section 4.2 our identification requires the death of KBM to be exogenous and unanticipated. Furthermore, districts in Balochistan need to exhibit parallel pre-treatment trends in violence compared to the rest of Pakistan.

We directly test for violations of the pre-trends and anticipation effect using an event-study with the full set of baseline controls and the number of other active groups. Figure VI reports diff-in-diff coefficients and their corresponding point-wise 95% confidence intervals, using incidents as the dependent variable.

The statistically insignificant coefficients in the pre-treatment period in Figure VI lead us to not reject the null of common trends or no pre-trends. Furthermore, it highlights that our effect is rather short-term. The point estimates we have obtained previously seem to be driven by competition between groups during the two years after treatment occurs. Note that the incident coefficients become statistically significant only from 2015 onward, which is not surprising given that KBM died in June of 2014 and the UBA split did not materialize for the first half year. The effect vanishes in $t+3$.

[32] Other active groups are the active number of groups excluding the BLA and UBA from the count. If we do not exclude the BLA and UBA from the count, our treatment would be perfectly colinear with the number of active groups. We use the number of active groups including the BLA and UBA in the instrumental variable regressions in Section 6.

FIGURE VI
Baseline effect



Notes: Reports the coefficient and their accompanying 95% CIs for an event study specification corresponding to column 1 in Table II. The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25), as well as dummies binning the endpoints of the event window.

Another issue is the correct timing of our treatment. As discussed in Section 2, the UBA faction formed during 2011, due to disagreements about strategy and competing leadership aspirations of two of the Mari sons.[33] Hence, it could be possible that the presence of the UBA faction already affects competition between groups that we currently neglect. We test for this possibility directly and include the following additional interaction term $Baloch_i \times UBAfaction_t$ that represents the counterfactual effect in Baloch districts from 2011 onward instead of after 2014. The results are reported in Table III.

In line with our argument, we find that the coefficient of $Baloch_i \times UBAfaction_t$ is statistically insignificant throughout all specifications. In turn, our coefficient of interest $Baloch_i \times Post_t$ is statistically significant and comparable in size to the estimates of our previous specifications and supports our previous reasoning in the incident specifications. The fatality specifications are more fickle.

[33]Note that if the split between the BLA and UBA has already really occurred in 2011 due to a dispute over leadership between the two brothers, this would have been an alternative identification strategy. At least as long as the struggle for leadership was not driven by the conflict itself, but by personal ambitions of the brothers.

TABLE III
UBA faction vs. UBA group

	Incidents (1)	Fatalities (2)	Ln Incidents (3)	Ln Fatalities (4)
$Baloch_i \times UBAfaction_t$	0.4716	-1.6000	0.2465	0.0506
	(0.6338)	(4.9944)	(0.3265)	(0.4155)
$Baloch_i \times Post_t$	1.8651	1.4571	1.0097	0.6398
	(0.5578)***	(3.5682)	(0.3243)***	(0.3825)*
No. other groups	2.5402	10.3968	1.7015	1.7181
	(0.2172)***	(1.4412)***	(0.2412)***	(0.2691)***
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144
Adj. Within-R^2	0.381	0.300	0.426	0.328

Notes: The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. The results are similar when using the Δ of No. other groups instead of the level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

6. Extensions and alternative channels

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 and to which part our results are explained by the government's response. Furthermore, we scrutinize how the BLA itself was affected by the split of the UBA and if the government changes its response to the BLA compared to other groups. Lastly, we obtain semi-elasticities of an additional armed group on the quantity and quality of violence within districts.

6.1. Targets of armed groups

We think that our effect runs via an increase in the number of armed groups within Baloch districts that compete for attention, as visibility facilitates the attraction of financiers and recruits. We find this effect to be particularly strong in terms of the number of incidents and weaker regarding the number of fatalities. Consistent with this we do not find a statistically significant result when using fatalities per incident as the dependent variable.[34] This indicates that there seems to be no strategy change towards more lethal attacks. However, our findings so far could also be a result of increased infighting between groups, i.e., 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).

[34]Results available upon request.

Our data allow us to test this alternative explanation directly. The GTD list the target type of an 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 core specification.

TABLE IV
Targets

	Incidents vs. other Groups (1)	Incidents vs. Gov. (2)	Incidents vs. public Infrastructure (3)	Incidents vs. Business (4)	Incidents vs. citizens (5)
$Baloch_i \times Post_t$	-0.0527	1.4151	0.9838	0.4582	0.8959
	(0.0416)	(0.5738)**	(0.3688)***	(0.1365)***	(0.3394)***
Controls	Yes	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144	3144
Adj. within-R^2	0.246	0.482	0.222	0.327	0.295

Notes: The definition of the dependent variables (GTD target type) are listed in Table A-3. The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Column 1 of Table IV shows that our treatment does not affect infighting in the treatment vs. the control group. Hence, KBM's death does not seem to have increased infighting between groups. To further verify that group competition is, in fact, not directed against each other, we examine the target of Baloch insurgency groups. In contrast to the GTD, the GED systematically lists both conflict parties. Before the BLA splits in 2014, there is no recorded incident of a Baloch group against another Baloch insurgency group. However, after 2014, we find that the BLA and UBA did attack each other once. The incident reassures us that the groups' real alienation started only after the death of KBM, while at the same time infighting is not the driver of our result.

We use the target type information provided by the GTD to qualify further what kind of competition effect we are observing, specifically if the type of political violence has changed. Columns 2 to 5 in Table IV replicate our baseline specification, using incidents against the government, against public infrastructure, business, and private citizens as the dependent variable.[35] We find that the coefficient on attacks against the government is most sizable. However, we cannot reject that the increase of attacks against the government, public infrastructure, businesses, and private citizens are not different from one another. In summary, it seems that the type of competition relating to the targets that groups pick, has not changed due to the treatment.

[35] Table A-3 provides the specific definitions for each of the measures.

6.2. Government response: Counter-insurgency

Is our observed effect of increased political violence in fact driven by counter-insurgency efforts of the government? One alternative explanation of our proposed mechanism is that the government sees the death of KBM as a chance to crack down on separatist groups within Balochistan. Thus, the increase in political violence could simply be increased counter-insurgency activity by the government.

We test for this possibility by analyzing if our effect sizes change if we explicitly include counter-insurgency efforts of the government of Pakistan. To do so, we need a proxy for incidents instigated by the government against armed groups. Obtaining a suitable proxy is not without problems. Recall that the GTD only codes terrorists 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, i.e., there is no indicator variable informing us if the incidents was initiated by the government or an armed group.

To proxy for counter-insurgency activity by the government, we use the incidents between the government and armed groups that are reported in GED but not in 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 to deal 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.[36] Events that are flagged as potential double counts are excluded form the analysis, which leaves us with a set of incidents that will use as our counter-insurgency proxy (roughly 60% of government incidents). Our count of counter-insurgency incidents obtained from this procedure is roughly 47% of all incidents in which the government is involved (46 % within the Baloch province and 48% outside of the Baloch province).

Column 1 of Table V replicates our baseline reduced form effect (column 1 of Table I) adding counter-insurgency efforts of the government as an additional variable. While counter-insurgency is positively correlated with violence (e.g., revenge) this does not alter our initial finding. Note that the effect of government counter-insurgency on political violence is theoretically ambiguous (Bueno De Mesquita, 2005). In column 2 we change our dependent variable. We now explicitly look at the determinants of counter-insurgency. We see that our treatment does not affect government-induced actions. In column 3,

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

TABLE V
Government response

	Incidents (1)	Incidents Gov. (2)	Incidents vs. Gov. (3)
$Baloch_i \times Post_t$	3.5422	0.1442	1.5593
	(0.9647)***	(0.4342)	(0.7340)**
Inci Gov.	0.2050		
	(0.0574)***		
Controls	Yes	Yes	Yes
District-FE	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes
Obs	3144	3144	3144
Adj. Within-R^2	0.291	0.0106	0.0525

Notes: Column 1 uses our baseline dependent variable. Column 2 uses our counter-insurgency proxy as the dependent variable. Column 3 uses the count of incidents against the government (column 2 Table IV) and counter-insurgency measures by the government as the dependent variable. The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

we add counter-insurgency events to the attacks against the governments by organized groups. We get a significant treatment effect but if we compare it to column 2 of Table IV, we see that adding counter-insurgency leaves the treatment effect virtually unchanged. We take the results reported in Table V as evidence that the increase in violence is primarily driven by increased activity of the armed groups, and not by counter-insurgency efforts in the wake of KBM's death.

6.3. Capacity effect and strategy changes

Our reduced-form effect has multiple potential channels that can explain the increase in political violence. First and foremost, the split changes the absolute and relative capacity of the BLA itself, as well as the relative capacity of other groups within the districts. Furthermore, the changes in relative capacity could lead the government to treat the BLA and UBA differently post treatment, i.e., targeting them more or less. We test for both possibilities. First, we analyze how the BLA as well as BLA and UBA jointly compare to other well established separatist groups within Balochistan. Second, we test if the government targets its counter-insurgency efforts more towards the BLA (and UBA) compared to other well established separatist groups in the Baloch province.

Ex-ante, we would expect that the amount of violence perpetrated by the BLA should decrease compared to other Baloch separatists groups (at least in relative terms). That is, if the UBA split did indeed significantly reduce BLA's assets. However, it could also simply change the composition of political violence perpetrated by the BLA without changing the absolute violence level, e.g., switching from symmetric to asymmetric

warfare, i.e., targeting civilians (Bueno de Mesquita, 2013). We test for both possibilities, running a diff-in-diff specification on the sample of Baloch separatist groups active within districts, defined as:

$$Y_{ijt} = \delta(BLA_{ij} \times PostKBM_t) + \eta_{ij} + \gamma_{it} + \epsilon_{ijt} \tag{3}$$

where Y_{ijt} is political violence committed by group j within district i at time t , $(BLA_{ij} \times PostKBM_t)$ is the interaction of out post-treatment period with the BLA, η_{ij} are district-group fixed effects controlling for the time-invariant capacity a group has within a district, γ_{it} are district-year fixed effects controlling for competition between groups, as well as government capacity within a district-year, and ϵ_{ijt} is the error term. δ is the diff-in-diff coefficient between the BLA and the other well established Baloch separatist groups, specifically the BRA, BLF, and BLUF. Note that the UBA is excluded for now, but we return to the UBA momentarily. Note further that we restrict the control group to the other big Baloch separatist groups since they are present both in the pre- and post-treatment periods and often operate in the same districts as the BLA. These last two points are of importance since our fixed effects restrict our identifying variation to differences in the outcome of within-district groups over time while also partialling out any district-specific time variation common to all groups. Thus a district-group-year observation is only informative if the BLA and at least one of the other groups operate in a district during the same year.[37]

TABLE VI
DiD: Within Baloch separatist groups

	Inci	Inci all	Inci civil	Fatal	Fatal all	Fatal civil
	(1)	(2)	(3)	(4)	(5)	(6)
$BLA_{ij} \times Post_t$	-0.0795	-0.2202	-0.1073	0.0348	-0.1143	-0.0802
	(0.0565)	(0.0917)**	(0.0519)**	(0.0788)	(0.1815)	(0.1056)
District-Group-FE	Yes	Yes	Yes	Yes	Yes	Yes
District-Year	Yes	Yes	Yes	Yes	Yes	Yes
Obs	7896	7896	7896	7896	7896	7896
Adj. within-R^2	0.374	0.378	0.295	0.291	0.328	0.237

Notes: Columns (1) and (4) include our standard GTD measure, (2) and (5) combine GTD and GED, (3), and (6), combines GTD and GED events against civilians. Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The results, presented in Table VI, provide some evidence that the BLA reduces its activity compared to the other armed Baloch groups. While we find no statistically significant effect using our standard measure (column 1), the combined GTD and all

[37]Since we change the treatment and control groups, we have to reconfirm the common trend assumption. We do so using an event study as before. The results are reported in Figure A-6 and show that also in this setting the null of no pre-treatment trends cannot be rejected.

GED events (column 2), and violence against civilians exclusively (column 3) show a reduction.[38] Note that the reduction in attacks against civilians (column 3) is only about half the magnitude compared to the reduction in all attacks (column 2), which is consistent with a strategy shift by the BLA due to reduced capacity. With respect to inflicted fatalities we cannot reject that the ratio of fatalities between the BLA and other armed Baloch groups remains constant.[39] We take this as evidence that the additional violence in our baseline results is likely not exclusively driven by an increase in BLA activities, but also from the increasing activity of other well established armed groups competing for the top spot. The within-group analysis highlights that the BLA commits somewhat fewer incidents compared to other established groups within Balochistan. However, this does not necessarily mean that BLA commits fewer incidents overall. If we run a first-difference analysis on the BLA exclusively, we cannot find robust evidence that the BLA changed the number of incidents it committed.[40]

Moreover, we find that the UBA is compensating for the decrease in BLA activities somewhat. Table A-11 in the Appendix replicates Table VI with the sole difference that incidences from the UBA are added to the BLA, simulating that the groups did not split. The negative effect on all incidents is both lower in magnitude and statistically significance, while the effect on violence against civilians vanishes. We also find some evidence that the BLA and UBA jointly increase their activity compared to all other groups.[41]

We now test if the government focuses its counter-insurgency efforts to a higher degree on the BLA (and UBA) following its split. The assumption would be that the government sees the death of KBM as a chance to finish the group off. In such a case our reduced form effect would exhibit an alternative channel besides the competition effect we have in mind.

To test for this possibility we generate group specific proxies for counter-insurgency efforts by the government and regress these proxies on the $BLA \times Post_t$ interaction. Note that the same restrictions apply as in Section 6.2. We only use incidents from the GED as proxies for counter-insurgency efforts that do not plausible appear in the GTD database, following our introduced double counting thresholds. We also include the respective amount of incidents and fatalities committed by each group as additional

[38] In column 3 we add up GTD incidents and GED incidents against civilians.

[39] Even when estimating non-state incidences and fatalities, e.g., group against another group, the results are largely unchanged. That is, non-state incidences are statistically lower in the post-treatment period for the BLA and insignificant for fatalities.

[40] Results not reported but available upon request.

[41] To further verify that it is indeed the BLA and other Baloch separatist groups that primarily seem to increase their violence, we flip the specification of Table VI and use all other groups as a control group. Note that since the comparison group is now more heterogeneous, we control for group ideology following Kis-Katos et al. (2014). We find some evidence that the UBA and BLA jointly commit more violence compared to other groups following the split (see Table A-12). Note, however, that the common support of incidents in districts during the same years is much more limited as mentioned above.

TABLE VII
DiD: Government action against BLA/UBA

	BLA only		*BLA & UBA*	
	Inci	Fatal	Inci	Fatal
	(1)	(2)	(3)	(4)
$BLA_{ij} \times Post_t$	0.0036	0.0034	0.0030	-0.0108
	(0.0035)	(0.0250)	(0.0034)	(0.0302)
Inci (GTD)	0.0187		0.0144	
	(0.0047)***		(0.0074)*	
Fatal (GTD)		0.1901		0.1185
		(0.1161)		(0.0779)
District-Group-FE	Yes	Yes	Yes	Yes
District-Year	Yes	Yes	Yes	Yes
Obs	7896	7896	7896	7896
Adj. within-R^2	0.0973	0.136	0.0702	0.0851

Notes: The dependent variable is counter-insurgency. Inci(GTD) and Fatal(GTD) are terror attacks and fatalities by major Baloch groups. Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

control variables, in order to partial out any government response driven by the time-varying activity levels of the groups.

Table VII reports the results of this specification using both counter-insurgency incidents as well as fatalities as the dependent variable. Column 1 and 2 show that we cannot reject the null that the government does not change its focus towards the BLA following KBM's death compared to the other separatist groups within Balochistan. The same holds true in columns 3 and 4 in which we treat the UBA and BLA as the same group. Hence, it seems unlikely that the government changed its behavior towards our core groups of interest following KBM's death. Finally, more terror attacks by the Baloch terror groups positively relates to government counter-insurgency efforts.

The within group analysis shows that there seem to be capacity effects that one would expect following Bueno de Mesquita (2013) and that the government does not seem to be particularly affected by the split itself. Rather, it seems the case that the government reacts to group activity in general, which poses no challenge for our preferred interpretation of the effect, as competition between groups.

6.4. Semi-elasticity of armed groups on political violence

Our previous findings are consistent with competition effects between armed groups resulting from group proliferation. Thus, we are interested in the marginal effect of an additional armed group on the amount of political violence. We estimate 2SLS models where we use our previous interaction as an instrument to predict the number of active groups within a district. Note that by definition, we get at least one additional active

group within most districts located in the Baloch province, namely the UBA.

Before we present the results, let us be transparent about a caveat. Using our interaction as an instrument assumes that the death of KBM only has a differential effect within Baloch districts compared to districts outside of Balochistan via an increase in the number of armed groups, specifically the UBA. Hence, we need to assume that the additional violence committed by the other well-established groups within Balochistan is a reaction to the additional competition and not driven by some other issue related to KBMs death. While our previous analysis is in line with this assumption, we cannot rule out alternative explanations entirely. Apart from this drawback, there is also a major benefit of the 2SLS analysis, namely addressing potential measurement error in the number of active groups.

TABLE VIII
2SLS evidence

	Second Stage			
	Incidents	Fatalities	Ln Incidents	Ln Fatalities
	(1)	(2)	(3)	(4)
No. active groups	4.0887	6.3049	2.3991	1.7993
	(0.8452)***	(2.5344)**	(0.3652)***	(0.2851)***
Avg. no. act grps in act districts:	*Treatment group:* 2.3509		*Control group:* 1.7669	
Std dev no. act groups:	*Treatment group:* 1.6395		*Control group:* 1.6595	
	First Stage DV: No. of active groups			
$Baloch_i \times Post_t$	0.8781			
	(0.2035)***			
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144
Adj. Within-R^2	0.3095	0.3004	0.3789	0.3576
F-stat IV	18.59			
First stage adj. R^2	0.698			

Notes: Reports the 2SLS results using $Baloch_i \times Post_t$ as an instrument for the number of active groups operating within a district. The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table VIII presents the results form the 2SLS estimation. The first stage result in the lower panel shows that, on average, the number of active groups rises by just short of one. A point estimate of below one is to be expected, given that the UBA does not operate in all districts of Balochistan. Instrument power also seems fine, given that our first stage F-stat passes the conventional threshold comfortably.

The second stage results in column 1 and 2 are similar to our reduced form effects. The semi-elasticities reported in columns 3 and 4 are also close to our reduced form. An

additional group increases the number of incidents by roughly 240% and fatalities by around 180%. On average, we have 2.4 active groups within a district[42]. This means that on average, the number of groups increases by roughly 37 %. As in the reduced form setting, the incident coefficients decrease if we use non-linear instrumental variable specifications Table A-13. Poisson and negative binomial methods coupled with control functions, provide estimates of between a 128% and 176% for incidents. In the case of fatalities, Poisson estimations obtain point estimates corresponding to an increase of 118%, while the negative binomial results point to an increase of 280%. We trust the conservative incidents estimates the most, since they are consistent with our reduced-form evidence and less volatile compared to the fatality estimates, nor as dependent on the exact model specification. Furthermore, we only include population as control variable in the negbin models due to convergence issues. Nevertheless, even the conservative estimates are quite sizeable given that an increase by one group – an increase in armed groups of on average 37% – leads to an increase in political violence of 135%. Additionally, the first stage residuals in Table A-13 point towards a sizable downward bias if one does not control for the endogeneity of the number of groups.

TABLE IX
2SLS evidence: Heterogeneous treatment effects?

	Second Stage			
	Incidents	Fatalities	Ln Incidents	Ln Fatalities
	(1)	(2)	(3)	(4)
No. active groups	4.6398	10.1287	1.7409	1.5004
	(1.0095)***	(3.0336)***	(0.3502)***	(0.3111)***
	First Stage DV: No. of active groups			
$Treatshare_i \times Post_t$	5.2300			
	(0.8993)***			
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144
Adj. within-R^2	0.219	0.282	0.429	0.335
F-stat IV	33.77			
First stage adj. R^2	0.710			

Notes: Reports the 2SLS results using $Treatshare_i \times Post_t$ as an instrument for the number of active groups operating within a district. The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

How heterogeneous is the effect across districts? Did other groups see the BLA's split as a chance to enter its areas of operation? We test this hypothesis by exploiting the

[42]Conditional on having at least one attack in our sample.

intensive margin of BLA and UBA activity within our instrument. Table IX shows a similar 2SLS regression as in Table VIII. The first stage's interaction term is the post-treatment indicator times the share of BLA and UBA presence in a particular district. That is $treatshare_i$ ranges between 0 and 1. A $treatshare_i$ value of 1, indicates that the BLA or UBA launched at least one attack in the particular district in every year of the sample period. The maximum value of $treatshare_i$ is about 0.5. This corresponds to attacks every other year, on average. In these districts, the model predicts two additional groups, on average. The results support the notion that other groups might have seen the split of the BLA as a chance to enter districts with a high BLA presence.

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. Section B-1 shows that using both incidents from GTD and GED increases the point coefficients of our baseline specification, although they are not statistically different from one another. We 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. Section B-2 investigates the problem of "potential" double-counting, which is introduced if one uses information from two datasets that, in theory, have an overlap of events that they code. Our probability-based approach to assess the likelihood of potential double counts suggests that double counting is around 10% for realistic scenarios in our case, which might explain the higher point coefficients in Section B-1. However, our results do not change if we exclude those potential double counts from our estimation.

We also further probe our concept of active armed groups (see Appendix C). Section C-1 introduces the concept of potential active groups, defined as groups that are active anywhere in the country, and have been active at least once in a specific district. This approach loosens the requirement that a group needs to commit an attack in a specific district to be counted as active there, and in theory, allows for 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 reasonably well correlated (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 the potential active armed groups measure to cover all districts falling within the convex hull of a groups incidents. Again our results remain stable. Section C-2 test if 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 armed groups and armed groups counts since they commit only a single incident. Once more, our results remain stable.

Finally, we test if our effects are driven by the ethno-political representation of the ethnic groups living across our treatment and control districts (see Appendix D). Given that the conflict in Balochistan has an ethnic origin, one might expect that our estimates could be biased by changing political representation of ethnic groups across our treatment and control districts (Bormann et al., 2019). Proxying for the ethno-political representation of both districts and provinces using data from the geocoded version of the Ethnic Power Relations dataset (GeoEPR) (Wucherpfennig et al., 2011; Vogt et al., 2015), we can reject this threat to our identification.

8. Conclusion

This paper studies the causal effect of an increase in the number of independent armed groups on 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 causal evidence on the matter.

Our estimates predict that one additional active armed group (corresponding to a 37 % increase in the number of groups within our sample) more than doubles organized political violence. The result highlights that the elasticity between armed groups and political violence is about two. These significant effects are consistent with 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."

Exploiting group mergers and splits is a promising avenue to analyze how the number of armed groups involved in a conflict interacts with other determinants of conflict. The main advantage of exploiting group splits and mergers, is that the number of combatants (or foot soldiers) should remain constant in the short-run and that there are no selection effects in and out of the combat zone. Hence, researchers can better isolate how economies of scale and increased competition affect the amount of political violence perpetrated by armed groups. This is not to say that the situation in which a new group enters into a combat zone are of no interest. Quite the opposite, we believe it vital to compare our results to situations in which a new actor actively selects into a combat zone, such as the Islamic State starting operations within Afghanistan.

Finally, mergers and spits of armed groups have in general received little attention so far, which makes them a promising topic of future studies. Future research needs to trace the reasons 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 kind 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 holistically analyze political violence of armed groups and increase coverage.

References

Abrahams, A., E. Berman, P. Khadka, E. F. Klor, and J. Powell (2019). Mostly deterred: An episodic analysis of the Israel-Gaza conflict. *Available at SSRN 3465438*.

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.

Berman, E., M. Callen, J. H. Felter, and J. N. Shapiro (2011). Do working men rebel? Insurgency and unemployment in Afghanistan, Iraq, and the Philippines. *Journal of Conflict Resolution 55*(4), 496–528.

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. and M. Couttenier (2015). External shocks, internal shots: The geography of civil conflicts. *Review of Economics and Statistics 97*(4), 758–776.

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.

Borusyak, K. and X. Jaravel (2017). Revisiting event study designs. *Available at SSRN 2826228*.

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.

Campante, F. R., Q.-A. Do, and B. Guimaraes (2019). Capital cities, conflict, and misgovernance. *American Economic Journal: Applied Economics 11*(3), 298–337.

Chabé-Ferret, S. (2015). Analysis of the bias of matching and difference-in-difference under alternative earnings and selection processes. *Journal of Econometrics 185*(1), 110–123.

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).

Colella, F., R. Lalive, S. O. Sakalli, and M. Thoenig (2019). Inference with arbitrary clustering. IZA Discussion Paper No. 12584.

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.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics 92*(1), 1–45.

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.

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.

Fortna, V. P. (2015). Do terrorists win? Rebels' use of terrorism and civil war outcomes. *International Organization 69*(3), 519–556.

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.

Hegre, H. and N. Sambanis (2006). Sensitivity analysis of empirical results on civil war onset. *Journal of Conflict Resolution 50*(4), 508–535.

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 (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.

Kydd, A. H. and B. F. Walter (2006). The strategies of terrorism. *International Security 31*(1), 49–80.

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

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.

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.

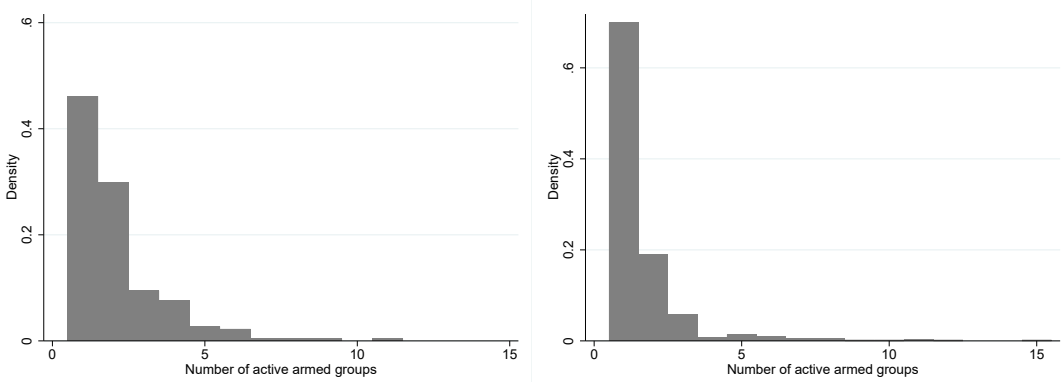
Wucherpfennig, J., N. B. Weidmann, L. Girardin, L.-E. Cederman, and A. Wimmer (2011). Politically relevant ethnic groups across space and time: Introducing the GeoEPR dataset. *Conflict Management and Peace Science 28*(5), 423–437.

A. Additional Figures and Tables

A-1. Figures

FIGURE A-1
Distribution: Number of armed group in and outside of Balochistan

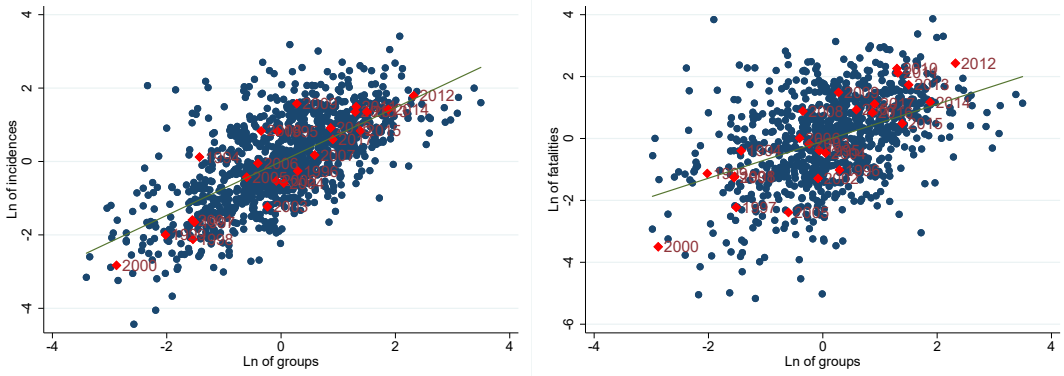
(A) Baloch province (B) Outside Baloch province



Notes: Panel A depicts the distribution of the number of *active* groups in district-years with at least one active group in Balochistan. Panel B depicts the distribution of the number of *active* groups in district-years with at least one active group in outside of Balochistan. *Active* means conducting at least one attack in a given year and district.

FIGURE A-2
Armed groups and political violence (demeaned by country & year)

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

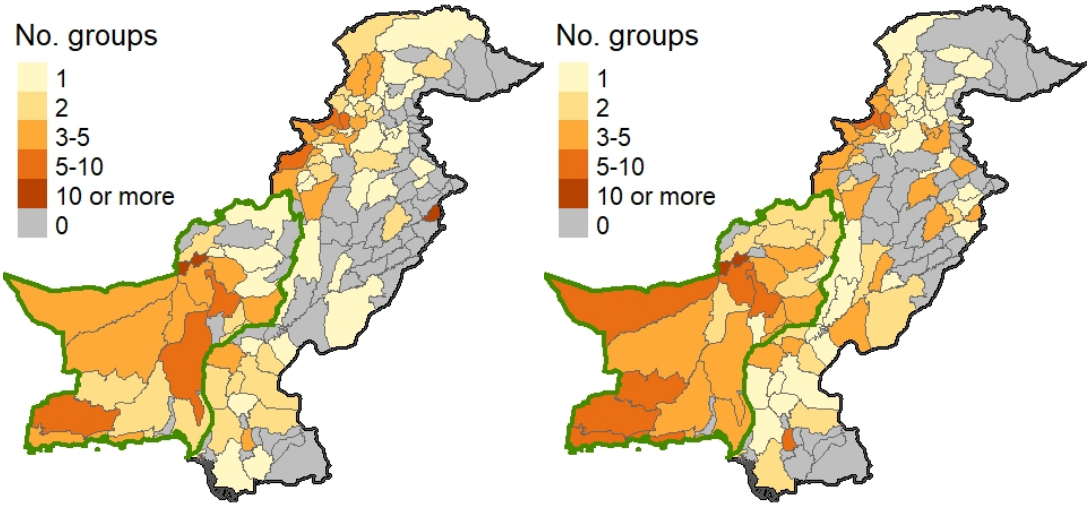


Notes: Depicts a scatter plot of the (log of) groups vs (log of) incidents & fatalities created by these groups, demeaned by country and year. The unit of observation is country-year. Pakistan is represented in red. The line illustrates the best linear fit using the global GTD sample between 1994 and 2018.

FIGURE A-3
Number of armed groups across districts

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

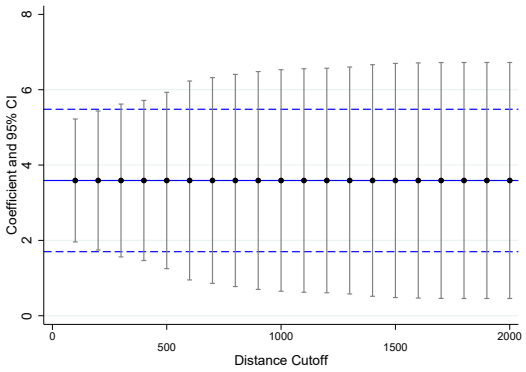
(B) Number of groups within districts (post-treatment)



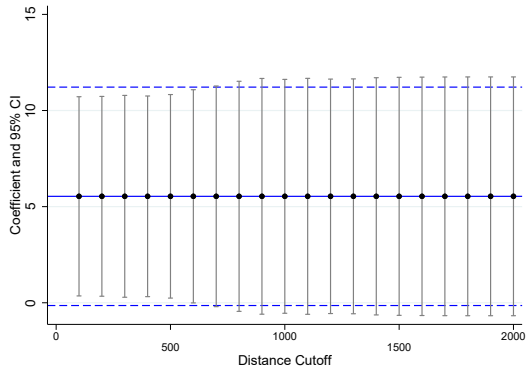
Notes: Panel A reports the number of groups active in districts in the pre-treatment period. Panel B plots the number of groups active within a district in the post-treatment group. The bold green line represents the borders of the Baloch province.

FIGURE A-4
Baseline estimate: Arbitrary spatial clustering

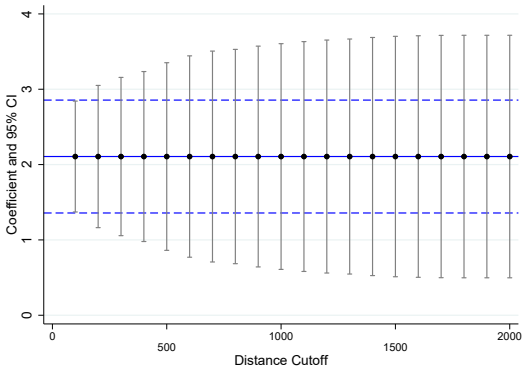
(A) Incidents



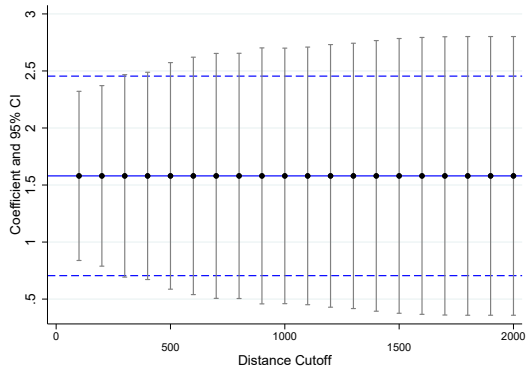
(B) Fatalities



(C) Ln Incidents



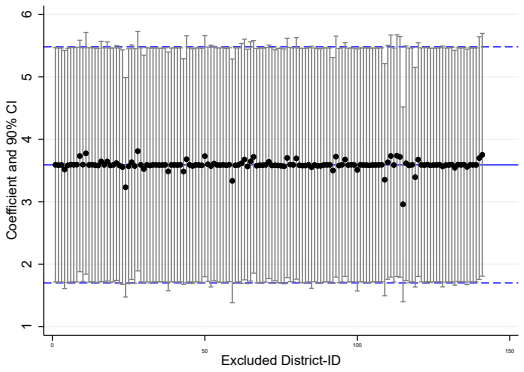
(D) Ln Fatalities



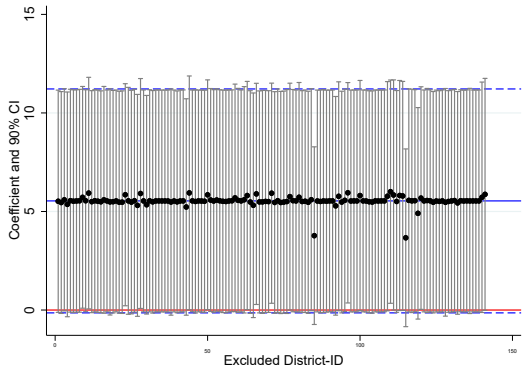
Notes: Reports the point coefficient and their accompanying 95% CIs for different distance thresholds of spatial clustering corresponding to columns 1 to 4 of Table I. The horizontal line depict that actual baseline estimate and its accompanying 95% CIs from clustering on the district level.

FIGURE A-5
Leave one out test: Districts

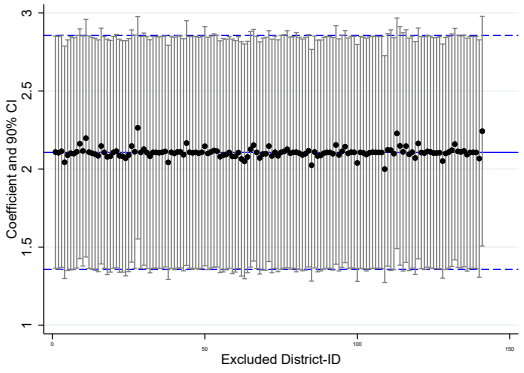
(A) Incidents



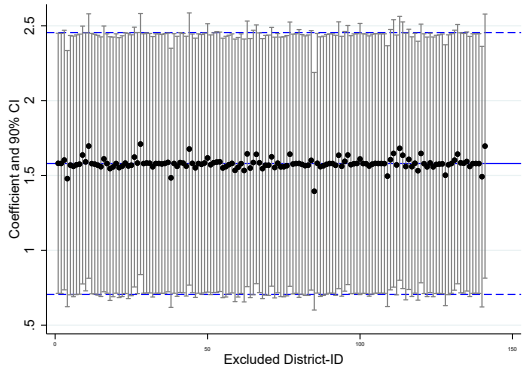
(B) Fatalities



(C) Ln Incidents

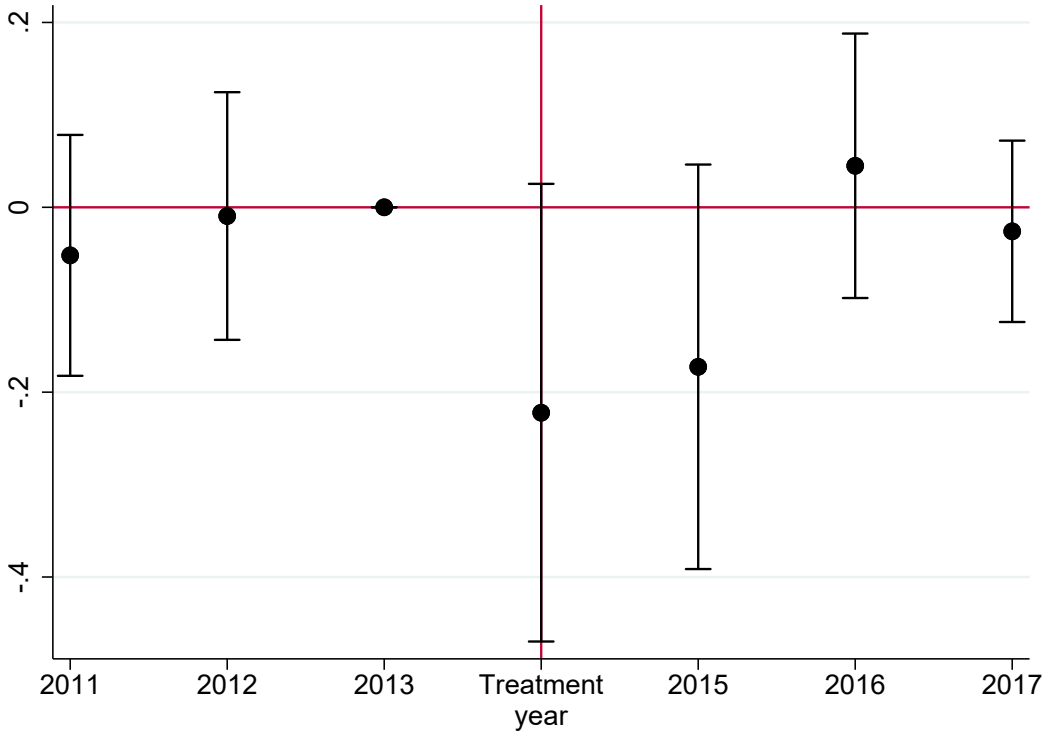


(D) Ln Fatalities



Notes: Reports the point coefficients of our baseline specifications in Table I, dropping one district at the time together with their 95% CI. The blue horizontal line is the respective baseline coefficient (full sample) with its 95% CIs (dashed lines).

FIGURE A-6
Event Study: Baloch secession groups



Notes: Reports the point coefficient and their accompanying 95% CIs for column 1 of Table VI. The controls include dummies binning the endpoint of the event window.

A-2. Tables

TABLE A-1
Treated groups (mergers and splits) and location of incidents and fatalities

Group name	Rest of Pakistan		Balochistan	
	Incidences	Fatalities	Incidences	Fatalities
Baloch Liberation Army (BLA)	11	17	195	433
United Baloch Army (UBA)	7	35	84	138
Harakat ul-Mujahidin Al-Almi (HuMA)	1	3	0	0
Harakat ul-Mujahidin (HuM)	1	6	0	0
Jaish-e-Mohammad (JeM)	3	5	0	0
Jamaat-ul-Ahrar (JuA)	19	93	0	0
Jundullah	22	311	3	14
Lashkar-e-Balochistan (LeB)	2	5	29	3
Lashkar-e-Jhangvi (LeJ)	97	738	81	805
Lashkar-e-Islam (LeI)	129	422	0	0
Taliban	1,442	6,330	88	640
Sipah-e-Sahaba/Pakistan (SSP)	15	59	1	1

TABLE A-2
Descriptive statistics

	Mean	SD	Min	Max	N
Dependent variables					
Incidents	0.93	4.02	0.00	71.00	3,144
Incidents (at least one wounded)	-3.59	2.20	-4.61	4.26	3,144
Incidents (at least one fatality)	3.27	17.34	0.00	398.00	3,144
Counter insurgency	-3.63	2.43	-4.61	5.99	3,144
Ln fatalities	13.50	78.01	0.00	1,543.00	3,144
Fatalities GED+GTD	-3.10	3.03	-4.61	7.34	3,144
Ln Fatalities GED+GTD	0.65	3.00	0.00	61.00	3,144
Fatalities per incidence	0.55	2.58	0.00	45.00	3,144
Incidents vs. other Groups	0.63	4.46	0.00	124.00	3,144
Incidents vs. Gov.	0.06	0.55	0.00	22.00	3,144
Incidents vs. Infrastructure	0.40	2.04	0.00	43.00	3,144
Incidents vs. Business	0.11	0.89	0.00	31.00	3,144
Incidents vs. Citizens	0.08	0.49	0.00	9.00	3,144
	0.25	1.40	0.00	27.00	3,144
Treatment variables					
Treatshare (UBA or BLA)	0.04	0.09	0.00	0.52	3,144
Control variables					
Nr. active groups	0.41	1.06	0.00	15.00	3,144
Nr. other groups	0.34	0.93	0.00	14.00	3,144
Δ nr. other groups	0.01	0.49	-4.00	5.00	3,013
Nr. potential other groups	1.11	1.64	0.00	20.00	3,144
Δ nr. potential other groups	0.05	0.50	-5.00	6.00	3,013
Incidence concentration	0.15	0.33	0.00	1.00	3,144
Fatalities concentration	0.12	0.31	0.00	1.00	3,144
Ln population	13.49	1.11	8.66	16.07	3,144
Airport presence	0.07	0.25	0.00	1.00	3,144
Ln ruggedness	10.45	2.14	6.03	14.07	3,144
Ln elevation (in meters)	5.83	1.47	2.20	8.50	3,144
Ln road density	-9.50	0.80	-11.60	-6.57	3,144
Area share of discriminated ethnic groups (district)	0.15	0.30	0.00	1.00	2,921
Area share of junior partner ethnic groups (district)	0.25	0.39	0.00	1.00	2,921
Area share of powerless ethnic groups (district)	0.20	0.37	0.00	1.00	2,921
Area share of senior partner ethnic groups (district)	0.40	0.47	0.00	1.00	2,921
Group-level variables					
Incidents GTD	0.08	0.82	0.00	37.00	9,870
Incidents GTD + GED	0.13	1.22	0.00	47.00	9,870
Incidents asymmetric	0.09	0.87	0.00	37.00	9,870
Incidents non-state	0.08	0.82	0.00	37.00	9,870
Fatalities GTD	0.10	1.12	0.00	36.00	9,870
Fatalities GTD + GED	0.23	2.35	0.00	89.00	9,870
Fatalities asymmetric	0.12	1.36	0.00	51.00	9,870
Fatalities non-state	0.10	1.12	0.00	36.00	9,870

TABLE A-3
Definition of the dependent variables in Table IV

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-4
Baseline results: Alternative standard errors

	Incidents (1)	Fatalities (2)	Ln Incidents (3)	Ln Fatalities (4)
$Baloch_i \times Post_t$	3.5904	5.5364	2.1067	1.5800
	(1.1315)***	(2.8108)*	(0.3803)***	(0.4671)***
	[1.5985]**	[3.1686]*	[0.8213]**	[0.6232]**
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144
Adj. Within-R^2	0.235	0.164	0.136	0.101

Notes: Reports the baseline results in Table I. Standard errors clustered at the division level in parenthesis, and allowing for spatial dependence with a distance cutoff of 2000 in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A-5
Baseline results: PPML

	Incidents			*Fatalities*		
	(1)	(2)	(3)	(4)	(5)	(6)
$Baloch_i \times Post_t$	0.9903	1.0780	0.6847	0.9216	1.0364	0.4163
	(0.2283)***	(0.2284)***	(0.1639)***	(0.3191)***	(0.3120)***	(0.2514)*
Δ No. other groups		0.2256			0.2855	
		(0.0461)***			(0.0650)***	
No. other groups			0.4341			0.4779
			(0.0718)***			(0.0685)***
Controls	Yes	Yes	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	2640	2640	2640	2448	2448	2448
R^2	0.7806	0.7779	0.7878	0.6552	0.6565	0.6756

Notes: Reports the baseline results in Table I using Poisson estimation method. Standard set of controls (population and geographical characteristics interacted with year dummies, see footnote 25) included but not reported. Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A-6
Baseline results: Negative Binomial

	Incidents		*Fatalities*	
	(1)	(2)	(3)	(4)
$Baloch_i \times Post_t$	1.2092	1.2504	0.9838	0.9900
	(0.1710)***	(0.1584)***	(0.2161)***	(0.2097)***
Δ No. other groups		0.3271		0.5389
		(0.0599)***		(0.0909)***
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	2592	2592	2328	2328
Log likelihood	-1697	-1655	-2237	-2188

Notes: Reports the baseline results in Table I using negative binomial estimation method. Standard set of controls is reduced to log of population due to convergence issues. Other group level specification is left out due to convergence issues. Standard errors are based on 250 bootstraps. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A-7
Log DV intensive margin

	Ln Incidents			*Ln Fatalities*		
	(1)	(2)	(3)	(4)	(5)	(6)
$Baloch_i \times Post_t$	0.9577	0.9891	0.8326	0.7246	0.7892	0.5649
	(0.1781)***	(0.1743)***	(0.1385)***	(0.2655)***	(0.2754)***	(0.2786)**
Δ No. other groups		0.0983			0.1294	
		(0.0300)***			(0.0622)**	
No. other groups			0.2488			0.2468
			(0.0410)***			(0.0653)***
Controls	Yes	Yes	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	561	561	561	434	434	434
Adj. within-R^2	0.192	0.205	0.267	0.0371	0.0457	0.0712

Notes: Reports the baseline results in Table I using a logarithmic transformation of the dependent variable. The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A-8
Baseline results: Additional fixed effects

	Incidents	Fatalities	Ln Incidents	Ln Fatalities
	(1)	(2)	(3)	(4)
$Baloch_i \times Post_t$	3.4631	3.4597	1.5282	1.2294
	(0.9192)***	(3.8167)	(0.3961)***	(0.4398)***
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
District-Decade-FE	Yes	Yes	Yes	Yes
District-trend	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144
Adj. within-R^2	0.0302	-0.0240	0.0459	0.0279

Notes: Reports the baseline results in Table I but includes additional fixed effects. The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A-9
Baseline results: no controls

	Incidents (1)	Fatalities (2)	Ln Incidents (3)	Ln Fatalities (4)
$Baloch_i \times Post_t$	3.9227	6.2560	2.4174	1.9353
	(1.2703)***	(3.2534)*	(0.3948)***	(0.4625)***
Controls	No	No	No	No
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144
Adj. within-R^2	0.0339	0.00453	0.0570	0.0280

Notes: Reports the baseline results in Table I but excludes all control variables. Standard errors clustered at the district in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A-10
Baseline results: severe and fatal incidents

	Incidents severe (1)	Incidents fatal (2)	Ln Incidents severe (3)	Ln Incidents fatal (4)
$Baloch_i \times Post_t$	1.6786	1.2365	1.5031	1.2911
	[0.7169]**	[0.6236]**	[0.3805]***	[0.3677]***
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144
Adj. Within-R^2	0.270	0.254	0.114	0.111

Notes: Reports the baseline results in Table I using only incidents in which at least one person got wounded (severe) or died (fatal). The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A-11
DiD: Within Baloch separatist groups (BLA & UBA)

	Inci (1)	Inci all (2)	Inci civil (3)	Fatal (4)	Fatal all (5)	Fatal civil (6)
$BLA_{ij} \times Post_t$	-0.0211	-0.1618	-0.0630	0.1721	0.0230	0.0528
	(0.0679)	(0.0850)*	(0.0535)	(0.1056)	(0.1875)	(0.1084)
District-Group-FE	Yes	Yes	Yes	Yes	Yes	Yes
District-Year	Yes	Yes	Yes	Yes	Yes	Yes
Obs	7896	7896	7896	7896	7896	7896
Adj. within-R^2	0.374	0.424	0.337	0.288	0.357	0.291

Notes: BLA and UBA are treated as one group. Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A-12
DiD: Within Non-Baloch separatist groups (BLA & UBA)

	Inci	Inci all	Inci civil	Fatal	Fatal all	Fatal civil
	(1)	(2)	(3)	(4)	(5)	(6)
$BLA_{ij} \times Post_t$	0.0741	-0.0952	-0.0145	0.1404	-0.0366	-0.0298
	(0.0308)**	(0.0810)	(0.0420)	(0.0872)	(0.1894)	(0.1139)
Ideology Controls × year	Yes	Yes	Yes	Yes	Yes	Yes
District-Group-FE	Yes	Yes	Yes	Yes	Yes	Yes
District-Year	Yes	Yes	Yes	Yes	Yes	Yes
Obs	57246	57246	57246	57246	57246	57246
Adj. within-R2	0.0529	0.299	0.353	0.273	0.248	0.223

Notes: BLA and UBA are treated as one group. Ideology controls are extensive dummies created from the three ideology dimension in Kis-Katos et al. (2014). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A-13
Nonlinear instrumental variable results

	Second Stage			
	Poisson		*Negative Binomial*	
	Incidents	Fatalities	Incidents	Fatalities
	(1)	(2)	(3)	(4)
No. active groups	1.280***	1.186**	1.763***	2.834***
	(0.440)	(0.525)	(0.475)	(0.847)
Residual 1st-stage	-0.942**	-0.749	-1.239***	-0.731
	(0.407)	(0.580)	(0.452)	(0.760)
	First Stage DV: No. of active groups			
$Baloch_i \times Post_t$	0.9569		0.9031	
	(0.2117)***		(0.2214)***	
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3384	3384	3384	3384
Log likelihood	-1495	-5580	-1915	-2526
F-stat IV	20.432		16.652	
First stage adj. $Within - R^2$	0.4330		0.2783	

Notes: Standard errors of the second stage are based on 250 bootstraps. First stage is estimated with OLS. Standard set of controls is reduced to population due to convergence issues. Note that we use population instead of ln(population) due to the nonlinear transformation introduced by Poisson and negative binomial second stages. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B. Counting incidents and fatalities

We take a closer look at how the incidents and fatalities reported by the different databases 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 fatalities 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 their definition of terrorism. Thus, if following our treatment, all groups would switch to committing more violence fitting the GTD definition, it could be that GTD provides us the best coverage, without substantially changing the overall violence level. For example, following treatment, the actors could be involved in fewer events that fulfill the criteria of internal armed conflict, which is the primary focus of the GED database. We test for the aggregated effect across databases in Table B-1, where we replicate our baseline table including all events committed by armed groups from the GTD, and GED.[43]

Table B-1
All incidents and fatalities

	Incidents	Fatalities	Ln Incidents	Ln Fatalities
	(1)	(2)	(3)	(4)
$Baloch_i \times Post_t$	4.7563	19.4801	1.7329	1.5295
	(1.3502)***	(14.4061)	(0.3549)***	(0.4600)***
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144
Adj. within-R^2	0.303	0.110	0.496	0.460

Notes: Replicates our 2SLS specifications combining incidents and fatalities form GTD and GED. Standard set of controls (population and geographical characteristics interacted with year dummies, see footnote 25) included but not reported. Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B-1 shows that the size and statistical significance of our point estimates remains

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

roughly constant across our specifications.[44] 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.

In Table B-2, we replicate our core specification on incidents and fatalities, using all events included in the GTD in column 1 and 2, those claimed by a group in column 3 and 4, and finally all events included in the GTD and GED in column 5 and 6.

TABLE B-2
All events

	All GTD Events		*Claimed GTD Events*		*All GTD & GED Events*	
	Incidents	Fatalities	Incidents	Fatalities	Incidents	Fatalities
$Baloch_i \times Post_t$	6.6975	6.7272	3.4439	4.7646	7.8635	20.6708
	(2.0678)***	(3.3198)**	(0.9636)***	(2.8339)*	(2.2933)***	(14.8812)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144	3144	3144
Adj. Within-R^2	0.202	0.166	0.296	0.178	0.158	0.0144

Notes: The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Again, we obtain economically more significant effects in columns 1, 2, 5, and 6, which suggests that our competition effect also influences less organized actors. Furthermore, columns 3 and 4 show comparable estimates to columns 1 and 2 of Table I, highlighting that the armed groups that we identify seem to claim a lot of the incidents they commit publicly. Summing up, our results seem not to be driven by GTD specific coding rules.

B-2. Double counting organized political violence

Our larger estimates using both GTD and GED events of organized political violence could be a result of double-counting. Double counting arises if both the GED and GTD

[44] An exception is the fatality count, which tends to be less stable across different specifications.

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 straight forward, 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 levels of time and location precision 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 with, 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 a 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

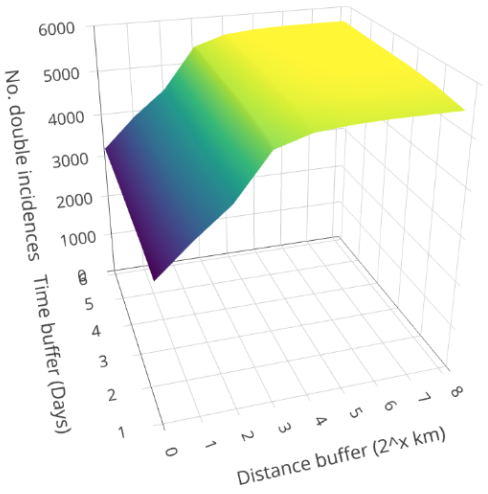
[45]The specific matching of group names between the GTD and GED dataset is documented in Appendix E

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

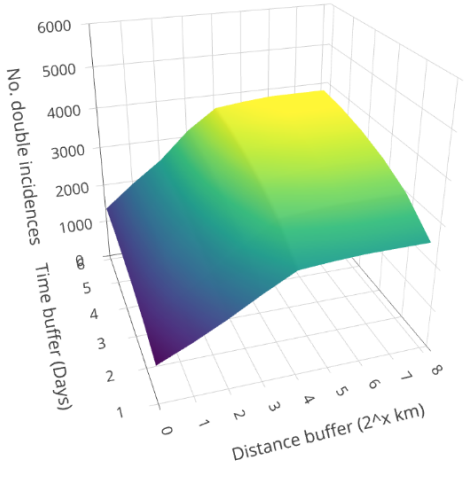
not within. We do not find this assumption problematic since both databases conduct internal quality control, and it seems unlikely that they systematically misinterpret the set of sources they judged to be meaningful. Finally, the approach may also be performed without matching groups and fatalities 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 criterion.

FIGURE B-1
Nr. of incidences of GTD and GED double coding

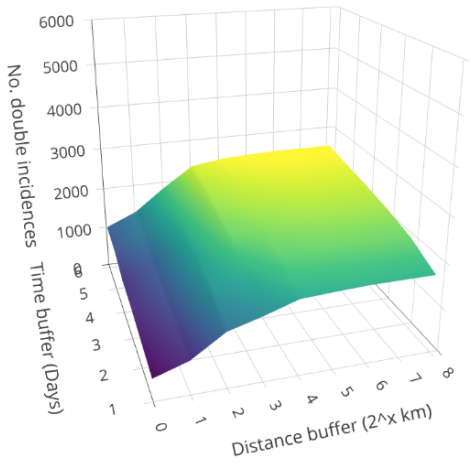
(A) Without group name matching



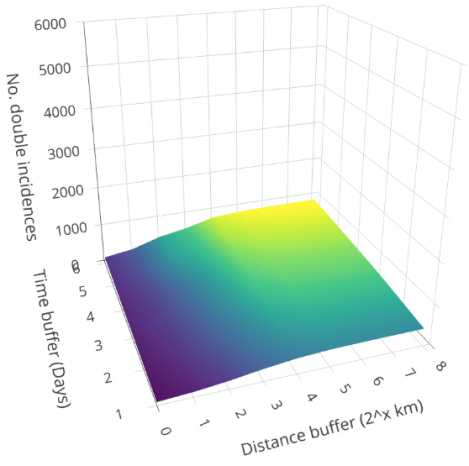
(B) With group name matching



(C) With fatality matching



(D) With group name and fatality matching



We illustrate the impact on the double-counts for each criterion. Panel A of Figure B-

1 shows the 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 is 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.

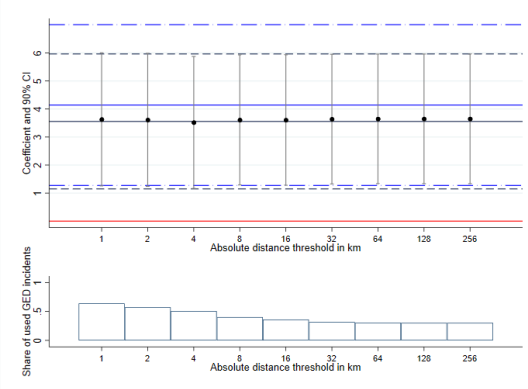
We conclude that the naive approach may involve too many type 1 errors. Type 1 errors are likely to be high if one ignores the same name criterion, given a high number of armed groups active within Pakistan. Hence, the name-matching seems a necessary condition for a meaningful application of the approach, 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 start with the naive approach for illustrative purposes.

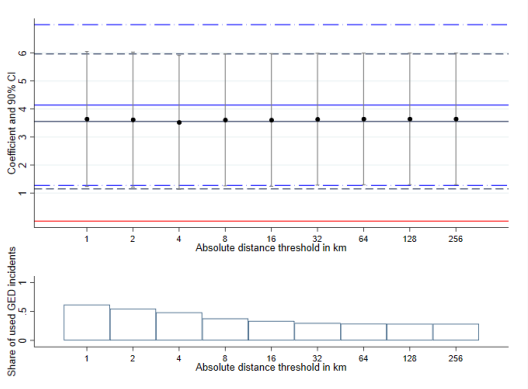
Panels A to F of Figure B-2 plot the point coefficients of our baseline specification (column 1 Table I) for increasing distance thresholds across different time buffers, starting with $+/-$ one day in Panel A, in the upper part of each panel and the fraction of GED events in use (not flagged as a potential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

FIGURE B-2
Potential double coding: GTD & GED

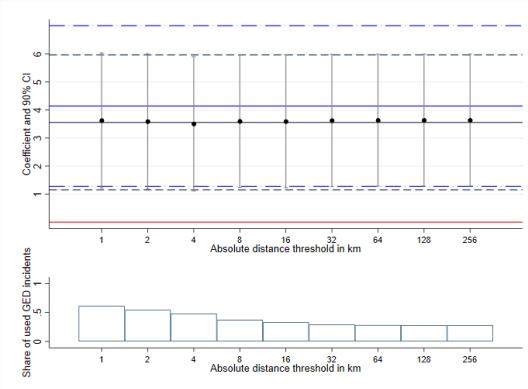
(A) A: +/- 1 days



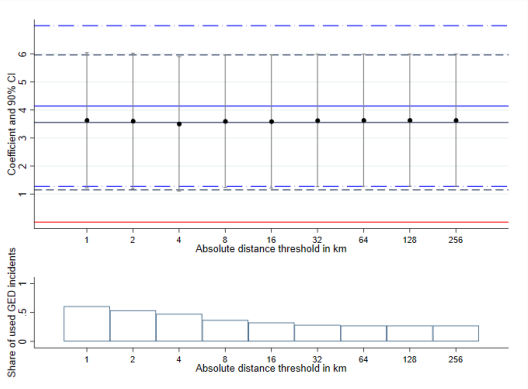
(B) B: +/- 2 days



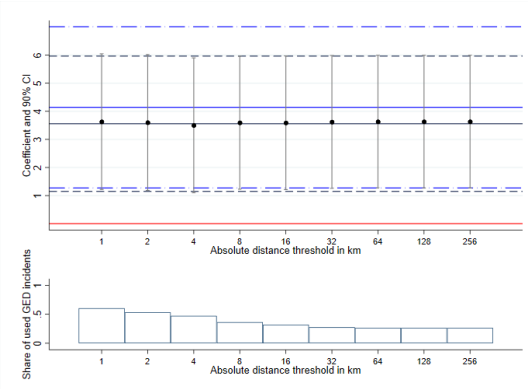
(C) C: +/- 3 days



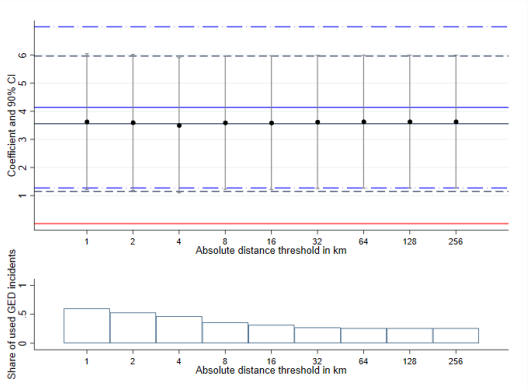
(D) D: +/- 4 days



(E) E: +/- 5 days



(F) F: +/- 6 days



Notes: Reports the point coefficient 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 are the accompanying 95% CIs. The horizontal blue solid line is the coefficient using all GED and GTD incident ignoring potential double counting. The two horizontal blue dashed lines are the accompanying 95% CIs.

GTD and GED (corresponding to column 1 Table B-1) with its 95% CIs (blue solid and dashed lines).

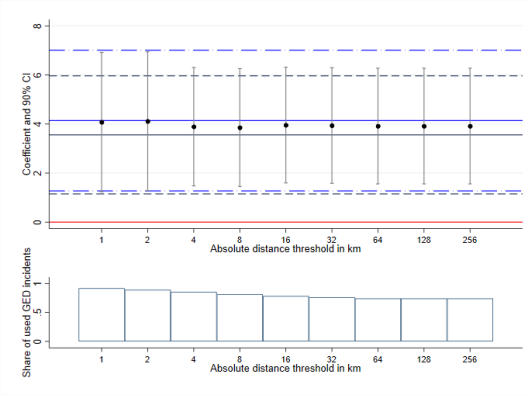
Figure B-2 shows that our core results seem relatively robust to double-counting since we cannot reject that they are identical to our baseline effect of the effect from Table I. Note that we also fail to reject that the baseline effect is the same when using only GTD compared to using both GTD and GED events.

Panels A to F of Figure B-3 report the results for the more nuanced approach relying on the same-name criterion. Again the results remain remarkably stable. The major difference between the two approaches is that the fraction of included GED events is much more stable in Figure B-3 compared to Figure B-2. The stable result is to be expected since the same-name criterion drastically reduces the set of potential double counts regardless of the buffers. This pattern does not change if one adds the fatality criteria to either approach. In fact the only major difference is that both approaches use a higher fraction of the GED events, while the decline in usable GED events in the distance and temporal buffer remain similar.[47] In summary, the results make us confident that our combined results using the full set of available events of organized political violence, is not biased by some systematic measurement error due to double counting.

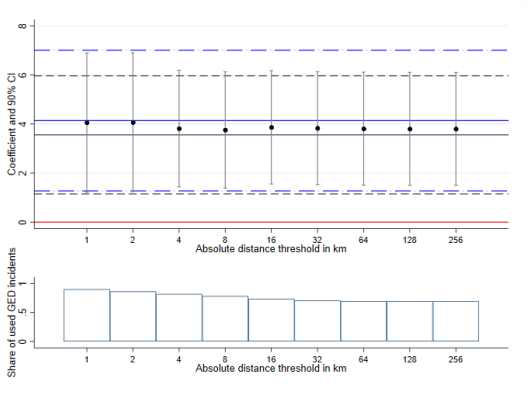
[47]Results nor reported but available upon request.

FIGURE B-3
Potential double coding (Same Names): GTD & GED

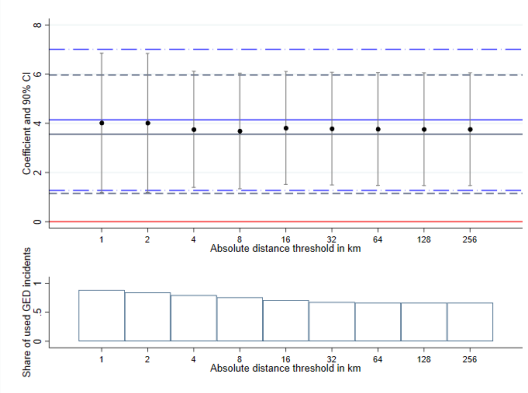
(A) A: +/- 1 days



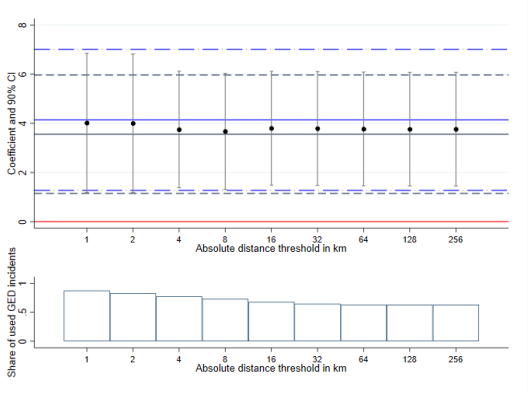
(B) B: +/- 2 days



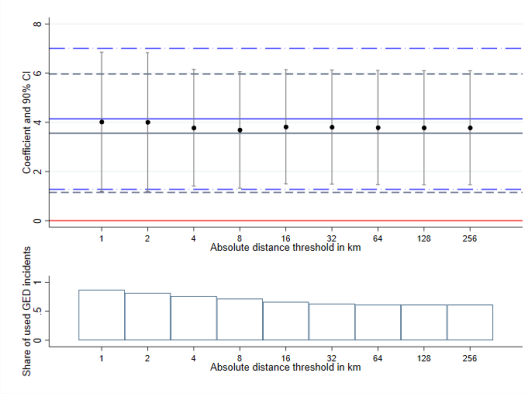
(C) C: +/- 3 days



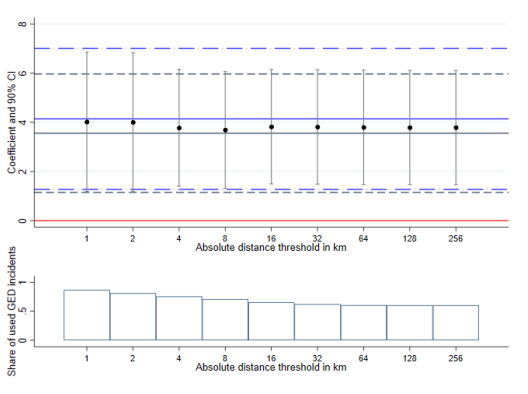
(D) D: +/- 4 days



(E) E: +/- 5 days



(F) F: +/- 6 days



Notes: Reports the point coefficient 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 incident 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 local market of violence, 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 practice 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.

C-1. Active armed groups & potential active armed groups

In this section, 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.

Table C-1 replicates Table II with the difference that the change of other groups and the number of other groups in the upper and lower panel is replaced by the change and number of potential active groups. Once more, our results are remarkably stable. None of the coefficients is statistically different from the once using the active number of groups instead of the potential one, either in changes or levels.

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

TABLE C-1
Controlling for potential other groups

	Incidents	Fatalities	Ln Incidents	Ln Fatalities
	(1)	(2)	(3)	(4)
$Baloch_i \times Post_t$	2.2487	0.3628	1.6310	1.1495
	(0.7815)***	(2.1562)	(0.3555)***	(0.4063)***
Δ No. potential other groups	1.4739	5.7693	0.5294	0.4879
	(0.2418)***	(1.3040)***	(0.1012)***	(0.1219)***
Adj. Within-R^2	0.325	0.241	0.187	0.134
	Potential other groups in levels			
$Baloch_i \times Post_t$	3.6176	5.6771	2.1174	1.5930
	(0.9699)***	(2.9068)*	(0.3807)***	(0.4448)***
No. potential other groups	0.6692	1.9767	0.1643	0.0797
	(0.1454)***	(1.2948)	(0.0883)*	(0.0840)
Adj. Within-R^2	0.240	0.167	0.137	0.101
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144

Notes: The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

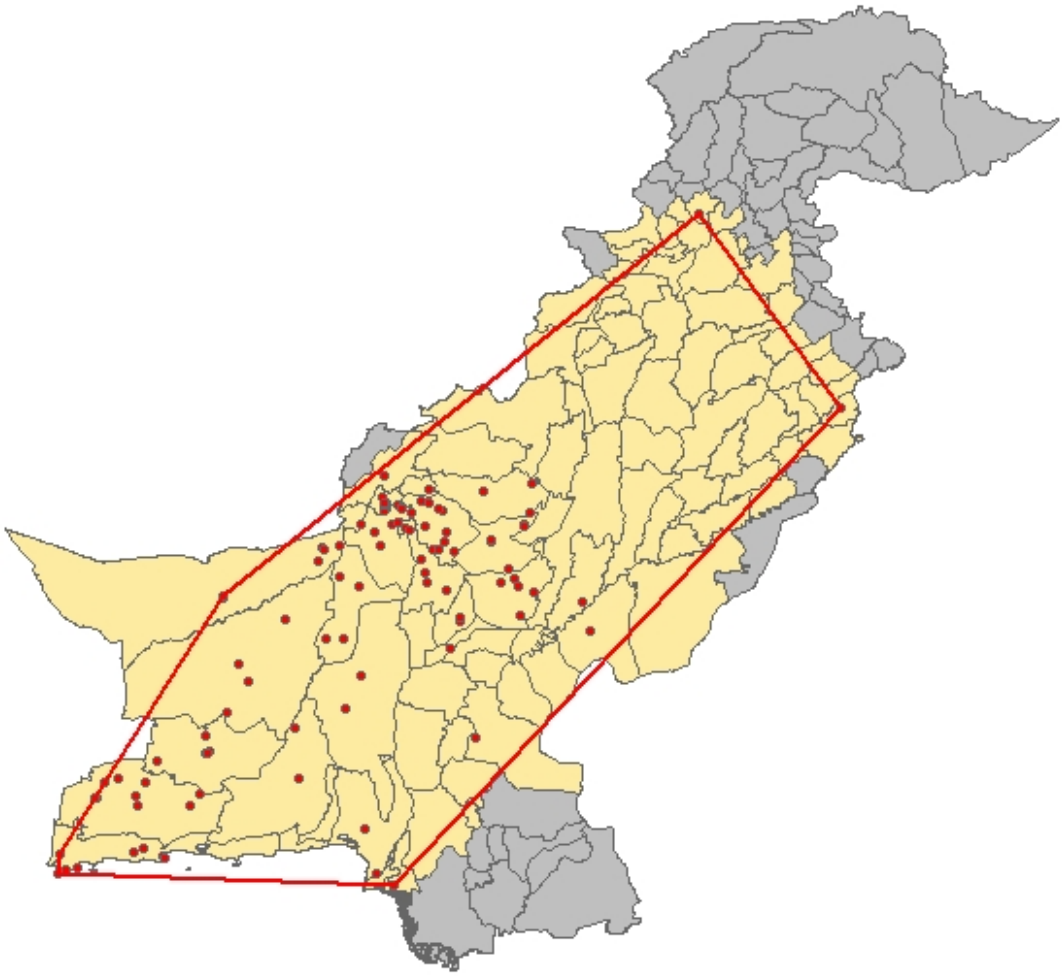
Instead of counting a group as active in all districts in which it ever operates, we can also define an area of operation in which we count a group as active whenever it is active somewhere. We follow König et al. (2017) and define a group's area of operation as the convex hull drawn around its incidents over the sample period. Specifically, we treat all districts as belonging to the group's area of operation if the convex hull of incidents intersects 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 BLA presence.

Table C-2 replicates Table C-1 using the area of operations instead of the districts in which a group ever operates to count the number of potential other groups. As before, we cannot reject that the coefficients are the same. Note, however, that the coefficient of the other group count is no longer statistically significant, which is likely driven by the districts that fall within the convex hull of a group but that never experience organized political violence perpetrated by that group.

C-2. Active armed groups: One-hit wonders

We now include "one-hit wonders" Blomberg et al. (2010) into our measure of potential active groups. Note that one-hit wonders affect the measure of active and potential active

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.

groups in the same way since, by definition, they only commit one incident in one district in one year. Table C-3 reports the results of another replication of Table II, including the one-hit-wonders in the number of potential groups. Columns 1 to 4 of Table C-3 show that our results are not sensitive to the inclusion of one-hit wonders.

TABLE C-2
Controlling for potential other groups area of operation (convex hull)

	Incidents (1)	Fatalities (2)	Ln Incidents (3)	Ln Fatalities (4)
$Baloch_i \times Post_t$	3.6119	4.9624	2.1067	1.5583
	(0.9979)***	(2.7818)*	(0.3831)***	(0.4444)***
Δ No. potential other groups	0.1099	-1.5900	0.0016	-0.0801
	(0.1914)	(1.1137)	(0.0829)	(0.0944)
Adj. Within-R^2	0.234	0.165	0.135	0.101
	Potential other groups in levels			
$Baloch_i \times Post_t$	3.1656	5.0426	1.8002	1.4120
	(0.7836)***	(2.3005)**	(0.3816)***	(0.4380)***
No. potential other groups	0.1862	0.2288	0.1402	0.0805
	(0.1561)	(0.7824)	(0.0640)**	(0.0703)
Adj. Within-R^2	0.236	0.164	0.140	0.102
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	3144	3144	3144	3144

Notes: The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE C-3
Controlling for other groups: Counting one-hit wonders

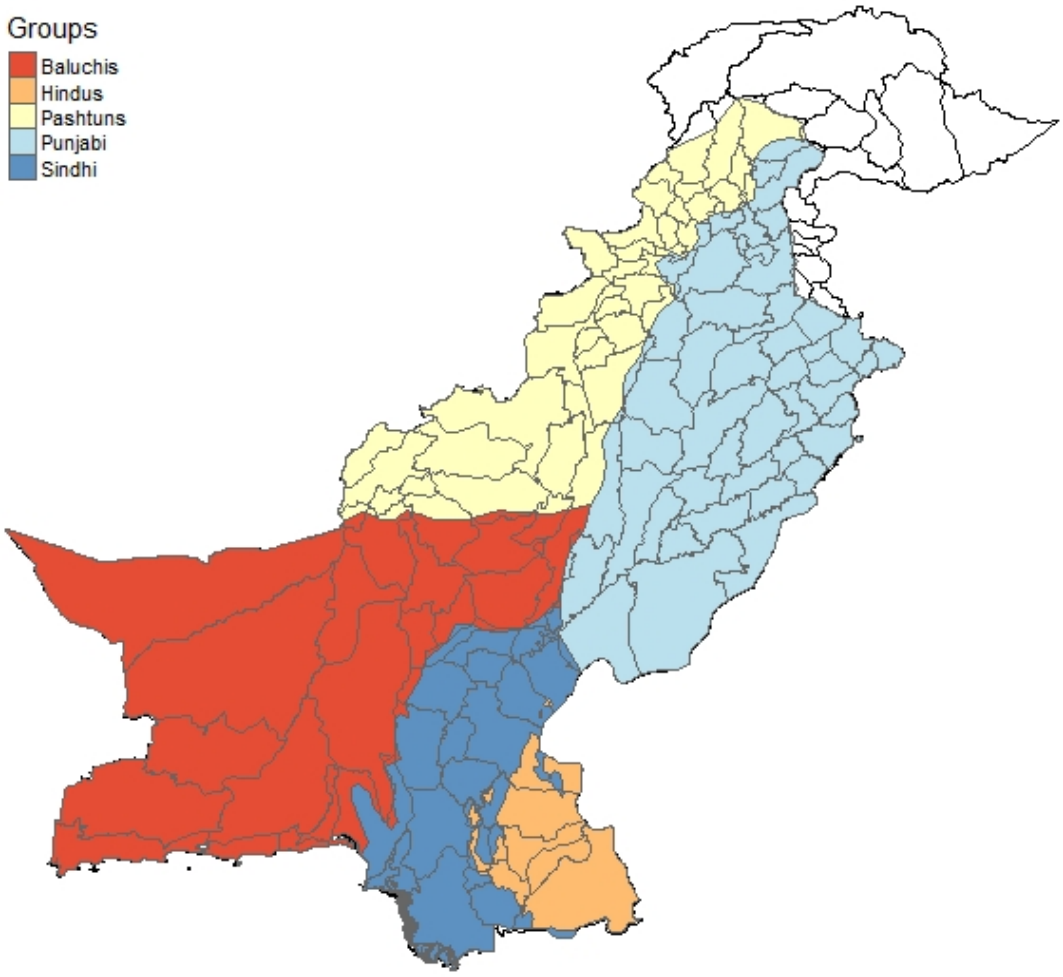
	Incidents (1)	Fatalities (2)	Ln Incidents (3)	Ln Fatalities (4)
$Baloch_i \times Post_t$	3.6128	5.6977	2.1590	1.6355
	(0.9773)***	(2.9259)*	(0.3865)***	(0.4460)***
Δ No. potential other groups	0.9368	3.5604	1.2074	1.0956
	(0.2036)***	(0.5533)***	(0.1375)***	(0.1542)***
Adj. Within. R^2	0.248	0.175	0.235	0.164
	Potential other groups in levels			
$Baloch_i \times Post_t$	2.2923	0.3220	1.2348	0.7108
	(0.7544)***	(2.1246)	(0.2839)***	(0.3166)**
No. potential other groups	2.6158	10.6708	1.7814	1.7923
	(0.2472)***	(1.3799)***	(0.2395)***	(0.2664)***
Adj. Within. R^2	0.374	0.294	0.421	0.324
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	0.248	0.175	0.235	0.164

Notes: Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

D. Ethnopolitical representation

Could our effects be driven by differing ethno-political representations of the people living across Pakistani districts (Bormann et al., 2019)? Given that the conflict in Balochistan is historically a conflict between the tribes of a specific ethnicity (i.e., the Baluchi with the central government), it seems a plausible source for omitted variable bias. We address this issue directly and control for the representation of the ethnic groups within districts over time using the geocoded version of the Ethnic Power Relations (GeoEPR) (Wucherpfennig et al., 2011; Vogt et al., 2015).

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



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

The EPR tracks the political access of political 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 junior or 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.

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 in the respective districts. Hence we have 3 variables ranging between 0 and 1 as additional control variables for our specifications. Note that the senior partner category is left out since it is time invariant. We are aware that this procedure is not suitable to estimate the effect of political representation on violence directly or provide meaningful estimates for specific ethnic power relations. Yet, it should be sufficient to proxy for potentially unobserved changes in the political representation across districts over time.

TABLE D-1
Ethnopolitical representation across districts

	Incidents (1)	Fatalities (2)	Ln Incidents (3)	Ln Fatalities (4)
$Baloch_i \times Post_t$	2.5522	3.9648	1.9981	1.4050
	(0.8023)***	(2.4389)	(0.3788)***	(0.4550)***
Area share discriminated	0.0970	-3.3771	0.2111	0.0634
	(0.4437)	(1.4641)**	(0.3119)	(0.3463)
Area share powerless	-0.4502	-2.5694	-0.5743	-0.3272
	(0.3596)	(0.9233)***	(0.1725)***	(0.2018)
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	2921	2921	2921	2921
Adj. within-R^2	0.388	0.172	0.151	0.110

Notes: The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D-1 replicates Table I adding the EPR-based variables. Note that we use the junior partner status as the omitted category. The main results remain stable. We refrain

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

interpreting our power dummies. Interpretation of the power dummies is not straight forward due to potentially opposing effects. Consider that the share of discriminated increases. This might make the discriminated group more likely to support an armed group within a districts. At the same time, it reduces the share of the junior partner group which is a potential target, in turn suppressing violence within a district. Hence, what we observe is a net correlation.

Given that our treatment group consists of all districts located within the Baloch Province we can also weight the political representation of ethnic groups on the Province level (see Table D-2). Again, our coefficient of interest remains stable throughout all specification.

TABLE D-2
Ethnopolitical representation across provinces

	Incidents (1)	Fatalities (2)	Ln Incidents (3)	Ln Fatalities (4)
$Baloch_i \times Post_t$	4.1426	3.6403	2.0263	1.5053
	(0.9919)***	(2.0383)*	(0.3789)***	(0.4455)***
Area share discriminated	-0.4056	-3.2752	-0.2037	-0.3093
	(0.6529)	(2.2391)	(0.3770)	(0.4134)
Area share powerless	-0.1994	-2.6305	-0.4414	-0.1108
	(0.3921)	(1.1759)**	(0.1992)**	(0.2175)
Controls	Yes	Yes	Yes	Yes
District-FE	Yes	Yes	Yes	Yes
Year-FE	Yes	Yes	Yes	Yes
Obs	2921	2921	2921	2921
Adj. Within-R2	0.243	0.172	0.146	0.109

Notes: The controls include log of population and geographical characteristics interacted with year dummies (see footnote 25). Standard errors clustered at the district level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In summary we confident that changing political representation of ethnic groups across districts and provinces does not bias our estimates.

E. Matching groups between GTD and 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
28		Baloch Liberation Front (BLF)	BLF
29	BLF		BLF

Continued on next page

Appendix E-1 – continued from previous page

No.	GTD Group Name	GED Group Name	Matched Group Name
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 E-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

Continued on next page

Appendix E-1 – continued from previous page

No.	GTD Group Name	GED Group Name	Matched Group Name
95	Bangesh		NA
96		Bhittani tribe	NA
97		Brelvi Muslims	NA
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 E-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

Continued on next page

Appendix E-1 – continued from previous page

No.	GTD Group Name	GED Group Name	Matched Group Name
163		Tehreek-e-Nafaz-e-Shariat-e-Mohammadi (TNSM)	TNSM
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)

F. Splits & mergers of armed groups within Pakistan

TABLE F-1
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	The UBA split from the BLA after the father of both group leaders died. The group leaders are brothers and could not agree who to 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 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 F-2
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)	Reconcilition 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 organisation Lashkar-i-Islam (LI) has merged with the TTP as part of a re-organisation 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.	