The geo-temporal evolution of violence in civil conflicts

Arzu Kibris

a.kibris@warwick.ac.uk

University of Warwick

Department of Politics and International Studies

Abstract

Existing works on diffusion fail to account for the incapacitating effects conflict events may have on the operational capability of the combatant sides and how these effects may determine the evolution a conflict. I hypothesize that it is those events with losses on the state side that are likely to be associated with geo-temporal spillovers whereas events with insurgency losses are less likely to be associated with future mayhem in their vicinity. To test my arguments, I first introduce a new, comprehensive and detailed event dataset on the long-running civil conflict in Turkey. The Turkey-PKK Conflict Event Database (TPCONED) includes the exact date and county level location for the fatal events of the armed conflict between the Turkish state and the rebel organization PKK since its very beginning in 1984 with detailed information on combatant casualties. I then employ a split population biprobit model which allows me to comprehensively depict the geotemporal evolution of the conflict by acknowledging, estimating and accounting for the variation in the underlying conflict proneness across locations as a latent variable that shapes the diffusion of events. The results of the statistical analyses offer support for my hypotheses and reveal that how events evolve over space and time is conditioned by the damages suffered by the combatant sides. I demonstrate the robustness of these results on a matched sample I obtain by employing the Coarsened Exact Matching (CEM) on the data.

Keywords: civil conflict, spatial analysis, conflict event dataset, split population model
9996 words.

Introduction

How do civil conflicts evolve over time and space? Are they chaotic episodes of armed violence in fragile states or is there a predictable pattern to the diffusion of events that we can decipher? Given the prevalence and destructiveness of civil conflicts it is imperative that we answers these questions, because if we can figure out the spatial and temporal dynamics, we can then hope to devise pre-emptive actions and policies to prevent or at least contain these humanitarian disasters.

The literature on conflict diffusion is dominated by works on international contagion and focuses on the factors that facilitate the spread of civil conflicts between countries (Ward & Gleditsch, 2010; Buhaug & Gleditsch, 2008; Cederman, Girardin & Gleditsch, 2009; Cederman et al., 2013; Lake & Rothchild, 1998; Lane, 2016; Salehyan & Gleditsch, 2007; Weidmann, 2015). Compared to this rich literature on cross-country diffusion, the literature on diffusion of events within conflicts is still in its early stages. Nevertheless, existing studies have already firmly established that conflict events exhibit spatial and temporal interdependencies (Townsley, Johnson & Ratcliffe, 2008; Hegre, Ostby, & Raleigh, 2009; Lyall, 2009; Raleigh et al., 2010; Weidmann & Ward, 2010; Schutte & Weidmann, 2011; O'Loughlin & Witmer, 2012). This strong result has then led scholars to explore the mechanisms that explain this spatio-temporal clustering of violence. Recent works have theorized about and provided empirical evidence for the role played by the relative capabilities of combatants (Holtermann, 2016; Beardsley, Gleditsch & Lo, 2015), accessibility of locations via road networks (Zhukov, 2012), environmental conditions (Carter & Veale, 2015), and retaliatory motives (Braithwaite & Johnson, 2012; Linke, Witmer & O'Loughin, 2012) in determining the time and location of conflict events. Surprisingly however, none of these works pays any attention to the incapacitating effects these events may have on the warring factions and how these effects may determine the evolution of the conflict. Even those studies that emphasize the tit-for-tat nature of conflicts seem to have forgotten what Machiavelli reminded Lorenzo Di Pierro De Medici 400 years ago: '*(men) can avenge themselves of lighter injuries, of more serious ones they cannot'* (p.7).

In this study, I argue that how an insurgency spreads within a country is associated with how conflict events affect the operational capabilities of the warring sides, or in Machiavelli's words, with the amount of *injury* each side suffers as a result of each event. Employing a split population bi-probit model I test my arguments on a new event dataset I introduce on the long running civil conflict in Turkey. The results offer support for my arguments. Confirming previous diffusion studies, I find that once a civil conflict starts geotemporal interdependencies play an important role in determining how events evolve over time and space. Significantly contributing to our understanding of these interdependencies, and in line with my theoretical arguments, my findings indicate that the geotemporal evolution of the conflict is conditioned by the damages suffered by the combatant sides, and that it is events with losses on the state side that are likely to be associated with geotemporal spillovers, whereas events with insurgency losses are less likely to be followed by further violence in the neighborhood.

In the next section, I discuss diffusion patterns in civil conflicts and introduce my theoretical arguments. Then I introduce my case, my data, my statistical model, the results and robustness checks in the following sections.

Theory and Literature

The conflict diffusion literature has two main branches. The first branch focuses on international diffusion and analyzes the mechanisms behind transnational spillovers of violence. The works under this branch are in fact cross-country studies of conflict onset that assess a country's risk of experiencing a civil conflict based on its local characteristics and its interactions with the outside world. Among the interactions identified in this literature as transborder carriers of conflict risk are refugee flows (Salehyan & Gleditsch, 2007; Braithwaite, 2010; Rüegger, 2018), communication networks (Weidman, 2015; Beiser, 2016), circulation of arms and combatants (Lane, 2016; Bara, 2018; Braithwaite & Chu, 2018), learning and strategic emulation (Buhaug & Gleditsch, 2008; Maves & Braithwaite, 2013; Hill, Rothchild & Cameron, 1998; Forsberg, 2008; Ayres & Saideman, 2000), and external sponsorship of insurgencies (Gleditsch, Salehyan, & Schultz, 2008; Cederman, Girardin & Gleditsch, 2009; Schultz, 2010).

The findings of this impressive literature provides theoretical and methodological insights for the second branch which takes a micro-level approach to analyze the within-conflict diffusion of violence. This literature is still in its early stages; however, existing works have already firmly established that conflict events exhibit a strong spatial and temporal interdependency (Townsley, Johnson & Ratcliffe, 2008; Hegre, Ostby & Raleigh, 2009; Lyall, 2009; Raleigh et al., 2010; Weidmann & Ward, 2010; Schutte & Weidmann, 2011; O'Loughlin & Witmer, 2012). Recent works have investigated the mechanisms that explain this spatio-temporal clustering and provided empirical evidence for the role played by factors such as the accessibility of locations via road networks (Zhukov, 2012), environmental conditions (Carter & Veale, 2015), internal displacements (Bohnet, Cottier & Hug, 2018), and state coercion (Duffy Toft & Zhukov, 2012) in determining the time and location of conflict events. Interestingly however, attention is yet to be paid to the incapacitating effects conflict events may have on the warring sides, and how these effects may influence the likelihood of geotemporal spillovers. This, I argue, is a serious shortcoming. Civil conflicts involve strategic actors making strategic decisions on whether, when and how to act. It is, thus, not possible to construct an accurate understanding of how conflict events evolve over time and space without taking into account those factors that shape these decisions. Operational capability is one of the most important of such factors. Operational capability refers to the capacity of conflict actors to carry out successful combat operations against the adversary on the battlefield, and is mainly determined by resources available to them and their ability to successfully use those resources (Tellis et al., 2000). Note that conflict events can inflict damages on both of those determinants of operational capability and as such can have incapacitating effects on conflict actors.

The argument that the evolution of a conflict is associated with how conflict actors are affected by events has in fact already been raised by those works that emphasize the tit-for-tat nature of conflicts and the retaliatory motives of conflict actors (Jaeger & Paserman, 2008; Lyall, 2009; Haushofer, Biletzki & Kanwisher, 2010; Linke, Witmer & O'Loughlin, 2010; Kocher, Pepinsky & Kalyvas, 2011; O'Loughlin & Witmer, 2012; Braithwaite & Johnson, 2012; Schutte & Donnay, 2014). However, even those works implicitly assume that combatants will always have the capacity to react against all instances of violence. In this study, I relax this assumption by controlling for combatant casualties in each violent event as a measure of the damage to the operational capability of warring sides.

The size of the military force is an important resource in all armed conflicts, but especially so for insurgents fighting against powerful state adversaries. In fact, the opportunity cost theories of conflict onset posit the availability of labor as the binding constraint on the production of violence by insurgencies (Grossman, 1991; Mikulascheck & Shapiro, 2018). Most civil conflicts are fought between organized, well-armed, and sizable state military forces and relatively much smaller and ill-equipped insurgent groups[1]. Balcells & Kalyvas (2014) use the term 'irregular conflicts' (p.1391) to refer to civil conflicts with such power asymmetry. The asymmetry in resources in irregular civil conflicts means that, compared to state forces, casualties are expected to be more costly and debilitating for insurgents since each combatant corresponds to a higher share of their operational capability. The relative difficulty of recruiting and training replacements inflates this cost further. Heavy losses against state forces may also

[1] According to the Technologies of Rebellion dataset, during the Cold War period, 66.34% of all major civil conflicts were asymmetric (Kalyvas & Balcells, 2010).

have a deterrent effect on both the insurgents and their civilian support base which may then render recruitment even more difficult and may even lead to defections.

The asymmetry in resources not only renders them more valuable for insurgents but also shapes the way they utilize them. Due to state's material advantages, insurgent groups stand a high risk of defeat if they try to fight the state with conventional tactics in consistent theaters of combat. They thus favor mobility and guerilla warfare. Staying mobile helps them evade attacks. It also gives them an opportunity to compete with the state's armed forces by varying targets and using the element of surprise to their advantage (Beardsley, Gleditsch & Lo, 2015). McColl's (1969) influential account of how revolutionary insurgencies evolve emphasizes the need for mobility as well.

Note that this type of irregular warfare creates a pattern in which insurgents proactively stage hit-and-run attacks which then drive state forces into reactive counterinsurgency operations[2]. But, the insurgents' capacity to sustain such a pattern depends on whether they can hit their targets and run afterwards without getting hit themselves. Casualty counts give us a grim account of their success in doing so. While security force casualties provide a measure of the *hit* the state side takes, insurgency casualties provide a measure of their (in)ability to run. It follows that with each insurgent casualty this pattern becomes less sustainable, and the likelihood of future hit and run attacks in the vicinity goes down. On the other hand, state casualties can be expected to bolster this pattern. It is reasonable to expect damages on state forces to make it easier for insurgents to escape after an attack. Successful attacks may also boost morale among rebels, help them in their efforts to gain public support and find new recruits (Kalyvas, 2006) thereby allowing them to increase the geotemporal scope of their activities.

Casualties are important in determining the actions of state forces as well. The size advantage and the relative ease of recruitment shield the operational capability of state forces against losing soldiers, thus state casualties are less likely to have a dampening effect on conflict activity. On the contrary, in many cases, losses on the state side carry heavy political costs for state leaders (Kibris, 2011), and lead them to resort to retaliatory actions that perpetuate violence (Jaeger & Paserman, 2008). Holtermann (2016) argues that state casualties can also be seen as a measure of the relative capacity of rebels and as such they can indicate higher likelihood of future conflict events in the neighborhood. Relatedly, losses in one area may lead

[2] Confirming this pattern, Linke, Witmer & O'Loughlin (2012) indicate that most attacks around Baghdad in the 2004-2009 period were initiated by insurgents (p.6).

to a transfer of state military resources from nearby locations and may leave those areas vulnerable.

If these are valid mechanisms then we should expect events with insurgency casualties to curb the potential for future events in nearby locations and events with state security force casualties to be harbingers of geotemporal spillovers. The next two hypotheses are derived from these expectations:

Hypothesis 1: In irregular civil conflicts, casualties on the state side are positively associated with future conflict events in other locations in the neighborhood.

Hypothesis 2: In irregular civil conflicts, casualties on the insurgency side are negatively associated with future events in other locations in the neighborhood.

I test these hypotheses on a new and detailed event dataset I introduce on the long running civil conflict in Turkey. The Turkey-PKK Conflict Event Database (TPCONED) includes the exact date and county level location for the fatal events of the armed conflict between the Turkish state and the rebel organization PKK in the 1984-2018 period with detailed information on combatant casualties.

Statistical analysis of conflict diffusion offers an empirical challenge because it comes with very specific and demanding data requirements. In order to be able to track the geotemporal path of violence, one needs a comprehensive and complete event dataset with detailed information on the time, location and characteristics of conflict events. Comprehensiveness, which is full geotemporal coverage of the conflict, is important to make sure that any observed association is not specific to a period or a location. Completeness, which is not having any missing observations, is even more important. Missing observations introduce a serious selection bias in any statistical analysis, but they are even more problematic in diffusion studies since with each missing observation another step in the geotemporal progression of events gets lost and the accuracy of the data set in reflecting the diffusion of violence weakens. TPCONED is a comprehensive and complete event data set on the armed conflict between the rebel organization PKK and the Turkish state, and as such, it avails the armed conflict in Turkey as a rich case study for understanding the geotemporal dynamics of civil conflicts.

The conflict

Since late 1984, Turkey has been suffering from an insurgency campaign led by the Kurdish separatist guerrilla organization the Kurdistan Workers' Party (PKK). The organization was first founded with the goal of establishing an independent Kurdish state in south-eastern Turkey, though later on in the 1990s, it appeared to have rolled back on its goal to a federational structure that would grant more autonomy to the large Kurdish minority in the country. The armed conflict between the PKK and the Turkish security forces (TSF) has been geographically concentrated in south-eastern and eastern regions which have traditionally been inhabited by ethnic Kurds. First dismissed by Turkish governments as the acts of a handful of outlaws, this irregular conflict has been going on for more than 35 years[3]. Financially, it has cost the country billions of dollars. But more importantly, it claimed the lives of tens of thousands.

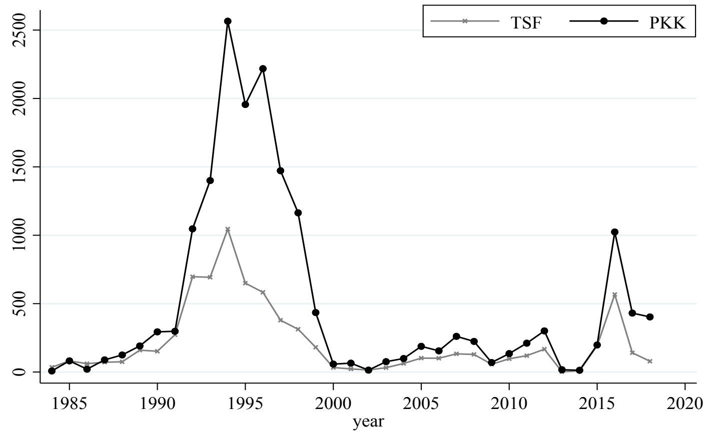

Figure 1. Combatant casualties, 1984 - 2018

Figure 1 depicts the total number of TSF and PKK casualties over time. As can be seen, the 90s has been the bloodiest period of the conflict. The insurgency announced a unilateral ceasefire after its leader Abdullah Öcalan was captured in 1999, and ceased its attacks in the early 2000s. Unfortunately, peace in the area did not last long and violence flared up again after 2004. The latest ceasefire was announced in March 2013 as part of a peace process which unfortunately broke down in July 2015 only to bring more bloodshed. Figure 2 depicts the geographical distribution of combatant casualties at the province level and reveals their geographical concentration.


Figure 2. Total combatant casualties, 1984- 2018

[3] The Turkish prime minister of the time referred to the PKK as a handful of outlaws after their first attacks in 1984 (Pulur, 2010).

The Turkey-PKK conflict event database - TPCONED

I test my hypotheses with a new, high-resolution event dataset that I introduce in this study on the armed conflict between the rebel organization PKK and the Turkish state. This is one of the longest running civil conflicts in the world, and it plays an important role in the political turmoil of the Middle East, and as such it has a lot to tell us about civil conflicts. However, due to scarcity of reliable data, it has been analyzed by only a handful empirical studies so far. I release the Turkey-PKK Conflict Event Database (TPCONED) with the hope of motivating those much-needed studies that will bring this deadly conflict into the spotlight.

TPCONED (https://wrap.warwick.ac.uk/138227) is a dataset on the fatal events of the armed conflict between the rebel organization PKK and the Turkish state that took place on Turkish soil and in which there was at least one combatant casualty. It contains detailed information on 7063 conflict events with 17308 PKK casualties and 7514 state casualties over the course of the conflict in the 1984-2018 period[4]. For each event, the exact date, location at the county (town) level, location at the province level, number of TSF and PKK casualties, and the data sources are listed. The dataset has been in the making since 2009, and relies on a wide range of local sources which include reports and publications by the Turkish Ministry of Defence, Turkish General Staff, the Directorate General of Press and Information of Turkish Prime Ministry and by various other state offices and local administrations in Turkey; the archives of the Gendarmerie Museum in Ankara; the digital and hard copy archives of five major Turkish daily newspapers (Cumhuriyet, Milliyet, Hürriyet, Sabah, and Zaman); PKK publications; daily and yearly reports by the Turkish Human Rights Association; scholarly works (Tezcür, 2016); and personal contacts and interviews with the families of the casualties. I provide detailed information on these sources in section A1 in the online appendix.

Note that my sources belong to different political affiliations and views. I cross checked contentious events across these sources to make sure that the resulting dataset is as comprehensive and accurate as possible. A very important thing to note here is that combatant casualties do not only serve as a binary decision rule for an event to be included in TPCONED but more importantly they provide a solid confirmation of the validity of each event in the dataset.

Turkey has a draft army and a mandatory military service system that requires each Turkish man, when he comes of age, to serve in the army for about a year. This mandatory system is supported by a military and religious culture that glorifies the army and service to the country.

[4] The dataset is continuously updated. Currently, events of 2020 are being coded.

When a soldier or a police officer dies in service he or she is considered a *martyr,* both legally and culturally. A state funeral is held for each of them in their hometowns. These funerals have always been important public events attended by high-level military and state officials and thousands of locals. The national and the local media cover the story of the fallen soldier or officer, his grief-stricken family, and the funeral ceremony, and by so extensively publicizing the casualties on the state side, they create the most credible way of tracking down the events and the geotemporal evolution of the conflict.

A similar glorification of combatant casualties is also present on the PKK side. They also refer to their casualties as martyrs and commemorate them by special publications. Although very limited compared to the information on the Turkish side, such resources present valuable information on conflict events.

Hence, all the events recorded in TPCONED have been confirmed by information on the identity of the each of the 7514 TSF casualties involved and of a good portion of the PKK casualties[5].

To demonstrate the accuracy, precision and comprehensiveness gains that TPCONED accomplishes, I present a systematic and detailed comparison (á la Donnay et al., 2018) of TPCONED and the UCDP-GED database (Sundberg & Melander, 2013), which is one of the most commonly used conflict event databases by researchers, in section A2 in the online appendix. Note that the GED does not have any data on the conflict for its first five years, and then has only a total of 4682 conflict events recorded for the 1989-2018 period. Moreover, of the recorded events, only 54%, have location information at the county level, that is, location at the county level is missing in 46% of the observations. With nearly half the observations missing, it becomes quite impossible to track the evolution of the conflict at the county level with the GED data. Unfortunately, the problem with missing observations persists at larger geographical aggregations as well. For 7.5% of the observations, reported location covers almost the entire conflict zone[6].

Another major problem that diminishes the quality of existing conflict event databases is that most of them rely on international news agencies. Going back to the GED example, nearly

[5] No personal information on the casualties is included in TPCONED but anonymized information is available on request.

[6] The Global Terrorism Database (GDT) which is also commonly used in conflict studies, includes even fewer observations with only 2208 events recorded for the conflict between the Turkish state and the PKK. Moreover, the database's coverage remains very lacking for the 1990's which, as Figure 1 depicts, is the most intense period of the conflict. A more serious shortcoming for my purposes is that the GTD reports fatalities in aggregate without any breakdown, thus, it does not allow one to assess the losses of the sides. Finally, data source is missing for 850 (38.5%) observations and location is coded as unknown for 14% (309) of the recorded events.

all the observations under the Turkey-PKK conflict are coded from sources like the BBC, Reuters, Agence France Press, etc. Admittedly, these international news agencies themselves rely on local sources for information, but they tend to rely more on state sources or agencies which, in many cases, feed them with biased information. In the Turkish case for example, they mostly rely on the Anadolu (News) Agency which is in fact owned and controlled by the Turkish state. Moreover, they tend to be biased towards more substantial events in their reporting while ignoring smaller ones. They also tend to summarize the information they receive from local sources and leave out certain details, like precise location of the event, which they deem irrelevant for their international readers. Moreover, in many cases, a lot gets lost in translation.

Statistical model and analyses

Simply stated, diffusion refers to the temporal and geographical spread of conflict events. In this study, I define a conflict event as any conflict related lethal violence that results in the death of at least one combatant. I use casualties as a measure of the size of losses in terms of operational capability of each warring faction in each conflict event.

In most civil conflicts, violence remains geographically clustered, and a great deal of places never experiences any events while a small subset of locations accounts for nearly all the violence in a recurring fashion. Civil conflicts are violent struggles about state formation (Weinstein, 2007; Tull, 2004; Pegg, 1998), in other words, insurgencies are would-be states. Note that no such would-be-state-maker takes up arms without a targeted territory. Insurgents may target to take over a country as a whole or some parts of it, either because it is the ethnic homeland, the *promised territory*, or because it is rich in economic resources, or home to ideologically sympathetic populations, etc. The fact that insurgents target certain territories implies that there is an underlying, base risk distribution over space, and that the geotemporal evolution of the conflict is correlated with this distribution. This implication is consistent with the observed geographical clustering of conflict events. It is also consistent with McColl's (1969) account of how insurgencies evolve. The targeted base areas become the hubs of conflict events, the areas in their periphery experience relatively fewer incidents as insurgents try to expand, and the rest of the territory remains outside the conflict zone. But if particular territorial units are not likely to experience any conflict events, then their inclusion in the sample may lead the diffusion analysis astray. Braumoeller & Goertz (2002) and Mahoney & Goertz (2004) argue that including irrelevant cases where the outcome of interest is impossible induces erroneous inference. Having said that trying to exclude them can also be very problematic since it requires distinguishing between the relevant and irrelevant areas when relevancy is unobservable. Any ad hoc identification risks excluding relevant areas or including irrelevant ones. Note that a similar problem had been discussed in the study of international conflicts. There the argument is that some country dyads, because they lack any opportunity to fight, are not relevant for an empirical study on the determinants of international war. To deal with this problem, Clark & Regan (2003) and Xiang (2010) propose split population models in which they estimate relevance as a latent variable. In this study, I follow their proposition and employ a split population bi-probit model to analyze the geotemporal evolution of conflict events.

In a split population model an additional binary choice regression—used as a selection step—is added to a standard statistical model to capture the idea that there are two data generating processes behind the observed data. In the context of modelling civil conflict diffusion, the selection step is meant to estimate the underlying risk distribution over territorial units and identify the potential base areas. A split population model, in that sense, can be used to bring together the civil conflict onset literature which explains the occurrence of violence in relation to specific conditions existing prior to the conflict, and the diffusion literature which emphasizes the importance of geotemporal interdependencies[7].

My dependent variable $Y_{i,t}$ is a binary incidence variable that takes on the value 1 if a conflict event took place in location i in time t, and 0 if not, and

$$Y_{i,t} \sim 0 \text{ with probability } 1 - p_i \text{ and}$$

$$Y_{i,t} \sim F_{i,t} \text{ with probability } p_i$$

where $F_{i,t}$ is a cumulative distribution function for a binary choice model.

Let $p_i = G_i$ where G_i is also a cumulative distribution function for a binary variable. Then, because the zero outcome is generated by both the binary choice models G_i and $F_{i,t}$, we have

$$Y_{i,t} = 0 \text{ with probability } (1 - G_i) + G_i(1 - F_{i,t}) \text{ and}$$

$$Y_{i,t} = 1 \text{ with probability } G_i F_{i,t}$$

under the assumption that the two distributions are independent. Here, G_i determines the distribution of underlying conflict risk and $F_{i,t}$ determines the distribution of event risk. Note that the underlying risk is not observable. We only observe whether a location experiences a conflict event at a given time or not.

The independence assumption is not a reasonable one in a civil conflict context. Locations with high underlying risk will most probably have high probability of experiencing conflict events. I therefore adopt a bivariate standard normal distribution to model the two correlated cumulative distribution functions. As a result, the above model becomes

[7] The identification restrictions are discussed in Section A3 in the online appendix.

$$Y_{i,t} = 0 \text{ with probability } [1 - \Phi_2(\beta X_{i,t}, \gamma Z_{i,t}; \rho)]$$

$$Y_{i,t} = 1 \text{ with probability } \Phi_2(\beta X_{i,t}, \gamma Z_{i,t}; \rho)$$

where Φ_2 is the bivariate standard normal cumulative distribution function and ρ is the correlation coefficient. Z is the vector of covariates that affect the base risk, and X is the vector of covariates that affect the probability of diffusion. This is very similar to the binary choice model Xiang (2010) employs in the context of international war.

The subscripts *i* and *t* designate time and space units to be used. In terms of designating space, the literature hosts two approaches. The first approach makes use of administrative divisions (Holtermann, 2016; Ward & Weidman, 2010; O'Louglin & Witmer, 2012; Zhukov, 2012) to identify *distinct spaces*. Scholars justify the *distinctiveness* of administrative units by referring their socioeconomic differences. However, as Schutte & Weidmann (2011) argue 'administrative boundaries may have little relevance in civil wars, since they can be crossed easily by armed forces' (p.147). But this concern is even more valid for the second approach which divides the area of study into grid cells of equal size and treats each grid cell as a *distinct space* (Schutte & Weidman, 2011; Raleigh & Hegre, 2009; Townsley, Johnson & Ratcliffe, 2008). Unfortunately, there is not much discussion in the works that employ geographical grids on what makes two adjacent grid cells distinct spaces from the perspective of conflict actors and/or for the sake of conflict dynamics. To deal with the ad hoc nature of this gridding exercise two alternatives have been offered so far. The more commonly adopted alternative is to repeat the analyses with multiple grid sizes. While this practice may help ease concerns about the robustness of estimated associations, it does not provide any theoretical justification for the *relevance* of resulting geographical units. Moreover, exact event coordinate information is very hard to come by. Most conflict event databases, including the GED, report the coordinates of the administrative unit in which the event takes place. The second alternative is offered in a recent study by Schutte (2017) where he employs a point process model (PPM) in which, rather than relying on predefined spatial units, suitable spatial windows are heuristically defined from the distribution of events over territory. As Schutte (2017) argues, these windows offer a better alternative to ad hoc spatial grids but are so far limited to cross-sectional analysis as 'the introduction of a temporal dimension provides additional challenges' (p.454).

Schutte & Weidman (2011) argue that spatially aggregating conflict activity leads to a loss of information, and they favor using as fine a spatial resolutions as possible with the data to hand. In this study, I follow up on their advice and estimate my model under the highest spatial resolution possible with my data, namely, the county level. This is quite a high resolution especially considering that I am able to sustain it across all events throughout the whole span of the conflict and that it cannot be attained by any of the other available data sets. Moreover, my socioeconomic controls are able to match the geographical resolution of my event data. To make sure that my results are not specific to any time aggregation, I estimate my model under monthly, quarterly, half-yearly and yearly time denominations.

My X vector of diffusion covariates contains within conflict dynamics. To control for the geotemporal interdependencies, I include the length of the last peace spell; the percentage of past periods with conflict incidences; one-period lagged state casualties; one-period lagged insurgent casualties; one period lagged inverse-road-distance-weighted state casualties in other counties; and one period lagged inverse-road-distance-weighted insurgent casualties in other counties. These last two are the control variables of interest for this study as they are the spatiotemporal lags whose association with the dependent variable will inform us about the extent and nature of geotemporal spillovers. Unlike existing studies, I do not limit geotemporal interdependencies to immediately neighbouring administrative units or units within a fixed distance. Instead, my spatio-temporal lags take into account all events of the previous period after assigning them weights according to distance. Accordingly, the one period lagged inverse-road-distance-weighted insurgent casualties for county i at time t is

$$\Sigma_{k\neq i}\ (1/d_{ki})(\text{state casualties})_{k,t-1}$$

where d_{ki} is the road distance between county k and county i in kilometers. Similarly,

$$\Sigma_{k\neq i}\ (1/d_{ki})(\text{insurgency casualties})_{k,t-1}$$

is the spatio-temporally lagged insurgency casualties for county i at time t[8]. I use road distances to calculate these spatial lags in order to account for the role of the road networks in facilitating the spread of violence (Zhukov, 2012).

I also control for area, border status and percentage of rural mountainous terrain across counties. Finally, I include season and year dummies to capture the seasonal and yearly variations.

My Z vector of covariates for base risk includes a rich set of controls which depict the socioeconomic situation at the start of the armed conflict. I control for percentage of farmers

[8] Simply put, for each county i, these expressions add up the casualties that took place in all other counties in the previous period after dividing the casualty count from each county by the road distance between that county and county i. This way, casualties elsewhere are weighted according to their distance from county i. Consequently, casualties from nearby counties get assigned a higher weight in the calculation of the weighted average, whereas casualties that took place in faraway counties are divided by their higher distances and so get assigned much smaller weights.

with no land; percentage of farmers with more than 100 acres of land; number of mosques per village; percentage of villages with no drinking water; percentage of villages with no electricity; unemployment rate; literacy rate and level of urbanization across counties as measures of state capacity and reach, and as indicators of economic development and welfare. Economic grievances and low state capacity create fertile environments for insurgencies (Holtermann, 2012). I also control for the percentage of villages whose names were changed by the state as a measure of ethnic discriminatory policies. Ethnic discrimination is identified as an important factor that increases the vulnerability of a country to experience conflict (Metternich, Minhas & Ward, 2017; Cederman et al., 2013; Cederman, Wimmer & Min, 2010). Relatedly, I also control for the ethnic distribution of the population across counties to account for the potential support base of the PKK.

Finally, I control for the percentage of mountainous area; border status and population across counties (Fearon & Laitin, 2003; Do & Iyer, 2010; Weidman & Ward, 2010; Zhukov, 2012; Holtermann, 2016; Cederman, Girardin & Gleditsch, 2009; Cederman et al., 2013).

I provide a detailed discussion on the control variables in Section A4 in the online appendix along with descriptive statistics and a visual representation of their predictive power in terms of the base risk of conflict across counties.

Results

Table I below presents the estimated parameters of the bivariate probit model under monthly, quarterly, half-yearly and yearly time aggregations.

The estimated coefficients for the within-conflict-dynamics equation reveal the importance of geotemporal interdependencies. The longer a county stays peaceful the less likely it becomes to experience violence. Similarly, counties with more troubled histories, measured by percentage of past times with conflict events and with casualties in the county itself in the previous period, are more likely to experience conflict events.

As hypothesized, results indicate that geotemporal spillovers are conditioned by the losses of the sides. The estimated parameters for the inverse-distance-weighted-lagged-casualties indicate that it is the state force casualties in an area that are heralds of conflict events in nearby locations. On the other hand, events with insurgent casualties have a significant dampening impact on future events in the vicinity especially in the longer runs. These results are consistent with the asymmetric power structure in civil conflicts and with a rebel strategy of hit and run attacks whose likelihood of spilling over to nearby locations increases as state forces incur more military losses and decreases as insurgents themselves get hit.

TABLE I. Partial observability bivariate probit results				
	Unit of observation: **County-month (1984-2018)** Number of obs.: 266845	Unit of observation: **County-quarter year (1984-2018)** Number of obs.: 88734	Unit of observation: **County-half year (1984-2018)** Number of obs.: 44367	Unit of observation: **County-year (1984-2018)** Number of obs.: 21862
Within conflict dynamics				
Length of last peace spell	-0.004** (-13.16)	-0.012** (-12.35)	-0.054** (-5.63)	-0.037* (-2.13)
Percentage of past time periods with conflict events	0.035** (10.06)	0.030** (16.60)	0.017** (10.59)	0.012** (9.42)
Inverse- distance-weighted **state casualties in other counties** in t-1	1.284** (6.43)	0.619** (6.25)	1.280** (4.82)	0.827** (5.92)
Inverse-distance-weighted **insurgent casualties in other counties** in t-1	0.084† (1.90)	-0.069** (-2.09)	-0.119** (-2.74)	-0.158** (-4.27)
State casualties in the county in t-1	0.102** (6.08)	0.047** (5.34)	0.060* (2.19)	0.158** (2.73)
Insurgent casualties in the county in t-1	0.020* (2.41)	0.012** (2.99)	0.025** (3.21)	0.067** (2.79)
County area in km squares	0.0001** (4.04)	0.0001** (3.52)	0.0001* (2.50)	0.0001* (2.39)
Percentage of mountainous terrain in rural area	0.002* (2.02)	0.003** (3.07)	0.001 (0.22)	0.001 (0.20)
Border status	0.037 (0.56)	0.028 (0.55)	0.056 (0.45)	0.115 (0.88)
Estimated parameters for seasonal and year dummies are not reported.				
Base risk				
Percentage of mountainous terrain in rural area	0.002 (0.80)	0.001 (0.27)	0.006 (0.99)	0.004 (1.09)
Percentage of villages whose names were changed by the state	0.007** (2.66)	0.008* (2.27)	0.006* (2.02)	0.004* (1.97)
Percentage of landless farmers	0.013** (3.02)	0.008 (1.45)	0.011† (1.88)	0.007** (2.94)
Percentage of farmers with more than 100 acres of land	-0.023** (-3.45)	-0.021** (-3.09)	-0.014* (-2.17)	-0.010** (-2.98)
Number of mosques per village	-0.590** (-4.32)	-0.260† (-1.68)	-0.423* (-2.47)	-0.223* (-2.49)
Percentage of villages with no drinking water	0.005 (1.15)	-0.002 (-0.70)	0.003† (1.93)	0.002 (1.27)
Percentage of villages with no electricity	0.004† (1.81)	0.003 (1.46)	0.005† (1.86)	0.003† (1.95)
Border status	0.478 (1.32)	0.798** (3.26)	0.341 (1.09)	0.159 (0.88)
Percentage of ethnically Kurdish population	0.031** (4.49)	0.081** (4.98)	0.013* (1.98)	0.010** (3.21)
Urbanization rate	-0.009 (-1.39)	-0.004 (-0.56)	-0.009 (-1.46)	-0.002 (-0.62)
Literacy rate	0.008 (0.95)	-0.013 (-1.03)	0.008 (1.06)	0.004 (0.87)
Unemployment rate	0.107* (2.51)	0.094* (2.22)	0.085 (1.54)	0.041 (1.74)
Population in 10 thousand	0.016** (2.77)	0.012** (2.84)	0.022** (2.99)	0.012** (3.05)
Wald Chi-square	3786.69	3604.00	687.50	269.11
Correlation (rho) between the two equations (robust standard error)	0.341 (0.191)	0.684 (0.131)	-0.231 (0.325)	-0.912 (0.282)
Wald test of rho=0: Chi-square	2.70*	11.56**	0.47	10.45**
Standard errors are adjusted for 648 counties. z-values in parenthesis. † $p<0.1$, * $p<0.05$, ** $p<0.01$.				

The estimated parameters of the base risk equation point towards the importance of state reach and control. Mosques as preachers of state ideology, and big land owners as the agents of the state in the rural areas seem to make an important dampening impact on the underlying conflict risk.

As expected, base risk is significantly higher for counties with higher Kurdish population percentages. Those locations which had been subject to discriminatory state policies and had their names Turkified by the state have a higher likelihood of being part of the conflict area.

Difficult terrains are not significantly associated with the base risk but they facilitate the diffusion of violence in the short run.

To make sure that these results are not specific to the biprobit model, I replicated the above analyses with a simple logit model as well. The estimated associations between the incidence probability and the control variables remain similar in direction but get statistically stronger[9]. I present the results in Table A5 in the online appendix.

Figure 3 plots the average marginal effects of spatio-temporally lagged casualties in order to give a better understanding of their relative substantive significance. The marginal effect of a variable corresponds to the average expected change in the estimated probability of observing a fatal conflict event in county i at time t given a marginal change in that control variable while other control variables are evaluated at their mean values. As can be seen, the spatio-temporal lag of state casualties has a high marginal effect on the probability of diffusion. The average marginal effect of a unit-increase in inverse-distance-weighted state casualties on the probability of observing a fatal conflict event at county i in time t ranges from 2.5% to 7.1% across different temporal aggregations. On the other hand, the average marginal effect of a unit-increase in inverse-distance-weighted insurgency casualties ranges from (-1.3) % to 0.0% depending on the time unit[10].

[9] I also reran the model after controlling for the vote share of the ethnically Kurdish parties across counties in the 1995, 1999, 2002, 2015 and 2018 general elections. The inclusion of this control either in the base risk or the diffusion equation does not create any substantive difference in the results which are available upon request.

[10] I report the estimated average marginal effects for all control variables in Table A4 in the online appendix.

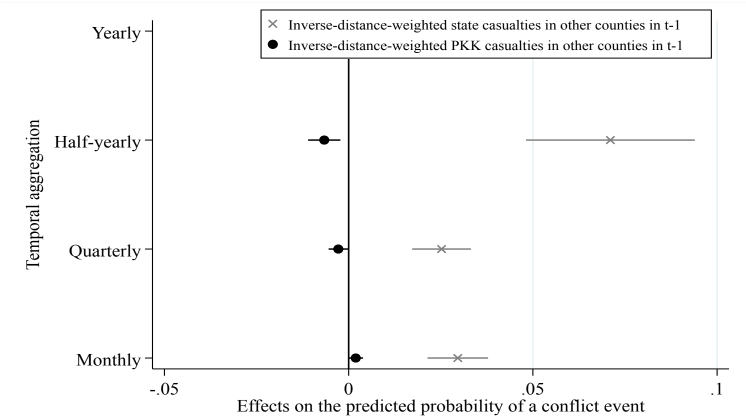

Figure 3. Average marginal effects with 95% CIs

Robustness of Results

In this section, I employ the Coarsened Exact Matching (CEM) (Blackwell et al., 2009) technique as a robustness check on my results.

Matching is a nonparametric method of controlling for some or all of the confounding influence of pretreatment control variables. The key goal of matching is to prune observations in order to better approximate experimental conditions in observational data. More specifically, a statistical matching criterion based on the similarity of confounding factors or their effect on the probability of treatment is used to pair treated and untreated observations, and only those matched observations are retained in the dataset while the unmatched ones are discarded. I argue that matching is a good robustness check for the results I obtain from the above split population biprobit model because it in fact offers another way of defining the *relevancy* of observations based on their similarity in terms of treatment likelihood. Those unmatched observation are left out as *irrelevant* for the purpose at hand.

The predicted base risk gives me a composite pre-treatment variable subsuming the socioeconomic pre-conflict variables across which I can match my observations. Note that matching techniques are usually employed on cross-sectional data with binary treatments, whereas I have a panel data set and my treatment, i.e. the spatio-temporally lagged casualties, is by design a continuous variable. The panel nature of the data is easily incorporated by treating groups of observations under each time unit as a separate cross section subset and by conducting matching within each of these subsets[11]. The continuous treatment variable is a more challenging obstacle. Matching methods have rarely been applied to a continuous treatment. Researchers have proposed techniques to generalize the binary case to continuous treatments (Hirano & Imbens, 2004; Bia & Mattei, 2007; Fong, Hazlett & Imai, 2018); nevertheless, these

[11] Otherwise, for each location, observations over time get matched to each other.

techniques come with some untrivial assumptions (Fong, Hazlett & Imai, 2018) which are not suitable for my purposes.

Instead, I adopt the more common approach in applied social sciences and binarize my treatment. My binary treatment variable takes on the value 1 if location i at time t has a high inverse-road-distance-weighted average of conflict event incidences in the neighbourhood at time t-1 and zero otherwise. Formally, I define my treatment variable as

$\text{Treatment}_{i,t} = 1 \quad \text{if } \Sigma_{k\neq i}\ (1/d_{ki})(\text{incidence})_{k,t-1} > 90^{th}$ percentile

$\text{Treatment}_{i,t} = 0 \quad$ otherwise

where d_{ki} is the road distance between locations k and i in kilometers.

The choice of the cutoff value is based on the distribution of the treatment variable and is chosen to distinguish between observations with high and low spatiotemporal incidence lags. To make sure that my results do not depend on my choice of cutoff value, I repeated the exercise with different cutoff percentiles such as the 75^{th} and the 95^{th}, and obtained substantively similar results[12].

I then use the CEM to match my observations. I then run a probit regression with clustered errors at the county level on my matched sample. Table II below presents the results[13]. As can be seen, the estimated coefficients still associate events with state force casualties with higher risks of future events in the vicinity, while events with insurgency casualties are associated with lower risks of mayhem in the neighborhood in the longer runs.

TABLE II. Results of the probit regression on matched sample

	Unit of observation: **County-month** **(1984-2018)** Number of obs.: 94206	Unit of observation: **County-quarter** **(1984-2018)** Number of obs.: 33029	Unit of observation: **County-halfyear** **(1984-2018)** Number of obs.: 15739	Unit of observation: **County-year** **(1984-2018)** Number of obs.: 11528
Length of last peace spell	-0.006** (-4.52)	-0.015** (-5.96)	-0.063** (-4.82)	-0.076** (-4.22)
Percentage of past time periods with conflict events	0.041** (8.81)	0.034** (12.13)	0.020** (15.64)	0.019** (18.12)
Inverse-distance-weighted **state casualties** in **other counties** in t-1	1.387** (4.99)	0.648** (3.97)	0.890** (9.33)	0.671** (8.90)
Inverse-distance-weighted **insurgent casualties** in **other counties** in t-1	0.182 (1.51)	0.064 (1.14)	-0.107** (3.86)	-0.103** (4.16)
State casualties in the county in t-1	0.096** (3.04)	0.011 (0.98)	0.041** (3.57)	0.035** (3.16)
Insurgent casualties in the county in t-1	0.017 (1.40)	0.015* (2.20)	0.012** (3.20)	0.022** (4.19)
County area in km squares	0.0001** (3.24)	0.0002** (5.85)	0.0001** (4.71)	0.0001** (3.67)
Percentage of mountainous terrain in rural area	0.001 (0.79)	0.002† (1.75)	0.0003 (0.36)	0.001† (1.73)
Border status	-0.086	-0.091	0.040	0.096†

[12] Results available upon request.

[13] The results are very similar when socioeconomic controls are included (and they are available upon request) but please note that in this subsample they are already controlled for by matching.

	(-0.94)	(-0.91)	(0.68)	(1.74)
Estimated parameters for seasonal and year dummies are not reported. Robust errors clustered at the county level. z-values in parenthesis. † p<0.1, * p<0.05, ** p<0.01.				

Conclusion

Kalyvas (2008) applauds the emergence of the literature on micro-dynamics of civil conflicts as a very exciting development that deepens our understanding of political violence. But he also points out some recurrent flaws in the literature stemming from 'insufficient theorization, superficial engagement with the case at hand and reliance on off-the-shelf datasets' (p.398). My starting point in this study echoes Kalyvas' criticisms specifically for the emerging literature on within-country diffusion of civil conflicts.

I argue that existing works fail to acknowledge that conflict events can have incapacitating effects on warring sides. I tackle this shortcoming by hypothesizing that the geotemporal interdependency among conflict events is conditioned by the impact these events have on the operational capability of the sides of the conflict. I then test my hypotheses on a new and detailed event dataset on the long running civil conflict in Turkey. TPCONED has been in the making for more than a decade. It relies on a wide range of local sources from different political affiliations and views, and as such offers comprehensive, accurate, high resolution and unbiased coverage of the long running civil conflict between the Turkish state and the rebel organization PKK. I release TPCONED with the hope that it will be a valuable resource especially for in-depth studies on micro-level dynamics of civil conflicts.

One important point to note here is that the hypotheses tested in this study are about the diffusion patterns of irregular civil conflicts, and not about their outcomes or their level of violence. While the first hypothesis does imply that an irregular civil conflict is likely to continue and geographically expand as long as rebels are able to inflict damages on state forces, this implication does not allow us to conclude on a rebel victory in such conflicts. In fact, we know from cross-country studies that irregular civil conflicts do last long, but, because of the power asymmetry that characterizes them, are mostly won by incumbents (Balcells & Kalyvas, 2014; Kalyvas & Balcells, 2010).

Similarly, while the second hypothesis does imply that states can curb the geographical spread of conflict events by inflicting damages on the insurgents, this implication does not mean that states can resolve conflicts by going on a killing rampage. Military coercion, as the theoretical arguments and the empirical results here indicate, can be effective in containing or even extinguishing violence in the short run, but it must be remembered that civil conflicts are population-centric contests with social, political and economic dimensions. Insurgent organizations will have the ability to generate political support and continue their violent campaign as long as states fail in their counterinsurgency efforts to address those social, economic and/or political problems and to secure the loyalties of civilians. The Turkish case itself is a good example. Turkish security forces dealt a major blow to the PKK through their use of extensive military coercion in the 90s and even captured its leader Abdullah Öcalan in 1999. The PKK then declared a unilateral ceasefire. Many commentators heralded this as the PKK's defeat and as the end of the conflict (Bila, 2000; Cemal, 1999). However, because the Turkish governments failed to take the necessary steps to address the political, economic and social problems underlying the conflict, the organization easily recovered in the early 2000s and violence resumed.

Finally, results must be tempered by the study's limitations. First, it must be admitted that, rather than absolute numbers, the ratio of casualties to group size in each location at each time period provides a better measure of the extent of the damage on the operational capabilities of the sides. Not surprisingly though, for strategic purposes, no army or insurgent organization reveals such information, hence working with ratios remains as an ideal.

Second, it must be emphasized that this is a single-case study. While the media coverage of ongoing conflicts around the world offers ample anecdotal evidence of the generalizability of results[14], a comparative study is yet to be conducted.

Replication data

The dataset, codebook, and do-files for the empirical analysis in this article, along with the online appendix, can be found at http://www.prio.org/jpr/datasets. All statistical analyses were conducted using Stata15. TPCONED is available at the University of Warwick Research Archive Portal https://wrap.warwick.ac.uk/138227.

Acknowledgments

[14] A BBC article, for example, reports Afghan security forces killing 5 Taliban insurgents near Kabul and seizing their vehicles which were packed with explosives for planned attacks in the city (https://www.bbc.com/news/world-asia-19090416). On the other hand, a recent CNBC article on the Rohingya insurgency in Myanmar mentions how attacks by insurgents in the northwestern parts of the country prompted a big military sweep by security forces and marked a major escalation of the conflict (https://www.cnbc.com/2017/08/25/at-least-32-killed-in-myanmar-as-rohingya-insurgents-stage-major-attack.html).

I thank the anonymous reviewers, and the Editor of JPR, Alex Braithwaite, and Vincenzo Bove and Kristian Skrede Gleditsch for helpful feedback on previous versions of the article.

References

Ayres, R William & Stephen Saideman (2000) Is Separatism as Contagious as the Common Cold or as Cancer? Testing International and Domestic Explanations. *International Interactions* 6(3): 91–113.

Balcells, Laia & Stathis N Kalyvas (2014) Does Warfare Matter? Severity, Duration, and Outcomes of Civil Wars. *Journal of Conflict Resolution* 58 (8):1390-1418.

Bara, Corinne (2018) Legacies of Violence: Conflict-specific Capital and the Postconflict Diffusion of Civil War. *Journal of Conflict Resolution* 62(9): 1991-2016.

Beardsley, Kyle; Kristian Skrede Gleditsch & Nigel Lo (2015) Roving Bandits? The Geographical Evolution of African Armed Conflicts. *International Studies Quarterly*. 59:503-516.

Beiser, Janine (2016) International information flows, government response and the contagion of ethnic conflict. In: Alan Wilson (ed.) *Global Dynamics: Approaches from Complexity Science*. Chichester: John Wiley & Sons, 214-230.

Bia, Michela & Alessandra Mattei (2007) Application of the generalized propensity score. Evaluation of public contributions to Piedmont enterprises. *POLIS Working Paper* 80. University of Eastern Piedmont.

Bila, Fikret (2000) Bu nasıl diplomasi? (What kind of a diplomacy is this?). *Cumhuriyet*. 23 February.

Blackwell, Matthew; Stefano Iacus, Gary King & Guiseppe Porro (2009) CEM: Coarsened Exact Matching in Stata. *The Stata Journal* 9(4): 524-546.

Bohnet, Heidrun; Fabien Cottier & Simon Hug (2018) Conflict-induced IDPs and the Spread of Conflict. Journal of Conflict Resolution 62(4): 691-716.

Braithwaite, Alex (2010) Resisting Infection: How State Capacity Conditions Conflict Contagion. *Journal of Peace Research* 47(3): 311-319.

Braithwaite, Alex & Tiffany S Chu (2018) Civil Conflicts Abroad, Foreign Fighters, and Terrorism at Home. *Journal of Conflict Resolution* 62(8): 1636-1660.

Braithwaite, Alex & Shane D Johnson (2012) Space–Time Modeling of Insurgency and Counterinsurgency in Iraq. *Journal of Quantitative Criminology* 28:31-48.

Braumoeller, Bear & Gary Goertz (2002) Watching Your Posterior: Comment on Seawright. *Political Analysis* 10 (2): 198–203.

Buhaug, Halvard & Kristian Skrede Gleditsch (2008) Contagion or Confusion? Why Conflicts Cluster in Space. *International Studies Quarterly* 52 (2): 215-33.

Carter, Timothy Allen & Daniel Jay Veale (2015) The Timing of Conflict Violence: Hydraulic Behavior in the Ugandan Civil War. *Conflict Management and Peace Science* 32(4): 370-394.

Cederman, Lars-Erik; Luc Girardin & Kristian Skrede Gleditsch (2009) Ethnonationalist Triads: Assessing the Influence of Kin Groups on Civil Wars. *World Politics* 61 (3): 403-37.

Cederman, Lars-Erik; Andreas Wimmer & Brian Min (2010) 'Why Do Ethnic Groups Rebel? New Data and Analysis. *World Politics* 62 (1): 87-119.

Cederman, Lars-Erik; Kristian Skrede Gleditsch, Idean Salehyan & Julian Wucherpfennig (2013) Transborder Ethnic Kin and Civil War. *International Organization* 67 (2): 389-410.

Cemal, Hasan (1999) Yol ayrımı (Crossroads). *Milliyet.* 4 March.

Clark, David & Patrick Regan (2003) Opportunities to Fight: A Statistical Technique for Modeling Unobservable Phenomena. *Journal of Conflict Resolution* 47 (1): 94–115.

Do, Quy-Toan & Lakshmi Iyer (2010) Geography, Poverty and Conflict in Nepal. *Journal of Peace Research* 47 (6): 735-48.

Donnay, Karsten; Eric T Dunford, Erin C McGrath, David Backer & David E Cunningham (2018) Integrating Conflict Event Data. *Journal of Conflict Resolution* 63(5): 1137-1364.

Duffy Toft, Monica & Yuri M Zhukov (2012) Denial and punishment in the North Caucasus: Evaluating the effectiveness of coercive counter-insurgency. *Journal of Conflict Resolution* 49(6): 785-800.

Fearon, James D & David D Laitin (2003) Ethnicity, Insurgency and Civil War. *American Political Science Review* 97(1): 75-90.

Fong, Christian; Chad Hazlett & Kosuke Imai (2018) Covariate Balancing Propensity Score for a Continuous Treatment: Application to the Efficacy of Political Advertisement. *The Annals of Applied Statistics* 12(1): 156-177.

Forsberg, Erika (2008) 'Polarization and Ethnic Conflict in a Widened Strategic Setting. *Journal of Peace Research* 45 (2): 283-300.

Gleditsch, Kristian Skrede (2017) Ornithology and Varieties of Conflict: A Personal Retrospective on Conflict Forecasting. *Peace Economics, Peace Science and Public Policy* 23(4).

Gleditsch, Kristian Skrede; Idean Salehyan & Kenneth Schultz (2008) Fighting at Home, Fighting Abroad. *Journal of Conflict Resolution* 52 (4): 479-506.

Global Terrorism Database (2020) National Consortium for the Study of Terrorism and Response to Terrorism. https://www.start.umd.edu/data-tools/global-terrorism-database-gtd.

Grossman, Herschell (1991) A General Equilibrium Model of Insurrections. *American Economic Review* 81 (4): 912-21.

Haushofer, Johannes; Anat Biletzki & Nancy Kanwisher (2010) Both Sides Retaliate in the Israeli-Palestinian Conflict. *Proceedings of the National Academy of Sciences* 107(42): 17927-32.

Hegre, Havard; Gudrun Østby & Clionadh Raleigh (2009) Poverty and Civil War Events: A Disaggregated Study of Liberia. *Journal of Conflict Resolution* 53(4): 598-623.

Hill, Stuart; Donald S Rothchild & Colin Cameron (1998) Tactical Information and the Diffusion of Peaceful Protests. In: David A. Lake and Donald S. Rothchild (ed.) *The International Spread of Ethnic Conflict: Fear, Diffusion, and Escalation*. Princeton, NJ: Princeton University Press, 61–88.

Hirano, Keisuke & Guido W Imbens (2004) The propensity score with continuous treatments. In: A. Gelman and X.-L. Meng (ed.) *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*. West Sussex, England: Wiley InterScience, 73–84.

Holtermann, Helge (2012) Explaining the Development-Civil War Relationship. *Conflict Management and Peace Science* 29(1): 56-78.

Holtermann, Helge (2016) Relative Capacity and the Spread of Rebellion: Insights from Nepal. *Journal of Conflict Resolution* 60(3):501-529.

Jaeger, David A & M Daniele Paserman (2008) The Cycle of Violence? An Empirical Analysis of Fatalities in the Palestinian-Israeli Conflict. *American Economic Review* 98(4): 1591-1604.

Kalyvas, Stathis (2006) *The Logic of Violence in Civil War*. Cambridge, UK: Cambridge University Press.

Kalyvas, Stathis (2008) Promises and Pitfalls of an Emerging Research Program: The Microdynamics of Civil War. In: S. Kalyvas, I. Shapiro, & T. Masoud (ed.) *Order, Conflict and Violence*. Cambridge: Cambridge University Press 2008, 397-421.

Kalyvas, Stathis & Laia Balcells (2010) International System and Technologies of Rebellion: How the End of the Cold War Shaped Internal Conflict. *American Political Science Review* 104 (3): 415-429.

Kibris, Arzu (2011) Funerals and Elections: The Effects of PKK Terrorism on Voting Behavior in Turkey. *Journal of Conflict Resolution* 55(2): 220-247.

Kocher, Matthew Adam; Thomas B Pepinsky and Stathis Kalyvas (2011) Aerial Bombings and Counterinsurgency in the Vietnam War. *American Journal of Political Science* 55(2): 201-218.

Lake, David A & Donald Rothchild (1998) Spreading fear: The genesis of transnational ethnic conflict. In: David A Lake & Donald Rothchild (ed.) *The International Spread of Ethnic Conflict*. Princeton, NJ: Princeton University Press, 3–34.

Lane, Matthew (2016) The Intrastate Contagion of Ethnic Civil War. *Journal of Politics* 78(2): 396-410.

Linke, Andrew M; Frank DW Witmer & John O'Loughlin (2012) Space-Time Granger Analysis of the War in Iraq: A Study of Coalition and Insurgent Action-Reaction. *International Interactions* 38:4, 402-425.

Lyall, Jason M (2009) Does Indiscriminate Violence Incite Insurgent Attacks? Evidence from Chechnya. *Journal of Conflict Resolution* 53(3): 31-362.

Machiavelli, Nicolo. *The Prince*. [Online].
Available at: https://www.gutenberg.org/files/1232/1232-h/1232-h.htm [Last accessed 11 August 2020].

Maves, Jessica & Alex Braithwaite (2013) Autocratic Institutions and Civil Conflict Contagion. *The Journal of Politics* 75: 478-90.

Mahoney, James & Gary Goertz (2004) The Possibility Principle: Choosing Negative Cases in Comparative Research. *American Political Science Review* 98 (4): 653–69.

McColl, Robert W (1969) The insurgent state: territorial bases of revolution. *Annals of the Association of American Geographers* 59(4): 613-631.

Metternich, Nils W; Shahryar Minhas & Michael D Ward (2017) Firewall? or Wall on Fire? A Unified Framework of Conflict Contagion and the Role of Ethnic Exclusion. *Journal of Conflict Resolution* 61(6): 1151-1173.

Mikulaschek, Christoph & Jacob N Shapiro (2018) Lessons on Political Violence from America's Post-9/11 Wars. *Journal of Conflict Resolution* 62(1): 174-202.

O'Loughlin, John & Frank DW Witmer (2012) Diffusion of Violence in the North Caucasus of Russia, 1999-2010. *Environment and Planning A: Economy and Space* 44(10).

Pegg, Scott M (1998) *International Society and the De Facto State*. Aldershot: Ashgate.

Pulur, Hasan (2010) Özal 26 yıl önce 'üç beş çapulcu' demişti. (Özal said "a handful of outlaws" 26 years ago) *Milliyet*. 5 August.

Raleigh, Clionadh & Havard Hegre (2009) Population size, concentration, and civil war. A geographically disaggregated analysis. *Political Geography* 28: 224-238.

Raleigh, Clionadh; Andrew Linke, Havard Hegre & Joakim Karlsen (2010) Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research* 47(5), 651:660.

Rüegger, Seraina (2018) Refugees, ethnic power relations, and civil conflict in the country of asylum. *Journal of Peace Research* 56 (1): 42-57.

Salehyan, Idean & Kristian Skrede Gleditsch (2007) Refugees and the Spread of Civil War. *International Organization* 60 (2): 335-66.

Schutte, Sebastian (2017) Regions at Risk: Predicting Conflict Zones in African Insurgencies. *Political Science Research and Methods* 5(3): 447-465.

Schutte, Sebastian & Karsten Donnay (2014) Matched wake analysis: Finding causal relationships in spatiotemporal event data. *Political Geography* 41:1-10.

Schutte Sebastian & Nils B Weidmann (2011) Diffusion Patterns of Violence in Civil Wars. *Political Geography* 30:143-152.

Schultz, Kenneth A (2010) The Enforcement Problem in Coercive Bargaining: Interstate Conflict over Rebel Support in Civil Wars. *International Organization* 64 (2): 281-312.

Sundberg, Ralph, & Erik Melander (2013) Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research* 50(4): 523-532.

Tellis, Ashley J; Janice Bially, Christopher Layne, Melissa McPherson & Jerry M Sollinger (2000) *Measuring National Power in the Post-Industrial Age.* Rand Corporation.

Tezcür, Güneş Murat (2016) Ordinary People, Extraordinary Risks: Joining an Ethnic Rebellion. *American Political Science Review* 110(2).

Townsley, Michael; Shane D Johnson & Jerry H Ratcliffe (2008) Space Time Dynamics of Insurgent Activity in Iraq. *Security Journal* 21:139-146.

Tull, Denis (2004) *The Reconfiguration of Political Order in Postcolonial Africa: A Case Study from North Kivu (DR Congo)*. Ph.D. Dissertation. University of Hamburg.

Turkish Ministry of Agriculture and Rural Affairs (1981) *Köy Envanter Etüdü* (Village Inventory). Ankara. All issues.

Xiang, Jun (2010) Relevance as a Latent Variable in Dyadic Analysis of Conflict. *The Journal of Politics* 72(2): 484-498.

Ward, Michael D & Kristian Skrede Gleditsch (2010) Location, Location, Location: An MCMC Approach to Modelling the Spatial Context of War and Peace. *Political Analysis* 10(3): 244-260.

Weidmann, Nils B (2015) Communication Networks and the Transnational Spread of Ethnic Conflict". *Journal of Peace Research* 52(3): 285-296.

Weidmann, Nils B & Michael D Ward (2010) Predicting Conflict in Space and Time". *Journal of Conflict Resolution* 54(6):883-901.

Weinstein, Jeremy M (2007) *Inside Rebellion: The Politics of Insurgent Violence*. Cambridge: Cambridge University Press.

Zhukov, Yuri M (2012) Roads and the Diffusion of Insurgent Violence: The Logistics of Conflict in Russia's North Caucasus. *Political Geography* 31 (3): 144-56.

ARZU KIBRIS, PhD. In Political Science (Sabanci University, 2010); Associate Professor, University of Warwick (2017 -); Principal Investigator, EXPOVIBE: Exposure to Political Violence and Individual Behavior, ERC-Starting Grant; Alan Turing Fellow (2018 -); Executive Coordinator, Network of European Peace Scientists.

The geo-temporal evolution of violence in civil conflicts

Online appendix

A1. The data

TPCONED (https://wrap.warwick.ac.uk/138227) is a dataset on the fatal events of the armed conflict between the rebel organization PKK and the Turkish state that took place on Turkish soil and in which there was at least one combatant casualty. It contains detailed information on 7063 conflict events with 17308 PKK casualties and 7514 state casualties over the course of the conflict in the 1984-2018 period[15]. For each event, the exact date, location at the county (town) level, location at the province level, number of TSF and PKK casualties, and the data sources are listed. Table A1 presents some basic descriptive statistics for the numeric variables.

Table A1. Descriptive statistics

Temporal coverage: 1984 - 2018

Geographical coverage: Turkey

Number of observations: 7063

Variable	Mean	Standard deviation	Minimum	Maximum
TSF casualties	1.06	2.06	0	66
PKK casualties	2.45	6.34	0	174

TPCONED relies on a wide range of local sources of information as listed below. The main source for each observation is listed under the variable "*mainsource*" in the dataset, and the complete reference for each source is available on the *References* page in the TPCONED excel file (https://wrap.warwick.ac.uk/138227).

1- The digital and hard copy archives of five major Turkish daily newspapers, Cumhuriyet, Milliyet, Hürriyet, Sabah, and Zaman. The digital archives of Milliyet, Hürriyet, Sabah and Zaman were available online on the websites of these newspapers. The online archive of Cumhuriyet was accessed through Sabancı University library. The hard copy archives were accessed and studied at the Ataturk Library in Istanbul.

2- A 5-volume publication, named "Our Martyrs" (Şehitlerimiz), published by the Turkish Ministry of Defense in late 1998. The publication lists the names, ranks, and place of death of all Turkish soldiers that died in combat since 1918. The fifth volume contains the list of soldiers who died in the fight against the PKK in the August 1984-September

[15] The dataset is continuously updated. Currently, events of 2020 are being coded.

1998 period. In total, the list contains information on 5554 soldiers. This publication actually provides the most credible, and accurate information on Turkish military casualties of the period as the information is directly from the personnel records of the Turkish General Staff.

3- Publications by local administrations, especially district governorships, commemorating their local security force casualties. Most of these "martyr albums" were published in 1998 on the occasion of the 75th anniversary of the Turkish Republic. They include information on local security force casualties and on the conflict events that claimed their lives. The list of these publications is available on the References page in the TPCONED excel file (https://wrap.warwick.ac.uk/138227).
4- Monthly reports of the Directorate General of Press and Information of Turkish Prime Ministry. These reports (*Ayın Tarihi*) were available online however, the Directorate (and hence its website) was abolished with the switch to a presidential regime in Turkey in 2017.
5- Issues of Serxwebun, a PKK publication (Serxwebun.org).
6- Daily and yearly (1990-2016) *Human Rights Reports* by the Turkish Human Rights Association (tihv.org.tr).
7- Web searches using the keywords *terror, martyr, PKK, armed attack, insurgent, insurgency*. Some 359 websites were referred. The list is available on the References page in the TPCONED excel file (https://wrap.warwick.ac.uk/138227). Most of these are websites of local administrations like district governorships, district and town municipalities, and village associations with pages commemorating and providing information on their local security force casualties. There are also websites by civil society organizations that are completely devoted to commemorate casualties.
8- I have also contacted and worked with the Associations of Families of Martyrs (*Şehit Aileleri Dernekleri)* across Turkey. These associations are very common in Turkey. There is one in almost every city. They are civil society organizations founded by families of soldiers and police officers who died on duty. The aim is to help each other cope with the situation. Some of them are very active, and have their own websites containing information on those family members they had lost to the conflict. I contacted 28 of these associations and worked with them in verifying the conflict events and casualties.

9- The Kurdish Insurgency Militants database (Tezcür, 2016). This database includes biographical information on 9,196 Kurdish militants who died while fighting in the ranks of the PKK between 1984 and 2016.
(available at https://sciences.ucf.edu/politics/kps/kim-dataset/).

10- The personnel archives of the General Directorate of Security (Turkish police forces). The public relations office of the Directorate was kind enough to provide the list of their casualties.

11- I visited the "martyr cemeteries" in Ankara, Istanbul, and Bursa, three major cities in Turkey with more than a thousand security force casualties resting in their cemeteries. Note that the tombstones contain information on the date and place of death.

12- The Gendarmerie Museum in Ankara proved to be an invaluable source of information with its extensive database on gendarmerie casualties. The museum also hosts an impressive monument at its entrance in the form of a great wall on which names, places and dates of deaths of gendarmerie "*martyrs*" were carved.

A2. Comparison of TPCONED with the UCDP-GED data

Donnay et al. (2018) proposes using temporal and spatial windows to assess the overlap between data sets. In this section I adopt their proposed data set comparison methodology and systematically compare TPCONED with the UCDP-GED data on the Turkish-Kurdish conflict for the 1984-2018 period. I used a 10-day temporal window and considered adjacent (border sharing) counties as my spatial window. Table A2 summarizes how TPCONED compares to the UCDP-GED dataset.

Table A2. Comparison with the UCDP-GED data	TPCONED	UCDP-GED dataset
Number of observations	7063	4682
Time span	1984-2018	1989-2018
Spacial unit precision	County level for all observations	Only 20.6% of observations have exact location (a known point like a village, county or city); 17% have a location within a 25 km radius around a known point; 17.5 % have a reference to a second order administrative division; 40% of observations have only a provincial reference; remaining 5.3% have spatial references to large and fuzzy regions or to the whole country.
Time unit precision	Exact date (dd/mm/yy) for all observations.	88% of observations have exact date (dd/mm/yy). For the remaining 12%, the date is only known within a number of days ranging from 2 to 180.
Overlapping observations within a 10-day temporal, and neighboring (border sharing) county spatial window	3480	
Unmatched observations	3583	1202

Nonmatches with explanation	287 (the 1984-1988 period which is not covered by the GED)	1202 (explanations below)

Explanations for the 1202 unmatched GED observations:

a. 685 of the 1202 unmatched GED observations (57%) have very broad and imprecise (4 or above in GED time and place precision coding) spatial and/or temporal references. It is thus impossible to match them with reasonable accuracy with my data.

b. 92 of the remaining 517 unmatched GED observations do not have any province or county listed even though their location precision coding indicates relatively precise location (4 or less in GED coding).

c. 42 of the remaining 425 unmatched GED observations refer to a location outside Turkey (Syria, Iraq or Iran).

d. 87 of the remaining 383 unmatched GED observations refer to events with no combatant casualties. TPCONED only includes events with combatant casualties.

e. 4 of the 296 unmatched GED observations refer to clashes between the PKK and paramilitaries. My data set only includes encounters with the state forces and the PKK. I deliberately left out encounters between paramilitaries and the PKK from my data set because paramilitaries (they are called village guards in Turkey) are in fact civilians and their identities are of course not public. As such it is very difficult if not impossible to distinguish between civilian casualties and paramilitary casualties. And not surprisingly, it has not been uncommon for both the PKK and the Turkish state to proclaim civilian killings as paramilitary deaths. So, there is significant uncertainty surrounding such events.

f. 12 of the 292 unmatched GED observations refer to terrorist attacks by other armed organizations like DEV-SOL or DHKPC. So, they are <u>not</u> part of the armed conflict between the Turkish state and the PKK.

g. 12 of the 280 unmatched GED observations are double entries, i.e. they are either duplicates or parts of other matched GED observations.

h. 4 of the 268 unmatched GED observations refer to killings of public servants (like prosecutors or teachers who in fact are civilians) by the PKK.

i. 5 of the 264 unmatched GED observations refer to a location (village) that does not exist.

j. 18 of the remaining 259 unmatched GED observations refer to a mountain or a river as location even though their location precision coding is less than 4. Thus, it is not possible to decipher the county information for these observations.
k. Of the remaining 241 unmatched GED observations 117 have no "original source" listed. 105 observations have only the Turkish authorities (Turkish state news agency, security forces, general staff, governors, etc.) listed as their original source, 7 of them list Turkish media sources as their original source of information. Only 12 observations list Kurdish media or PKK sources as their original source. Not surprisingly, while most of these events involve PKK casualties (as high as 76 casualties in some of them), Turkish security forces are reported to incur casualties in only 33 of them which indicates that there might be a source bias concerning these observations. Moreover, it is not possible to confirm any of those 33 events via the identity of TSF casualties because no such casualties at those dates and locations are listed or mentioned in any of the official and media resources that TPCONED refers to.

A3. Model identification

The biprobit model employed in this study is mathematically equivalent to a bivariate probit model with partial observability. This implies that identification requires a restriction that affects only one of the equations (Xiang, 2010; Poirier, 1980). In his model of interstate conflict Xiang (2010) argues that the control variables that are uncommon in the two equations serve as identification restrictions (p. 491). I have a similar but a much stronger argument for identification in my model. Note that the base risk equation (Z vector of covariates) includes a rich set of controls which depict the socioeconomic and political situation <u>at the start of the armed conflict</u>, and these (cross sectional) controls are unique to the base risk equation. These variables are all measurements taken at early 1980s, and thus, are geared towards taking the socioeconomic picture of the country (at the county level) right around the start of the armed conflict in order to understand the conditions that render certain areas ripe for the conflict. In contrast, the diffusion equation is mainly composed of within conflict dynamics, aimed at understanding those factors that facilitate geotemporal spillovers, and accordingly, the measurements for these panel variables all belong to the period after the start of the armed conflict. And similarly, these controls are unique to the diffusion equation. In total, the model includes some 55 control variables and only 2 of them (percentage of mountainous terrain in rural area, and border status) are common to both equations. The remaining 53 control variables, with 11 of them uniquely in the base risk equation and cross-sectionally measured right around the start of the conflict in 1984, and 42 of them uniquely in the diffusion equation with panel measurements belonging to the period after the start of the conflict, create the necessary restriction for identification.

A4. Socioeconomic controls and the base risk of conflict

The vector of covariates for base risk in my split population bi-probit model includes a rich set of controls which depict the socioeconomic situation at the start of the armed conflict. The data comes from a "village inventory" study conducted in 1981 by the Turkish Ministry of Agriculture and Rural Affairs, and the general census of 1985 which coincides roughly with the start of the armed conflict in August 1984. I control for percentage of mountainous rural area; percentage of farmers with no land; percentage of farmers with more than 100 acres of land; number of mosques per village; percentage of villages with no drinking water; percentage of villages with no electricity; and percentage of villages whose names were changed by the Turkish state; the unemployment and literacy rate, border status, Kurdish population percentages, and urbanization and population across counties.

Land ownership is both an indicator of wealth and an indicator of political participation in rural areas. The Turkish state has a long tradition of using big land owners as its agents in the periphery. Since the foundation of the Republic, large landowners have been co-opted in politics by political parties (Tachau, 1973). The political dominance of *agas* (large landowners) within the Parliament and local party structures have been frequently noted by scholars of Turkish politics (Kudat, 1975; Meeker, 1972). Hence, I expect the state to establish better control and authority over rural areas where big land owners reside and thereby to reduce any political and military opportunity for organizing an insurgency (Holtermann, 2012).

On the other hand, landless peasants can also be an important determinant of the underlying risk distribution as their presence is associated with economic grievances which create a more fertile recruitment ground for the insurgency (Holtermann, 2012).

I also control for the number of mosques per village as mosques in Turkey have always been instrumental in establishing state control and disseminating state ideology over populations. In the context of the Turkish-Kurdish conflict, religious institutions have been used by the state to promote an encompassing Muslim identity over ethnic ones. Note that all imams in Turkey are public employees appointed from the center by the state. Not only that, their sermons are also prepared at and sent from the Directorate of Religious Affairs.

The percentage of villages with no water and electricity are included to control for the infrastructure and services provided by the state.

Another control I include is the percentage of villages whose names were changed by the state. The Turkish state had replaced the official names of a great deal of administrative units which originally had names in the native languages of local inhabitants from different ethnic groups, with new Turkish names as part of a *Turkification* strategy. I use the percentage of such villages across geographical units as a proxy for the intensity of ethnic discriminatory state policies. Ethnic discrimination is identified as an important factor that increases the vulnerability of a country to experience conflict (Metternich, Minhas & Ward, 2017; Cederman et al., 2013; Cederman, Wimmer & Min, 2010). I make a similar argument for within country distribution of conflict risk.

Apart from the controls I derive from the 1981 village inventory study, I also control for the unemployment and literacy rate, border status, Kurdish population percentages, and urbanization and population across locations. The data for these controls comes from the 1985 census except for the Kurdish population percentages which is an estimate I calculated based on prior scholarly work on the ethnic distribution of population in Turkey (Mutlu, 1996).

Unemployment and literacy rates are indicators of economic development and welfare. I expect unemployment rate to be positively associated with the base risk of conflict since it is an indicator of economic grievances and the opportunity costs of joining a rebellion. Literacy rates can cut both ways. On one hand, illiterate populations are expected to be more religiously conservative, less politically aware, and less sensitive to social issues. On the other hand, one can expect opportunity costs of joining a rebellion to be lower for less educated people.

I expect the distribution of ethnically Kurdish populations to be one of the most important determinants of base risk of conflict since the PKK is an ethnic insurgency founded with the goal of establishing a Kurdish state. Naturally, those areas inhabited by ethnically Kurdish populations are expected to make up the targeted "ethnic homeland". Relatedly, ethnically Kurdish populations constitute the main support base for the insurgency.

Border locations are also more likely to be targeted by insurgencies especially in cases where ethnic kins live on both sides of the border (Cederman, Girardin & Gleditsch, 2009;

Cederman et al., 2013), or neighbours are sympathetic to the insurgents for various reasons (Salehyan & Gleditsch, 2007). In the Turkish case, neighbouring countries like Syria and Iraq host large Kurdish populations. Moreover, these countries have weak states with limited ability to control their territories and borders.

I expect urban areas to carry less risk since insurgencies mostly target rural areas with limited state reach. I also control for population as more populous locations are expected to have a higher risk of conflict (Weidman & Ward, 2010).

Finally, I control for the percentage of mountainous rural areas. There exists ample empirical evidence that difficult terrains provide fertile grounds for civil conflicts (Fearon & Laitin, 2003; Do & Iyer, 2010; Weidman & Ward, 2010; Zhukov, 2012; Holtermann, 2016). Table A3 presents the descriptive statistics for the socioeconomic controls included in the analyses.

TABLE A3. Descriptive statistics of socioeconomic controls	Mean	Standard deviation	Range
County area in km squares	1025.9	704.1	[11-6585]
Percentage of mountainous terrain in rural area	50.4	27.2	[0-100]
Border status	0.07	0.26	[0-1]
Percentage of villages whose names were changed by the state	32.7	30.7	[0-100]
Percentage of landless farmers	27.4	14.4	[0-76]
Percentage of farmers with more than 100 acres of land	10.1	12.7	[0-77]
Number of mosques per village	0.96	0.44	[0-4.3]
Percentage of villages with no drinking water	12.9	13.8	[0-87]
Percentage of villages with no electricity	46.3	36.1	[0-100]
Percentage of ethnically Kurdish population	16.2	24.5	[0.1- 85]
Urbanization rate	32.7	19.4	[6-100]
Literacy rate	71.8	12.5	[29-93]
Unemployment rate	3.5	2.6	[0.2-20]
Population in 10 thousand	6.9	9.6	[0.4-100]

Figure A1 plots the base risk across counties predicted by these variables along with the actual distribution of total combatant casualties. As the two maps reveal, the model accurately points to the southeastern part of the country as the conflict zone.

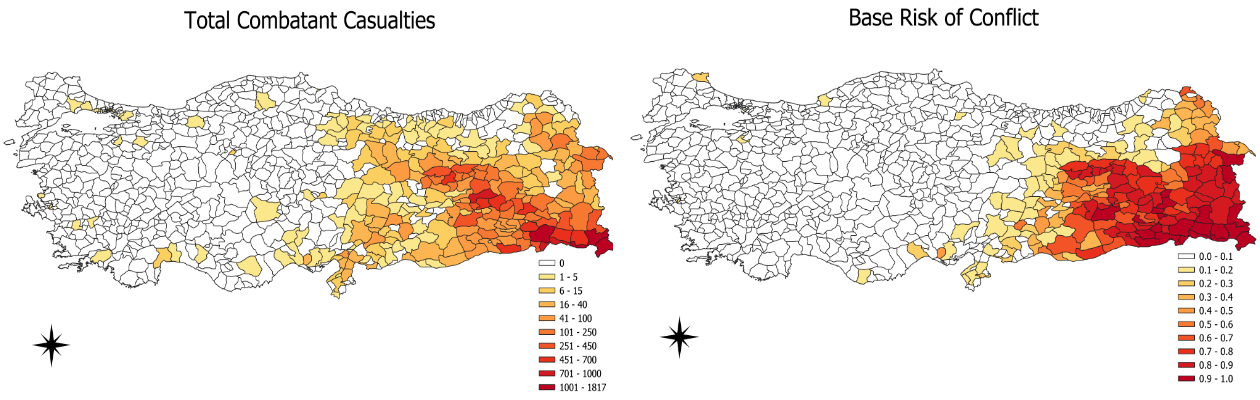

Figure A1. Comparison of the estimated base risk with the actual distribution of combatant casualties

Table A4 reports the estimated average marginal effects from the main split population biprobit model in order to give a better understanding of the relative substantive significance of the estimates parameters.

TABLE A4. Marginal effects in percentages	**County-month** dy/dx	**County-quarter year** dy/dx	**County-half year** dy/dx	**County-year** dy/dx
Length of last peace spell	-0.0001** (-10.97)	-0.0005** (-11.65)	-0.003** (-5.65)	-0.003* (-2.19)
Percentage of past time periods with conflict events	0.001** (9.88)	0.001** (16.73)	0.001** (10.96)	0.001** (8.62)
Inverse-distance-weighted **state casualties** in **other counties** in t-1	0.030** (7.06)	0.025** (6.20)	0.071** (6.08)	0.069** (8.41)
Inverse-distance-weighted **insurgent casualties** in **other counties** in t-1	0.002† (1.90)	-0.003* (-2.09)	-0.007** (-2.96)	-0.013** (-5.24)
State casualties in the county in t-1	0.002** (6.65)	0.002** (5.26)	0.003* (2.18)	0.013** (2.82)
Insurgent casualties in the county in t-1	0.0005* (2.53)	0.0005** (2.99)	0.001** (3.10)	0.006** (2.69)
County area in km squares	0.000002** (4.00)	0.000003** (3.52)	0.000004* (2.51)	0.000006** (2.65)
Percentage of mountainous terrain in rural area	0.0001** (2.86)	0.0001** (3.21)	0.0002** (2.66)	0.0003** (2.82)
Border status	0.003** (2.65)	0.005* (2.49)	0.012** (2.97)	0.019** (3.45)
Percentage of villages whose names were changed by the state	0.00003* (2.24)	0.00004* (2.34)	0.0002* (2.32)	0.0002* (2.17)
Percentage of landless farmers	0.00006* (2.26)	0.00004 (1.46)	0.0003** (3.02)	0.0004** (3.26)
Percentage of farmers with more than 100 acres of land	-0.001* (2.42)	-0.0001* (3.15)	-0.0004* (3.00)	-0.0006* (3.25)
Number of mosques per village	-0.003* (2.41)	-0.001† (1.66)	-0.011** (3.03)	-0.013** (2.94)
Percentage of villages with no drinking water	0.00002 (1.01)	-0.00001 (-0.71)	0.0001 (-0.98)	0.0001 (1.30)
Percentage of villages with no electricity	0.00002† (1.69)	0.00002 (1.47)	0.0001* (2.50)	0.0002* (2.21)
Percentage of ethnically Kurdish population	0.0001** (5.05)	0.0004** (5.86)	0.0003** (3.93)	0.0006** (4.26)
Urbanization rate	-0.00004 (1.09)	-0.00002 (0.56)	-0.0002 (1.47)	-0.0001 (0.64)

Literacy rate	0.00004 (0.83)	-0.00007 (1.01)	0.0002 (1.17)	0.0002 (0.93)
Unemployment rate	0.0005* (1.97)	0.0005* (2.20)	0.002* (2.25)	0.002* (1.97)
Population in 10 thousand	0.00008† (1.67)	0.00006** (2.81)	0.0006* (1.98)	0.0007** (3.44)
Estimated parameters for seasonal and year dummies are not reported. † p<0.1, * p<0.05, ** p<0.01. z-values in parenthesis.				

A5. Results from a logit model

Table A5 presents the estimated odds ratios under a simple logit model with clustered errors at the county level. The model includes the same controls as the main biprobit model.

TABLE A5. Logit results, odds ratios	Unit of observation: **County-month (1984-2018)** Number of obs.: 266845	Unit of observation: **County-quarter year (1984-2018)** Number of obs.: 88734	Unit of observation: **County-half year (1984-2018)** Number of obs.: 44367	Unit of observation: **County-year (1984-2018)** Number of obs.: 21862
Length of last peace spell	0.988** (-13.77)	0.971** (-11.72)	0.892** (-8.28)	0.875** (-4.54)
Percentage of past time periods with conflict events	1.056** (10.11)	1.042** (13.39)	1.025** (10.52)	1.018** (8.18)
Inverse-distance-weighted **state casualties** in **other counties** in t-1	6.275** (9.13)	2.996** (6.89)	5.256** (10.08)	3.030** (7.75)
Inverse-distance-weighted **insurgent casualties** in **other counties** in t-1	1.156* (2.33)	0.867** (-2.77)	0.878** (-2.63)	0.854** (-3.50)
State casualties in the county in t-1	1.116** (5.48)	1.071** (4.72)	1.060** (2.83)	1.078** (3.48)
Insurgent casualties in the county in t-1	1.022** (3.39)	1.020** (2.71)	1.023** (3.20)	1.038** (3.78)
County area in km squares	1.0001† (1.87)	1.0001* (2.27)	1.0001* (2.52)	1.0002** (3.29)
Percentage of mountainous terrain in rural area	1.005† (1.93)	1.005* (2.23)	1.006* (2.454)	1.007** (2.94)
Border status	1.236* (1.96)	1.288* (2.32)	1.435** (3.20)	1.617** (3.88)
Percentage of villages whose names were changed by the state	1.004† (1.93)	1.006* (2.54)	1.006** (3.08)	1.008* (3.14)
Percentage of landless farmers	1.008** (3.27)	1.009** (3.20)	1.009** (3.07)	1.012** (3.49)
Percentage of farmers with more than 100 acres of land	0.986** (-2.84)	0.987** (-2.82)	0.988** (-2.84)	0.986** (-3.07)
Number of mosques per village	0.681** (-4.11)	-0.683 (-4.30)	0.664** (-4.33)	0.649** (-3.93)
Percentage of villages with no drinking water	1.003 (0.80)	1.002 (0.85)	1.003 (0.95)	1.003 (0.91)
Percentage of villages with no electricity	1.008** (3.27)	1.008** (3.39)	1.008 (3.39)	1.006** (2.85)
Percentage of ethnically Kurdish population	1.018** (6.77)	1.017** (5.98)	1.011** (3.89)	1.014** (4.19)
Urbanization rate	0.986** (-3.28)	0.986** (-3.12)	0.988* (-2.61)	0.989* (-2.01)
Literacy rate	1.009† (1.68)	1.009† (-1.70)	1.008 (1.55)	1.011† (1.84)
Unemployment rate	1.107** (4.55)	1.116** (4.60)	1.106** (4.00)	1.104** (3.57)
Population in 10 thousand	1.022** (6.79)	1.022** (6.99)	1.024** (6.95)	1.024** (6.14)
Estimated parameters for seasonal and year dummies are not reported. Standard errors are clustered at the county level. † p<0.1, * p<0.05, ** p<0.01. z-values in parenthesis.				

References

Cederman, Lars-Erik; Luc Girardin & Kristian Skrede Gleditsch (2009) Ethnonationalist Triads: Assessing the Influence of Kin Groups on Civil Wars. *World Politics* 61 (3): 403-37.

Cederman, Lars-Erik; Andreas Wimmer & Brian Min (2010) 'Why Do Ethnic Groups Rebel? New Data and Analysis. *World Politics* 62 (1): 87-119.

Cederman, Lars-Erik; Kristian Skrede Gleditsch, Idean Salehyan & Julian Wucherpfennig (2013) Transborder Ethnic Kin and Civil War. *International Organization* 67 (2): 389-410.

Do, Quy-Toan & Lakshmi Iyer (2010) Geography, Poverty and Conflict in Nepal. *Journal of Peace Research* 47 (6): 735-48.

Donnay, Karsten; Eric T Dunford, Erin C McGrath, David Backer & David E Cunningham (2018) Integrating Conflict Event Data. *Journal of Conflict Resolution* 63(5): 1137-1364.

Fearon, James D & David D Laitin (2003) Ethnicity, Insurgency and Civil War. *American Political Science Review* 97(1): 75-90.

Holtermann, Helge (2012) Explaining the Development-Civil War Relationship. *Conflict Management and Peace Science* 29(1): 56-78.

Holtermann, Helge (2016) Relative Capacity and the Spread of Rebellion: Insights from Nepal. *Journal of Conflict Resolution* 60(3):501-529.

Kudat, Ayşe (1975) Patron–Client Relations: The State of the Art and Research in Eastern Turkey. In: Engin Deniz Akarlı & Gabriel Ben-Dor (eds) *Political Participation in Turkey.* Istanbul: Boğaziçi University Press, 61–87.

Meeker, Michael M (1972) *The Great Family Aghas of Turkey.* In: Antoun & Harik (ed.) Rural Politics and Social Change in the Middle East, Indiana University Press.

Mutlu, Server (1996) Ethnic Kurds in Turkey: A Demographic Study. *International Journal of Middle East Studies* 28: 517-541.

Metternich, Nils W; Shahryar Minhas & Michael D Ward (2017) Firewall? or Wall on Fire? A Unified Framework of Conflict Contagion and the Role of Ethnic Exclusion. *Journal of Conflict Resolution* 61(6): 1151-1173.

Poirier, Dale J (1980) Partial Observability in Bivariate Probit Models. *Journal of Econometrics* 12 (2): 209–17.

Salehyan, Idean & Kristian Skrede Gleditsch (2007) Refugees and the Spread of Civil War. *International Organization* 60 (2): 335-66.

Tachau, Frank (1973) Turkish Provincial Party Politics. In: Kemal Karpat (eds) *Social Change and Politics in Turkey*. E.J Brill, Leiden

Tezcür, Güneş Murat (2016) Ordinary People, Extraordinary Risks: Joining an Ethnic Rebellion. *American Political Science Review* 110(2).

Turkish Ministry of Agriculture and Rural Affairs (1981) *Köy Envanter Etüdü* (Village Inventory). Ankara. All issues.

Turkish Ministry of Defense (1998) *Şehitlerimiz.* Ankara.

Weidmann, Nils B & Michael D Ward (2010) Predicting Conflict in Space and Time". *Journal of Conflict Resolution* 54(6):883-901.

Xiang, Jun (2010) Relevance as a Latent Variable in Dyadic Analysis of Conflict. *The Journal of Politics* 72(2): 484-498.

Zhukov, Yuri M (2012) Roads and the Diffusion of Insurgent Violence: The Logistics of Conflict in Russia's North Caucasus. *Political Geography* 31 (3): 144-56.