

Explaining Ethnic Violence on China's Western Frontier:

The Ethnic Violence in China (EVC) database and an Initial Test on the Effects of Inter-ethnic Inequality and Natural Resources in Xinjiang

Xun Cao, Haiyan Duan, Chuyu Liu, James Piazza, and Yingjie Wei¹

October 28, 2017

Forthcoming, *China Review*

Abstract: Unrest in the Xinjiang region of China currently poses the most imminent threat to the internal security of China and to central government control over peripheral regions. Instability in Xinjiang, furthermore, has ramifications for the wider security environment in Central Asia as the conflict becomes linked with jihadist groups in other security hotspots, like Pakistan and Syria. However, our understanding on important potential factors affecting political instability in Xinjiang is limited by the lack of systematically collected event data of ethnic violence. In this paper, we introduce the first effort to fill this gap in data collection, that is, the Ethnic Violence in China (EVC) Database: the Xinjiang Region. This is a geo-coded database of yearly incidents of ethnic violence at the county level in Xinjiang from 1990 to 2005. Using the EVC database, we demonstrate some initial results modelling ethnic violence in Xinjiang. We find that ethnic violence is positively associated with inter-ethnic inequality; resources such as oil and cotton, on the other hand, are unrelated to chances of ethnic violence.

Key words: ethnic violence; event data; China; Xinjiang.

¹ We want to thank the journal reviewers and editors for their helpful comments to improve the paper. All errors are ours. For the event data and replication files for the empirical analysis in the paper, please contact the corresponding author, Yingjie Wei, at yingjie0991@gmail.com.

Why Study Ethnic Violence in Xinjiang, China?

China is a prominent, essential country on the world stage and the status of its domestic stability has wider regional and global implications. A key factor affecting the internal security – and the political stability – of China is persistent political violence in the Xinjiang Autonomous Region in Western China. This resource-rich but underdeveloped province is home to the Muslim Uyghur ethnic minority group. In the past several decades, Xinjiang has seen a large in-migration of Han Chinese and has been the site of acute ethnic violence, often in the form of Uyghur riots and antigovernment terrorism perpetrated by Uyghur separatists.² Both Chinese government and international communities, furthermore, worry about growing transnational ties between Uyghur militants in Xinjiang and al Qaeda-oriented terrorists in South and Central Asia and the Middle East. During a recent visit to China, former U.S. National Security Advisor Susan Rice alleged that Chinese Uyghurs from Xinjiang had travelled to Iraq and Syria to fight in the ranks of the Islamic State (ISIS) terrorist insurgency (Page 2014).

Unrest in this region currently poses the most imminent threat to internal security and to Chinese central government control over peripheral regions (Mackerras 2012).³ Though no published study to date provides comprehensive figures, we recorded 213 ethnic violent events between 1990 and 2005 in the Ethnic Violence in China (EVC) database. In a 2004 speech, former Chinese President Hu Jintao identified Uyghur, "...separatism, extremism and terrorism" as the paramount domestic security question facing contemporary China (Davis 2010). The Chinese central government has implemented various policies to address violent unrest in Xinjiang, ranging from traditional policing and counterterrorism tactics such as those featured in the various "Strike Hard" campaigns to a recent move to send 200,000 civil servants to villages in Xinjiang to better provide local public goods.⁴ These strategies, however, have not successfully quelled the unrest or dampened tensions between ethnic minorities and the government.

The conflict in Xinjiang has security implications globally. Jihadist groups in South Asia and the Middle East, including the Iraq and Syria-based Islamic State movement, have issued statements in support of Xinjiang's independence from China and integration into a global Islamic Caliphate, while Uyghur militants from China and Central Asia have joined and fought within a host of terrorist movements such as the Taliban and the Al Qaeda core (Potter 2013). The East Turkestan Islamic Movement (ETIM), based in Pakistan, appears on the U.S. State

² One should note the difference between terrorism and separatism, especially given the fact that they have sometimes been used interchangeably by the government and media in China. Terrorism can be defined, for example, by the UN as criminal acts intended or calculated to provoke a state of terror in the general public, a group of persons or particular persons for political purposes. The defining characteristics are often first, no distinction between civilian and military targets, and second, the victim is not the target. Separatists, on the other hand, do not have to engage in terrorist activities. Separatism is often considered the advocacy or practice of separation of a certain group of people from a larger body on the basis of ethnicity and religion.

³ See Jay Ulfelder's recent post on Hong Kong protests and Xinjiang's insurgency: <http://dartthrowingchimp.wordpress.com/2014/07/01/another-chicken-little-post-on-china/>, accessed on November 3, 2016. Also, see O'Brien and Deng (2017) for a recent study on protest and repression in China.

⁴ Some believe that at least some of the civil servants were associated with conducting intelligence and surveillance work (Leibold and Zenz 2016).

Department's Foreign Terrorist Organizations list, and is believed to have been implicated in attacks within China and in third countries.⁵ Understanding what drives ethnic violence in Xinjiang helps provide leverage on the global jihadist picture.

With regard to contributions to academic literatures on ethnic violence and civil conflicts, Xinjiang, China is an ideal case which allows us to examine what we suspect is a key process in fostering an environment in which political instability and ethnic conflict is likely to occur in a large number of developing countries: liberal economic reform coupled with continued authoritarian rule. Though China embarked on free market economic reforms starting in the early 1980s, and has applied those policies in Xinjiang as an economic development strategy and lure for Han migration to the region, it has engaged in little political reform. This situation of liberal economic policies with illiberal rule in Xinjiang can be identified in countries in the Middle East (for example, Jordan and Gulf states like Qatar and Bahrain) and in some Southeast Asian countries like Singapore and Vietnam. Studying violence and conflict in Xinjiang helps us to better understand civil order in such pluralistic societies.

The literature on conflict in Xinjiang speculates on several causal factors for ethnic violence in the region: economic inequality between ethnic minority groups and Han Chinese; a legacy of suppression of local cultural and religious life by Chinese authorities; political repression and underrepresentation of ethnic minority groups in provincial affairs and political life; excessive use of force by government agents during policing and counterterrorism efforts; and growing Islamist sentiment, some of which is fueled by radical Muslim movements in the neighboring Central Asian republics and in Pakistan and Afghanistan (Zhao 2010). However, scholarly and expert understanding of the contours and causes of ethnic violence in Xinjiang have been limited by one key factor: existing work on Xinjiang is often qualitative and has not systematically evaluated patterns and precipitants of violence using data and other forms of analysis, leaving us with little understanding of how and why ethnic violence occurs, and what can be done about it.⁶

In this paper, we introduce the first effort to fill this gap in data collection, that is, the Ethnic Violence in China (EVC) Database: the Xinjiang Region, 1990-2005. This is a geo-coded database of yearly incidents of ethnic violence (e.g., riots, terrorist incidents, acts of inter-ethnic violence) occurring at the county level in Xinjiang from 1990 to 2005. The specific collection and coding technique is described in detail in the next section. Incidents of ethnic violence, such as riots, and political violence, such as acts of terrorism credited to extremists, are included in this event data set as information such as location and number of casualties, to measure the “intensity” of the event.⁷ Moreover, using the EVC event data, we demonstrate some initial

⁵ See, for example, <https://www.cfr.org/backgrounder/east-turkestan-islamic-movement-etim>, for the Council on Foreign Relations' description of the ETIM, accessed on August 2, 2017. Also see Chung (2002) and Kan (2010).

⁶ However, see Hong and Yang (Forthcoming) for how natural resources and local mosques affect violence in the region; Liu (2017) for how local public goods spending decreases ethnic violence in Xinjiang; Cao et al. (2017) for how local religious institutions mediate the effect of horizontal inequality on ethnic violence. The main empirical analyses in Liu (2017) and Cao et al. (2017) use the EVC data introduced in this paper.

⁷ A “county-year” version of panel data will accompany this database to best capture the granularity and local nature of the events. Such a unit of analysis is crucial, we argue, as we expect local factors to be of importance in determining patterns of violent events in the region.

results modelling ethnic violence in Xinjiang as a function of factors such as horizontal inequality between the Han majority and ethnic minorities, resources such as oil and cotton, income per capita, population density, local ethnic composition, and government fiscal spending. We find that ethnic violence is positively associated with horizontal inequality; resources such as oil and cotton, however, are unrelated to chances of ethnic violence.

Collection and Coding Ethnic Violence for Xinjiang, China

In this section we discuss the construction of the Ethnic Violence in China (EVC) database. Our data collection and analysis is focused on the county-year level: Xinjiang has 15 prefectures that are further divided into 100 counties. We have coded violent events at the county level from 1990 to 2005. Various types of events are included such as bombing, assassination, riot, arson, and armed attacks. We also have coded variables on the time, location, perpetrators and targets, casualties, and event types for each event and also include detailed notes on the event.

Definition of Ethnic Violence: for the purposes of the study, we consider ethnic violence to be the intentional execution of violent acts, perpetrated by individuals or groups of ethnic minority status, with political motivations. Acts of ethnic violence include but are not limited to: 1) acts of political terrorism – for these we use the operational definition of terrorism provided by the Global Terrorism Database (GTD) hosted by the University of Maryland START center; 2) acts of political insurgency – for these we use the operational definition provided by the Peace Research Institute Oslo (PRIO); 3) riots; 4) violent street demonstrations and protests; and 5) assassinations.

Note that this conceptualization associates intentionality with the committing of the act itself, irrespective of the outcome it produces. We accordingly define violent attempts also as violence in our database. For instance, bombing and assassination attempts that were prevented by the police, often via violent means (e.g., arrests followed by siege and gun fights) are counted as violent events.⁸ Moreover, because our focus is on predictors of violence rooted in the Xinjiang crisis, we only include violent events that at least involved one ethnic minority. In other words, we do not include violent events that occur between members of the Han majority or events occurring between Han civilians and the government.⁹

Variables in the EVC Xinjiang Database: the EVC database is an event data set. We first include variables that measure the timing and duration of an event. *Year Original* is the exact description of time from the original document. Second, we use three variables, *Day*, *Month*, and *Year* to further describe the date(s)/month(s)/year(s) of an event onset, if available. For example, a value of 1 represents the first day of a month/January for the *Day/Month* variable; 1990 represents year 1990. Finally, the variable *Duration* counts the number of days that an event lasted.

⁸ We also assume that arms manufacturing is a step towards violent attacks, therefore weapons production captured by the government is counted as violence. There are total six such events between 1990 and 2005.

⁹ Based on our definition, Han violence on minorities is part of our data. However, we found no record of Han violence on minorities. We suspect that there are two reasons. First, as Cote (2011) argues, the Han Chinese political mobilization was generally rare and limited in Xinjiang. Han Chinese tends to perceive themselves as the ethnic group in power; thus they are more likely to press the government to address their grievances. Second, it is possible that there are reporting biases that the media and the government tend to treat Han violence on minorities as crimes rather than politically charged ethnic violence.

We also coded variables measuring the location of an event. Most events do not have location information below the county level; therefore the most disaggregated spatial precision for our event data is county. The variable *Location* is the name(s) of the place(s) where an event occurred from the original source: note that there is only one report that mentioned multiple social unrest events in multiple counties, beginning since March 5, 1992 and lasted 4 days; the five counties are Khotan, Kashgar, Quca, Cocak, and Bortala (Bovingdon 2010). *County* is the name(s) of the county(ies) where an event occurred. *County Precision* describes the level of precision when we coded the event at the county level. There are two values for this variable: “confirmed” – either the original document provides county name(s) or we are able to locate county name(s) by our archival or Internet search; “missing” – the original document does not provide county name(s) and we are unable to locate county name(s) by our archival or Internet search.

Among the total 213 events coded between 1990 and 2005, there are 35 events for which despite our best efforts, the county location is missing. For instance, Bovingdon (2010) reports that on November 8, 1999 that “Turghun Aqniyaz, a policeman, was killed by violent terrorists on his way home.” We only know that this event happened in the Hetian prefecture and we are unable to find information on the exact county location of the event.¹⁰ Finally, the variable *Prefecture* gives the name(s) of the prefecture(s) where an event occurred.

Furthermore, we provide variables on perpetrators and targets of an event. *Perpetrator* provides information on the perpetrator of an event – Uyghur, others (i.e., members of non-Han and non-Uyghur ethnic groups), mixed (i.e., Uyghur and other non-Han ethnic groups), and missing (i.e., whether information regarding the perpetrator side is missing). Similarly the variable *Target* the nature of target of an event following the same coding rule: Uyghur, Han, others, mixed, government, and missing. Finally, a variable *Target Govt* shows whether government agencies (e.g., police station) were directly targeted: we identify three categories – government, civilians, government and civilians.

We also include variables that describe event types. *Type* includes event types such as bombing, assassination, riot, arson, gunfight, and attack on local governments. *Death* is the number of people killed in the event, if available; *Injury* the number of people injured in the event, if available; *Number Involved* is the number of people involved in the event, if available.

Finally, variables *source1*, *source2*, *source3*, *source4*, and *source5* list the data sources: 58 events were reported in more than one source. Note that different sources were used to collect event data. This is partly motivated by the advice to address the common concern of reporting bias when it comes to event data collection, which is to triangulate different sources (Davenport 2010; Weidmann 2015). These sources can be categorized into three broad categories: 1) government documents (Xinjiang Public Security Gazette, government white papers, and county gazettes of Xinjiang), 2) online event data sets and news search engines (The Global Terrorism Database (GTD), Minorities at Risk (MAR), and the WiseNews), and 3) secondary data from existing scholarly work (Bovingdon 2010 and Ma 2002).¹¹

¹⁰ In the regression analysis, we assign the events with only prefecture location to the prefecture capital city as the county-level location. However, excluding these events from the regression analysis does not change our main results.

¹¹ None of the sources that we use is considered classified material. More details about each data source are provided in the appendix of this paper.

We acknowledge the limitation of gathering information on ethnic violence in Xinjiang and in China in general. For instance, local officials and state media might seek to hide these incidents from public view.¹² While social media has increased our access to these incidents in the recent years, there is no certainty in our ability to know whether our event data has captured most, if not all, incidents that occurred during 1990-2005. However, even though it is impossible to completely rule out media bias, the fact that in addition to media reports, we also heavily rely on government documents – some were written for circulation within the government, e.g., government white papers – helps to address the concern of potential reporting biases.

In addition to the aforementioned variables, for each event, we also include a *Note*, which is a detailed description on an event from the original source(s). We provide this so that any researcher can check our coding of variables and also draw information that is not included in our coded variables. For example, one note includes an event reported by both the WiseNews and the Global Terrorism Database (GTD) report that on August 4, 2008:

“On Monday, at an unknown time, in Kashgar, Xinjiang, China, two unidentified assailants drove a truck into a group of policemen jogging near the barracks. The attack was claimed by the Eastern Turkistan Islamic Movement. There were no reported casualties in this incident.”

Data Collected, 1990-2005: we have finished coding ethnic violent events between 1990 and 2005 in Xinjiang. According to our data, there were 213 ethnic violent events in the region. Figure 1 shows the spatial distribution of the cumulated number of events between 1990 and 2005 for each county, with darker colour indicating higher number of violent events. The Yecheng county of the Kashgar prefecture (the darkest polygon, southwest of the region) experienced 17 ethnic violent events during this 16 year period, closely followed by the Kashgar county of same prefecture (14 events) and Yining county of the Ili prefecture (14 events). On the other hand, a number of counties in the east and southeast of Xinjiang, such as Yiwu, Barkol Kazakh, Ruoqiang, and Qiemo, experienced no violence at all. Figure 2 further displays the temporal distributions of violent events. The top panel shows the provincial annual sums and we saw the highest number of events in 1997. The bottom panel presents the patterns at the prefecture level. Violent events tend to be geographically concentrated in four prefectures: Akesu, Hetian, Kashgar, and Yili.

Insert Figure 1 and 2 here.

An Application of the EVC Database

In this section, we demonstrate one application of the EVC database by modelling ethnic violence at county-year level in Xinjiang. We believe that such a unit of analysis is crucial because we expect local factors to be of vital importance in determining patterns of violent events in the region. Regarding the dependent variable, we create a binary variable that takes the value of 1 in county-years where one or more ethnic violent events occurred; and 0 otherwise. This binarization process loses some information because there are county-years that experienced more than one event. However, multiple counts within one county-year are likely to

¹² See, for example, Leibold (2016).

be realizations of the same conflict process (Pierskalla and Hollenbach 2013). Because of missing values from the right-hand side variables, our regression analysis covers 1996-2005.

We have collected many potential predictor variables. This enables us to run logistic regressions to evaluate factors that might affect chances of ethnic violence. However, the exercise here is to demonstrate what we can learn by using the EVC event data rather than exhausting all choices of potential explanatory variables. We focus our theoretical discussion on the effects of inter-ethnic inequalities and natural resources. We control for the effects of many relevant local factors on violence such as mosque density, local government public spending, local state capacity, and ethnic fractionalization and polarization.

Inter-ethnic Inequalities: inter-ethnic inequalities are often considered a salient factor driving grievances not only in Xinjiang, but also other parts of the world (Fuller and Lipman 2004; Cederman et al. 2011; Piazza 2011). Indeed, grievances thesis has a long tradition and posits that civil conflicts are often associated with ethnic or religious hatred, political repression, political exclusion, and economic inequalities. Conflict therefore is explained as a result of widely felt grievances among the relatively disadvantaged in the society. Inequality has often been the focus of this tradition including theories of ethnic conflict and structural inequality and relative deprivation (Galtung 1964; Feierabend and Feierabend 1966; Hechter 1975; Gurr 1970, 1993 and 2000).¹³ Many studies of inequality and conflict indeed relate to the theory of relative deprivation which posits that while absolute poverty may lead to apathy and inactivity, comparisons with those in the same society who do better may inspire radical action and even violence (Gurr 1970).¹⁴

Recent studies have pointed out the fact that to properly test the grievances argument, one needs inequality measures based on differences between collective actors. Moreover, given the fact that armed conflicts and ethnic violence are often local events, it is also important to measure inequalities at the subnational level. Part of the ongoing effort in the study of civil conflicts and local violence is to come up with better measures using geocoded data newly available to researchers such as Demographic and Health Surveys (DHS), spatial distribution of ethnic groups and GDP data, and census data (Østby 2013).¹⁵

In the case of Xinjiang, socioeconomic inequality between the ethnic minority groups and the Han populations is perhaps the most frequently mentioned source of political violence in the

¹³ According to the greed model of conflicts, rebellion can be conceptualized as an industry that generates profits from looting, so that the insurgents are indistinguishable from bandits or pirates (Grossman 1999). For instance, Collier and Hoeffler (2004) emphasize the importance of factors affecting opportunities for financing rebellion, such as extortion of natural resources, donations from diasporas, and subventions from hostile foreign governments. They also highlight opportunities arising from atypically low costs for rebellion such as low opportunity costs of enlisting as a rebel, unusually cheap conflict-specific capital such as military equipment, and weak government military capability.

¹⁴ Others, on the other hand, disagree and view mobilization as the key factor affecting chances of civil unrest, because they believe that grievance are always present in every society while conflicts are not (Tilly 1978); while grievances might provide motivations, there is still often the collective action and mobilization issue.

¹⁵ Many find empirical support for the grievances approach (Østby 2008; Barron et al. 2009; Fjelde and Østby 2014; Cederman et al. 2011).

region by scholars (Fuller and Lipman 2004). For instance, economic, social, health and educational outcomes for Uyghurs and Han in Xinjiang exhibit stark differences (Wu and Song 2014; Wu and He 2016).¹⁶ Schuster (2009) finds that the infant mortality rate for Han Chinese is 13 per 1,000, but 102 per 1,000 for Uyghurs. According to Hasmath (2012), relatively to minorities, Han Chinese have a tendency to settle in wealthier urban areas, where Uyghurs tend to constitute the majority in rural areas or the poorer urban areas of southern Xinjiang. Cao (2010) demonstrates that Xinjiang experienced an increasing urban-rural income disparity in the 1990s. In 2000, counties that are characterized by worst urban-rural income gap are overwhelmingly concentrated in the south-west Xinjiang (26 of 27 counties) where ethnic minorities usually account for more than 90% of the total population. In addition, these counties are far laggard behind the average level of per capita GDP in Xinjiang (Chaudhuri 2010). In sum, we expect that:

H₁. Counties in Xinjiang with higher levels of inter-ethnic inequality experience higher levels of ethnic violence.

Natural Resources: Xinjiang is China's most crucial energy producing region. Around 60% of the gross regional product is produced by oil and gas extraction and ancillary industries, and the Chinese government has made large-scale investments in pipelines connecting Xinjiang with coastal cities like Shanghai (Charles 2005). Though oil and gas are Xinjiang's premier heavy industries, the central government has also promoted industrial-scale cotton production. Many believe that the significant orientation of the regional economy towards natural resource production has significantly contributed to conflict in Xinjiang. This is due to several reasons. First, the oil, gas and cotton boom in Xinjiang has attracted large numbers of Han migrants, exacerbating ethnic conflict.¹⁷

Second, the majority Uyghur population of Xinjiang expects local resource revenues to be spent on Uyghur priorities. Large scale investment in oil, coal and other natural resource extraction by Han-dominated government firms are perceived as unjustly exploiting the wealth of the local, non-Han population, fostering sharp resentments (Millward 2004). The benefits of government investment in the oil and gas industries have also been more heavily reaped in urban areas where Uyghurs are less dominant (Becquelin 2004). Finally, natural resource exploitation has generated environmental costs that sharpen conflict. Only 4.3 percent of Xinjiang's territory is inhabitable. The in-migration of Han, governmental agricultural policies, and oil and gas extraction have put a significant strain on water and arable land resources (Bhattacharji 2012). The combination of the migration-fueled population boom and the despoliation of the environment have served to deepen grievances in the region (Clarke 2008).¹⁸

These conclusions are often supported in the wider literature linking resource wealth and resource extraction with increased risks of internal political conflict and violence (Ross 2004 and 2012), including recent empirical work showing that when coupled with ethnic minority

¹⁶ See Tang et al. (2016) on the role played by language (Mandarin) efficiency in creating socioeconomic inequality between the Han majority and the Uyghur minority.

¹⁷ Kurlantzic (2004) argues that the settlement of Han in Xinjiang was part and parcel of a deliberate Chinese central government plan to maintain control over the region as it became more economically valuable. This conclusion is seconded by a U.S. Congressional study on China (CECC 2014).

¹⁸ See Shen and Xu (2016) for an overview of migration and development in China.

grievances and political disenfranchisement, oil greatly increases the chances of armed conflict (Asal et al. 2014). Indeed, the natural resource-conflict nexus is one of three branches of the resource curse literature (Ross 2015).¹⁹ In theory, there are multiple causal mechanisms connecting natural resources to chances of violence. For instance, some argue that rebels from an ethnically marginalized region could be motivated by the prospect of establishing an independent state, so that local resource revenues would not be shared with the rest of the country (Dal Bo and Dal Bo 2011). In sum, this leads us to expect:

H₂. Counties in Xinjiang with more natural resource wealth experience higher levels of ethnic violence.

Coding Horizontal Inequality and Resources: educational attainment is widely used to measure inter-ethnic inequality. For example, Fjelde and Østby (2014) use education years provided by the DHS to calculate the relative wealth of the poorest and richest group at the subnational level; using census data, Barron, Kaiser and Pradhan (2009) construct a district-level horizontal inequality index based on the ratio of group-level average education indicators. We use education attainments rather than direct measures on wealth also because the Chinese census data that we use does not contain information on wealth.

To measure the horizontal inequality, we use the 1% sample of the Chinese National Population Census of 1990.²⁰ It is reasonable to use horizontal inequality measures from 1990 to explain conflicts in 1996-2005. First, since we measure the value of inequality in 1990, it makes this indicator exogenous to the following conflicts between 1996 and 2005.²¹ Second, there is a widely shared assumption that the temporary changes in horizontal inequalities are relatively slow (Tilly 1999; Stewart and Langer 2009; Deiwiks et al. 2012).²² The 1% sample of 1990 census allows us to construct prefecture level inequality measures along ethnicity lines using information on individual-level educational attainments: the census data do not have county information so we have to measure horizontal inequality at the higher, prefecture, level.²³

We limit our sample to adults who were above 18 years old in 1990. There are five categories of educational attainments: less than primary completed, primary completed, lower

¹⁹ The other two branches concern the effects of resources on development and institutions.

²⁰ Minnesota Population Center (2015): <https://international.ipums.org/international-action/variables/samples?id=cn1990a>, accessed in June 2014.

²¹ For the pre-1996 period, the only publicly available Chinese census sample data are from 1982 and 1990. We choose not to use the 1982 data in the main text because 1990 is much closer to the period covered by our study (1996-2005).

²² In fact, we created the same horizontal inequality measure using the 1982 1% census sample and checked its correlations with the one using the 1990 data. The correlation is 0.773: therefore, at least between 1982 and 1990, inter-ethnic horizontal inequality changes slowly.

²³ The 2000 census data became publicly available in June 2017. The data were downloaded from the IPUMS-International (<https://international.ipums.org/international/>), accessed on June 4, 2017). This enables us to conduct a robustness check to see if we can replicate the main findings of this paper when using a new horizontal inequality measure based on both the 1990 and 2000 census data. An online Appendix of this paper (https://sites.psu.edu/xuncao/files/2016/01/EVC-research-note_sep_2017_online_app-uoglyf.pdf) discusses the details of the variable construction and presents the results of this robustness check: briefly, our main findings remain unchanged.

secondary school completed, high school completed, and university completed. Our horizontal inequality measure is the prefecture-level difference between Han and the largest minority group in terms of the percentage of individuals who at least completed the lower secondary school. The choice of lower secondary school completion as the threshold for education attainment fits the Chinese context. Despite China's "nine years" compulsory education system which requires a student to finish lower secondary school, the implementation of this national policy has been far from being ideal, especially in periphery regions like Xinjiang. Therefore, measures using lower secondary school completion as the threshold for education attainment better serve as proxies for wealth and income.

For the choices of ethnic minority groups, we compare the Han majority to the largest ethnic minority group in a prefecture. Despite media's focus on the Uyghur population when it comes to violence in Xinjiang, ethnic violent events are not limited to this ethnic group. Among the 213 events between 1990 and 2005, there are 166 events (77%) that we are certain did involve the Uyghur ethnic group. Moreover, in 5 among 15 prefectures in Xinjiang, the largest ethnic minority group is not Uyghur. In prefectures such as Tacheng, Changji, Shihezi, and Aletai, Uyghurs are less than 5% of the local population. If we want to understand all ethnic violence, we should focus on the differences between Han and the largest ethnic minority group.

Many studies have found a robust association between natural resources, especially oil wealth, and the prevalence of civil violence (De Soysa 2002; Collier and Hoeffler 2004). The redistribution of natural resource bounties often exacerbates the grievances of local minorities since they are usually excluded from sharing these benefits. We have collected county-level data regarding the geographic distribution of oilfields in Xinjiang. These data are taken from the *General Chronicles of Xinjiang: Oil Industry* (1999). We code this variable as 1 when there was at least one oilfield in a county-year and 0 otherwise. Finally, in addition to oil wealth, cotton is of great importance to the local economy in Xinjiang. It is also an economic sector where we see high level of wealth extraction by the government. Therefore, we include a variable for county-year level cotton production (in tons) per capita.²⁴

Control Variables: violence is more likely to happen in areas with higher population intensity (Hegre and Sambanis 2006; Raleigh and Hegre 2009), because population pressure exacerbates resource scarcity and worsens intergroup competition (Urdal 2008). *Xinjiang 50 Years* provides detailed annual statistics on county-level socio-economic variables, including population density data (Qiao 2005). Chances of ethnic conflict also depend on the size of local minority groups because large groups often are equipped with more resources for mobilization. We therefore include the county-year level largest ethnic minority group size, measured as a percentage of total county population.²⁵ Past studies have shown that income matters greatly for conflicts (Murshed and Gates 2005; Bohlken and Sergenti 2010). For example, Buhaug et al. (2011) find strong evidence for a negative correlation between per capita income and political violence. To control for the impact of low income and poverty, we include county-year GDP per capita.²⁶

²⁴ Data are from the county statistics of China data online: <http://chinadataonline.org/member/county/>, last accessed November 9, 2016.

²⁵ Data are also from *Xinjiang 50 Years*.

²⁶ Data are from *Xinjiang 50 Years*; we adjust GDP per capita based on 1990 Xinjiang price index to account for inflation.

To control for ethnic fractionalization and polarization, we use county ethnic composition data from the *Statistical Yearbook of Xinjiang* 1990-2005. We construct the fractionalization index following Fearon and Laitin (2003): $fractionalization = \sum_{i=1}^N \pi_i (1 - \pi_i)$; and polarization indicator following Montalvo and Reynal-Querol (2005): $polarization = 4 \sum_{i=1}^N \pi_i^2 (1 - \pi_i)$; in both cases, π_i is the percentage of people who belong to ethnic group i in a given county-year, and N is the number of ethnic groups in that county-year. Intuitively, ethnic fractionalization captures the probability that two randomly selected individuals do not belong to the same group. The polarization index, on the other hand, aims to capture how far the distribution of ethnic groups is from the bipolar distribution which represents the highest level of polarization.

Recent studies on religious institutions reveal their potential pacifying effects. For instance, religious institutions often assist local population by providing public goods (Dhingra and Becker 2001; Caputo 2009; Davis and Robinson 2012; Warner et al. 2015).²⁷ Public goods and emergency aid provided by local religious institutions help to address grievances and to prevent desperate citizens from using extreme actions against the government and other ethnic groups. Moreover, local public goods can improve living conditions of the population. We collect data on the numbers of mosques for each county. We use mosque density, standardizing the number of mosques by the 1990 county-level non-Han population.²⁸ The mosque data are from 81 Xinjiang official county gazettes published between the late 1980s and the early 2000s.²⁹ Since our regression analysis covers 1996-2005, we use mosque density of 1996 when the number of mosques in 1996 is available; for counties without 1996 mosque counts, we use the value from the most recent year between 1985 and 1995.³⁰ This strategy allows us to minimize missing data: if we only use the numbers of mosques in 1996, we would lose a large number of observations.³¹

²⁷ Religions such as Catholicism and Islam generate substantial amounts of charitable donations and volunteer work that help to sustain themselves as organizations with important public goods provision functions such as health clinics, local schools, and natural disasters relief efforts. For recent research on religion and politics, see Cao (2017) and McCarthy (2017).

²⁸ The overwhelming majority of non-Han population is Muslim, while few Han Chinese affiliate with Islam.

²⁹ The Spatial Religion Explorer from the China Data Online also provides mosque data (<http://chinadataonline.org/religionexplorer>). However, the Explorer data quality varies greatly from county to county. E.g., for counties in southern Xinjiang where the Muslim population concentrates, few mosques are reported, which contradicts what we found from county gazettes.

³⁰ For example, if the mosques data are only available for a county in 1985, 1990, and 1993, we use the value from 1993.

³¹ This is a crude measure for the density of local religious institutions. For instance, the data from county gazettes do not allow us to differentiate mosques attended by different Muslim ethnic groups such as Uyghur mosques vs. Kazak mosques. More importantly, there is the potential risk that the spatial distribution of mosques changed during the investigation period of 1996-2005. However, Li (2014) shows that since September of 1990, the government enforced a regulation on religious activities which strongly restricted, if not completely prohibited, the construction of new mosques in the region. As a result, the change in the number of mosques since late 1990 has been really small. For instance, there were 22.9 thousand mosques in Xinjiang in 1995, which is only about 600 (or 2.3%) more than in 1991 (Li 2014).

We also include a number of geographic variables: the distance to the provincial capital, the distance to the prefectural capital, and a dummy variable indicates whether the county is a border county. These variables often serve as proxies for state capacity (Fjelde and Østby 2014). For instance, the reach of the state declines as we move away from the center and into peripheral regions. Given the large size of Xinjiang, the distance to the provincial capital, Urumqi, may not fully capture state reach: this is the reason why we also include the distance to the prefectural capital.³² Furthermore, the Xinjiang Production and Construction Corps (XPCC) is a unique organization that combines functions of government, military and production. It has a hierarchical structure composed of 14 divisions and 175 regiments and these administrative domains of XPCC spread across the whole area of Xinjiang. XPCC can serve as an instrument of government control in the region. We include a dummy variable to indicate whether there was a XPCC administrative unit in a given county: note this is a time-invariant variable.³³

National border has been identified as a determinant of civil conflict (Buhaug and Rod 2006). State boundaries can offer porous exits for insurgents to find sanctuaries in neighboring countries. Moreover, in the case of Xinjiang, border counties are often more exposed to the transnational diffusion of radical Islam movement that overwhelmed many central Asian states such as Afghanistan and Pakistan. To further control for the roles of local governments, we also calculate government expenditure as a percentage of GDP.³⁴

Finally, spatial dependence between the units of observations often poses a challenge to the analysis of conflict. For a binary dependent variable, it is often computationally challenging to address this issue (Weidmann and Ward 2010). Following Pierskalla and Hollenbach (2013), we construct a temporally lagged spatial lag of the dependent variable by dividing the number of neighboring counties with violent events at year $t - 1$ over the total number of neighboring counties. To account for temporal dependence, we follow Carter and Signorino (2010) and add cubic polynomial approximation (t, t^2, t^3): t is the number of years since last violent event. We have also included year fixed effects in all model specifications to take into account common exogenous shocks. We use logit models and all standard errors are clustered by county. To weaken the problem of reverse causality, all time-variant independent variables lagged by one year; this does not affect the time-invariant variables included in the analysis though (Distance to the Prefecture's Capital, Distance to Ürümqi, Border County, Mosques per 1000 Non-Han population, XPCC, and Horizontal Inequality). Descriptive and correlation statistics are presented in Table 1 and 2.

Insert Table 1 and 2 here.

Empirical Results: Table 3 presents the empirical results. Our theoretical focus is on the effects associated with inter-ethnic education differences and natural resources. Across all four model specifications, *Horizontal Inequality*, which is constructed as the difference, between Han and the largest minority group, in the percentage of individuals who at least completed lower secondary

³² To calculate the distance between each county and provincial prefectural capital, we use *Google Map* to find the latitude and longitude of each county office building in Xinjiang.

³³ The list was downloaded from the following link <http://baike.baidu.com/view/38528.htm>, accessed on September 21, 2016.

³⁴ The government spending variable is from the *National Prefecture and County Finance Statistics Compendium*, 1994–2005.

school, has a positive effect on the chances of ethnic violence. On the other hand, the oil variable is negatively associated with violence; but the statistical significance level of this association varies by model specifications – the only statistically significant effect is found in the last model specification when we include government expenditure as a percentage of GDP in the regression analysis. County-level cotton production per capita also has no effect on violence. However, we need to be cautious in interpreting these non-results: we use the location of oilfields and county-level cotton production to capture resource extraction; but the location of oilfields and total cotton production might not reflect how jobs and resources are distributed in these extractive industries and especially to ethnic minority groups, if at all.³⁵

Insert Table 3 here.

To get a sense of the substantive effects, based on the third model specification in Table 3, we simulated the change in probability of a county-year experiencing ethnic violence given a one standard deviation increase from the mean of an independent variable, when holding all other variables at their mean levels.³⁶ The rope ladders in Figure 3 are 95% confidence intervals of the simulated probability changes. Horizontal inequality stands out in terms of the substantive effect: holding all other variables at their mean levels, one standard deviation increase from the mean of the horizontal inequality variable results in an increase in the probability of ethnic violence for a county-year by about 0.03. This is a substantively important effect because ethnic violent events do not occur often in the region; indeed, between 1996 and 2005, the mean of our binary dependent variable is only 0.08.

Insert Figure 3 here.

Other variables that affect the chances of ethnic violence include GDP per capita, population density, and border county status. Interestingly, we find that GDP per capital increases the chances of violence, which contradicts some recent studies on political violence. Many believe that economic prosperity reduces chances for violence because it alleviates ethnic competition by offering more resources; higher per capita income also presents higher opportunity cost for violence; economic development may also mitigate the level of grievances among marginalized groups (Buhaug et al. 2011). This positive effect of GDP per capita agrees

³⁵ Empirically, however, it is extremely difficult to measure the amount of natural resource wealth distribution to local ethnic minorities. Until recently, only a small fraction of oil revenue was kept by local governments while the rest went to central state-owned oil companies and the central government. For example, focusing on the city of Korla whose economy is built around oil extraction, Cliff (2016) shows that SOEs in the oil sector such as the local Tarim Oilfield Company predominantly employ Han Chinese. Moreover, even for the small fraction of oil revenue that was kept local, no data exists regarding what proportion was used to provides services and create jobs for local population.

³⁶ For year fixed effects, it is hard to interpret the results if we use the mean level, which is 0.1, for the simulations in Figure 3. We therefore choose to use the latest year for our investigation period, 2005, for the simulations. Using other years changes the magnitudes of predicted probabilities, but the shapes of the rope ladders do not change much. Simulation figures using other years are available upon request from the authors.

with other recent studies though. For instance, in the context of the collapse of Soviet Union, Beissinger (2002) finds that the level of urbanization (a proxy of economic development) increases the frequency of protests over ethno-nationalist issues because economic development often fosters the very condition for local minorities to form nationalist networks which increase the chances of violence (Gellner and Breuilly 2008).

Furthermore, we find that population density increases ethnic violence, confirming previous findings from earlier conflict studies. Finally, border counties are also associated with higher chances of conflicts; this is also similar to recent studies that has identified national border as a determinant of civil conflict (Buhaug and Rod 2006).

Conclusion and Discussion

In this paper, we introduce the Ethnic Violence in China (EVC) Database: the Xinjiang Region, 1990-2005. This is a geo-coded database of yearly incidents of ethnic violence occurring at the county level in Xinjiang from 1990 to 2005. Moreover, using the EVC event data, we further demonstrate some initial results explaining ethnic violence in Xinjiang. We find that ethnic violence is positively associated with horizontal inequality which is measured by education attainment difference between Han and the largest ethnic minority group in a prefecture. This suggests that grievances as a function of inter-ethnic inequalities need to be addressed to reduce ethnic violence in the region. Recent literature has also suggested that resource wealth and resource extraction increase risks of internal political conflict and violence (Ross 2004 and 2012; Asal et al. 2014). However, we do not find evidence supporting such a resource curse argument; neither oil nor cotton production is related to chances of ethnic violence in Xinjiang.

The results here are preliminary because there are other potential factors not controlled in the analysis. For instance, what is the role of marketization on violence and conflicts? Do free-market economic reforms – such as those implemented by China in Xinjiang – have an impact on political stability and the prospects for violence and armed conflict? These are unanswered questions not only in the study of Xinjiang unrest, but also in the broader literature of conflict studies.³⁷

Indeed, one of the most transformative politico-economic phenomena in the world in the past few decades has been the process of economic liberalization (marketization), especially in the developing world. Yet, we know little about how marketization affects domestic political stability in many countries of the world, especially those with histories of domestic conflict. Roland Paris, discussing the effectiveness of immediate democratization and marketization as a common strategy for post-conflict peacebuilding, argues that such political and economic liberalization is inherently tumultuous, and thereby undermines the prospects for stable peace.³⁸ Fearon and Laitin (2003) also point to the role of economic modernization in enhancing the positive effect of ethnic diversity on civil war. They argue that this is because more modernization should imply more discrimination and more nationalist contention in culturally divided countries. However, other scholars have argued the opposite, associating economic

³⁷ Note that we find that GDP per capita increases the chances of violence. In the Chinese context, wealth is highly correlated with marketization (Fan, Wang, and Zhu 2011). This result seems to suggest possible connections between marketization and ethnic violence.

³⁸ Paris (2004) cites research by Walton and Seddon (1994) that focuses on the relationship between widespread popular unrest and the promotion of free market through structural adjustment policies in many developing countries in the 80s and 90s.

liberalization with increased efficiency and welfare, and therefore lower chances of conflicts. For instance, recent research suggests that states that trade frequently experience fewer onsets of civil wars (de Soysa 2002) and lower chances of mass protests and political violence (Bussmann, Scheuthle, and Schneider 2003).

As far as we know, no existing research has studied the distributive effects of market liberalization in an ethnically diverse society and how it affects chances of ethnic violence and political stability. Aggregated, country level analyses often have difficulties to reveal nuanced differences in the ways that various aspects of marketization affect a society. This is especially relevant when we look at localized violent events instead of large scale civil wars. Another complication and empirical challenge is that democratization and marketization are often intertwined so that it is really difficult to separate the effects. Regarding both empirical challenges, the Xinjiang case provides a unique testing ground to study the effects of marketization on ethnic and political violence: our disaggregated event data approach allows us to better specify local precipitants of ethnic violence; the fact that China has engaged in substantial economic reform without substantial political reform allows us to examine effects of economic factors on political violence while holding political factors constant.

The aforementioned case of testing the marketization-violence nexus using the EVC data is just one example of what we can potentially learn by using this event data set. Accompanying this paper, we will make the database and replication code publicly available. One temporal limitation of the current event data is that it only covers 1990 to 2005.³⁹ There are questions concerning more recent events that cannot be answered using our data. For example, have the dynamics of interethnic conflict changed since the 7.5 Urumqi riots in 2009?⁴⁰ Therefore, we plan to extend event coding to cover the 2006-2016 period.⁴¹ Our hope is that this data set, as an

³⁹ There are a few reasons why we only collected data for 1990-2005 at this point. First, the collection of the data is very time-consuming: for example, one of the most challenging tasks is to find county-level locations of the events because many events from news reports do not provide precise location information such as county names for the events; we had to conduct online searches and consult secondary sources to find county names. Second, in addition to online event data sets, news search engines, and secondary data from existing scholarly work, we heavily rely on government documents. These government documents help to address the concern of potential reporting biases by the mass media (Ortiz et al. 2005). Moreover, many government documents are based on police sources and administrative archives, which significantly improve the scope of coverage (McCarthy et al. 1996; Barranco and Wisler 1999). However, one trade-off of using government documents is the fact that most of the government documents are only available till the early 2000s – this is the second reason why given the resources we had, we decided to code the 1990-2005 period first.

⁴⁰ Our empirical analysis shows that horizontal inequality is associated with violence during the 1996-2005 period. Without additional analysis covering the post-2005 period, we cannot answer questions like this. However, our main finding seems to be consistent with a few recent studies focusing on the post-“7.5” period (Brox and Bellér-Hann 2014; Hillman and Tuttle 2016). Our main finding that inter-ethnic hostility in Xinjiang was driven by local grievances is also consistent with recent studies on social exclusion (Cliff 2016), social inequalities (Fisher 2014), and unbalanced urbanization (Cappelletti 2015).

⁴¹ We will first be working with some of the same sources that we used for the 1990-2005 period. Moreover, for more recent years such as late 2000s, more online data sources become

important first step, would help scholars in fields such as civil armed conflicts, ethnic violence, and Chinese politics to not only better study and understand this important region of China, but also use Xinjiang as a test case for theories that can be generalized to other regions of the world.

available: we can turn to English databases such as Lexis-Nexis – this is the approach employed by the AidData project to code China’s aid projects in developing countries; we can use key word search on public search engines such as Google and Baidu to code events – the same method has been successfully applied by recent research on sensitive topics in Chinese politics such as Wallace and Weiss (2015) that use online search to collect data on recent mass anti-Japanese protests in Chinese cities. Finally, in addition to general search engines, there are a number of websites that closely follow the Xinjiang issue in recent years (e.g., www.ifeng.com, and <http://www.people.com.cn/>). For example, a key work search at <http://news.qq.com/a/20140302/005359.htm> provides detailed reports (in Chinese) of 16 violent events in Xinjiang from mid-2011 to early 2014.

References

- Asal, Victor, Michael Findley, James A. Piazza and James I. Walsh. 2014. "Political Exclusion, Oil and Ethnic Armed Conflict." *Journal of Conflict Resolution*. Forthcoming.
- Barron, Patrick, Kai Kaiser, and Menno Pradhan. 2009. Understanding Variations in Local Conflict: Evidence and Implications from Indonesia. *World Development* 37(3): 698–713.
- Barranco, J. and Wisler, D., 1999. Validity and systematicity of newspaper data in event analysis. *European sociological review* 15(3): 301-322.
- Becquelin, Nicolas. 2004. "Staged Development in Xinjigan." *China Quarterly* 178: 369-378.
- Beissinger, Mark. 2002. Nationalist mobilization and the collapse of the Soviet State. Cambridge University Press.
- Bhattacharji, Preeti. 2012. "Backgrounder: Uighurs and China's Xinjiang Region." *Council on Foreign Relations* 29. Available online at: <http://www.cfr.org/china/uighurs-chinas-xinjiang-region/p16870>.
- Bohlken, Anjali Thomas, and Ernest Sergenti. 2010. Economic Growth and Ethnic Violence: An Empirical Investigation of Hindu-Muslim Riots in India. *Journal of Peace research* 47(5): 589-600.
- Bovington, Gardner. 2010. *The Uyghurs: Strangers in Their Own Land*. Columbia University Press.
- Brox, T. and Bellér-Hann, I. eds., 2014. On the Fringes of the Harmonious Society: Tibetans and Uyghurs in Socialist China. Nias Press.
- Buhaug, Halvard, Lars-Erik Cederman and Kristian Skrede Gleditsch. 2014. "Square Pegs in Round Holes: Inequalities, Grievances, and Civil War." *International Studies Quarterly* 58(2): 418-431.
- Buhaug, Halvard, Kristian Gleditsch, Helge Holtermann, Gudrun Østby, and Andreas Forø Tollefsen. 2011. It's the local economy, stupid! Geographic wealth dispersion and conflict outbreak location. *Journal of Conflict Resolution* 55(5): 814-840.
- Buhaug, Halvard, and Jan Ketil Rød. 2006. Local determinants of African civil wars, 1970–2001. *Political Geography* 25(3): 315-335.
- Bussmann, Margit, Harald Scheuthle and Gerard Schneider. 2006. "Trade Liberalization and Political Instability in Developing Countries". In *Programming for Peace: Computer-Aided Methods for International Conflict Resolution and Prevention*, Robert Trapp ed. Springer: Berlin/Heidelberg.
- Cederman, Lars-Erik, Nils Weidmann and Kristian Gleditsch. 2011. "Horizontal Inequalities and Ethno-Nationalist Civil War: A Global Comparison." *American Political Science Review* 105(3): 478-95.
- Cao, Huhua. 2010. "Urban–rural Income Disparity and Urbanization: What is the Role of Spatial Distribution of Ethnic Groups? A Case Study of Xinjiang Uyghur Autonomous Region in Western China." *Regional Studies* 44(8): 965-982.
- Cao, Nanlai. 2017. Spatial Modernity, Party Building, and Local Governance: Putting the Christian Cross-Removal Campaign in Context. *China Review* 17(1): 29-52
- Xun Cao, Haiyan Duan, Chuyu Liu, and Yingjie Wei. 2017. Religious Institutions and the Impact of Inter-Ethnic Inequality on Conflict: the Case of Xinjiang, China. Working paper.

- Caputo, Richard. 2009. Religious Capital and Intergenerational Transmission of Volunteering as Correlates of Civic Engagement. *Nonprofit and Voluntary Sector Quarterly* 38: 983–1002.
- Cappelletti, A., 2015. Socio-economic disparities and development gap in Xinjiang, in Anna Hayes and Michael Clarke (ed.), *Inside Xinjiang: Space, Place and Power in China's Muslim Far Northwest*: 151 -182.
- Carter, David, and Curtis Signorino. 2010. Back to the future: Modeling time dependence in binary data. *Political Analysis* 18(3): 271-292.
- CECC, Congressional-Executive Commission on China. 2014. “Annual Report.” Available online at: <http://www.cecc.gov/>.
- Chaudhuri, Debasish. 2010. “Minority Economy in Xinjiang—A Source of Uyghur Resentment.” *China Report* 46(1): 9-27.
- Charles, Alain. 2005. *The China Business Handbook*. 8th Edition.
- Chung, Chien-peng. 2002. China's ‘War on Terror’: September 11 and Uighur Separatism. *Foreign Affairs* 81(4): 8–12.
- Clarke, Michael. 2008. “China’s War on Terror in Xinjiang: Human Security and the Causes of Violent Uighur Separatism.” *Terrorism and Political Violence* 20(2): 271-301.
- Cliff, Tom. 2016. *Oil and Water: Being Han in Xinjiang*. Chicago University Press.
- Cliff, T 2016, 'Lucrative Chaos: Inter-Ethnic Conflict as a Function of the Economic 'Normalisation' of Southern Xinjiang', in Ben Hillman and Gray Tuttle (ed.), *Ethnic Conflict and Protest in Tibet and Xinjiang*, Columbia University Press, New York, US, 122-150.
- Collier, Paul, and Anke Hoeffler. 2004. Greed and grievance in civil war. *Oxford economic papers* 56(4): 563-595.
- Côté, I., 2011. Political mobilization of a regional minority: Han Chinese settlers in Xinjiang. *Ethnic and Racial Studies* 34(11): 1855-1873.
- Dal Bó, Ernesto and Pedro Dal Bó. 2011. Workers, warriors, and criminals: social conflict in general equilibrium. *Journal of the European Economic Association* 9(4): 646-677.
- Davenport, Christian. 2010. *Media Bias, Perspective, and State Repression: The Black Panther Party*. New York: Cambridge University Press.
- Davis, Elizabeth V. W. 2010. “Uyghur Muslim Ethnic Separatism in Xinjiang, China.” *Asian Affairs* 35(1):15-30.
- Davis, Nancy and Robert Robinson. 2012. *Claiming Society for God: Religious Movements and Social Welfare*. Indiana University Press: Bloomington, IL.
- De Soysa, Indra. 2002. Paradise is a bazaar? Greed, creed, and governance in civil war, 1989-99. *Journal of Peace Research* 39(4): 395-416.
- Deiwijs, Christa, Lars-Erik Cederman, and Kristian Gleditsch. 2012. Inequality and Conflict in Federations. *Journal of Peace Research* 49(2): 289-304.
- Dhingra, Pawan and Penny Edgell Becker. 2001. Religious Involvement and Volunteering: Implications for Civil Society. *Sociology of Religion* 62: 315–35.
- Fan, Gang, Wang Xiaolu and Zhu Hengpeng. 2011. *NERI Index of Marketization of China's Provinces 2011 Report*. Economic Science Press, National Economic Research Institute, China Research Foundation.
- Fearon, James D., and David D. Laitin. 2003. Ethnicity, Insurgency, and Civil War. *American political science review* 97(01): 75-90.

- Feierabend, Ivo K., and Rosalind L. Feierabend. 1966. Aggressive Behaviors within Polities, 1948-1962: a Cross-national Study. *Journal of Conflict Resolution* 10(3): 249-271.
- Fischer, A. 2014. Labour Transitions and Social Inequalities in Tibet and Xinjiang: a comparative analysis of the structural foundations of discrimination and protest, in Brox, T. and Bellér-Hann, I. eds., 2014. *On the Fringes of the Harmonious Society: Tibetans and Uyghurs in Socialist China*, Nias Press, pp.29-68.
- Fjelde, Hanne, and Gudrun Østby. 2014. Socioeconomic Inequality and Communal Conflict: A Disaggregated Analysis of Sub-Saharan Africa, 1990–2008. *International Interactions* 40(5): 737-762.
- Fuller, Graham and Jonathan Lipman. 2004. "Islam in Xinjiang" in S. Fredrick Starr (ed.) *The Xinjiang Problem*. Washington D.C.: Central Asia-Caucasus Institute.
- Galtung, Johan. 1964. A structural Theory of Aggression. *Journal of Peace Research* 1(2): 95-119.
- Gellner, Ernest and John Breuilly. 2008. *Nations and nationalism*. Cornell University Press.
- Gurr, Ted Robert. 1993. *Minorities at Risk: a Global View of Ethnopolitical Conflicts*. Washington, DC: United States Institute of Peace Press.
- Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton, NJ: Princeton University Press.
- Gurr, Ted Robert. 2000. *Peoples versus States: Minorities at Risk in the New Century*. US Institute of Peace Press.
- Hasmath, Reza. 2012. "Migration, Labour and the Rise of Ethno-religious Consciousness among Uyghurs in Urban Xinjiang." *Journal of Sociology*: 1-15.
- Hechter, Michael. 1975. *Internal Colonialism: The Celtic Fringe in British National Development, 1536-1966*. University of California Press.
- Hegre, Håvard, and Nicholas Sambanis. 2006. Sensitivity Analysis of Empirical Results on Civil War Onset. *Journal of conflict resolution* 50(4): 508-535.
- Hillman, B. and Tuttle, G. eds., 2016. *Ethnic Conflict and Protest in Tibet and Xinjiang: Unrest in China's West*. Columbia University Press.
- Hong, Ji Yeon and Wenhui Yang. Forthcoming. Oilfields, Mosques, and Violence: Is There a Resource Curse in Xinjiang." *British Journal of Political Science*.
- Kan, Shirley A. 2010. U.S.-China Counterterrorism Cooperation: Issues for U.S. Policy. Congressional Research Report: <https://fas.org/sgp/crs/terror/RL33001.pdf>.
- Leibold, James. 2016. Interethnic Conflict in the PRC: Xinjiang and Tibet as Exceptions? In *Ethnic Conflict and Protest in Tibet and Xinjiang: Unrest in China's West*, edited by Ben Hillman and Gray Tuttle, Columbia University Press.
- Leibold, James and Adrian Zenz. 2016. Beijing's Eyes and Ears Grow Sharper in Xinjiang. *Foreign Affairs*, 23 December.
- Li, Xiaoxia. 2014. Changes in the Number of Mosques and Regulatory Policies in Xinjiang. *Sociology of Ethnicity* 164: 40-48.
- Liu, Chuyu. 2017. Local Public Goods Expenditure and Ethnic Conflict: Evidence from China. Working paper.
- Lu, Xiaobo and Pierre Landry. 2014. "Show Me the Money: Interjurisdiction Political Competition and Fiscal Extraction in China." *American Political Science Review* 108(3): 706-722.
- Ma, Dazheng. 2002. *National Interests the Highest*. Urumqi: Xinjiang People's Publishing House.

- Mackerras, Colin. 2012. "Causes and Ramifications of the Xinjiang July 2009 Disturbances." *Sociology Study* 2(7): 496-510.
- McCarthy, J.D., McPhail, C. and Smith, J., 1996. Images of protest: Dimensions of selection bias in media coverage of Washington demonstrations, 1982 and 1991. *American sociological review*: 478-499.
- McCarthy, Susan K. 2017. In Between the Divine and the Leviathan: Faith-based Charity, Religious Overspill and the Governance of Religion in China. *China Review* 17(2): 65-93.
- Millward, James. 2004. *Violent Separatism in Xinjiang: A Critical Assessment*. Washington: East-West Center.
- Montalvo, José G. and Marta Reynal-Querol. 2005. "Ethnic Polarization, Potential Conflict, and Civil Wars." *American Economic Review* 95(3): 796-816.
- Murshed, S. Mansoob, and Scott Gates. 2005. Spatial–horizontal Inequality and the Maoist Insurgency in Nepal. *Review of Development Economics* 9(1): 121-134.
- O'Brien, Kevin J. and Yanhua Deng. 2017. Preventing Protest One Person at a Time: Psychological Coercion and Relational Repression in China. *The China Review* 17(2): 179–201.
- Ortiz, D., Myers, D., Walls, E. and Diaz, M.E., 2005. Where do we stand with newspaper data?. *Mobilization: An International Quarterly* 10(3): 397-419.
- Østby, Gudrun. 2008. 'Polarization, Horizontal Inequalities and Violent Civil Conflict.' *Journal of Peace Research* 45 (2): 143–162.
- Østby, Gudrun. 2013. Inequality and Political Violence: A Review of the Literature. *International Area Studies Review* 16(2): 206–231.
- Page, Jeremy. 2014. "U.S. Seeks China's Backing for Coalition against Islamic State." *Wall Street Journal*: September 9.
- Paris, Roland. 2004. *At War's End: Building Peace After Civil Conflict*. New York: Cambridge University Press.
- Piazza, James A. 2011. "Poverty, Minority Economic Discrimination and Domestic Terrorism." *Journal of Peace Research* 48(3): 339-353.
- Pierskalla, Jan H. and Florian M. Hollenbach. 2013. "Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa." *American Political Science Review* 107:207–224.
- Potter, Philip. 2013. Terrorism in China: Growing Threats with Global Implications. *Journal of Strategic Studies* 7(4): 70-92.
- Qiao, Yuanrao. 2005. *Xinjiang 50 Years*. China Statistics Press.
- Qin, Bei, David Strömberg and Yanhui Wu. 2014. "Media Bias in Autocracies: Evidence from China." Working paper: <http://www-bcf.usc.edu/~yanhuiwu/MediaBias1.pdf>.
- Raleigh, Clionadh, and Håvard Hegre. 2009. Population Size, Concentration, and Civil War: A Geographically Disaggregated Analysis. *Political Geography* 28(4): 224-238.
- Ross, Michael. 2012. *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton, NY: Princeton University Press.
- Ross, Michael. 2004. "What Do We Know about Natural Resources and Civil War?" *Journal of peace research* 41(3): 337-356.
- Ross, Michael. 2015. What Have We Learned about the Resource Curse? *Annual Review of Political Science* 18:239-259.

- Schuster, Brenda L. 2009. "Gaps in the Silk Road: An Analysis of Population Health Disparities in the Xinjiang Uyghur Autonomous Region of China." *The China Quarterly* 198: 433-441.
- Shan, Wei, and Cuifen Weng. "China's New Policy in Xinjiang and its Challenges." *East Asian Policy* 2: 58-66.
- Shen, Jianfa, and Wei Xu. 2016. Migration and Development in China: Introduction. *China Review* 16(3): 1-7.
- Statistical Communique of the Provinces on National Economic and Social Development. 2013. Available online at: <http://www.tjcn.org/plus/view.php?aid=26117>.
- Stewart, Frances and Arnim Langer. 2009. Horizontal Inequalities: Explaining Persistence and Change. In: Frances Stewart (ed.) *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*. Houndmills: Palgrave Macmillan, 54–84.
- Tang, Wenfang, Yue Hu and Shuai Jin. 2016. Affirmative Inaction: Education, Language Proficiency, and Socioeconomic Attainment Among China's Uyghur Minority. *Chinese Sociological Review* 48(4): 346-366.
- Tilly, Charles. 1999. *Durable Inequality*. Berkeley, CA: University of California Press.
- Tyler, Tom R. 2006. "Psychological Perspectives on Legitimacy and Legitimation." *Annual Review of Psychology* 57: 375-400.
- Urdal, Henrik. 2008. Population, Resources, and Political Violence A Subnational Study of India, 1956–2002. *Journal of Conflict Resolution* 52(4): 590-617.
- Wallace, Jeremy L. and Weiss, Jessica Chen. 2015. The Political Geography of Nationalist Protest in China: Cities and the 2012 Anti-Japanese Protests. *The China Quarterly* 222: 403–429.
- Walton, John K. and David Seddon. 1994. *Free Markets and Food Riots: The Politics of Global Adjustment*. Wiley-Blackwell: Cambridge, MA.
- Warner, Carolyn, Ramazan Kilinc, Christopher Hale, Adam Cohen, and Kathryn Johnson. 2015. Religion and Public Goods Provision: Experimental and Interview Evidence from Catholicism and Islam in Europe. *Comparative Politics* 47(2): 189-209.
- Weidmann, Nils B. 2015. On the Accuracy of Media-based Conflict Event Data. *Journal of Conflict Resolution* 59 (6): 1129-1149.
- Weidmann, Nils B. and Ward, Michael D. 2010. Predicting Conflict in Space and Time. *Journal of Conflict Resolution* 54(6): 883-901.
- Wu, Xiaogang, Xi Song, 2014 "Ethnic Stratification amid China's Economic Transition: Evidence from the Xinjiang Uyghur Autonomous Region." *Social Science Research* 44:158-172.
- Wu, Xiaogang and Guangye He 2016. "Changing Ethnic Stratification in Contemporary China." *Journal of Contemporary China* 25(102): 938-954.
- Zhao, Tong. 2010. "Social Cohesion and Islamic Radicalization: Implications from the Uighur Insurgency." *Journal of Strategic Studies* 3(3): 39-52.

Appendix: Data Sources for the EVC

1. Chinese Government Documents:

- a. Xinjiang Public Security Gazette (XPSG: 《新疆通志 : 公安志》): This was published by the Xinjiang Uyghur Autonomous Regional Government Compilation Committee in 2004. It provides detailed information on ethnic conflicts in Xinjiang during the 1990s. The most valuable information in this book is presented in the 5th and 6th sections of the appendix (p.840-862). The title of 5th section is “Events of all types”, which covers protests, riots, and (armed) fights. In the 6th section, the gazette records important violent incidents from 1996 to 2000. XPSG offers valuable local sources for ethnic conflicts. This book is based on internally circulated government statistics, which enhances its reliability and credibility.
- b. The unpublished draft of Xinjiang Public Security Gazette (unpublished XPSG《新疆通志草稿油印本 : 公安志》): This is an unpublished draft of Xinjiang public security gazette discussed above. The draft provides a rare opportunity to access to original, internally circulated sources on violent events in Xinjiang. Main criminal activities and violent events are recorded annually in this draft. Different from the final published version, the draft covers episodes from 1949 to 1995. Moreover, the draft also presents each case thoroughly by articulating the time, locations and detailed descriptions. This source is highly reliable and authoritative owing to the nature of internal circulation. This draft gives us first-hand information regarding ethnic conflicts in Xinjiang after 1949. Same as the published version, the draft is informative about the location of each violent event, which makes the county-level coding feasible.
- c. Government white paper: the Information Office of the State Council of China made a white paper, “*East Turkistan Terrorist Forces Cannot Get Away with Impunity*”, publicly available via People’s Daily in 21 January 2002.⁴² This white paper includes events such as explosions, assassinations and attempted assassinations, attacks on police and government institutions, crimes of poison and arson, training and arms manufacturing bases, and plotting and organizing riots and creating terror. Whereas the white paper claims over 200 terrorist incidents between 1990 and 2001, it only reports details about 39 episodes. Moreover, many scholars doubt the reliability of this report (Clarke 2008b, p.282-283; Bovington 2010, p.115). However, this white paper is still the most important official document that reveals valuable statistics about ethnic tensions in Xinjiang during the 1990s. According to Clarke (2008b), the “*report is the sole source of material regarding incidents of assassinations and attacks involving poison and arson.*”
- d. County gazettes of Xinjiang (新疆县志) : <http://www.fudan.edu.cn/englishnew/library/library.html>. These gazettes record most detailed socioeconomic information regarding counties in Xinjiang. Strikingly, only a few of them mention ethnic violence and these records tend to only include the most salient cases such as the 1997 Ghulja incident.

2. Online Databases and News Archives and Search Engines.⁴³

⁴² See <http://www.china.org.cn/english/2002/Jan/25582.htm>.

⁴³ We have also searched popular English news online archives such as *LEXIS-NEXIS*. However, we found no addition event. Therefore, we do not list them in the main text of this proposal.

- a. The Global Terrorism Database (GTD, <http://www.start.umd.edu/gtd/>): This is a global database covering both international and domestic terrorist attacks. It is widely used in empirical research on terrorism. GTD records 39 events associated with Uyghur participants in Xinjiang from 1996 to 2011. Within these 39 events, however, 23 of them are listed repetitively: there was a bombing occurred on July 20, 1996 which is recorded 23 times in the GTD. We do not include other 22 records in our dataset. For each case, GTD provides information on timing, location, group involved, fatalities, target type, attack type, and weapon type.
 - b. Minorities at Risk (MAR, <http://www.cidcm.umd.edu/mar/chronology.asp?groupId=71003>): the MAR is a university-based research project that monitors the status and conflicts of politically active communal groups in all countries with a current population of at least 500,000. The project is designed to provide information in a standardized format that aids comparative research and contributes to the understanding of conflicts involving relevant groups. Although MAR's quantitative data do not disaggregate the geographic distribution of ethnic violent events in Xinjiang, the chronology of Turkmen in Xinjiang (Turkic minorities including Uyghurs, Kazakhs and Tajiks) is useful. After scrutinizing this chronology, we recorded 44 events.
 - c. WiseNews (<http://libguides.ust.hk/wiseneews>): WiseNews is a Hong Kong based online newspaper data search engine. It covers over 600 key newspapers, magazines, and websites from mainland China, Hong Kong, Macau, Taiwan, and the United States. WiseNews provides digital archives of 259 newspapers based in Mainland China from 2000 to date (Qin, Strömberg, and Wu, 2014). Among them, 125 are general-interest newspapers. Geographically, these newspapers cover 29 out of 31 provinces. In terms of government affiliation, ownership, and the type of readership, the WiseNews sample represents the whole newspaper industry in Mainland China. We use the relevant strings of Chinese characters (e.g., 新疆 (Xinjiang) and 暴力 (violence)) to search for the mass media coverage of political violence in the area. After ruling out overlapped events already reported by other sources, we found 5 new events between 2000 and 2011.
3. *Secondary Datasets from Existing Scholarly Work:*
 - a. Bovington (2010): This book offers very comprehensive quantitative evidence on the amount and frequency of political violence in Xinjiang between 1949 and 2005. This book carefully balances between Chinese official statistics and messages from transnational Uyghur organizations and human rights organizations.⁴⁴ The appendix of this book, entitled “organized protests and violent events in Xinjiang, 1949-2005”, provides the year, duration, location, type, number of people involved, and comments (which provides detailed descriptions of coded events) for mass protests and violent events. This appendix records 158 political violent events and organized protests from 1949 to 2005. Within these episodes, 142 had clear ethnic component. However, these events are recorded case by case without disaggregating at the county level. Our dataset reorganize the structure of this appendix by geo-coding each case at the county level.

⁴⁴ On page 174: “Bibliographic references for each of these reported events available on request. As discussed in chapter 4, all but a handful of these events are described in multiple sources. Wherever possible, I compared the Chinese and foreign sources when preparing the descriptions.”

- b. Ma (2002): this is a highly valuable internally circulated source regarding ethnic conflicts in Xinjiang. According to Bovington (2010), “Ma’s importance as both a researcher and a policy adviser on Xinjiang cannot be overestimated.” The research that he and his assistants conducted for the book was commissioned by top party leaders in Beijing and written in consultation with high officials in Xinjiang. This book offers detailed information on 28 ethnic violent events in Xinjiang from the 1950s to the 1990s. Furthermore, the book highlights the evolution of ethnic violence in Xinjiang during the 1990s.

Figure 1: Distribution of Ethnic Violent Events in Xinjiang, between 1990 and 2005.

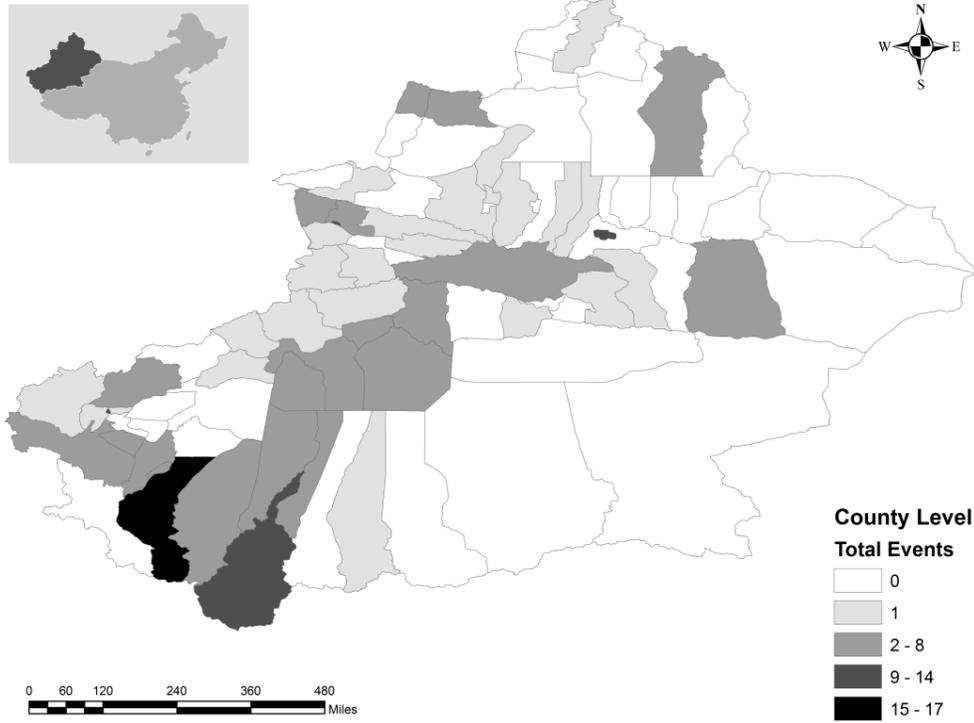
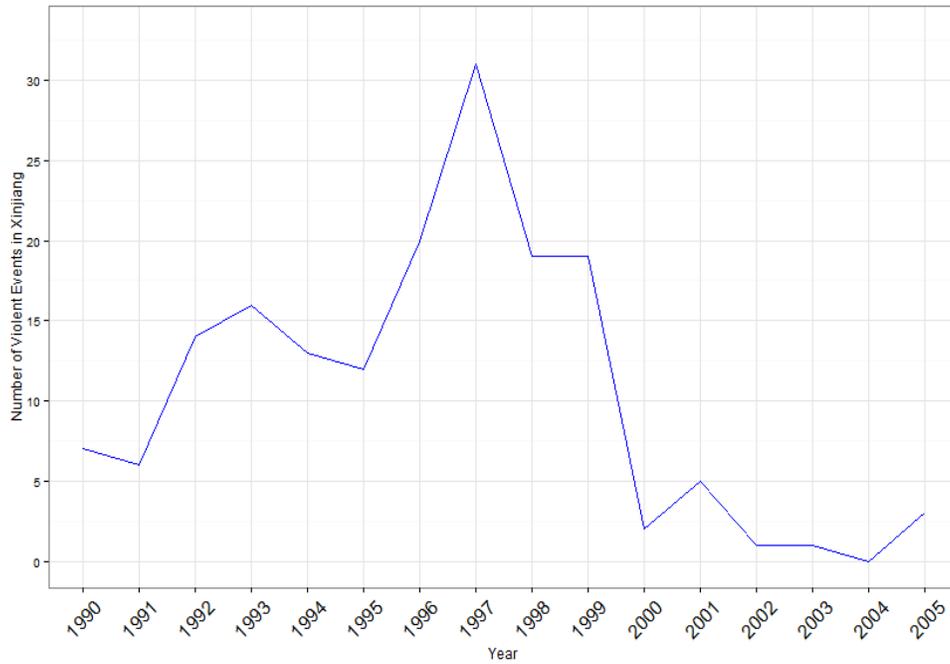
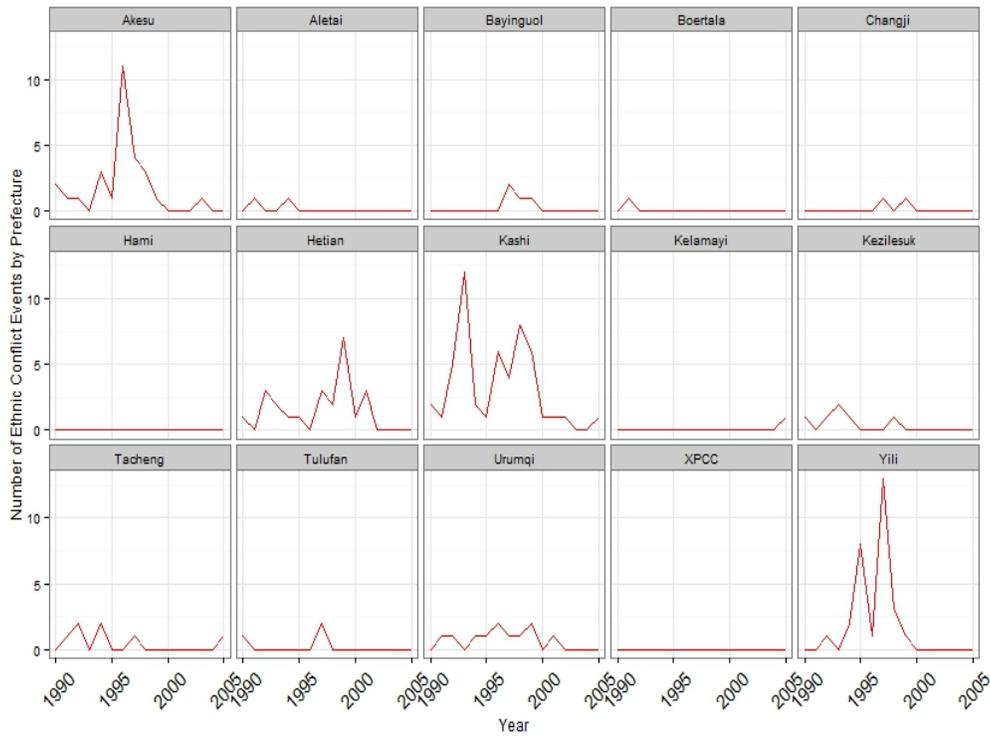


Figure 2: Annual Violent Events in Xinjiang, Provincial and Prefectural Trends.



(a) provincial level trend



(b) prefecture-level trends

Figure 3: Change in Probability of Ethnic Violence Given One Standard Deviation Increase from the Mean of a Variable (based on the third model specification from Table 3).

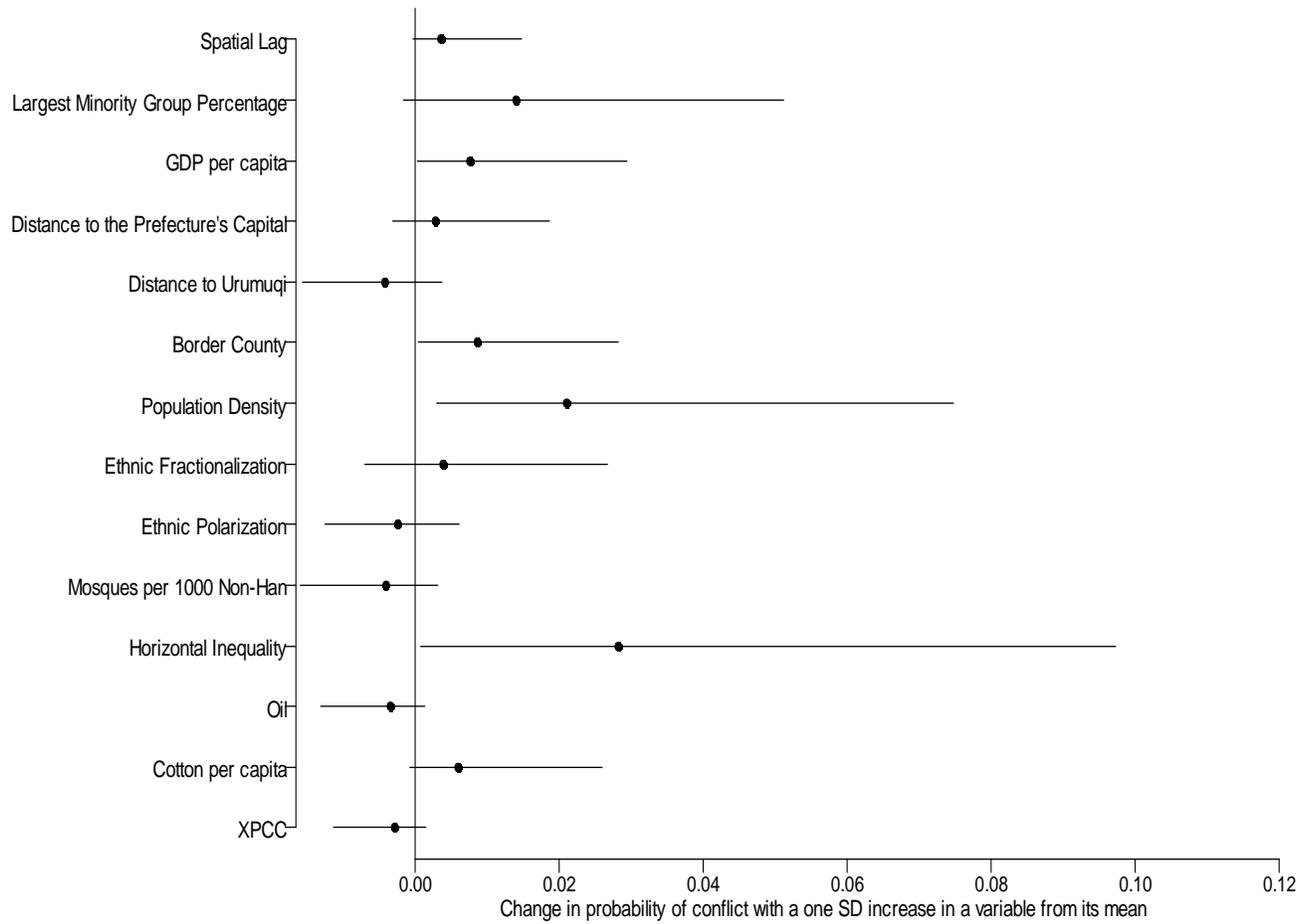


Table 1: Descriptive Statistics.

Statistic	N	Mean	St. Dev.	Min	Max
Conflict Event	957	0.090	0.286	0	1
Spatial Lag	979	0.084	0.162	0.000	1.000
Largest Minority Group Percentage	940	49.612	34.655	0.150	99.453
GDP per capita	848	0.309	0.425	0.052	5.780
Distance to the Prefecture's Capital	1,001	96.593	93.590	0	406
Distance to Üümqi	1,012	540.826	338.210	0	1,241
Border County	1,122	0.304	0.460	0	1
Population Density	932	588.821	2,156.332	1.339	26,000
Ethnic Fractionalization	951	0.388	0.165	0.011	0.754
Ethnic Polarization	951	0.550	0.213	0.022	0.970
Mosques per 1000 Non-Han	712	1.635	1.116	0.128	5.398
XPCC	1,122	0.098	0.298	0	1
Horizontal Inequality	1,056	29.871	21.330	-8.659	62.850
Oil	1,122	0.137	0.344	0	1
Cotton Production per capita	804	0.063	0.100	0.000	0.869
Government Expenditure /GDP	850	0.184	0.149	0.020	1.051

Note: XPCC stands for the Xinjiang Production and Construction Corps.

Table 2: Correlation Statistics.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1:Conflict Event	1															
2:Spatial Lag	0.28	1														
3:Minority Group Percentage	0.08	0.08	1													
4:GDP per capita	-0.06	-0.15	-0.28	1												
5:Distance to the Prefecture's Capital	-0.09	-0.03	0.02	-0.14	1											
6:Distance to Ürümqi	0.09	0.12	0.63	-0.35	0.03	1										
7:Border County	-0.10	-0.10	-0.18	-0.07	0.00	0.15	1									
8:Population Density	0.20	0.00	0.09	-0.04	-0.17	0.19	-0.16	1								
9:Ethnic Fractionalization	-0.06	-0.08	-0.64	0.15	-0.11	-0.49	0.21	-0.03	1							
10:Ethnic Polarization	-0.03	-0.03	-0.46	0.18	-0.05	-0.41	0.14	0.00	0.79	1						
11:Mosques per 1000 Non-Han	0.03	0.09	0.75	-0.30	-0.11	0.77	-0.10	-0.06	-0.50	-0.36	1					
12:XPCC	-0.06	-0.04	-0.20	0.00	-0.23	-0.16	0.00	0.02	-0.06	-0.04	-0.19	1				
13:Horizontal Inequality	0.14	0.19	0.40	-0.31	0.04	0.57	-0.20	0.16	-0.49	-0.43	0.54	-0.19	1			
14:Oil	-0.04	-0.01	0.07	0.18	0.22	-0.15	-0.08	-0.08	-0.04	0.01	-0.07	-0.13	0.10	1		
15:Cotton Production per capita	-0.01	0.00	0.03	0.10	0.19	-0.11	-0.31	-0.11	-0.10	0.06	0.04	0.02	0.07	0.13	1	
16:Government Expenditure /GDP	-0.12	-0.07	0.18	-0.29	0.33	0.49	0.22	-0.10	-0.19	-0.23	0.34	-0.19	0.22	-0.10	-0.18	1

Note: XPCC stands for the Xinjiang Production and Construction Corps.

Table 3: Binary Logistic Regression Models Predicting Ethnic Violence in Xinjiang

	(1)	(2)	(3)	(4)
Spatial Lag	1.448 [*] (0.842)	1.540 [*] (0.909)	1.586 [*] (0.926)	1.368 (0.917)
Largest Minority Groups Percentage	0.017 (0.012)	0.011 (0.011)	0.021 (0.013)	0.025 (0.015)
GDP per capita	1.545 ^{***} (0.504)	1.202 [*] (0.601)	1.251 ^{**} (0.611)	0.816 (0.708)
Distance to the Prefecture's Capital	0.002 (0.003)	-0.0001 (0.002)	0.001 (0.002)	0.002 (0.002)
Distance to Ürümqi	-0.002 (0.001)	-0.001 (0.001)	-0.002 (0.001)	-0.001 (0.001)
Border County	0.931 [*] (0.526)	0.839 [*] (0.468)	1.141 ^{**} (0.520)	1.101 ^{**} (0.491)
Population Density	0.572 ^{***} (0.173)	0.565 ^{***} (0.165)	0.603 ^{***} (0.168)	0.451 ^{***} (0.164)
Ethnic Fractionalization	0.381 (2.473)	0.780 (2.329)	1.282 (2.374)	1.733 (2.246)
Ethnic Polarization	-1.167 (1.758)	-1.316 (1.646)	-1.450 (1.682)	-1.708 (1.468)
Mosques per 1000 Non-Han	-0.354 (0.341)	-0.370 (0.304)	-0.465 (0.319)	-0.466 (0.319)
XPCC	-1.635 (1.093)	-1.046 (0.941)	-1.233 (0.942)	-1.018 (0.958)
Government Expenditure/GDP				-0.067 (0.051)
Horizontal Inequality	0.071 ^{**} (0.032)	0.049 [*] (0.027)	0.065 ^{**} (0.030)	0.072 ^{**} (0.034)
Oil	-0.350 (0.471)		-0.903 (0.557)	-1.108 [*] (0.559)
Cotton Production per capita		2.838 (2.005)	2.945 (2.044)	1.778 (2.009)
t	-0.451 (0.301)	-0.281 (0.325)	-0.264 (0.327)	-0.246 (0.333)
t^2	0.058 (0.055)	0.024 (0.060)	0.023 (0.060)	0.021 (0.062)
t^3	-0.002 (0.003)	-0.0004 (0.003)	-0.0004 (0.003)	-0.0002 (0.003)
Constant	-12.518 ^{***} (2.271)	-9.870 ^{***} (2.351)	-11.050 ^{***} (2.543)	-8.954 ^{***} (2.465)
Observations	682	580	580	580
Log Likelihood	-120.628	-108.010	-106.909	-105.081
Akaike Inf. Crit.	295.255	268.021	267.818	266.162

Note:

* p<0.1; ** p<0.05; *** p<0.01