

# Trading Hard Hats for Combat Helmets: The Economics of Rebellion in Eastern Ukraine\*

Yuri M. Zhukov

Department of Political Science, University of Michigan

forthcoming in *Journal of Comparative Economics*

November 1, 2015

## Abstract

Using new micro-level data on violence in Eastern Ukraine, this article evaluates the relative merits of ‘identity-based’ and ‘economic’ explanations of civil conflict. The first view expects rebellion to be most likely in areas home to the geographic concentration of ethnolinguistic minorities. The second expects more rebel activity where the opportunity costs of insurrection are low. Evidence from the armed conflict in Ukraine supports the second view more than the first. A municipality’s prewar employment mix is a more robust predictor of rebel activity than local ethnolinguistic composition. Municipalities more exposed to trade shocks with Russia experienced a higher intensity of rebel violence throughout the conflict. Such localities also fell under rebel control earlier – and took longer for the government to liberate – than municipalities where the labor force was less dependent on exports to Russia. (*JEL* D74, F14, F51, F52, H56, O24)

---

\*I am grateful to the Weiser Center for Emerging Democracies for financial support, to Ralph Clem, Michael Colaresi, Ruben Enikolopov, Evgeny Finkel, Zuhaib Mahmood, Vera Mironova, Lauren Pinson, Alina Polyakova, anonymous reviewers at JCE and workshop participants at the University of Michigan, George Washington University, Michigan State University, University of Toronto, Yale University and the Association for the Study of Nationalities for helpful comments, and to Bogdan Belei, Christina Benson, Maura Drabik, Laura Marsh, Daniella Raz and Alyssa Wallace for excellent research assistance.

*Before the war, I worked in a mine for three years. Last year, the price of coal started to drop and many mines started to close. I lost my job. Then, with what happened during the spring, I decided to go out and defend my city.*

1

Rebel fighter with Vostok battalion  
Pisky, Donetsk oblast, July 2015

In March and April 2014, angry mobs and armed med stormed administrative buildings and police stations in Eastern Ukraine. Waving Russian flags and condemning the post-revolutionary government in Kyiv as an illegal ‘junta,’ the rebels proclaimed the establishment of ‘Peoples’ Republics’ of Donetsk and Luhansk and organized a referendum on independence. Despite initial fears that the uprising might spread to other provinces, the rebellion remained surprisingly contained. While 61 percent of municipalities in Donetsk and Luhansk fell under rebel control during the first year of the conflict, just 20 percent experienced any rebel violence.

What explains local variation in rebellion? Why might two municipalities in the same region experience different levels of separatist activity? Why do some towns remain under government control while others slip away? Why might residents of one municipality be more receptive to foreign fighters and support than another?

This paper uses new micro-level data on violence and economic activity in Eastern Ukraine to evaluate the relative explanatory power of ‘identity-based’ and ‘economic’ drivers of civil conflict. The first category places an emphasis on ethnicity and language. According to this view, the geographic concentration of an ethnolinguistic minority – in this case, Russians or Russian-speaking Ukrainians – helps local rebels overcome collective action problems (Bates, 1983; Weidmann, 2009), while attracting an influx of fighters, weapons and economic aid from co-ethnics in neighboring states

---

<sup>1</sup>Langston (2015)

(Gleditsch, 2007; Salehyan, 2007). The second group of explanations privileges more traditional economic factors, like negative shocks and reduced opportunity costs of rebellion (Miguel, Satyanath and Sergenti, 2004; Ciccone, 2011; Iyer and Topalova, 2014). According to this framework, violence should be most pervasive in areas potentially harmed by trade openness with the EU, austerity and trade barriers with Russia.

The conflict in Eastern Ukraine would appear to be a hard test for economic shocks, which have received far less attention in the public debate than nationalism or identity (see Gentile 2015). Of all international news items on the conflict published in its first year, 55 percent mentioned Russian ethnicity, language or military support, while just 23 percent mentioned any economic issues.<sup>2</sup> If economic conditions can explain variation in a conflict so widely perceived to have been fought along ethnic lines, we should expect to see similar patterns unfold in less unlikely circumstances.

The challenge in empirically testing these competing models of conflict is that, at the regional level, Ukraine's ethnic and economic profiles overlap. The Donets coal basin – or Donbas – is both highly dependent on trade with Russia, and has deep cultural and linguistic ties to Russia. This overlap has made it difficult to parse out the independent effects of language and trade. A regional empirical focus also masks potentially important variation in local conditions. While the Donbas is home to numerous enterprises almost entirely dependent on export to Russia – such as locomotive plants and certain machine-building firms – other local industries, like steel, have been far more competitive on the European market.

This paper confronts these challenges with a disaggregated approach, using new micro-level data on violence, ethnicity and economic activity to explain sub-regional variation in conflict behavior. The event data on

---

<sup>2</sup>Statistics are from a search of all sources in the ProQuest News & Current Event database, with date range 4/1/2014-3/31/2015. The query (donbass OR donbas OR donetsk OR luhansk) yielded 66,729 hits. Sub-searches of these results with keywords (ethnic OR language OR nationalist) and (trade OR industry OR economic) yielded 36,545 (55 percent) and 15,259 results (23 percent), respectively.

violence are based on human-assisted machine coding of incident reports from multiple sources, including Ukrainian and Russian news agencies, government and rebel press releases, daily ‘conflict maps’ released by both sides, and social media news feeds. The data include 10,567 unique violent events in the Donbas, at the municipality level, recorded between President Viktor Yanukovich’s departure in February 2014 and the second Minsk ceasefire agreement of February 2015.

Using Bayesian model averaging, I show that a municipality’s prewar employment mix is a stronger and more robust predictor of rebel activity than local ethnolinguistic composition. Municipalities more exposed to economic shocks – like ‘monocities’ dominated by the machine-building industry – experienced a higher intensity of rebel violence, weekly and overall. They also fell under rebel control earlier – and stayed under their control longer – than municipalities where the labor force was less dependent on exports to Russia.

These findings contribute to research on the relationship between economic conditions and conflict (Collier and Hoeffler, 1998, 2004; Miguel, Satyanath and Sergenti, 2004; Bazzi and Blattman, 2014), and particularly recent work on the effects of economic shocks on the opportunity costs of fighting (Brückner and Ciccone, 2010; Dal Bó and Dal Bó, 2011; Besley and Persson, 2011; Dube and Vargas, 2013).

From a policy standpoint, this evidence offers a corrective to the dominant ethno-nationalist narrative surrounding public debate on the Ukrainian conflict. It is tempting to associate terms like ‘pro-Russian’ and ‘pro-Ukrainian’ with the ethnic and linguistic groups that bear their name. Yet in the Donbas, those rebelling may have ‘pro-Russian’ sympathies for reasons other than ethnicity or language. Depending on the source of income, one’s livelihood may depend on maintaining close trade ties with Russia.

These findings highlight the national security risks of regional economic dependence on a single export market. If the outside country chooses to impose targeted trade restrictions and import substitution, this local de-

pendence can significantly increase the risk of civil conflict.

The paper proceeds as follows. Section 1 briefly reviews recent literature on ethnic conflict and economic shocks. Section 2 presents an overview of events leading up to the war in Eastern Ukraine, and the political economy of the Donbas. Section 3 introduces a new dataset on violence in the Donbas. Section 4 presents the empirical strategy used to test the competing theories of violence, and summarizes the paper's empirical findings. Section 5 offers a discussion of the broader significance of these results, and some concluding remarks.

## 1 IDENTITY, ECONOMICS AND CONFLICT

A growing literature in economics and political science has explored the relative influence of ethnolinguistic and economic conditions on the causes and dynamics of civil conflict (for review, see [Blattman and Miguel 2010](#)).

Rationalist theories of ethnicity expect conflict to be more likely and more intense in areas where ethnic groups are geographically concentrated, for several reasons. The 'exclusion hypothesis' emphasizes the greater motivation for rebellion among ethnic groups subject to systematic discrimination and exclusion, particularly where the excluded group is relatively large or has recently lost political power ([Wimmer, 2002](#); [Buhaug, Cederman and Rod, 2008](#); [Cederman, Wimmer and Min, 2010](#); [Cederman, Weidmann and Gleditsch, 2011](#)). Bargaining models of ethnic conflict emphasize the intrinsic – potentially indivisible – value of territory for ethnic groups ([Toft, 2003](#)), and commitment problems in which ethnic minorities expect their ability to secure a better deal by force to decline in the future ([Fearon, 1998, 2004](#)). Collective action perspectives ([Bates, 1983](#); [Weidmann, 2009](#)) expect territorially-concentrated groups with a shared language to better overcome problems of monitoring, punishment, coordination, and private goods distribution. Others have highlighted external support from trans-border diasporas and co-ethnic regimes as key predictors of armed con-

flict (Saideman and Ayres, 2000; Gleditsch, 2007; Salehyan, 2007; Salehyan, Gleditsch and Cunningham, 2011).

Economic opportunity cost models of political violence maintain that, as real income from less risky legal activities declines relative to income from criminal or rebellious behavior, participation in the illicit activity is expected to rise (Becker, 1968; Tullock, 1971; Grossman, 1991). This view has found empirical support in early cross-national studies on the outbreak of civil war (Collier and Hoeffler, 1998, 2004; Elbadawi and Sambanis, 2002), and in more recent work on exogenous shocks due to international commodity prices (Besley and Persson, 2008; Brückner and Ciccone, 2010; Savun and Cook, 2010; Dube and Vargas, 2013), terms of trade (Iyer and Topalova, 2014), rainfall (Miguel, Satyanath and Sergenti, 2004; Brückner and Ciccone, 2011; Ciccone, 2011; Miguel and Satyanath, 2011), floods (Chaney, 2013), natural disasters (Besley and Persson, 2011), climate change (Koubi et al., 2012; Bergholt and Lujala, 2012) and extreme weather (Maystadt and Ecker, 2014).

Although research has been vibrant in both areas, several gaps remain. First, with some notable exceptions (Buhaug, Cederman and Rod, 2008; Dube and Vargas, 2013; Iyer and Topalova, 2014), most of the literature in each camp – particularly on the economic side – has been cross-national in empirical focus. While macro-level analysis is essential to uncover broad international trends, its dominance in the field has left much local variation unexplained. Leading ‘ethnic’ and ‘economic’ perspectives make claims about the intensity and timing of violence at the micro level, but most of their empirical evidence has resided at a higher level of aggregation.

Second, few if any subnational studies have sought to explicitly evaluate the robustness of the two sets of explanations side by side. Efforts akin to Hegre and Sambanis (2006)’s cross-national sensitivity analysis – which found income levels and economic growth to be stronger predictors of civil war onset than ethnic differences – are relatively rare in disaggregated research. The current study addresses both gaps by conducting an overall

assessment of the competing theories, with municipality-level data rich in variation on all dimensions of interest – violence, ethnicity and economics.

## 2 POLITICAL ECONOMY OF THE DONBAS

Ukraine’s Donbas region – comprising the country’s eastern Donetsk and Luhansk provinces – offers an opportune test case for identity-based and economic theories of conflict. Due to its ethnic geography and employment mix, the Donbas is potentially exposed to both types of challenges. The Donbas is also a big place, with 6.45 million prewar residents – a sixth of Ukraine’s population – and 3,037 municipalities. These towns and villages vary significantly in their local linguistic and economic conditions, permitting a comparative empirical analysis at a disaggregated level. The current section presents historical context for this analysis.

### 2.1 THE RUSSIAN SPRING

When angry crowds took to the squares in Donetsk and Luhansk in March 2014, waving Russian flags and chanting ‘RO-SSI-YA!’, the dominant explanation of these events in the West was Russian nationalism. Indeed, Russian nationalist groups – local and foreign – had a visible presence at the protests.<sup>3</sup> Many in the square cited the post-revolutionary Ukrainian government’s abortive attempt to restrict Russian language rights as a threat to their way of life. Emboldened by Crimea’s annexation, many protesters were hopeful that Moscow would recognize their own referendum on independence, and even incorporate the region into the Russian Federation.

The Donbas is home to a large concentration of ethnic Russians. In the

---

<sup>3</sup>Although popular support for secession had been low prior to the war, several dozen pro-Russian separatist groups had established themselves in the Donbas since the Soviet collapse, including the Donbass People’s Militia, Donbass Russian Union, Monolit, National Unity, Oplot, Russian Block, Russkaya Obshchina, Youth Guard and various Cossack groups.

most recent census, ethnic Russians represented 38.2 of the population in Donetsk province, the highest share anywhere but Crimea. A much higher proportion of the region's population – 74.9 percent – claimed Russian as a native language, including 58.7 percent of ethnic Ukrainians ([State Committee on Statistics of Ukraine, 2001](#)). One of these Russian-speaking Ukrainians was the country's deposed President, Viktor Yanukovich, a former governor of Donetsk.

During the 2010 elections – in which Yanukovich won 88 to 90 percent of the region's votes ([Central Election Commission of Ukraine, 2010](#)) – the status of the Russian language became a major campaign issue. Prior to Ukraine's independence, Russian enjoyed official language status country-wide, along with Ukrainian. After the Soviet collapse, Ukrainian became the country's sole official language. Although Russian remained widely spoken, its diminished status occasionally caused difficulties in courts, hospitals and schools. While in office, Yanukovich signed a new law on regional languages, giving official status to any language spoken by at least 10 percent of a local population.

Yanukovich's language bill was popular in the Donbas, where jurisdictions rushed to declare Russian an official regional language. Elsewhere in Ukraine, the bill was highly controversial. It had been rushed through the Verkhovna Rada – Ukraine's parliament – amid fistfights and public protests. Opponents of the law claimed it was unconstitutional and undermined the status of the Ukrainian language. On 23 February 2014, the same day Yanukovich was overthrown in a popular revolution, the Verkhovna Rada passed a bill abolishing the language law, making Ukrainian the sole state language at all levels.

The anti-Russian language bill never became a law, and was blocked by Ukraine's interim authorities. But it received wide coverage in the Russian press and sparked protests in Ukraine's south and east. The bill – along with the prominent role of Ukrainian nationalist groups in the protest movement that unseated Yanukovich – gave local activists little



confidence that Kyiv's new authorities would respect minority rights in the future. At protest rallies in the east, 'fascist junta' became shorthand for Ukraine's post-revolutionary regime.

According to 'identity-based' theories of conflict, variation in rebel violence should reflect the prewar ethnic geography of the Donbas. The Donbas has a large Russian population. The post-revolutionary shift in relative power between Ukraine's eastern and western elites created commitment problems on language policy, and fears of cultural discrimination. These fears attracted co-ethnic nationalists from neighboring Russia, who were eager to defend their ethnic brethren from the specter of Ukrainian 'Nazis.' The pre-existing network of Russian nationalist groups in the Donbas helped the rebels solve many of the startup collective action problems associated with mounting a rebellion.

One problem with the ethnic story is that the Donbas conflict has not been fought primarily along ethnic lines. Russian has been the language of command in many units of the Ukrainian armed forces, and even in some ultra-nationalist Ukrainian volunteer battalions like 'Azov.' The Facebook page of Ukraine's Interior Minister Arsen Avakov – where he issues official press releases on operations by the National Guard – is in Russian. Many senior Ukrainian officials, including Prime Minister Arseniy Yatsenyuk, speak Russian at home (Popeski, 2014). The 'head of state' of the so-called Donetsk People's Republic, Aleksandr Zakharchenko, is an ethnic Ukrainian.

## 2.2 THE RUST BELT

As Blattman and Miguel (2010, 27) note, 'the finding that many civil conflicts are fought partially along ethnic lines alone is insufficient to [show] that ethnic-based grievances are driving the fighting.' An alternative explanation is that, amid Ukraine's revolutionary turmoil, the Donbas became uniquely exposed to several potential sources of negative economic shock:

Ukraine's Association Agreement with the EU, fiscal austerity, and Russian trade restrictions.

Ukraine's current crisis began on November 22, 2013, when Yanukovich decided not to sign an Association Agreement and Deep and Comprehensive Free Trade Area (DCFTA) agreement with the EU, and opted instead for a Russian proposal to join a Customs Union with Belarus and Kazakhstan. This decision triggered the 'Euromaidan' mass protest movement in Kyiv, which ultimately led to Yanukovich's ouster. Ukraine's interim authorities signed the political part of the EU agreement in March 2014, and the country's newly-elected President, Petro Poroshenko, signed the economic portion in May 2014.

The DCFTA gives Ukraine preferential access to the EU internal market, eventually setting 95 percent of tariff lines at zero, while committing Ukraine to a series of reforms aligning its technical norms and standards with those of the EU ([Delegation of the European Union to Ukraine, 2013](#)). The Customs Union, Russia's alternative, had immediate monetary benefits in the form of reduced gas prices and a 15 billion USD loan, but would also have transferred Ukraine's ability to set tariffs and trade policy to a supranational body dominated by Moscow, excluding the possibility of future tariff dismantlement with the EU. Each agreement would have created winners and losers in Ukraine, but the EU deal would disproportionately harm certain industries in the Donbas.

The Donbas has historically played a significant role in Ukraine's economy, representing a sixth of its gross-domestic product in 2013, 27.3 percent of industrial production, and 23.3 percent of exports – surpassing every other province in these categories ([State Statistics Service of Ukraine, 2015](#); [Koluntovych, 2014](#)). Much of this economic power is a legacy of the Soviet era, when the Donbas served as the industrial heartland of the world's second superpower, supplying it with coal, metals and heavy machinery. When Ukraine became independent in 1991, the formerly domestic Soviet market became a foreign Russian one, and the region's economy became

export-oriented. As of 2014, some 80 percent of production in the Donbas went to export, with Russia as the leading destination ([Makogon, 2014](#)).

Although Russia has traditionally been Ukraine's main trading partner, the collective share of trade claimed by EU member states has been steadily growing – accounting for 27 percent of exports and 37.1 percent of imports in 2013, compared with 34.8 and 36.3 percent for Russia and the Commonwealth of Independent States ([State Committee on Statistics of Ukraine, 2013](#)). Ukraine's trade with the EU has been dominated by exports of ferrous metals and cereals, and imports of European machinery, transportation equipment, clothing and consumer goods. Ukraine's main exports to Russia have included heavy machinery, metals and agriculture. Russian imports are overwhelmingly oil and gas ([Ministry of Economic Development of the Russian Federation, 2014](#)).

Producers in the Donbas have long worried that reduced tariffs with the EU would result in a flood of imports and unfair competition. The adoption of EU production standards carried significant short-term costs, while many enterprises that had received generous government subsidies and were shielded from competition under Yanukovich – and his predecessors – were about to lose these privileges.

A second, and potentially more devastating, effect of the EU agreement was Russia's response. During final negotiations on the agreement in July 2014, Moscow implemented a series of retaliatory trade measures, including stricter customs procedures and a partial export blockade. In September 2013, President Putin's chief economic adviser Sergey Glazyev explicitly warned of heavier trade sanctions if Ukraine signed the EU agreement. He also suggested that separatist movements could emerge in the industrial east and south of Ukraine, and that Russia would consider the bilateral treaty delineating the countries' borders to be void ([Walker, 2013](#)).

The 'big three' industries of the Donbas – metals, coal and machine-building – are differentially vulnerable to trade disruptions with Russia. Relatively efficient and competitive, metallurgy accounts for three quarters

of industrial output in the Donbas, with exports to 50 countries ([Anushevskaya, 2014](#)). Although metals represented 23.4 percent of Ukraine's exports to Russia in 2014 ([Ministry of Economic Development of the Russian Federation, 2014](#)), loss of the Russian market would not be catastrophic. Along with the agricultural sector, metals producers rank among the main beneficiaries of increased market access to the EU ([Delegation of the European Union to Ukraine, 2013](#)).

Coal mining, by contrast, has been a dying industry. Its dependence on Russia is indirect – a tenth of Donbas coal production is exported to Russia, and another third goes to Ukrainian factories dependent on Russian orders ([Luhn, 2014](#)). The coal industry's larger problem lies in its inefficiency and reliance on state subsidies. In response to a rapid succession of closures – 137 mines have shut down since 1996, leaving thousands out of work – Yanukovich kept mines in operation through direct subsidies. As a result, most mining enterprises operated at a loss, with many state-owned mines producing nothing, maintaining a staff only to keep the mine in working order, prevent flooding and explosions. As a condition of IMF loans received by Ukraine in 2014, Kyiv's new authorities agreed to an austerity program that would cut subsidies and close 50 to 70 mines – potentially leaving tens of thousands unemployed ([Guillory, 2014](#)).

The Donbas industry most exposed to Russian trade shocks is machine-building. In 2014, machinery represented 33.4 percent of exports to Russia – a higher share than any other industry ([Ministry of Economic Development of the Russian Federation, 2014](#)). 45.8 percent of Ukraine's machine-building is concentrated in the Donbas, specializing in capital goods like mining equipment, steel furnaces, industrial cranes, rolling mills and locomotives. 60 percent of this production is exported to Russia, with some enterprises – particularly in the locomotive industry – reaching 90 percent dependence on Russian orders ([Tarnavskiy, 2014](#)).

Uniquely among local industries, Donbas machine-builders have been unable to find an alternative market for their goods. Their production is

not competitive in the EU or Asia – it is expensive and highly specialized for the needs of Russian industrial clients. Domestic demand is virtually non-existent, given fiscal constraints in Kyiv and reduced tariffs on cheaper and higher-quality European machines.

Adding to the pain, machine-building has been the primary target of Russia's long-term strategy of import substitution. In 2013, Russia stripped several Ukrainian manufacturers of licenses to supply goods to the Russian market, while opening and expanding several locomotive plants and military-industrial facilities in the Urals. Facing increased competition and reduced demand for its wagons and engines, Ukrainian locomotive production fell by 47 percent in 2013 ([Tarnavskiy, 2014](#)).

Russian import substitution intensified in 2014, as a response to American and EU sanctions over Crimea. In December 2014, Russia's Minister of Industry and Trade Denis Mansurov estimated that Russia will be able to fully replace Ukrainian machinery imports with domestic production in three years ([TASS, 2014](#)).

According to economic explanations of the Donbas conflict, variation in rebel violence should reflect variation in the vulnerability of the local labor force to negative shocks. In the events surrounding Yanukovich's departure, the Donbas experienced a perfect economic storm, including a free trade deal with the EU, IMF-imposed austerity and a 'trade war' with Russia. These shocks threatened to put thousands out of work, particularly in machine-building 'company towns' with a single large enterprise and one primary customer. Left with few alternative lawful sources of income, local industrial workers saw the opportunity costs of rebellion rapidly decline.

Anecdotal evidence appears to support the opportunity cost narrative. Miners had a visible presence during March-April 2014 protests in Donetsk, helping to storm government buildings and build barricades ([Luhn, 2014](#)). After the Kyiv government stopped paying salaries of workers at state-owned mines under rebel control, much of the mining labor force reportedly dissolved into the ranks of the so-called Donetsk and Luhansk Peo-

ples' Republics (Guillory, 2014).

The Donbas has a long history of political activism at times of economic shock, with local unrest becoming a common occurrence after the Soviet collapse. Facing wage arrears, factory closings and declining terms of trade, the region's large industrial labor force – led by coal miners – mounted a series of major strikes in 1993, 1996 and 2007. The protesters' economic demands included increased government subsidies, (somewhat paradoxically) financial independence from Kyiv, and more barriers to foreign competition (Siegelbaum and Walkowitz, 1995).

In the Donbas, political demands during hard economic times have usually centered on regional autonomy. In a ballot initiative held during the 1994 presidential election, between 79 and 91 percent of the region's voters supported the federalization of Ukraine and economic union with Russia (Morin, 2014). Twenty years later, in March 2014, protesters again took to the streets of Donetsk and Luhansk. They began with an initially similar set of demands, calling for economic and political autonomy. On May 11, the region's new rebel authorities held a popular referendum, and declared independence from Ukraine.

The preceding narrative suggests that 'pro-Russian' forces in the Donbas may be 'pro-Russian' not due to ethnicity or language, but because their livelihood and prosperity depend on maintaining close trade ties with Russia. Whether this type of economic dependence better explains patterns of rebellion than ethnolinguistic geography is an open empirical question.

### 3 DATA

To test the relative explanatory power of identity-based and economic models of conflict, I use a new micro-level dataset on violence, ethnicity and economic conditions in the Donbas. The goal is to explain geographic and temporal variation in rebel violence and territorial control. The spatial units are 3,037 municipalities (i.e. cities, towns, villages) in Donetsk and

Luhansk provinces. The temporal domain ranges from February 28, 2014 – the week after Yanukovich’s departure – to February 22, 2015 – the week following the second Minsk ceasefire agreement.

### 3.1 REBEL VIOLENCE

The violent event data are based on human-assisted machine coding of news reports, press releases and blog posts from Ukrainian, Russian, rebel and international sources.<sup>4</sup> For each data source, I created a separate electronic text corpus that contained all incident reports published on the Donbas since February 2014 ( $N = 53,754$ ).

To determine the geographic locations of events mentioned in the reports, I ran an automated geocoding script that identified populated place names referenced in the text, and matched them against the U.S. National Geospatial Intelligence Agency’s GeoNames database.<sup>5</sup> Figure 1 shows the resulting spatial distribution of events, by source.

To determine the content of the incident reports, I used a supervised learning algorithm – Support Vector Machine – to classify each event into a series of pre-defined categories. These categories include event type, initiator, target, tactic, and casualties. The codebook, with complete list of categories and examples, is included in the online appendix.

The event of primary interest is a *rebel attack*. For a report to be classified as a rebel attack, it must involve a specific act of organized violence initi-

---

<sup>4</sup>Ukrainian sources include Channel 5, Espresso.tv, Information Resistance, and the newswire services Interfax-Ukraine and Ukrinform. Russian sources include the non-government news websites Gazeta.ru, Lenta.ru and BFM.ru, and the Interfax newswire service. Pro-rebel sources include Rusvesna.su. Also included are the Russian-language edition of Wikipedia, and daily briefings from the OSCE Special Monitoring Mission to Ukraine.

<sup>5</sup>I used a one-to-many mapping algorithm, to account for multiple events mentioned in the same report. To identify and correct geocoding errors and double-counts, each list of geocoded locations was referenced against a lookup table of regular errors (e.g. to ensure that ‘Donetsk oblast’ isn’t mis-coded as ‘Donetsk city,’ and that references to the ‘Shakhtar battalion’ are not mis-coded as ‘Shakhtarsk city’). I also performed manual inspection.

ated by any anti-Kyiv armed group.<sup>6</sup>

For each dataset shown in Figure 1, I and a team of research assistants read a randomly-selected training set of 130-600 reports (depending on the size of the corpus), in Russian, Ukrainian and/or English.<sup>7</sup> To account for potential disagreement between coders, at least two sets of eyes read each training set document, including the author and another member of the research team. Inter-coder reliability statistics, reported in the technical appendix, indicate a high and statistically significant level of agreement between coders on the relevant categories, including where coders read the same documents in different languages.

I used the randomly-selected reference texts in each training set to train a Support Vector Machine (SVM) classifier to predict the categories for all previously unseen corpus texts.<sup>8</sup> I created a separate document-term ma-

---

<sup>6</sup>A specific act of violence is a reference to a single ongoing or recent military operation, act of terrorism, targeted killing, detention, other violent event. Not included in this category are general summaries of war statistics or press statements. Anti-Kyiv groups include any forces explicitly labeled as ‘insurgents,’ ‘rebels,’ ‘terrorists,’ as well as specific formations like the Novorossiia Armed Forces, Donetsk People’s Republic (DNR), Lugansk People’s Republic (LNR), Vostok Battalion, Oplot, Kal’mus battalion, Bezler band, Zarya battalion, Russian Orthodox Army (RPA), People’s Militia of Donbass (NOD), Prizrak battalion, Army of the South East, Don Cossacks, Russian National Unity, Eurasian Youth Union, Yovan Sevic. References to actions by the Russian Armed Forces (mostly in Ukrainian media) are also labeled ‘rebel.’

<sup>7</sup>The author and one research assistant read all training set documents in their original languages. Research assistants not fluent in Russian or Ukrainian read training sets containing the same reports, machine-translated into English.

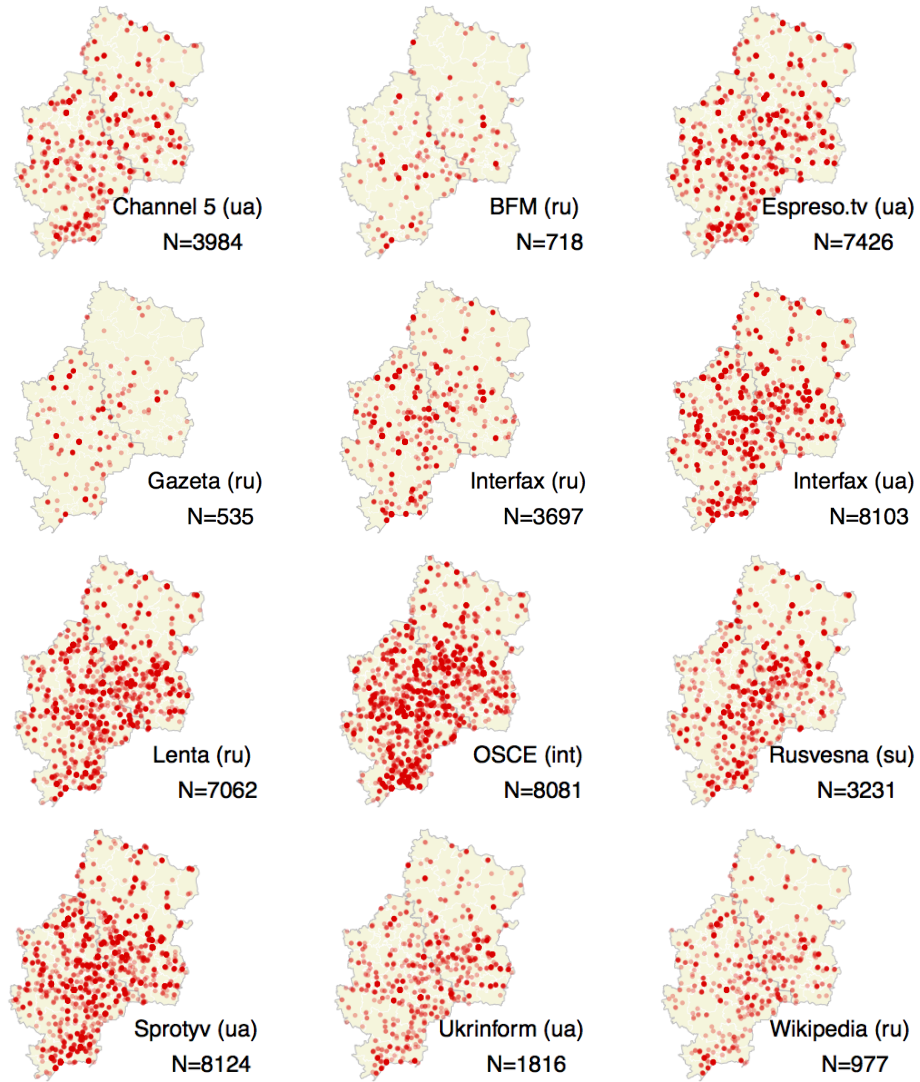
<sup>8</sup>The SVM classifies documents by fitting a maximally-separating hyperplane to a feature space, examining combinations of features that best yield separable categories. Formally, the SVM separates data points from each other according to their labels ( $y_{it} \in \{-1, 1\}$ ), and finds maximum marginal distance  $\Delta$  between the points labeled  $y_{it} = 1$  and  $y_{it} = -1$ , solving the optimization problem

$$\arg \max_{\Delta, \alpha, \phi} \Delta \text{ s.t. } y_{it}(\alpha + \phi(X_{it})) > \Delta$$

where  $y_{it}(\alpha + \phi(X_{it}))\beta$  is a functional margin,  $\phi(\cdot)$  is a function that maps the training data  $X$  to a high-dimensional space, and  $\mathbf{K}(x_i, x_j) = \phi(x_i)' \phi(x_j)$  is a kernel function. The advantage of the SVM is that it is well-suited to sparse, high-dimensional data, is highly robust, and can handle a low training-to-test data ratio.



Figure 1: EVENT LOCATIONS, BY DATA SOURCE.



trix for each corpus, and ran the SVM classifier separately for each.<sup>9</sup>

To pool the data across the sources shown in Figure 1, I used a one-a-day filter for each municipality-day.<sup>10</sup> This filter produced 10,567 unique violent events. Because the one-a-day filter implies a maximum of one event per municipality on a given day, I aggregated the atomic-level events to the level of municipality-week and municipality-year. This temporal aggregation permits a measure of intensity, ranging from 0 to 7 on the weekly level, and 0 to 361 on the annual level. Figure 2a visualizes the resulting overall intensity of violence across the 3,037 populated places of the Donbas.

### 3.2 TERRITORIAL CONTROL

The second outcome variable of interest is territorial control, particularly whether a populated place was under rebel or government control on a given day. For this information, I draw on three sets of sources. First are official daily situation maps publicly released by Ukraine’s National Security and Defense Council (RNBO). Second are daily maps assembled by the pro-rebel bloggers ‘dragon\_first\_1’ and ‘kot\_ivanov.’ Third are Face-

---

<sup>9</sup>In the document-term matrix, the rows are documents  $d \in \{1, \dots, D\}$ , columns are terms  $t \in \{1, \dots, T\}$ , cell entries are weighted term frequencies, and each row vector  $\mathbf{y}_d \in \mathbb{R}^T$  represents document  $d$  in a  $T$ -dimensional feature space. Features were weighted by term frequency - inverse document frequency,

$$tf.idf_{dt} = tf_{dt} \log \left( \frac{D}{df_t} \right)$$

where  $tf_{dt}$  is term frequency (number of times term appears in  $d$ ), and  $df_t$  is document frequency (# documents with term  $t$ ). A high  $tf.idf_{dt}$  weight indicates that a term appears a lot in document  $d$ , but rarely in the corpus.

In the preprocessing stage, I removed HTML tags, control characters, non-alphanumeric characters, capitalization, punctuation and stopwords for all corpora, but ran a stemming algorithm only on English-language texts, so as to preserve inflections in Ukrainian and Russian (i.e. tense, voice, aspect, person, number, gender and case) – which contain important information for differentiating between initiators and targets.

<sup>10</sup>For each of 3,037 unique populated places in Donetsk, Luhansk oblasts, on each day between February 28, 2014 and February 22, 2015, I coded a rebel attack as occurring if at least one of the twelve SVM-classified datasets reported it as occurring. This technique, common in event data research, is designed to eliminate double-counts.

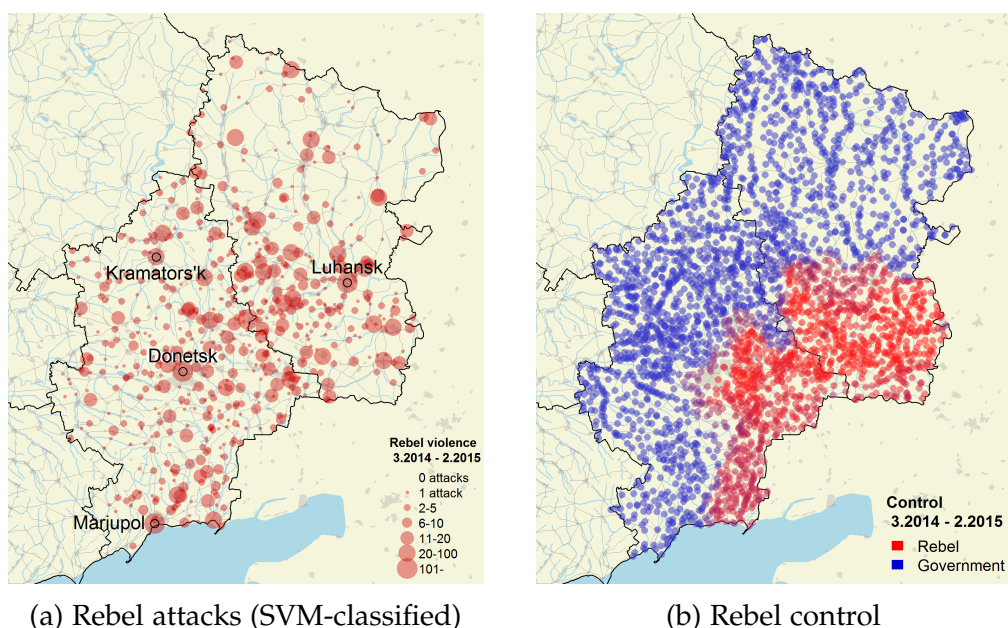


Figure 2: OUTCOME VARIABLES.

book posts on rebel checkpoint locations prior to June 18, 2014 – the earliest date for RNBO and pro-rebel maps. For RNBO and pro-rebel blog maps – released to the public as high-resolution image files – my research team georeferenced and vectorized each map into spatial polygons. To construct polygons from user-reported checkpoint locations, I used the geographic convex hull of the coordinates of observed checkpoints on each day.

I coded a municipality as being under *rebel control* if, on a given day, it fell inside the rebel control polygons from one of the map collections.<sup>11</sup> Figure 2b displays the distribution of rebel control over the full period of observation. The points are colored according to the proportion of time each municipality spent under rebel control since March 2014, with bright red indicating that a municipality spent almost the full period under rebel

<sup>11</sup>I created separate daily indicators for each of the three collections, and two combined indicators – rebel control according to at least one map collection, and control according to both RNBO and the bloggers. Prior to June 18, 2014, the two measure are equal, as there is only one source (Facebook rebel checkpoints.)

control, and blue indicating that a municipality was under government control for most of the period. Purple shades indicate that a municipality spent a significant duration of time under the control of each actor.

### 3.3 EXPLANATORY VARIABLES

The primary source for data on local languages is the 2001 Ukrainian Census ([State Committee on Statistics of Ukraine, 2001](#)). For each municipality, I created a measure of *proportion Russian-speaking* (Figure 3a).<sup>12</sup> For information on the local employment mix, I used Bureau van Dijk's Orbis database ([Bureau van Dijk Electronic Publishing, 2015](#)). The database includes records for 445,399 private and publicly-owned firms in Donetsk and Luhansk provinces, with names, addresses, industry and employment information. For each municipality, I calculated the *proportion of the local labor force employed in machine-building, mining and metals industries* (Figure 3b).<sup>13</sup> In the empirical analysis, I consider additive combinations of employment and language, as well as their interaction.

In addition to these covariates of primary theoretical interest, I included a series of other predictors common in subnational conflict research. These include characteristics of the local military geography, like population density ([CIESIN and Columbia University, 2005](#)), elevation ([U.S. Geological Survey, 1996](#)), forest cover ([Loveland et al., 2000](#)), distance to the nearest road ([Defense Mapping Agency, 1992](#)), and distance to the Russian border ([Global Administrative Areas, 2012](#)). The first of these is a proxy for the number of potential military targets in a municipality. The next three capture the mobility of armed forces, as well as – in the case of elevation and roads – the strategic value of a municipality for artillery firing positions

---

<sup>12</sup>The census uses respondents' self-reported 'native language' to measure this variable.

<sup>13</sup>'Local labor force' is defined as total employment at the 1 percent of companies geographically closest to a given municipality. Industry designations of individual firms (machine-building, mining and metals) are based on European industrial activity classification (NACE) codes and descriptions.

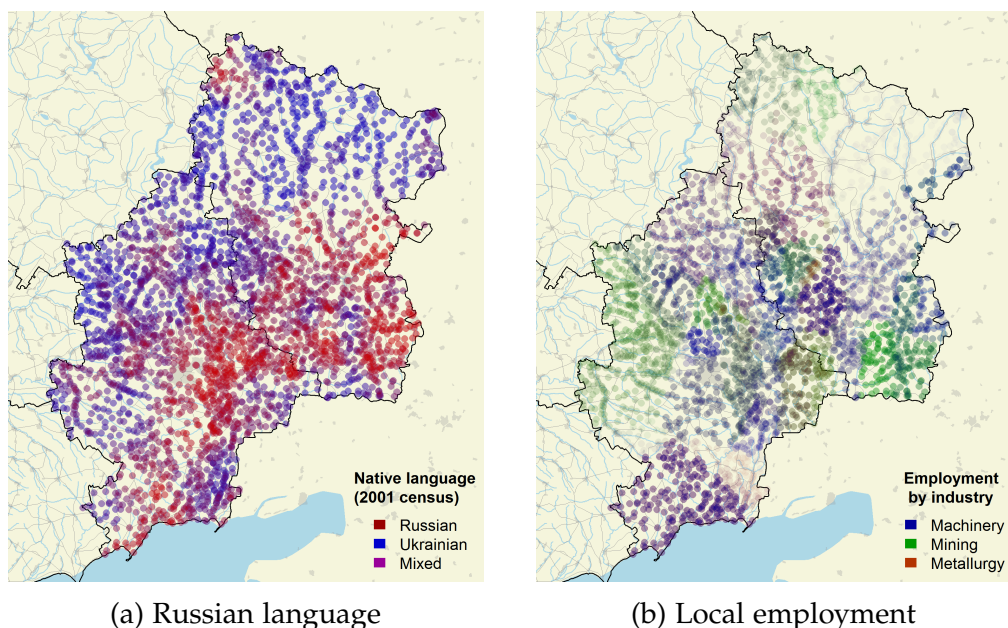


Figure 3: EXPLANATORY VARIABLES.

and logistics. The fifth – distance from the Russian border – is a proxy for the availability of Russian military support for the rebels. If military geography drives variation in fighting (Zhukov, 2012), one should expect rebel violence and control to be greatest in areas of high population density, elevation and forest cover, and proximate to roads and the Russian border.

I also include data on the prewar political loyalties of the local population. For each municipality, I recorded the percent of the popular vote received by Viktor Yanukovych in the 2010 presidential election (Central Election Commission of Ukraine, 2010). If prewar political loyalties are a driving force in the fighting (Balcells, 2011), one should expect violence and control to be greatest where Yanukovych had received more support.

Finally, I control for the persistence and spatial spillover of violence. In municipality-week level data, I include temporal lags of rebel violence during the previous week, and a spatio-temporal lag of the average number of

violent rebel incidents in neighboring municipalities in the past week.<sup>14</sup>

## 4 EMPIRICAL ANALYSIS

To evaluate the relative explanatory power of ethnic and economic explanations of violence in the Donbas, I use Bayesian Model Averaging (BMA). BMA accounts for uncertainty over the ‘true’ theory of violence by computing posterior probabilities over models with all possible combinations of theoretically-relevant covariates, and constructing a weighted average over the most likely models.<sup>15</sup>

In the current application, an ensemble approach like BMA is preferable to alternative empirical strategies for several reasons. First, results in empirical research on conflict are notoriously fragile to slight changes in specification (Hegre and Sambanis, 2006). BMA permits an overall assess-

---

<sup>14</sup>I define ‘neighboring municipalities’ with with a row-standardized spatial weights matrix  $W$ , based on a Delaunay triangulation of municipality coordinates. Matrix cell entries  $w_{ij}$  take nonzero values if the Delaunay triangles of two municipalities share a common border.

<sup>15</sup>Formally, my quantities of interest are model-weighted posterior distributions for the coefficients

$$P(\beta|y, X) = \sum_m^{2^K} P(\beta|M_m, y, X)P(M_m|y, X)$$

where  $M_m$  is the  $m$ -th model (e.g.  $y = g^{-1}(\alpha_m + X'_m\beta_m + \epsilon)$ ), and the model weights  $P(M_m|y, X)$  are based on posterior model probabilities

$$P(M_m|y, X) = \frac{P(y|M_m, X)P(M_m)}{P(y|X)} = \frac{P(y|M_m, X)P(M_m)}{\sum_s^{2^K} P(y|M_s, X)P(M_s)}$$

where  $P(y|M_m, X)$  is the marginal likelihood of model  $M_m$ , and  $P(M_m)$  is the prior model probability. I use a uniform distribution for model priors ( $1/2^K$ ), and BIC approximation for inclusion probability priors

$$P(\beta_1 \neq 0|y, X) = \sum_{j:\beta_1 \in M_j} \frac{\exp(BIC_j/2)}{\sum_i^K \exp(BIC_i/2)}$$

ment of performance across the full model space, without assuming that any one model is *a priori* preferable to another (Madigan and Raftery, 1994; Hoeting et al., 1999). This feature is particularly appealing in situations, like the current one, where the empirical validity of one theory need not imply that an alternative is incorrect.

Second, by calculating posterior distributions of coefficients and models, BMA enables one to assess not only the direction of an empirical relationship, but also whether a given variable consistently contributes to models' explanatory power (Montgomery and Nyhan, 2010). If a municipality's prewar employment mix positively correlates with violence, but models that include the variable have a low cumulative posterior probability of being correct, the averaging procedure will assign that variable a low weight.

Third, it may well be the case that neither ethnicity nor economics contribute much explanatory power, while military geographic factors like roads or distance to the border do most of the heavy lifting. To that end, BMA is a transparent model selection procedure, which guards against cherry picking models in which one's preferred coefficient is 'significant.'

Sensitivity to model specification is particularly salient given the ongoing nature of the conflict under investigation, the partisan nature of the public debate surrounding it, and the potentially high costs of misdiagnosing its origins. In addition to incorporating uncertainty over the 'best' specification, BMA produces results that are more accurate than any individual model under consideration (Madigan and Raftery, 1994). For this reason, BMA is increasingly used in policy-oriented research on conflict forecasting and early warning (Montgomery, Hollenbach and Ward, 2012; Ulfelder, 2012).

I estimate four BMA ensembles of models: two on the determinants of rebel violence and two on territorial control. First, I use an ensemble of logit models to explain the *incidence of any rebel violence* across municipal-

ities during the first year of the conflict.<sup>16</sup> Second, I model the *intensity of rebel violence* in a municipality-week, using an ensemble of spatiotemporal autoregressive GLMs with quasi-Poisson links.<sup>17</sup> Third, I model the *duration until a municipality falls under rebel control*, using Cox proportional hazards (CPH) models.<sup>18</sup> Finally, I consider the *duration until the loss of rebel control* to pro-government forces.<sup>19</sup>

Because economics and language may have an interactive effect on rebel behavior (i.e. if the ‘machine-building effect’ is stronger or weaker in Russian-majority municipalities), I include multiplicative interactive terms of the three employment mix variables with Russian language in all ensembles. To correct for the presence of interactive terms, I avoid averaging over models that exclude one or more of the constituent, lower-order terms (Montgomery and Nyhan, 2010).

#### 4.1 DETERMINANTS OF REBEL VIOLENCE

As war gripped Eastern Ukraine in 2014-2015, why were some municipalities overwhelmed by bloodshed, while others were completely spared? Of the 3,037 populated places in the Donbas, only 619 experienced any rebel

---

<sup>16</sup>The core model specification is  $y_i = g^{-1}(\alpha + X_i'\beta + \epsilon)$ , where  $g^{-1}(\cdot)$  is the inverse logit link, and  $X_i$  is a matrix of the covariates described in section 3.3. The level of analysis is municipality-year.

<sup>17</sup>The quasi-Poisson model – an extension of the Poisson GLM that leaves the dispersion parameter unrestricted – is appropriate because the outcome variable (number of rebel attacks in a municipality-week) is an overdispersed event count. The core model specification is  $y_{it} = g^{-1}(y_{i,t-1}\gamma + \mathbf{W}y_{i,t-1} + X_i'\beta + \epsilon)$ , where  $g^{-1}(\cdot)$  is an inverse quasi-Poisson link,  $y_{i,t-1}$  is a one-week time lag of the outcome,  $\mathbf{W}y_{i,t-1}$  is a spatial lag of the time lag, and  $X_i$  is a matrix of time-invariant covariates, which includes all predictors described in section 3.3.

<sup>18</sup>Duration is measured in days, with February 28, 2014 as the start date. The quantity of interest is the hazard ratio  $h(t|X_i) = h_0(t) \exp(X_i'\beta)$ , where  $X_i$  is a matrix containing various combinations of the covariates described in section 3.3.

<sup>19</sup>The design is identical to the previous ensemble of CPH models, with the rebel-organized independence referendum of May 11, 2014 as the start date, and liberation by Kyiv-aligned forces as the event. The sample of municipalities includes only the subset under rebel control ( $N = 1910$ ).



violence. Although the remainder doubtless felt the impact of the conflict in other ways, they somehow managed to avoid direct physical destruction. Figure 4a reports the results of a set of BMA analyses, examining the incidence of *any* rebel violence in a municipality between March 2014 and late February 2015.

The evidence in Figure 4a suggests that vulnerability to trade shocks is among the strongest predictors of whether a municipality experienced rebel violence or remained at peace. In 5120 model specifications considered by BMA, only three variables consistently predicted the incidence of rebel violence, with posterior inclusion probabilities greater than .97. These include the proportion of the local labor force employed in machine-building, and military-geographic variables like distance to road and population density. Neither the linguistic composition of a municipality, nor any interaction between language and economics explains the overall occurrence of violence.

All else equal, municipalities where machine-building represented a small share of local employment (2 percent, the lowest in the data) were 37.84 percent less likely to experience violence than municipalities where the industry was more dominant (45 percent, the highest in the data), with a 95 percent credible interval of (14.32, 56.79).<sup>20</sup> A standard deviation shift around the mean – from 4 to 26 percent of the local population employed in machine-building – yields an 18.22 percent increase (95% CI: 7.22, 27.57) in risk of rebel violence.<sup>21</sup>

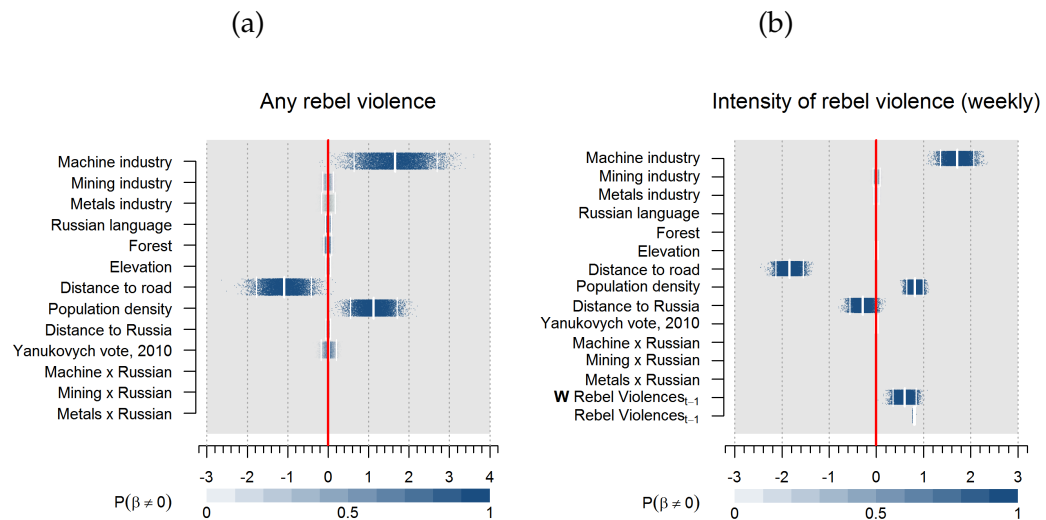
What explains the intensity of this violence from week to week? The first year of the conflict saw the emergence of several ‘hot spots’ of violence – like Donetsk, Shchastya, Pisky and Alchevsk – where rebels have engaged in almost daily ambushes, firefights and artillery duels to either harass the Ukrainian army and other pro-Kyiv forces, or to push them out of con-

---

<sup>20</sup>The mean posterior probabilities of rebel violence are 0.47 (95% CI: 0.42, 0.52) and 0.65 (95% CI: 0.53, 0.76), respectively. All other variables are held constant at median values.

<sup>21</sup>The mean posterior probabilities of rebel violence are 0.48 (95% CI: 0.43, 0.53) and 0.57 (95% CI: 0.49, 0.65), respectively.

Figure 4: DETERMINANTS OF REBEL VIOLENCE. Quantities reported are  $P(\beta|\beta \neq 0, y, X)$ . Each blue point on the plot is a single draw from the posterior distribution of the coefficient associated with each variable, with vertical white stripes denoting means and 95% credible intervals. The opacity of the points reflects posterior inclusion probabilities, with full transparency indicating  $P(\beta \neq 0|y, X) = 0$  and full opacity indicating  $P(\beta \neq 0|y, X) = 1$ .



tested territory. Outside of this ‘top 1 percent’, however, rebel violence has been relatively infrequent. The average Donbas municipality experienced less than one violent day per week. What drives this variation?

Figure 4b shows the conditional posterior distribution of coefficients from 20,480 quasi-Poisson models, weighted by their posterior inclusion probabilities. As in the previous BMA ensemble, the proportion of local population employed in machine-building is among the strongest predictors of rebel violence, with a posterior inclusion probability of 1. The share of a municipality’s Russian-speaking population had no discernible impact on the the intensity of rebel activity, with a posterior inclusion probability of less than 0.003.

For a median Donbas municipality, an increase in the machine-building labor force from one standard deviation below (4 percent) to one standard deviation above the mean (26 percent) yields a 44 percent increase (95% CI: 33.51, 55.83) in the frequency of rebel violence.<sup>22</sup> Increasing the machine-building employment share from its minimum (0.02) to its maximum observed value (0.45) yields a more than two-fold rise in the expected number of violent rebel events per week – from 0.72 (95% CI: 0.65, 0.79) to 1.53 (95% CI: 1.27, 1.82). In a hypothetical machine-building ‘mono-city’ – where everyone is employed in the industry – rebel violence would occur 3.9 days per week on average (95% CI: 2.74, 5.59).

There is little evidence of either a Russian language effect on violence, or an interaction between language and economics. The impact of prewar industrial employment on rebellion is the same in majority Russian speaking municipalities as it is in majority Ukrainian-speaking ones.

The remaining drivers of rebel violence, Figure 4b suggests, are largely military-geographic. There has been considerably more violence in localities accessible by road, and in towns with high population density. The

---

<sup>22</sup>The means of the corresponding posterior predictive distributions of rebel violence per week are 0.76 (95% CI: 0.68, 0.84) and 1.09 (95% CI: 0.95,1.24), respectively. All other variables are held constant at median values.

intensity of violence is also far higher in areas exposed to conflict during the previous week – within the same locality, and in neighboring areas.

## 4.2 DETERMINANTS OF REBEL CONTROL

What types of municipalities were likely to fall under rebel control early in the conflict? The establishment of territorial control is a far more challenging enterprise than an isolated skirmish. It requires a sequence of operations, in which rebels seize police stations and other government buildings, clear the town of pro-government forces, set up checkpoints and establish nascent governing structures. These tasks are difficult to accomplish without a local tactical advantage and some public acquiescence.

Figure 5a shows the conditional posterior distributions of hazard ratios, weighted by the cumulative posterior inclusion probability of models containing each variable. Hazard ratios higher than 1 (right of the red line) indicate that a variable increased the risk of rebel control, shortening the duration before forces aligned with the so-called Donetsk and Luhansk People’s Republics seized a municipality. Ratios below 1 indicate that a variable reduced the hazard of rebel control, lengthening the duration.<sup>23</sup>

Figure 5a shows that the determinants of territorial control are considerably more diverse than those of violence. Both economic and linguistic factors influence the hazard of rebel control, and they interact in telling ways. The hazard ratios for all higher-order interaction terms are negative, indicating that the two sets of risk factors act as substitutes for one another, rather than as amplifiers. The ‘dulling effect’ of industrial employment, however, generally outweighed that of Russian language.

As reported in Table 1, prewar industrial employment was strongly predictive of the timing of rebel control, in both minority- and majority Russian-speaking municipalities. An increase in Russian speakers, however, raised

---

<sup>23</sup>Because hazard ratios, like coefficients, are sensitive to a variable’s unit of measurement, all variables were rescaled between 0 and 1.

the hazard of rebel takeover only in municipalities less exposed to negative economic shocks. Where a higher-than-average proportion of the local labor force was employed in at-risk industries like machine-building or mining, Russian language had no impact on rebel control.

Rebels were able to assert control soonest in areas most vulnerable to economic shocks. On any given day, a municipality with higher-than-average employment in the beleaguered machine-building industry (26 percent, or one standard deviation above the mean) was about twice as likely to fall under rebel control as a municipality with below-average employment in the industry (4 percent, or one standard deviation below the mean). The mean percent change in posterior hazard was 195 percent (95% CI: 146, 252) in minority Russian-speaking localities, and a smaller, but still formidable 92 percent (95% CI: 28, 187) in majority Russian-speaking areas.

Mining towns also fell under rebel control earlier in the conflict. Compared to Russian-minority municipalities with below-average employment in the industry (0.1 percent, one standard deviation below the mean), otherwise similar municipalities with a higher-than average mining labor force (39 percent, one standard deviation above the mean) were 105 percent (95% CI: 70, 147) more likely to fall under rebel control. This increase in hazard was 60 percent (95% CI: 2, 140) in Russian-majority municipalities.

Localities where the prewar labor force was less harmed by trade disruptions with Russia saw the opposite pattern. A town where the relatively competitive metals industry was a major employer (22 percent, one standard deviation above the mean) was 40 percent less likely (95% CI: -53, -23) to fall under rebel control at any point in time than a town with less-than-average employment in the industry (0.4 percent, or one standard deviation below the mean), although this negative impact was far more pronounced in majority-Russian speaking municipalities.

While Russian language had no discernible impact on the incidence or intensity of rebel violence, this variable fared better as a predictor of rebel control. As Table 1 shows, a non-industrial, but majority Russian-speaking

town was highly likely to fall under rebel control on a given day, with a posterior hazard 90 percent (95% CI: 6, 241) to 112 percent (95% CI: 22, 258) higher than in a majority Ukrainian-speaking non-industrial town. However, the 'language effect' disappeared in municipalities where any one of the three industries had a major presence.

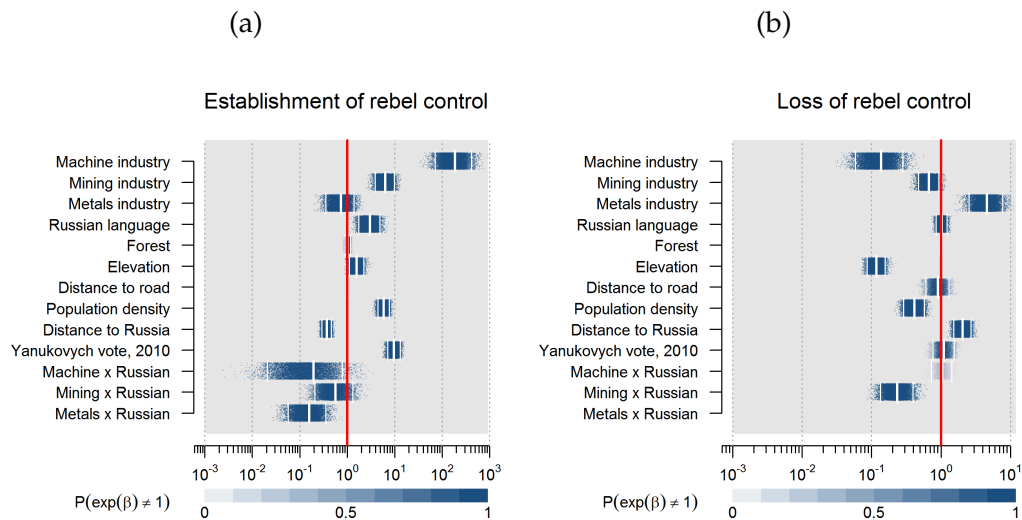
In sum, Russian language accelerated the pace of rebel takeover only in those areas, where economic motivations for (or against) rebellion were relatively weak. In municipalities whose employment structure did not necessarily favor closer integration with Russia, cultural factors like language policy were more salient drivers of public support for rebellion. In municipalities more directly exposed to economic shocks, local linguistic patterns fell to the background, and had no discernible impact on territorial control. Economic incentives, by contrast, were salient in both minority- and majority-Russian towns.

In addition to these ethnic and economic risk factors, BMA identified several other predictors of rebel control. Rebels were able to more quickly establish control in locations that were densely populated, situated at relatively high elevation, and close to the Russian border. Not surprisingly, these locations of high military-strategic value received top priority in the rebels' plans. Prewar political loyalties also seemed to play a role – strongholds of electoral support for former President Yanukovich were generally among the first to fall under rebel control.

The ability to take ground does not necessarily imply an ability to hold it. After rebels seize a territory, they must quickly establish defensive positions capable of withstanding repeated government assaults. Such a task requires that a population not turn on or abandon the rebels during a potentially protracted siege, and continue to provide active or tacit support at times of extreme hardship.

Figure 5b reports posterior hazard ratios for the loss of rebel control. Values greater than one (right of the red line) indicate a higher hazard, or shorter duration until liberation by pro-government forces. Values less

Figure 5: DETERMINANTS OF REBEL CONTROL. Quantities reported are  $P(\exp(\beta) | \exp(\beta) \neq 1, y, X)$ . Each blue point on the plot is a single draw from the posterior distribution of the hazard ratio associated with each variable, with vertical white stripes denoting means and 95% credible intervals. The opacity of the points reflects posterior inclusion probabilities, with full transparency indicating  $P(\exp(\beta) \neq 1 | y, X) = 0$  and full opacity indicating  $P(\exp(\beta) \neq 1 | y, X) = 1$ .



than one (left of the red line) indicate that a variable decreased the risk of liberation by Kyiv, extending the duration a municipality spent under rebel control.

Rebel control lasted longer in municipalities where a high proportion of the population was employed in the machine-building and mining industries prior to the war. Where the more competitive metals industry was a major employer – such as the port city of Mariupol – rebel control was far more brief. Russian language was far less predictive of the loss of rebel control than of its initial establishment.

At any point in time after the May 11, 2014 referendum on independence, rebel-held municipalities with a higher-than average machine-building labor force were 34 percent less likely (95% CI: -43, -23) to return to government control than rebel-held municipalities with below-average employment in the industry. On the opposite end of the spectrum, rebel-held municipalities with higher-than-average employment in the metals industry were 40 percent more likely (95% CI: 25, 60) to fall to government troops than ones where the industry was less dominant. The posterior distributions of these changes in hazard were the same in majority and minority Russian-speaking areas.

The only instance where language had an impact on the duration of rebel control was in mining towns. In minority Russian-speaking municipalities, a positive standard deviation shift in the mining employment share (from 0.1 to 39 percent of the labor force) yielded a 16 percent decrease (95% CI: -27, -4) in the hazard of liberation by Kyiv. This quantity dropped to -56 percent (95% CI: -67, -41) in majority Russian-speaking mining towns, like the rebel strongholds of Torez and Horlivka.

Other than economics, the only variables strongly and consistently predictive of the loss of rebel control were military-geographic. Pro-Kyiv forces were able to re-establish government control much sooner in municipalities at relatively low elevation, with low population density, and farther away from the Russian border. In this sense, the military geography



of the loss of rebel control is a near-mirror image of its initial establishment.

## 5 CONCLUSION

Data from the first year of the Donbas conflict show that a municipality's prewar employment mix is one of the strongest predictors of rebel activity. Where the opportunity costs of rebellion were low – like in machine-building company towns especially exposed to trade shocks with Russia – the risk of rebellion was greater overall, separatist violence was more frequent, the Ukrainian government lost control earlier, and rebels were able to hold on to their territory for longer. Where the opportunity costs of rebellion were higher – such as in centers dominated by Ukraine's relatively competitive metals industry – rebels had a much harder time establishing and maintaining control.

Economic preferences are not the only determinants of conflict. Rebels were able to seize ground much quicker where Russian speakers were geographically concentrated. Yet the results of Bayesian Model Averaging show that the 'language effect' asserted itself mostly in areas where economic incentives for (or against) rebellion were weak – where the industrial labor force was smaller, and the population was less directly exposed to trade shocks. In municipalities where a larger share of the population was employed in Russia-dependent industries, cultural and linguistic factors proved far less salient.

The seemingly rational economic self-interest at the heart of the conflict stands in sharp contrast with the staggering costs of war. In the twelve months since armed men began storming government buildings in the Donbas, over 6000 people have lost their lives, and over a million have been displaced. Regional industrial production fell by 49.9 percent in 2014, with machinery exports to Russia down by 82 percent ([Segodnya, 2015](#); [Stasenko, 2014](#)). Suffering heavy damage from shelling, many factories have closed. Others have been converted by rebels into military mainte-

nance facilities, refurbishing tanks and armored vehicles before they return to the battlefield ([Ostrovsky, 2015](#)). With airports destroyed, railroad links severed and roads heavily mined, a previously export-oriented economy has found itself isolated from the outside world.

This new economic reality has had a devastating impact on even the most dominant enterprises of the region. Established in 1896, Luhanskteplovoz was at one point the largest locomotive manufacturer in Eastern Europe. Its main customer in recent years has been Russian Railways, Russia's state-owned railroad giant. After the city of Luhansk lost electricity and the factory was hit by over 160 artillery shells in July 2014, the plant suspended operations ([Karpov, 2014](#)). Citing component shortages and logistical challenges, Luhanskteplovoz's Russian parent company ordered another halt to operations in January 2015, moving much of the production to factories inside Russia. By March 2015, the plant's 6500 workers had not received a paycheck in several months, and management had announced massive layoffs ([OstroVu, 2015](#)). In April 2015, Ukrainian prosecutors opened criminal proceedings against Luhanskteplovoz for 'terrorist financing,' after the Ukrainian Security Service (SBU) intercepted a locomotive attempting to cross the border for scheduled delivery to Russia ([CityNews, 2015](#)).

Rather than restoring the Donbas' status as the economic engine of the former Soviet Union, the war has returned many parts of the region to a pre-industrial state. The Donbas war, as much as any other, underscores the ex-post inefficiency of war ([Fearon, 1995](#)). If local machinists and miners had only known the scale of the destruction to come, the economic rationale for rebellion would surely have appeared less compelling. Yet their choice set at the time was between a high-risk attempt to retain their economic livelihood and an almost certain loss of income. Intent is difficult to empirically discern, even with the highest-resolution data, but the results presented here suggest that many people chose the first option.

Table 1: INTERACTION EFFECTS, ESTABLISHMENT OF REBEL CONTROL. Quantities reported are percent changes in posterior hazard of rebel control, associated with each counterfactual. Below- and above-average employment in industry defined as  $E[X] - sd(X)$  and  $E[X] + sd(X)$ , respectively.

COUNTERFACTUAL SCENARIO	MEAN	95% CREDIBLE INTERVAL
<b>Machine-building industry:</b>		
Below $\rightarrow$ above average employment in industry (minority Russian-speaking municipality)	+194.53	(+146.29, +251.68)
Below $\rightarrow$ above average employment in industry (majority Russian-speaking municipality)	+91.86	(+27.68, +186.66)
Minority $\rightarrow$ majority Russian-speaking (below-average employment in industry)	+111.56	(+22.37, +258)
Minority $\rightarrow$ majority Russian-speaking (above-average employment in industry)	+32.06	(-33.77, +171.32)
<b>Mining industry:</b>		
Below $\rightarrow$ above average employment in industry (minority Russian-speaking municipality)	+104.75	(+69.71, +146.53)
Below $\rightarrow$ above average employment in industry (majority Russian-speaking municipality)	+59.53	(+2.02, +140.02)
Minority $\rightarrow$ majority Russian-speaking (below-average employment in industry)	+90.26	(+6.12, +241.16)
Minority $\rightarrow$ majority Russian-speaking (above-average employment in industry)	+42.24	(-27.78, +182.54)
<b>Metals industry:</b>		
Below $\rightarrow$ above average employment in industry (minority Russian-speaking municipality)	-7.65	(-20.1, +6.07)
Below $\rightarrow$ above average employment in industry (majority Russian-speaking municipality)	-39.58	(-52.78, -23.05)
Minority $\rightarrow$ majority Russian-speaking (below-average employment in industry)	+92.97	(+4.52, +245.52)
Minority $\rightarrow$ majority Russian-speaking (above-average employment in industry)	+23.64	(-32.18, +131.31)

## REFERENCES

- Anushevskaya, Anastasiya. 2014. "Ekonomika Donbassa." *Argumenty i Fakty* .
- Balcells, Laia. 2011. "Continuation of politics by two means: direct and indirect violence in civil war." *Journal of Conflict Resolution* 55(3):397–422.
- Bates, Robert H. 1983. "Modernization, ethnic competition, and the rationality of politics in contemporary Africa." *State versus ethnic claims: African policy dilemmas* 152:171.
- Bazzi, Samuel and Christopher Blattman. 2014. "Economic Shocks and Conflict: Evidence from Commodity Prices." *American Economic Journal: Macroeconomics* 6(4):1–38.
- Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76(2):pp. 169–217.  
**URL:** <http://www.jstor.org/stable/1830482>
- Bergholt, Drago and Päivi Lujala. 2012. "Climate-related natural disasters, economic growth, and armed civil conflict." *Journal of Peace Research* 49(1):147–162.
- Besley, Timothy J and Torsten Persson. 2008. The incidence of civil war: Theory and evidence. Technical report National Bureau of Economic Research.
- Besley, Timothy and Torsten Persson. 2011. "The logic of political violence." *The Quarterly Journal of Economics* 126(3):1411–1445.
- Blattman, Christopher and Edward Miguel. 2010. "Civil War." *Journal of Economic Literature* 48(1):3–57.
- Brückner, Markus and Antonio Ciccone. 2010. "International Commodity Prices, Growth and the Outbreak of Civil War in Sub-Saharan Africa\*." *The Economic Journal* 120(544):519–534.
- Brückner, Markus and Antonio Ciccone. 2011. "Rain and the democratic window of opportunity." *Econometrica* 79(3):923–947.
- Buhaug, Halvard, Lars-Erik Cederman and Jan Ketil Rod. 2008. "Disaggregating Ethno-Nationalist Civil Wars: A Dyadic Test of Exclusion Theory." *International Organization* 62(3).
- Bureau van Dijk Electronic Publishing. 2015. "Orbis database."
- Cederman, Lars-Erik, Andreas Wimmer and Brian Min. 2010. "Why do ethnic groups rebel? New data and analysis." *World Politics* 62(01):87–119.
- Cederman, Lars-Erik, Nils B Weidmann and Kristian Skrede Gleditsch. 2011. "Horizontal inequalities and ethnonationalist civil war: A global comparison." *American Political Science Review* 105(03):478–495.

- Central Election Commission of Ukraine. 2010. "Vybory Prezydenta 2010 [Presidential elections 2010]." [http://www.cvk.gov.ua/vp\\_2010/](http://www.cvk.gov.ua/vp_2010/).
- Chaney, Eric. 2013. "Revolt on the Nile: Economic shocks, religion, and political power." *Econometrica* 81(5):2033–2053.
- Cicccone, Antonio. 2011. "Economic shocks and civil conflict: A comment." *American Economic Journal: Applied Economics* 3(4):215–227.
- CIESIN and Columbia University. 2005. "Gridded Population of the World, Version 3 (GPWv3) Data Collection." Center for International Earth Science Information Network (CIESIN), Centro Internacional de Agricultura Tropical (CIAT).
- CityNews. 2015. "Prokuratura otkryla proizvodstvo v otnoshenii PAO Luganskteplovoz za finansirovanie terrorizma [Prosecutor opens criminal proceedings against Luhanskteplovoz for terrorist financing]." *citynews.net.ua*.
- Collier, Paul and Anke Hoeffler. 1998. "On economic causes of civil war." *Oxford economic papers* 50(4):563–573.
- Collier, Paul and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers New Series* 56(563–595).
- 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.
- Defense Mapping Agency. 1992. "Development of the Digital Chart of the World." U.S. Government Printing Office.
- Delegation of the European Union to Ukraine. 2013. "The EU-Ukraine Association Agreement and Deep and Comprehensive Free Trade Area: What's it all about?" European External Action Service.  
**URL:** [http://eeas.europa.eu/delegations/ukraine/eu\\_ukraine/association\\_agreement/index\\_en.htm](http://eeas.europa.eu/delegations/ukraine/eu_ukraine/association_agreement/index_en.htm)
- Dube, Oeindrila and Juan F Vargas. 2013. "Commodity price shocks and civil conflict: Evidence from Colombia." *The Review of Economic Studies* 80(4):1384–1421.
- Elbadawi, Ibrahim and Nicholas Sambanis. 2002. "How much war will we see? Explaining the prevalence of civil war." *Journal of Conflict Resolution* 46(3):307–334.
- Fearon, James D. 1995. "Rationalist explanations for war." *International Organization* 49(3):379–414.
- Fearon, James D. 1998. Commitment problems and the spread of ethnic conflict. In *The international spread of ethnic conflict*, ed. David A. Lake and Donald Rothchild. Princeton University Press pp. 107–126.
- Fearon, James D. 2004. "Why do some civil wars last so much longer than others?" *Journal of Peace Research* 41(3):275–301.

- Gentile, Michael. 2015. "West oriented in the East-oriented Donbas: a political stratigraphy of geopolitical identity in Luhansk, Ukraine." *Post-Soviet Affairs* 31(3):201–223.
- Gleditsch, Kristian Skrede. 2007. "Transnational dimensions of civil war." *Journal of Peace Research* 44(3):293–309.
- Global Administrative Areas. 2012. "GADM Database of Global Administrative Areas."
- Grossman, H. 1991. "A General Equilibrium Model of Insurrections." *American Economic Review* (81).
- Guillory, Sean. 2014. "Donbass Miners and the People's Republics." *Warscapes* .
- Hegre, Håvard and Nicholas Sambanis. 2006. "Sensitivity analysis of empirical results on civil war onset." *Journal of Conflict Resolution* 50(4):508–535.
- Hoeting, Jennifer A, David Madigan, Adrian E Raftery and Chris T Volinsky. 1999. "Bayesian model averaging: a tutorial." *Statistical science* 14(4):382–401.
- Iyer, Lakshmi and Petia B Topalova. 2014. Poverty and Crime: Evidence from Rainfall and Trade Shocks in India. Technical Report 14-067 Harvard Business School BGIE Unit Working Paper.
- Karpov, Pyotr. 2014. "Zavod rabotaet v obychnom rezhime." *luganskteplovoz.com*.  
**URL:** <http://www.luganskteplovoz.com/News/OnlyNews/277>
- Koluntovych, Aleksandr. 2014. "Rol' Donbassa v ekonomike Ukrainy [Role of Donbas in Ukraine's economy]." *Ukrainskiy vybor* .
- Koubi, Vally, Thomas Bernauer, Anna Kalbhenn and Gabriele Spilker. 2012. "Climate variability, economic growth, and civil conflict." *Journal of Peace Research* 49(1):113–127.
- Langston, Henry. 2015. "On The DNR Frontline: Ukraine's Failed Ceasefire (Part 1)." *Vice News* .
- Loveland, TR, BC Reed, JF Brown, DO Ohlen, Z Zhu, LWMJ Yang and JW Merchant. 2000. "Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data." *International Journal of Remote Sensing* 21(6-7):1303–1330.
- Luhn, Alec. 2014. "East Ukraine protestors joined by miners of the barricades." *The Guardian* .
- Madigan, David and Adrian E Raftery. 1994. "Model selection and accounting for model uncertainty in graphical models using Occam's window." *Journal of the American Statistical Association* 89(428):1535–1546.
- Makogon, Yuri. 2014. "Ekonomika Donbassa ne smozhet sushchestvovat' bez Ukrainy [Donbass' economy cannot exist without Ukraine]." *Segondya* .

- Maystadt, Jean-François and Olivier Ecker. 2014. "Extreme weather and civil war: does drought fuel conflict in Somalia through livestock price shocks?" *American Journal of Agricultural Economics* 96(4):1157–1182.
- Miguel, Edward and Shanker Satyanath. 2011. "Re-examining economic shocks and civil conflict." *American Economic Journal: Applied Economics* 3(4):228–232.
- Miguel, Edward, Shanker Satyanath and Ernest Sergenti. 2004. "Economic shocks and civil conflict: An instrumental variables approach." *Journal of Political Economy* 112(4):725–753.
- Ministry of Economic Development of the Russian Federation. 2014. "Kratkie itogi vzaimnoy torgovli tovarami i uslugami Rossii i Ukrainy v janvare-sentyabre 2014 goga [Short summary of Russian-Ukrainian bilateral trade in January-September 2014]." Foreign trade information portal.  
**URL:** [http://www.ved.gov.ru/exportcountries/ua/ua\\_ru\\_relations/ua\\_ru\\_trade/](http://www.ved.gov.ru/exportcountries/ua/ua_ru_relations/ua_ru_trade/)
- Montgomery, Jacob M and Brendan Nyhan. 2010. "Bayesian model averaging: Theoretical developments and practical applications." *Political Analysis* 18(2):245–270.
- Montgomery, Jacob M., Florian M. Hollenbach and Michael D. Ward. 2012. "Improving predictions using ensemble Bayesian model averaging." *Political Analysis* 20(3):271–291.
- Morin, Evgeniy. 2014. "Donbass: zabytyi referendum-1994 [Donbass: the forgotten referendum of 1994]." *The Kyiv Times* .
- Ostrovsky, Simon. 2015. "Should the US Send Lethal Aid to Ukraine?" *Vice News* .
- OstroVu. 2015. "Luganskteplovoz okonchatel'no ostanovil vse raboty [Luhanskteplovoz permanently halted operations]." *OstroVu* .  
**URL:** <http://www.ostro.org/lugansk/economics/news/465675/>
- Popeski, Ron. 2014. "Kyiv at crossroad: 'All we can do is hope for the best'." *Winnipeg Free Press* .
- Saideman, Stephen M and R William Ayres. 2000. "Determining the causes of irredentism: Logit analyses of minorities at risk data from the 1980s and 1990s." *The Journal of Politics* 62(04):1126–1144.
- Salehyan, Idean. 2007. "Transnational rebels: Neighboring states as sanctuary for rebel groups." *World Politics* 59(02):217–242.
- Salehyan, Idean, Kristian Skrede Gleditsch and David E Cunningham. 2011. "Explaining external support for insurgent groups." *International Organization* 65(04):709–744.
- Savun, Burcu and Scott Cook. 2010. Exogenous shocks, bargaining problems, and the onset of civil war. In *American Political Science Association Annual Meeting*.

- Segodnya. 2015. "Ekonomika Donetskoy oblasti v upadke iz-za voyny - gubernator Kikhtenko." *Segodnya* .
- Siegelbaum, Lewis H and Daniel J Walkowitz. 1995. *Workers of the Donbass Speak: Survival and Identity in the New Ukraine, 1989-1992*. SUNY Press.
- Stasenko, Maria. 2014. "Novaya ekonomika Ukrainy budet stroit'sya bez Rossii i Donbassa [Ukraine's new economy will be built without Russia or the Donbas]." *Delo.ua* .
- State Committee on Statistics of Ukraine. 2001. *Ukrainian census 2001*. State Committee on Statistics of Ukraine.
- State Committee on Statistics of Ukraine. 2013. "Geografichna struktura zovnishn'oy torgovli tovaramy za 2013 rik [Geographic structure of foreign trade in goods in 2013]." URL: [http://www.ukrstat.gov.ua/operativ/operativ2013/zd/ztt/ztt\\_u/ztt1213\\_u.htm](http://www.ukrstat.gov.ua/operativ/operativ2013/zd/ztt/ztt_u/ztt1213_u.htm)
- State Statistics Service of Ukraine. 2015. "Statystichna informatsiya [Statistical information]." URL: <http://www.ukrstat.gov.ua/>
- Tarnavskiy, Viktor. 2014. "Vymirayushchaya promyshlennost' [Dying industry]." *MinProm* .
- TASS. 2014. "Manturov: Rossiya zamestit vse postavki iz Ukrainy za 3 goda [Manturov: Russia will substitute all Ukrainian imports within 3 years]." TASS .
- Toft, Monica Duffy. 2003. *The Geography of Ethnic Violence: Identity, Interests and the Indivisibility of Territory*. Princeton, New Jersey: Princeton University Press.
- Tullock, Gordon. 1971. "The Paradox of Revolution." *Public Choice* 11(1):89-99.
- Ulfelder, Jay. 2012. "Forecasting onsets of mass killing." Available at SSRN 2056306.
- U.S. Geological Survey. 1996. "Global 30-Arc-Second Elevation Data Set." Dni.ru.
- Walker, Shaun. 2013. "Ukraine's EU trade deal will be catastrophic, says Russia." *The Guardian* .
- Weidmann, N.B. 2009. "Geography as Motivation and Opportunity: Group Concentration and Ethnic Conflict." *Journal of Conflict Resolution* 53(4):526-543.
- Wimmer, Andreas. 2002. *Nationalist exclusion and ethnic conflict: shadows of modernity*. 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):144-156.



Figure 6: CORRELATION MATRIX, MUNICIPALITY-LEVEL DATA.

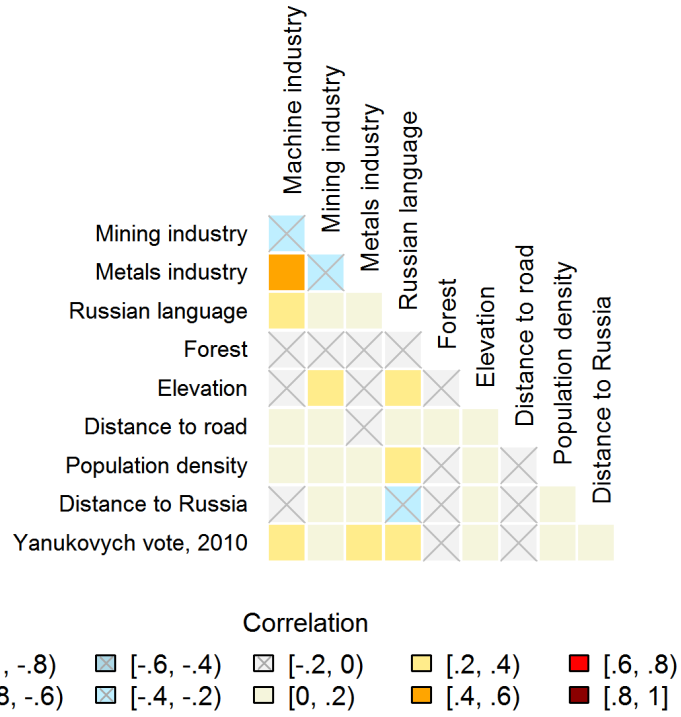


Table 2: SUMMARY STATISTICS. Municipality-level data, unless otherwise noted. Lags excluded. Covariates re-scaled to  $x \in [0, 1]$  for all analyses.

	Obs	Mean	Median	SD	Min	Max
<i>Outcome variables</i>						
Any rebel violence	3037	0.20	0.00	0.40	0.00	1.00
Duration until rebel control	3037	676.48	113.00	748.77	40.00	1650.00
Duration until rebel control (censoring dummy)	3037	0.63	1.00	0.48	0.00	1.00
Duration until loss of rebel control	3037	197.03	0.00	495.48	0.00	1576.00
Duration until loss of rebel control (censoring dummy)	3037	0.89	1.00	0.32	0.00	1.00
Intensity of rebel violence (municipality-week)	160,961	0.02	0.00	0.26	0.00	7.00
<i>Covariates</i>						
Machine industry	3037	0.15	0.15	0.11	0.02	0.45
Mining industry	3037	0.20	0.17	0.20	0.00	0.78
Metals industry	3037	0.11	0.07	0.11	0.00	0.41
Russian language	3037	39.82	38.00	22.41	2.38	95.41
Forest cover	3037	0.89	1.00	0.31	0.00	1.00
Elevation	3037	143.77	143.00	64.00	-2.00	343.00
Distance to road	3037	3.34	2.66	2.87	0.00	17.23
Population density	3037	188.01	36.53	362.44	12.28	2407.10
Distance to Russia	3037	69.51	60.64	48.85	0.02	197.05
Yanukovich vote, 2010	3037	0.75	0.76	0.05	0.61	0.84