ECONOMISTS
FOR UKRAINE

# WORKING PAPER SERIES

WORKING PAPER No. 20 MARCH 2026

# DIVERSIONARY ESCALATION: THEORY AND EVIDENCE FROM EASTERN UKRAINE

**Natalie Ayers, Christopher Blair, Joseph Ruggiero, Austin L. Wright, and Konstantin Sonin**

We kindly thank Alexandra Chinchilla, Taylor Damann, Zuheir Desai, Scott Gehlbach, Leonard Wantchekon, and participants at the 2023 Empirical Studies of Conflict Annual Meeting, the 2023 Conflict Research Society Conference, and the 2024 Peace Science Society Annual Meeting for helpful comments. For excellent research assistance, we thank Bakdaulet Baitan, George Beknazar-Yuzbashev, and Maria Milosh. Any remaining errors are our own.



**Economists for Ukraine (Econ4UA)**

Website: https://econ4ua.org/ Email: info@econ4ua.org

# ABSTRACT

## Diversionary Escalation: Theory and Evidence from Eastern Ukraine

When leaders face threats to their authority, escalating foreign conflict can help divert public attention away from domestic grievances. We develop a formal microfoundation for diversionary escalation rooted in a theory of regime change. Although the idea of diversionary escalation is classic, systematic quantitative evidence has been challenging to obtain. Using a new data set of 1.8 million conflict incidents, obtained from the Organization for Security and Co-operation in Europe's (OSCE) Special Monitoring Mission to Ukraine in 2015-2022, we find evidence that the Russian government strategically employed proxy-initiated separatist violence in Eastern Ukraine to divert attention from domestic unrest and opposition-led protest. We also find a positive link between opposition protest and inflammatory anti-Ukrainian coverage in the Russian media, complementary to battlefield escalation.

**JEL CLASSIFICATION: D74, D72, F51, F52, C72, P26**
**KEYWORDS: Conflict, Diversion, Domestic Politics, Ukraine, Protest**

Natalie Ayers
Harvard University
Department of Economics
Cambridge, Massachusetts
natalieayers@g.harvard.edu

Christopher Blair
Princeton University
Department of Politics
Princeton, New Jersey
chris.blair@princeton.edu

Joseph Ruggiero
University of Virginia
Department of Politics
Charlottesville, Virginia
jruggiero@virginia.edu

Austin L. Wright
The University of Chicago
Harris School of Public Policy
Chicago, Illinois
austinlw@uchicago.edu

Konstantin Sonin
The University of Chicago
Harris School of Public Policy
Chicago, Illinois
ksonin@uchicago.edu

# Introduction

The diversionary theory of war posits that leaders may wage foreign conflicts to stem domestic opposition. By fostering national cohesion and mobilizing public support (Mueller, 1973), demonstrating foreign policy competence (Richards et al., 1993; Smith, 1996), distracting mass attention (Alrababa'h and Blaydes, 2021), and scapegoating foreign adversaries (Oakes, 2006), conflict can help politically vulnerable leaders overcome domestic threats.

Although both academics and policymakers acknowledge this intuitive logic, the empirical evidence is surprisingly weak (Levy, 1989; Fravel, 2010). Mixed findings in the existing literature owe to a host of challenges, including strategic behavior of belligerent parties and the complexity and rarity of war onset. In addition, initiating interstate conflict entails high costs and risks (Miller, 1995), and leaders have numerous domestic policy options to address unrest, like repression or concessions (Gelpi, 1997). For these reasons, many scholars are skeptical that leaders would incite foreign conflict in response to domestic tumult, casting doubt on the diversionary theory of war (Leeds and Davis, 1997; Gowa, 1998; Chiozza and Goemans, 2004).

The first challenge for diversionary theorists is determining which mechanism is acting to produce the observed conflict and potential diversionary behavior (Mercer, 2005; Levy, 1989; Tir, 2010; Mehrl and Choulis, 2021; Haynes, 2017). Chiozza and Goemans (2004) classify these mechanisms into categories: scapegoating, rally-round-the-flag, and gambling for resurrection. Others have suggested modifications to this categorization, adding signaling competence and distraction as viable and distinct alternatives (Oakes, 2006). The specific reasons why leaders may engage in diversion are important beyond purely theoretical argument—they bear on design and empirical strategy, since different underlying logics may result in distinct diversionary targets (Haynes, 2017).

The question of which disputes are selected is another important consideration when evaluating diversionary behavior. By considering targets and their selection more precisely, recent empirical literature has begun to find greater alignment with theoretical argument for diversion. In particular, territorial conflicts are especially conducive for diversionary behavior because citizens have stronger emotional reactions to territory, and thus feel a greater sense of national identity and unity when faced with territorial threats (McLaughlin Mitchell and Thyne, 2010; Mehrl and Choulis, 2021; Tir, 2010). Oakes (2006) similarly argues that opportunity, in the form of low-cost conflicts and "domestically popular target[s]," is required for diversionary behavior. Long-standing territorial disputes often fit the combination of opportunity and incitement necessary for successfully diverting public attention. While the question of whether and when states engage in diversionary behavior is actively debated, this debate assumes another layer of complexity when proxy forces are involved, as during the recent separatist conflict along Ukraine's eastern border with Russia.

We argue that diversionary motives are more likely to influence the *conduct* than the *onset* of war. Leaders will often find it easier to address domestic turmoil by waging foreign violence if they can do so within existing crises. Escalation is an appealing strategy precisely because it yields the purported benefits of diversionary war without the costs of initiating conflict anew. If these additional costs are not outweighed by additional benefits, the intensification of hostilities in ongoing conflicts can offer useful diversion. This argument extends work by Mitchell and Prins (2004), Tir (2010), and Haynes (2016) about how opportunities for fighting shape the incidence of diversionary behavior. By focusing on the escalation of combat, we also expand literature on diversionary actions short of war initiation (Carter, 2020; Alrababa'h and Blaydes, 2021).

In this paper, we present a theory and supportive empirical evidence that diversionary motives can explain periodic intensification of combat during proxy wars. Our formal model adapts work in global games of regime change (Angeletos, Hellwig and Pavan, 2006, 2007; Bueno de Mesquita, 2010) to consider a regime that may employ diversionary tactics to thwart a citizen rebellion. These tactics include both conventional war, which suppresses domestic opposition, and diversionary escalation by proxy war, which manipulates citizen information through state propaganda. The model suggests that proxy war escalation serves as an effective diversionary tactic for two types of regimes: those on the verge of being overthrown, and those that would survive otherwise but nonetheless face significant opposition. Additionally, we show that the efficacy of information manipulation via diversionary escalation is diminishing in the quality of citizen information and, hence, depends on domestic news sources being sufficiently unreliable.

Evaluating Russian-led escalation in eastern Ukraine (i.e., Donbas), we find supportive evidence for our argument.[1] The Russian invasion of Ukraine in February 2022 sparked renewed interest in the dynamics of interstate war. We examine the origins of the contemporary conflict, with a particular focus on the period from April 2015–February 2022, when Russian-backed separatist strife morphed into a broader Russian-led proxy war in eastern Ukraine.[2] Conflict flared numerous times in this period, claiming more than 14,000 lives as Ukrainian and (pro-)Russian forces fought to a bitter stalemate.

To conduct our empirical analysis, we use newly-assembled dataset of ceasefire violations in eastern Ukraine, representing more than 1.87 million records of violence. We compiled this dataset from reports filed by the Organization for Security and Co-operation in Europe's (OSCE) Special Monitoring Mission to Ukraine (SMM). From September 2014–February 2022, SMM monitors were tasked with overseeing the Minsk Protocol, a fragile ceasefire between Ukrainian and (pro-)Russian forces in Donbas. To this end, SMM teams patrolled government and non-government controlled areas along the contact line, recording ceasefire

[1] Literature on the Donbas conflict, as in larger diversionary war literature, is indeterminate. Some experts attribute Russian actions to regime concerns about domestic opposition (Stoner and McFaul, 2015; Gomza, 2022), while others argue geopolitical and ideational factors drove Russian strategy (Marten, 2015; Götz, 2017).

[2] See section 3.1 for background on the conflict from 2014-2022. We refer to eastern Ukraine and Donbas interchangably. Donbas includes Donetsk and Luhansk, two regions of eastern Ukraine where (pro-)Russian forces are active.

violations using physical observation, aerial surveillance, and a network of sensors. Each record contains rich information on the location, target, source, and tactic of a ceasefire violation. Pairing this new data with granular information on anti-regime protests in Russia, we document evidence of diversionary Russian escalation.

We further trace the diversionary motivation behind this violence with a large-scale text analysis of Russian television, radio, and print media. We identify content related to the Donbas conflict using a corpus of government-controlled and (nominally) independent press sources. Our media analysis provides additional supportive evidence. Using natural language processing techniques, we find that opposition protests in Russia correspond with increasingly extreme and vitriolic coverage of the conflict in Ukraine. Growing reliance on inflammatory anti-Ukrainian content in state media reflected a deliberate regime strategy to justify the proxy conflict, divert attention from domestic political issues, and build mass support for Russian-backed escalation in Ukraine (Arel and Driscoll, 2024, p. 104-106).

In sum, this paper makes four contributions to the broader literature. First, we offer novel, microlevel evidence for diversionary theories of war. Our emphasis on escalation extends important strands of the qualitative (Oakes, 2006) and cross-national, quantitative literatures (Mitchell and Prins, 2004; Mitchell and Thyne, 2010; Tir, 2010; Haynes, 2016), which highlight how leaders capitalize on opportunities for diversion by engaging in belligerence within ongoing crises. More broadly, our analyses offer new empirical support for prominent, second-image models of war (Tarar, 2006; Jackson and Morelli, 2007).

Second, while existing theories assume that leaders choose between diversion and domestic repression (Gelpi, 1997), we show how autocrats can use diversionary escalation to complement domestic responses to unrest. In Russia, Vladimir Putin's regime responded to opposition protest with diversionary belligerence, repression (Götz, 2017), and disinformation (Stukal et al., 2022) in tandem. We also contribute to a related theoretical literature on regime change (Angeletos, Hellwig and Pavan, 2006, 2007; Bueno de Mesquita, 2010). Recent work in this domain has focused mainly on the coordination problem citizens face. For instance, Tyson and Smith (2018) consider two types of citizens—opponents and adherents—who affect the resiliency of the regime, while Bueno de Mesquita and Shadmehr (2023) study how different types of rebel motivations can lead to differing levels of vulnerability to repression, and Correa, Nandong and Shadmehr (2024) examine the role of grievances in facilitating coordination. Fewer studies have given direct attention to the regime's strategy. Most notable is Edmond (2013), which studies the ability of a regime to improve its position by manipulating the information of its citizens through propaganda. We directly extend these ideas and connect them with a theory of diversionary war.

Third, we draw attention to and provide microfoundations for diversionary behavior in a *proxy war*. To the best of our knowledge, we are the first to demonstrate how great powers can manipulate proxy conflicts for diversionary ends. Proxy wars are increasingly common and prone to intensification (Berman and Lake, 2019), rendering them important subjects of study. Shifting attention to the dynamics of the Donbas conflict helps illuminate the origins of the full-scale Russian invasion of Ukraine, and the extensive violence that occurred in prelude.

Finally, our study introduces a large, highly-detailed new dataset of combat in eastern Ukraine. These records will be of use to scholars of conflict, mediation, and peacekeeping, and to specialists on the former Soviet Union. The data offer particular promise for testing theories of Russian foreign policymaking developed in qualitative and area studies research (e.g., Treisman, 2011; Marten, 2015; Hale, 2022).

## 2 Theory

In this section, we present our theory of diversion. We build off of the literature on global games of regime change by considering a model in which individual citizens learn about the regime's quality and choose whether to lend their support or join the opposition. Citizens may collectively overthrow the regime, with the success of attempts depending on (i) how many citizens choose to oppose and (ii) the quality of the regime, with higher quality regimes being less easily overthrown. However, the regime seeks to retain power and may attempt to thwart public opposition through use diversionary tactics: either the initiation of a conventional war or escalation of a proxy war. The key difference between these tactics is that proxy war escalation serves to manipulate citizen information through state propaganda, while conventional war is an act of suppression that ensures survival at great cost.[3]

## 2.1 Setup

There is a regime and a continuum of citizens $i \in [0, 1]$. The regime's quality is randomly drawn from a normal distribution, $\theta \sim N(z, \tau^2)$, with larger types reflecting higher quality. Upon observing their type, the regime may choose whether to do nothing, escalate an existing proxy war, or start a diversionary conventional war, $r \in \{0, 1, 2\}$ respectively. On the other hand, each citizen forms beliefs about the regime's type given their information and chooses whether to support or oppose the regime, $a_i \in \{0, 1\}$.

First, a regime may choose to fight in a conventional war at a cost $w \in (0, 1)$, guaranteeing that any citizen attempts to oppose the regime are squashed. The important property is that the action is decisive: the regime knows that it will counteract attacks against the regime and guarantee survival ex ante. The negation of citizen opposition could be microfounded in a number of ways, most notably rally-round-the-flag effects (Tarar, 2006; Chatagnier, 2012). As a result, the regime receives a conventional war payoff of $\theta - w$, all citizens receive a payoff of zero (without loss of generality), and the game ends.

[3] Both in practice and in the model, the diversionary effects of conventional war can be grounded in logic such that it affects citizen information, as well. However, the important feature is that conventional war creates circumstances in which a rebellion would fail, and citizens are aware of this fact. As a result, we can focus on conventional war as a general act of suppression, in contrast to diversionary escalation.

Second, consider a regime that simply chooses to do nothing. This case resembles a standard global game (Carlsson and Van Damme, 1993; Morris and Shin, 1998). Specifically, each citizen observes a private signal $x_i = \theta + \varepsilon_i$ with noise that is independent of the regime's type as well as independent and identically distributed across citizens, $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$. We assume that this signal is sufficiently informative; specifically, $\sigma < \tau^2\sqrt{2\pi}$.[4] Given the realized signal, citizens update their beliefs about $\theta$ and choose whether to engage in an act of rebellion, $a_i = 1$, at a personal cost $c \in (0, 1)$. The size of the attack on the regime is given by $A := \int a_i di$ and the regime survives $(S = 1)$ if and only if $\theta > A$. The regime receives a payoff of zero in regime change and $\theta - A$ if rebellion is unsuccessful. Further, rebel citizens receive a payoff of $1 - c$ if the regime is overthrown or $-c$ if not, while passive citizens receive zero.

Finally, the regime may engage in diversion by escalating an ongoing foreign war, altering citizen information. The mechanism by which states alter citizen beliefs via diversionary escalation is most similar to information manipulation by propaganda in Edmond (2013), except that proxy war is more crude in the sense that states cannot fine-tune their level of manipulation – by engaging in escalation, they facilitate diversionary news coverage of the proxy conflict that is given. Specifically, if the regime chooses to employ proxy warfare at a cost $p \in (0, 1)$, each citizen observes a private signal $x_i = t + \varepsilon_i$ where $t \in \mathbb{R}$ is the mean signal on quality citizens receive from state propaganda. Importantly, while citizens take the possibility of propaganda into account when making their decisions, regimes can nevertheless use this technology to effectively reduce opposition: citizens do not know from which distribution their signal was drawn. If the regime engages in diversionary proxy war escalation at cost $p$, they receive $\theta - A$ if they survive $(S = 1)$ and zero otherwise.

Figure 1 summarizes the total payoffs for the regime, given their quality, diversionary choices, and citizen behavior. It is noteworthy that, holding fixed the number of citizens

[4] This condition is sufficient but not necessary for our results. See Lemma 4 in Supplementary Appendix C.

Figure 1: Regime type $\theta$ payoffs given size of attack $A$

| | **No diversion** | **Conventional war** | **Proxy escalation** |
|---|---|---|---|
| $A \leq \theta$ | $\theta - A$ | $\theta - w$ | $\theta - A - p$ |
| $A > \theta$ | $0$ | $\theta - w$ | $-p$ |

opposing the regime, the choice to not engage in diversion is always preferable to proxy escalation. The effect of engaging in proxy escalation is therefore only useful insofar as it affects the size of the attack through citizen beliefs about the world. Conventional war, on the other hand, serves to fully suppress attempts to overthrow the regime and may be preferable to no diversion if the size of citizen opposition is large enough. The virtue of this modeling choice is that it isolates the role of diversionary escalation by proxy war as a tool for information manipulation, and of conventional war as a costlier tool of suppression.

## 2.2 Equilibrium

Our solution concept is Bayesian Nash equilibrium, henceforth simply equilibrium. It is immediate to see that citizens will always prefer to oppose when their posterior belief that $\theta$ is negative is more than $c$ and support when their posterior that $\theta$ is greater than 1 is more than $c$. Then, we can focus on citizens with intermediate posterior beliefs about $\theta$. As a result, it is natural to follow previous work in global games of regime change in looking for equilibria that are symmetric and monotone, i.e., all citizens play the same strategy in which they oppose the regime if and only if they receive a sufficiently high signal. Further, we focus on equilibria where diversionary escalation occurs.

**Proposition 1.** *There is a symmetric monotone equilibrium in which each citizen* $i \in [0, 1]$ *opposes the regime if and only if* $x_i \leq x^*$ *for some threshold* $x^*$*, and the regime engages in diversionary escalation if and only if* $\theta \in (\tilde{\tilde{\theta}}(x^*), \tilde{\theta}(x^*))$ *for some* $\tilde{\tilde{\theta}}(\cdot), \tilde{\theta}(\cdot) \in \mathbb{R}$.

*Proof.* Let us then conjecture that there exists an $x^* \in \mathbb{R}$ such that citizens oppose the regime if and only if $x \leq x^*$. If the regime does not engage in diversionary escalation, the citizen receives an unbiased signal; however, the citizen receives manipulated information otherwise. Then, the size of attack can be expressed by $A(\theta, r) = P(x \leq x^* | \theta, r)$, which is

$$A(\theta, r = 0) = \Phi\left(\frac{x^* - \theta}{\sigma}\right) \quad \text{and} \quad A(\theta, r = 1) = \Phi\left(\frac{x^* - t}{\sigma}\right)$$

after doing nothing and escalating a proxy war, respectively. The highest quality regime that gets overthrown by doing nothing is thus given by type $\theta^* = A(\theta^*, r = 0)$. Given any citizen threshold $x^*$ and these corresponding expected attacks, a regime of type $\theta$ simply chooses the diversionary tactic that yields them the greatest expected payoff.

When citizens play a threshold strategy according to some $x^*$, a regime of type $\theta$ has a unique best response corresponding to the payoff

$$\max\left\{0,\, \theta - A(\theta, r = 0),\, \theta - A(\theta, r = 1) - p,\, \theta - w\right\},$$

with indifferences breaking in favor of the least aggressive action. By Lemmas 1, 2, and 3 in Supplementary Appendix C, we show that the set of regime types that engage in diversionary escalation tactics constitute an open interval of reals. Therefore, to see such a threshold exists, suppose the regime plays equilibrium strategies in which they engage in diversionary escalation for all types $\theta \in (\tilde{\tilde{\theta}}, \tilde{\theta})$ for any $\tilde{\tilde{\theta}}, \tilde{\theta} \in \mathbb{R}$ such that $\tilde{\tilde{\theta}} \leq \tilde{\theta}$.

Citizens only want to attempt to overthrow the regime if their signal is large enough that they expect enough other citizens also will join in the opposition. We can express the posterior probability that the regime is weaker than the strongest type that cannot avoid overthrow by doing nothing as

$$P(\theta \leq \theta^* | x) = \gamma \Phi\left(\frac{\theta^* - \mu(x)}{\nu}\right) + (1 - \gamma)\Phi\left(\frac{t - \mu(x)}{\nu}\right),$$

where $\mu(x)$ and $\nu^2$ are the posterior mean and variance, respectively, and we define the ex-ante probability of diversionary escalation by $\gamma := 1 - \Phi\big(\frac{\tilde{\theta}-z}{\tau}\big) + \Phi\big(\frac{\tilde{\tilde{\theta}}-z}{\tau}\big) \in (0, 1]$, taking the regime's strategy as given.

The optimal threshold $x^*$ then solves $P(\theta \leq \theta^* | x^*) = c$ or, equivalently, the highest quality regime to be overthrown without diversion $\theta^*$ solves

$$\gamma\Phi\left(\frac{\theta^* - \mu(\theta^* + \sigma\Phi^{-1}(\theta^*))}{\nu}\right) + (1-\gamma)\Phi\left(\frac{t - \mu(\theta^* + \sigma\Phi^{-1}(\theta^*))}{\nu}\right) = c. \tag{1}$$

Note that, because conventional war suppresses any attempted rebellion, citizens do not incur the costs of opposition and hence do not need to condition on the event that regimes do not launch a war of their own. By Lemma 4 in Supplemental Appendix C, $\sigma < \tau^2\sqrt{2\pi}$ is a sufficient condition for a unique solution to equality (1). $\square$

### 2.3 Analysis

As identified in Proposition 1, diversionary escalation may occur in equilibrium. There are several things to note about this behavior and how it arises.

First, as our proof reveals, if proxy war escalation is part of a regime's strategy, only intermediate quality regimes choose to engage in it. If the regime quality is greater than this range, they expect citizens to receive a strong enough signal in the absence of diversion as to make the costs of diversion not worthwhile. However, while proxy war escalation can be a useful diversionary tactic for improving the performance of regimes with lower quality, those of significantly lower quality beyond this range in fact prefer being overthrown, accepting their fate without incurring the additional costs.

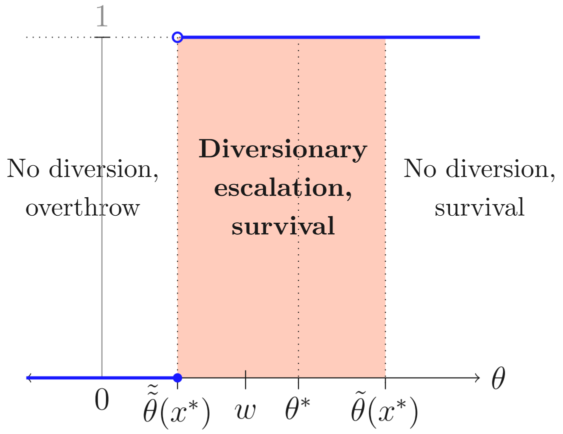
Second, if information manipulation via escalation is the preferred method of diversion for some measurable set of regime types then conventional war will not occur in equilibrium, and vice versa. In the model, conventional war onset and proxy war escalation are two distinct, costly technologies for altering the regime's opposition. However, one will always be preferred to the other, regardless of the realized regime type. This is due to the fact that the realized type is orthogonal to the net difference in payoffs between the two means of diversion. While this condition may not always hold, this feature nevertheless provides useful insight in the context of our stylized model: if information pertinent to citizen decisions to oppose does not affect the relative advantages of different diversionary tactics, all possible regimes will prefer one specific method. Moreover, this provides a credible explanation for why it has been difficult to find empirical evidence of diversionary conventional war: proxy war escalation may always be a preferable method.

Third, whenever diversionary escalation may occur in equilibrium, there are always some regimes that do so to avoid overthrow and there are always some regimes that do so simply to reduce their opposition. That is, the highest quality regime that would be overthrown in the absence of diversion necessarily partitions the set of regimes that engage in proxy war escalation. The result is stated formally in the following proposition. All remaining proofs are presented in Supplemental Appendix C.

**Proposition 2.** *In symmetric monotone equilibrium, if some types resort to diversionary escalation, and hence* $\tilde{\tilde{\theta}}(x^*) < \tilde{\theta}(x^*)$*, then* $\theta^* \in (\tilde{\tilde{\theta}}(x^*), \tilde{\theta}(x^*))$*.*

There are several key implications. First, it reveals that diversion is indeed an effective way for regimes to retain power. In particular, given an equilibrium in which citizens play a threshold strategy $x^*$, all types $\theta \in (\tilde{\tilde{\theta}}(x^*), \theta^*)$ survive only as a result of their ability to escalate the proxy war and consequently use state propaganda to manipulate citizen information. Second, when proxy war escalation may occur in equilibrium, it suggests that some regimes would be overthrown if conventional war was their only means for diversion. In fact, there exists a set of regime types, $\theta \in (\tilde{\tilde{\theta}}(x^*), w]$, that survive under diversionary escalation by proxy war but would fail under diversionary conventional war. Third, Proposition 2 implies

Figure 2: Probability of regime survival in regime type




that, if diversionary escalation may occur in equilibrium, the regimes that choose to employ this method constitute the weakest surviving regimes. Figure 2 illustrates these points.

Finally, we use our model to investigate the implications of the information environment on the regime ability to effectively use diversionary escalation. We find that, as citizen information becomes increasingly precise, proxy war escalation ceases to be a useful means of diversion for the regime.

**Proposition 3.** *Diversionary escalation does not occur in symmetric monotone equilibrium as citizen signal quality grows precise,* $\lim_{\sigma \to 0^+} \tilde{\tilde{\theta}}(x^*(\sigma)) > \lim_{\sigma \to 0^+} \tilde{\theta}(x^*(\sigma))$.

Proposition 3 demonstrates that the regime will be unable to effectively use proxy war escalation as a diversionary tactic if citizens are sufficiently informed. Then, societies with an information environment that has a high degree of accuracy will be fairly robust to state efforts to leverage escalatory proxy war to their domestic political gain. Importantly, this remains true even when the information citizens receive is fully distorted by the regime's decision to engage in proxy war: as the accuracy of new information becomes precise, diversionary escalation results in a signal of $x_i = t$ for all citizens, leaving them entirely uninformed about the true quality of the regime. However, upon observing $x_i = t$ in a highly accurate information environment, the citizen would be able to infer that they are receiving a false signal. Crucially, the effective use of proxy war escalation to manipulate citizen beliefs about the regime depends not only on the state's ability to change the news, but on citizen uncertainty about whether or not they have chosen to do so.

# 3 Context

As the more recent diversionary violence literature demonstrates, when the trigger and targets are appropriate, the threat or use of diversionary force can appear to be a reasonable course of action for leaders facing domestic difficulties. To be implemented successfully, this requires careful calibration of the use of violence by the military and is designed to shape public perception and media portrayal. Russia, as a regular user of misinformation and disinformation to suit its geopolitical aims, is well-placed to employ such careful perception-crafting in its foreign policy endeavors (*GEC Special Report: Russia's Pillars of Disinformation and Propaganda*, 2020). Research in the context of another media-savvy state, Israel, has found precisely this careful control of violence for the media during its conflict with the Palestinians. Durante and Zhuravskaya (2018) found that Israel timed its attacks to occur with predictable US news events which would distract international attention.

Russia has also been a frequent and successful employer of this technique in the context of conflict. Filippov (2009) argues that Russia employs a strategy of diversion in seeking to maintain a constant level of contention with the West in order to discredit Western ideas—thus insulating its political system from democratic influence—while still allowing relationships with Western countries for economic gain. Filippov proposes that post-Soviet states provide the perfect foil for this contest, as they can be portrayed as agents of the West, and he argues the conflict between Georgia and Russia in August of 2008 can be partially attributed to this diversionary agenda.

Russia's annexation of Crimea has also been described as a response to domestic challenges, as Putin was facing a weak economy and falling approval ratings at the start of his 2012 presidential term (Theiler, 2018; Pifer, 2020; Hale, 2022). Putin leaned heavily into nationalist rhetoric and imagery during this time to regain support, and as returning Crimea to Russian control was a popular nationalist goal, Putin saw his approval ratings rise in response (Nardelli, Rankin and Arnett, 2015). Russia's intervention in Syria in September 2015 has further been suggested to be partially driven by diversionary motives, perhaps in part to divert attention from the stalemated separatist movements in eastern Ukraine and to act against a revolutionary movement that had parallels with the Maidan revolutions in Ukraine (Baev, 2019; Borshchevskaya, 2015). In each of these cases, Russia's strategic propaganda campaigns has been employed in tandem with its military activities, providing strong suggestive evidence for Putin's diversionary motivations and capabilities (Bohush et al., 2016).

### 3.1 War in the Donbas from 2014–2022

To more formally examine diversionary Russian escalation, we study conflict in eastern Ukraine from 2014–2022. Prior to the full-scale invasion of Ukraine in February 2022, Russia waged an eight-year proxy conflict in Eastern Ukraine. This conflict began in the wake of large-scale Ukrainian protests in late 2013.[5] Specifically, in November 2013 Ukrainians took to the streets after their pro-Russian government moved to halt an association agreement with the European Union (Pifer, 2019). After police fired on protesters in Maidan Square in February 2014, killing nearly 100, the Ukrainian president fled to Russia amid the public outrage. In the chaos that ensued, Russian troops occupied the Crimean peninsula (Pifer, 2020). Though widely condemned by world powers, Russia was able to take Crimea without strong opposition, exploiting covert action and support from the significant ethnic Russian population residing in Crimea. After the annexation of Crimea, separatist violence broke

[5] Arel and Driscoll (2024) offer a systematic overview of the conflict.

out in areas of Ukraine's Donetsk and Luhansk oblasts, collectively known as the Donbas.

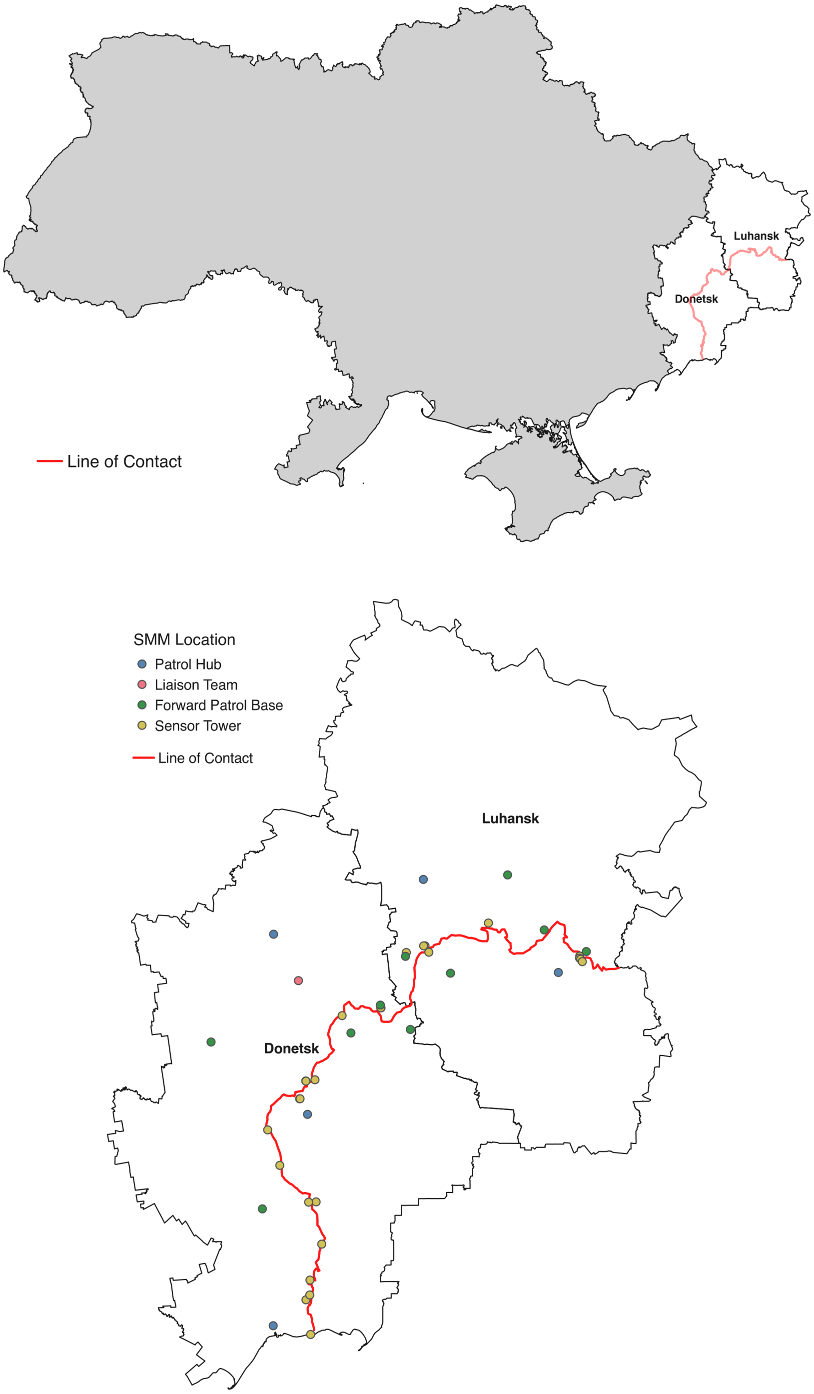
In Spring 2014, Russian-backed Donbas separatists formed the self-proclaimed Donetsk and Luhansk People's Republics (DNR, LNR) (Mankoff, 2022). Ukraine's military quickly mobilized to repulse this separatist threat, prompting Russia forces to mobilize combat support, including arms and reinforcements, for separatist forces. Over the ensuing eight years, more than 14,000 people were killed in this separatist conflict, despite attempts at international mediation and two ceasefire agreements, the Minsk 1 and 2 Accords, signed in 2014 and 2015 (Kramer, 2022). To monitor the ceasefire and facilitate conflict mediation, the Organization for Security and Co-operation in Europe (OSCE) deployed a Special Monitoring Mission to Ukraine (SMM) in March 2014. The SMM's stated aim was "reducing tensions and fostering peace, stability and security," and of monitoring and reporting on the situation (*Permanent Council Decision No. 1117*, 2014). To facilitate this mission, SMM observers established a network of patrol bases, liaison offices, and sensor towers throughout Donbas (Figure 3), and conducted daily patrols and drone flights over the line of contact between Ukrainian and Russian-backed forces. The OSCE Mission maintained an independent and neutral status, despite being criticized as a tool of the West by some Russian parties, and published daily reports on the status of the conflict, including cease-fire violations (Oberson, 2021). The OSCE's monitoring in Ukraine thus provides an invaluable source of information on the progression of the conflict.

# 4 Data

## 4.1 Ceasefire Violations

We evaluate Russian use of diversionary escalation in the Ukrainian conflict using data we compiled from systematic daily reports filed by OSCE SMM observers deployed in Donbas

Figure 3: Fighting Area and OSCE Presence in Eastern Ukraine




*Note*: The top panel shows the Donetsk and Luhansk oblasts, which broadly comprise the Donbas. The red line marks the Line of Contact between Ukrainian forces (west and north of the line) and Russian-backed forces (east and south of the line) in the period from 2015–2022. The bottom panel plots locations of the OSCE SMM, including key bases and sensor towers used for data collection.

between April 15, 2015–February 22, 2022.[6] Using multi-source inputs from ground patrols, sensor towers, and drone overflights, SMM monitoring teams compiled a daily record of ceasefire violations. We assemble these daily reports, which include information about the geolocation, timing, tactic of violence, and responsible party.[7] Overall, our dataset of OSCE-recorded ceasefire violations includes 250,001 exchanges of fire between Ukrainian and Russian-backed forces, including 1.87 million individual ceasefire violations. We map these violations in Figure 4, and plot variation overtime in Figure 5. The process of collecting and organizing the ceasefire violation data is described in detail in the Data Appendix. We view this newly compiled dataset as a major contribution of this project, and will release the data publicly upon publication.

#### 4.1.1 Identifying Escalatory Behavior

We also recognize that a large proportion of violence in conflict is reactionary and reciprocal. In many cases, one side retaliates after their opponent initiates an attack, or initiates an attack in anticipation of opponent-initiated violence. This explains the high degree of interdependence between Russian and Ukrainian-initiated ceasefire violations that we record (Pearson's $\rho = 0.655$). To isolate the variation in violations which theoretically may be driven by diversionary incentives from the violations we expect in response to opposition violence, we perform an additional set of analyses using the residuals from regressions of Russian violations on the remaining, non-Russian violations, and vice versa, as our response variables. This allows us to examine the impact of protests on violations which are not explained by reciprocal firing along the line of contact.

In addition to using residuals to identify violations beyond what would be expected from

[6] SMM observers were withdrawn from the region on February 23, 2022, one day prior to the full-scale Russian invasion.

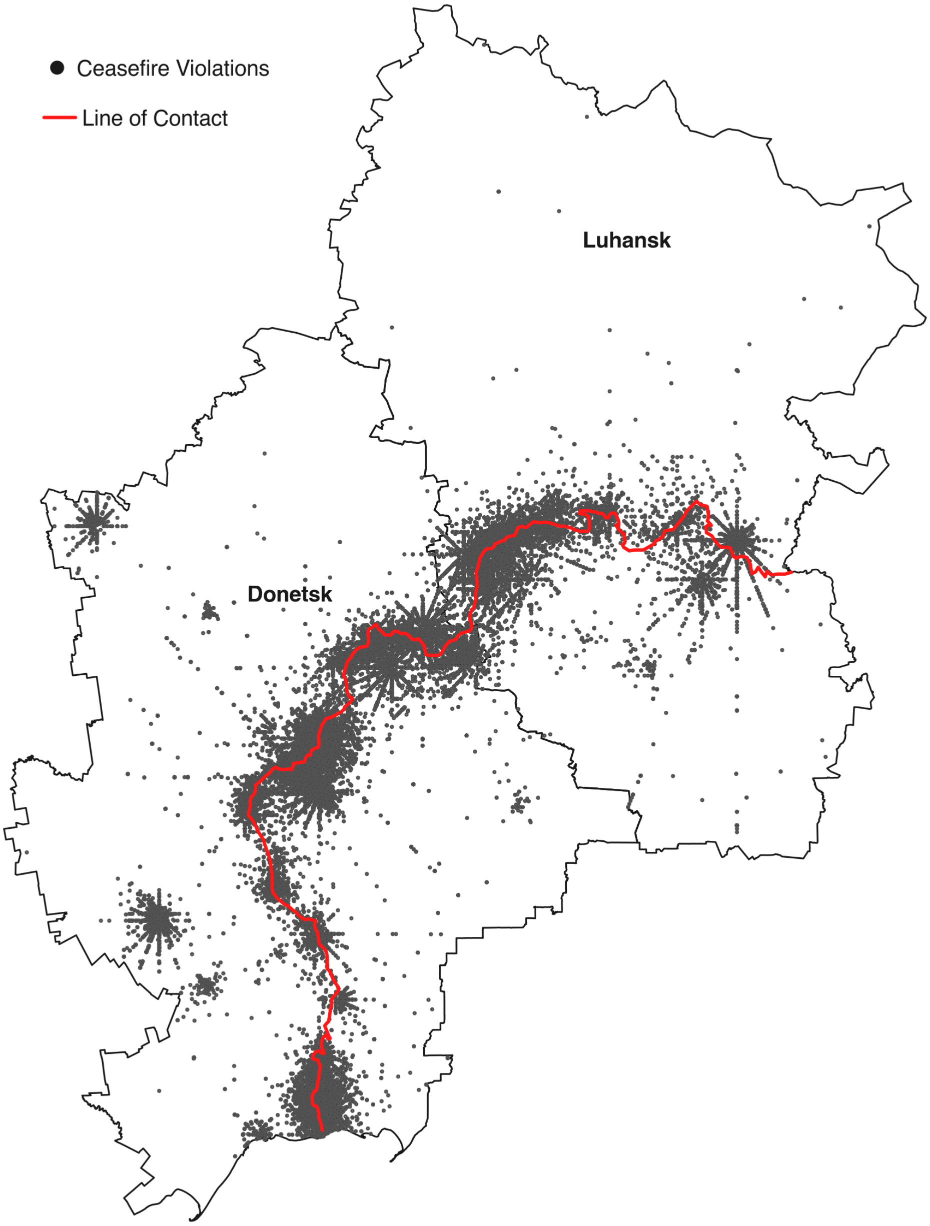
[7] Following the OSCE SMM, we attribute strikes using information on the direction of fires. Because the line of control was stable over the 2015–2022 period, this information allows us to identify the belligerent party responsible for each incident. In total we attribute roughly 90% of incidents to Ukrainian or Russian-backed forces.

simple reactionary conflict, we also separate the ceasefire violations to identify types that are more extreme, and thus more likely to be a diversionary rather than low-level, individual-driven violence. Specifically, we use the violation typology to create two additional measures of escalatory violence: non-gunshot ceasefire violations and explosions. Non-gunshot ceasefire violations include all types of violations except for gunshots. We would expect gunshots to typically reflect individual actions, when not accompanied by other types of attacks, and therefore they are the least likely to be driven by a broader diversionary strategy. The explosions-only category is even more restrictive, allowing only violations recorded as explosions into the analysis. While the explosions category may miss some other types of potentially diversionary violence, we expect that violations involving explosions are the most likely to be selected as part of a strategy to draw attention, and we use it as a secondary measure of escalatory violations.

## 4.2 Protests

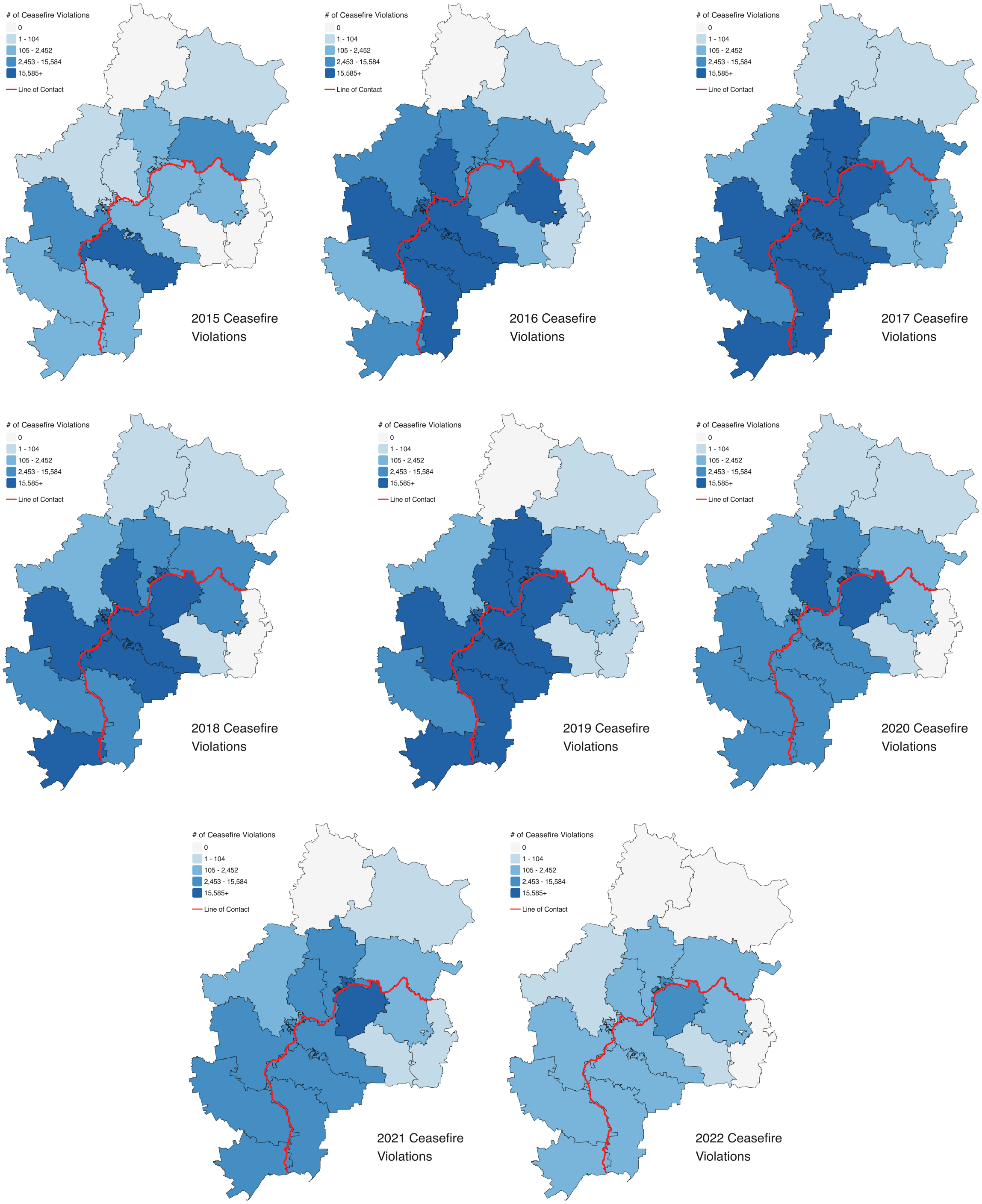
To measure protest activity in Russia, we use records from the LAruPed dataset, which documents protests within Russia directed against the ruling party from 2007–2016 (Lankina, 2018). These records are sourced from namarsh.ru, a site which aggregates news and records of Russian anti-government protests. For the purposes of this study, LAruPED provides a number of advantages over other protest data sources. The namarsh.ru site collected protest records from "regional correspondents, press and online reports", and other online and social media sites, making it less susceptible to overlooking protests which aren't covered in international news media, which are the sole source of many other protest datasets (Lankina and Tertytchnaya, 2020). There is no limit on the number of protesters which much participate for an event to be included, further adding to the comprehensive nature of these records. LAruPED protests also only include events which aren't sponsored by the government, making this well-suited to studying the government's reaction to shows of discontent (Lankina

Figure 4: Ceasefire Violations in Eastern Ukraine, 2015–2022




*Note*: The plot shows the geolocation of each ceasefire violation in Donbas that we record based on OSCE reports.

Figure 5: Annual Variation in Ceasefire Violations




*Note*: Each plot shades raions within the Donetsk and Luhansk oblasts by the intensity of ceasefire violations in each year. Data from 2015 cover April 15 – December 31, 2015. Data from 2022 cover January 1 – February 22, 2022.

Codebook, 2018).

In order to expand the time frame of our analysis beyond the 2016 date provided in the LAruPED data, in early 2022 we worked to collect additional protest data from namarsh.ru. However, there are periods of missing data after late 2018, and the namarsh.ru site has been taken down since the start of Russia's invasion of Ukraine, so we are currently unable to fill in the missing information. To confirm our results are not an artifact of missing data, we perform our main analysis using data from April 2015 through October 2018, which has uninterrupted high-quality data, and perform a secondary analysis using additional selected time frames with selected timeframes of high fidelity data through February 2021.[8] Figure 6 displays the full range of LAruPED data used in this analysis with the primary and supplemental time frames marked with solid and dashed-line boxes, respectively. We perform additional robustness checks using only the original LAruPED datasource, from 2015-2016.

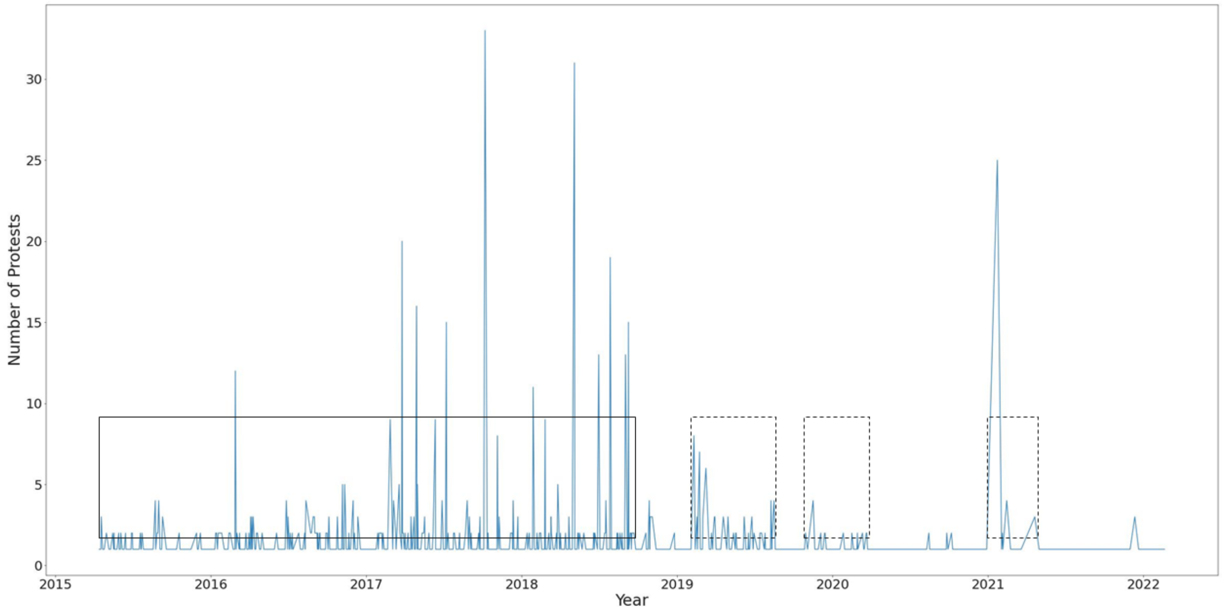
We estimate the effects of Russian protests on violence in the Ukrainian separatist conflict by regressing ceasefire violations on protests clustered by day with month and day of week fixed effects. We perform this analysis using both the preceding four days of protests and the preceding seven days of protests to allow for differential rates in diversionary response after protests.

### 4.3 Russian News Media

While the fine-grained analysis our ceasefire violations and protest datasets allow us can provide strong suggestive evidence of the employment of diversionary violence, tracing the precise mechanisms remains a challenge. One avenue which offers some potential for clarifying the diversionary mechanism is a textual analysis of Russian media, and advances in natural language processing have made text analysis feasible for this extensive corpus of

[8] Our primary date range, bounded with the solid box in Figure 6, is 4/14/2015 - 10/31/2018. The additional time frames selected for the supplemental analysis, represented with the dashed lines in Figure 6, are 2/01/2019 - 08/25/2019, 10/20/2019 - 03/25/2020, and 01/20/2021 - 02/15/2021.

Figure 6: Recorded Protest Events in Selected Timeframes




*Note*: LAruPED and `namash.ru` protest events from April 2015 – February 2022. The solid box identifies our primary analysis data, which has the highest contiguous degree of confidence. The dashed-line boxes identify the additional, high quality timeframes which are included in a supplemental analysis.

articles.

Through INTEGRUM Profi, we obtain the full archive of news articles from 8 Russian news sources: TASS, Vzglyad, RIA Novosti, NEWS.ru, Lenta, Gazeta-Pravda, Meduza.io, and Komsomolskaya Pravda. Of these, NEWS.ru was an independent source, Meduza.io is affiliated with the Russian political opposition, and the remainder are affiliated with the Russian government. We extract the headlines from each of these articles and group them into topics for analysis.

There are many possible methods for generating topic models, with one of the most popular being Latent Dirichlet Allocation (LDA). LDA, however, has shown mixed results in recent years, particularly for noisy datasets, and has some limitations including flexibility of the process and ease of analyzing and refining the output (Egger and Yu, 2022). We instead use BERTopic, a transformer-based topic modeling technique (Grootendorst, 2022).

This technique allows for a high degree of flexibility in each step of the modeling process, provides easy searching of topics related to desired themes, provides for hierarchical topic reduction, and more generally, produces relatively coherent, useful topic outputs.

To model our Russian news archive, we first convert the headlines into embeddings using a multilingual sentence transformers model, perform dimension reduction of these embeddings using Uniform Manifold Approximation and Projection (UMAP), and then perform hierarchical clustering using HDBSCAN (Reimers and Gurevych, 2019; McInnes, Healy and Melville, 2020; Campello, Moulavi and Sander, 2013). Topics were generated with a class-based term frequency inverse document frequency (c-TF-IDF) algorithm, which generates representations of each term within a cluster based on the term's importance in the cluster. All of the processing is done in the original Russian to avoid any negative impact on topic modeling due to translation errors, but the identified topics are translated post-processing to allow for easier evaluation.[9]

The topics produced by the model were manually sorted into two groups: those related to the Ukrainian separatist conflict, and those related to Russian protests. While there is a margin for error in any topic modeling process, which includes both incorporating articles which are not part of a topic and failing to identify articles which should be part of the topic, we believe the resulting topic categorizations should provide a fairly representative sample of both categories.

### 4.3.1 Russian TV/Radio Keyword Analysis

Given the higher consumption of TV and radio than print sources, we also sought to incorporate these sources into the analysis. Due to limitations in access to the transcripts of TV and radio sources, however, we were not able to classify transcripts of Russian TV and radio to obtain topic classifications, as in the print media. Instead, we perform a

[9] There are a small number of articles written in English, but the vast majority are Russian language articles.

carefully-curated keyword search of TV and radio sources using INTEGRUM Profi's search functionality. Specifically, we identify instances inflammatory, anti-Ukrainian terminology which the Russian government frequently uses to justify its behavior.[10] We obtain daily counts of the occurrence of each of these terms across all TV and radio sources, as well as among a selected set of the most-watched pro-Government TV and radio sources.[11] We use these counts of occurrences in our TV and radio analyses below.

# 5 Results

## 5.1 Primary Data Range: April 2015–October 2018

### 5.1.1 Total Ceasefire Violations

We find support for our hypothesis of diversionary conflict with both the preceding four and preceding seven days of protests contributing significant, positive impacts on the total number of ceasefire violations in the Ukrainian separatist conflict (Table 1). Since both the violations and protest data distributions contain only positive values and are right skewed (Empirical Appendix Figures A-6, A-8), we can also consider the same analysis using the natural log of our values and find even stronger results regressing the natural log of violations on the natural log of preceding protests. When these results are visualized across an entire seven-day lag period in Figure 7, with the impact of each lagged day's log protests on total log violations displayed, we see a clear pattern of strong, positive effects on violations in the

[10] The list of terms we include is: Banderites, fascists, Nazis, NATO, Russophobia, and genocide in Ukraine. The Russian government frequently refers to those in Ukraine who do not support Russian control as Banderites, referring to Stepan Bandera, a nationalist Ukrainian leader, Nazis, and fascists. It accuses Ukraine of Russophobia and attempting a genocide on Russian-speakers. These Ukrainians which it criticizes are those who want to join NATO, or whom the Russian government accuses NATO of courting.

[11] To obtain the counts for all TV and radio sources, we searched through three of INTEGRUM Profi's predefined lists: (1) Moscow TV and Radio, (2) Regional TV and Radio, and (3) TV and Radio monitoring online. For the counts for pro-Government sources, we searched through Channel One, Ren TV, Rossiya 1, and NTV transcripts.

first three to four days following a protest, with a diminishing effect after day four. This indicates that the presence of protests is connected to an increase in violations in the days following.

Table 1: Total Ceasefire Violations and Protests

April 2015 - October 2018

| | **Ceasefire Violations** | | | |
|---|---|---|---|---|
| | (1) Violations | (2) Violations | (3) Ln Violations | (4) Ln Violations |
| Previous 4 Days Protests | 17.880** | | | |
| | (8.041) | | | |
| Previous 7 Days Protests | | 15.080** | | |
| | | (7.213) | | |
| Previous 4 Days Ln Protests | | | 0.274*** | |
| | | | (0.099) | |
| Previous 7 Days Ln Protests | | | | 0.351*** |
| | | | | (0.128) |
| Observations | 1290 | 1290 | 1290 | 1290 |
| Clusters | 1290 | 1290 | 1290 | 1290 |

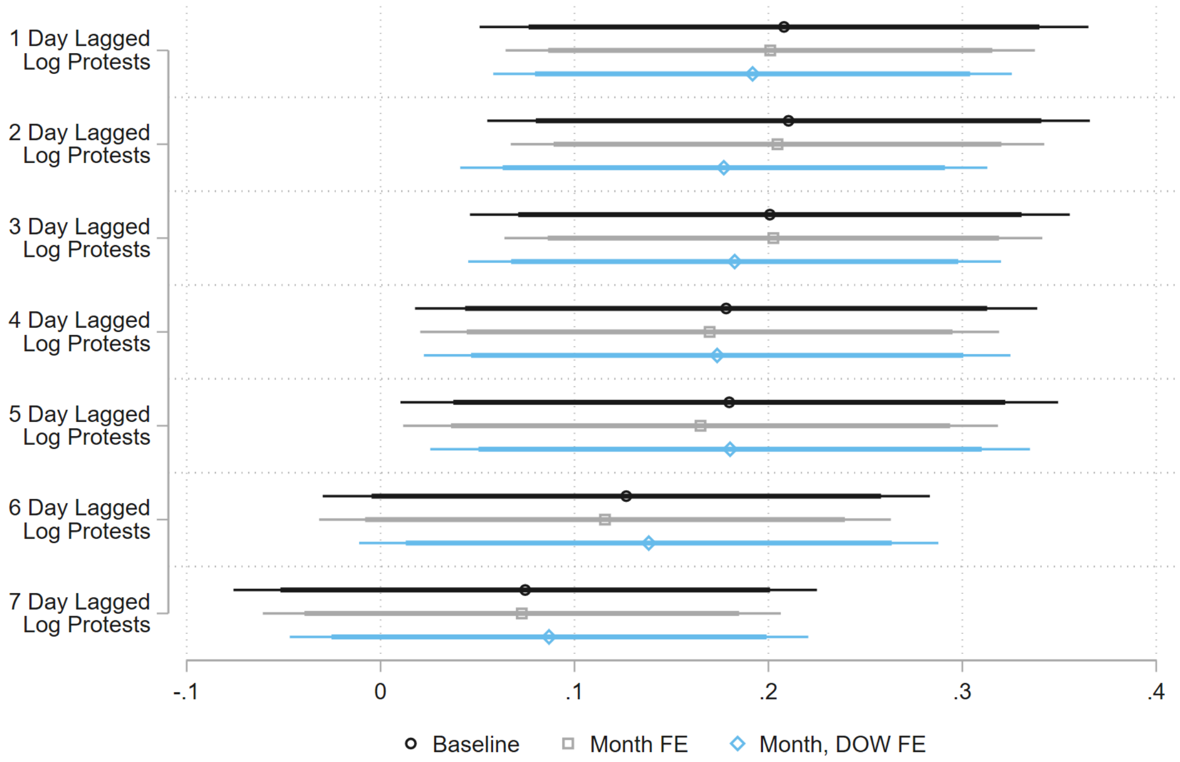
*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Combined ceasefire violations by LAruPED recorded protests in the preceding four (1) and seven (2) days, as well as the natural log of violations by the natural log of the preceding four (3) and seven (4) days of protests.

#### 5.1.2 Escalatory Violations with Attribution

In this context of frequent ceasefire violations by both sides, however, it is likely that many of the violations observed are retaliatory in nature. In order to better isolate violations which

Figure 7: Ceasefire Violations (Ln) by Lagged Protests

April 2015 - October 2018




*Note*: The impact of seven, lagged days of log protests on log total violations, with robust, time-clustered standard errors. Results are shown at the baseline level with no fixed effects (circle), with month fixed effects (square), and with month and day of week fixed effects (diamond).

may be driven by external factors, such as diversionary behavior, we next identify only those violations which (a) exceed the violations expected from retaliation alone and (b) are at a higher level of violence, and so more likely to draw attention.

To isolate these non-retaliatory events which represent additional escalation, we first identified how many violations from a given side, Russian or Ukrainian, were expected based upon the number of preceding violations perpetrated by their opponent. We estimated this using a standard OLS regression of violations by Russians (or Ukrainians) on all non-Russian (or non-Ukrainian) violations, where the non-Russian (or non-Ukrainian) violations are defined as all unattributed violations plus attributed Ukrainian (or Russian) violations. We were able to assign approximately 83 % of ceasefire violations to either Russian or Ukrainian forces, which allows us to generate expectations for retaliatory violations and identify any additional ceasefire violations beyond these expectations as likely driven by external factors.[12] These additional violations are estimated by the standardized residuals from the OLS regressions.[13]

To narrow our focus to escalatory violent behavior, of the kind more likely to be newsworthy and draw attention, we remove all gunshot violations, which represent instances of shots heard from individual weapons. This should allow us to better identify diversionary violence, designed to generate a response, for which attacks of higher severity such as with artillery or other heavy weaponry are more appropriate than individuals firing personal weapons.

Incorporating both of these components, attribution and escalatory violations, provides a measure closer to violence that is diversionary in nature than a simple measure of total ceasefire violations. To consider whether Russian protests may impact this potentially-diversionary behavior, we regress protests on Russian standardized residual violations, excluding gunshots, and we find significant positive impacts of both four and seven preceding days of protests (Table 2). This indicates that anti-government protests in Russia are correlated with an increase in high-impact, Russian-initiated violations in the Ukrainian east.

Importantly, for this theory of diversionary escalation to hold, we cannot see a similar pattern in Ukrainian-initiated, non-gunshot violations. In line with these expectations, when

[12] We are able to assign Russian or Ukrainian attribution to 82.99 % of violations: of the 1,855,570 violations recorded in our dataset, 741,475 are attributed to Russians and 798,498 are attributed to Ukrainians. While we have no reason to suspect the attribution of strikes follows some pattern, and thus we should be able to use differentiated Russian and Ukrainian strikes as a random sample of the total in our analysis, it is preferable to include both analyses with total violations and analyses with differentiated Russian and Ukrainian strikes for robustness.

[13] More precisely, to calculate the Russian standardized residuals, we first regress the log of Russian airstrikes on the log of the sum of Ukrainian and unknown airstrikes, then collect the residuals from this regression. Finally, to ease in the interpretation of these residuals, we standardize them (mean of 0, standard deviation of 1). Ukrainian standardized residuals are created with the same process, regressed on the sum of Russian and unknown airstrikes.

Table 2: Russian Residual Non-Gunshot Ceasefire Violations and Protests

April 2015 - October 2018

| | Russian Residual Non-Shot Ceasefire Violations | |
|---|---|---|
| | (1)<br>Ln Violations | (2)<br>Ln Violations |
| Previous 4 Days Ln Protests | 0.129**<br>(0.063) | |
| Previous 7 Days Ln Protests | | 0.167**<br>(0.081) |
| Observations | 1267 | 1267 |
| Clusters | 1267 | 1267 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Russian residual non-gunshot ceasefire violations (standardized residuals generated by regressing the natural log of Russian non-gunshot violations on the natural log of non-Russian non-gunshot violations) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests.

we perform a parallel analysis of Russian protest impact on Ukrainian escalatory violence, we find no significant pattern (Table 3). This confirms a difference in behavior between the Russian-backed separatists and the Ukrainians in line with the expectations of the diversionary theory, as there is no theoretical basis for an increase in Ukrainian escalatory behavior due to Russian protests.

As an alternate, even more conservative specification of the kind of escalatory violence we expect in diversionary behavior, we perform a series of robustness checks using only explosion-type violations. These are a more extreme type of violation than many others still included in our non-gunshot analysis, and while we hypothesize they may not be the only type of violation used for escalation, attacks resulting in explosions are more likely than most other types to indicate escalation.[14] Using only explosions, we find even stronger evidence linking Russian protests with Russian-initiated explosions, while still finding no evidence

[14]The use of OSCE-recorded explosions as a primary measure of violence in the separatist conflict is in line with reporting from groups like Crisis Group, who also employ the OSCE ceasefire monitoring and focus on explosion-type violations in their reports (*Conflict in Ukraine's Donbas: A Visual Explainer*, 2022).

linking Russian protests with Ukrainian-initiated explosions (Appendix Tables A-3, A-4).

Overall, these results which consider more extreme ceasefire violations separated by initiator provide even stronger suggestive evidence for diversionary proxy behavior on the part of the Russian-backed troops.

Table 3: Ukrainian Residual Non-Gunshot Ceasefire Violations and Protests

April 2015 - October 2018

| | **Ukrainian Residual Non-Shot Ceasefire Violations** | |
|---|---|---|
| | (1)<br>Ln Violations | (2)<br>Ln Violations |
| Previous 4 Days Ln Protests | 0.049<br>(0.077) | |
| Previous 7 Days Ln Protests | | 0.114<br>(0.093) |
| Observations | 1270 | 1270 |
| Clusters | 1270 | 1270 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Ukrainian residual non-gunshot ceasefire violations (standardized residuals generated by regressing the natural log of Ukrainian non-gunshot violations on the natural log of non-Ukrainian non-gunshot violations) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests.

## 5.2 Extended Data Range: April 2015–February 2021

To perform additional robustness testing on our data, we next consider the extended date range from April 2015 through February 2021, with some date ranges removed due to lack of protest data. As we are unaware whether this data is missing according to any pattern, and therefore whether the sample of dates we do possess can be considered a representative, random sample, we treat these findings as supplemental to our previous analyses, rather than stand-alone results. Broadly, they do support our previous findings, with the relationship between protests and total violations significant and positive for both the previous 4 and 7 days of protests (Table 4).

Table 4: Total Ceasefire Violations and Protests

April 2015 - February 2021, Selected Dates

| | Ceasefire Violations | | | |
|---|---|---|---|---|
| | (1) Violations | (2) Violations | (3) Ln Violations | (4) Ln Violations |
| Previous 4 Days Protests | 15.109* (7.754) | | | |
| Previous 7 Days Protests | | 11.970* (6.684) | | |
| Previous 4 Days Ln Protests | | | 0.179** (0.081) | |
| Previous 7 Days Ln Protests | | | | 0.244** (0.106) |
| Observations | 1680 | 1680 | 1680 | 1680 |
| Clusters | 1680 | 1680 | 1680 | 1680 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Combined ceasefire violations by LAruPED recorded protests in the preceding four (1) and seven (2) days, as well as the natural log of violations by the natural log of the preceding four (3) and seven (4) days of protests. Results given for the extended time period, which includes April 2015 - October 2018 with three additional time frames between 2018 and 2021 (see Figure 6).

The Russian- and Ukrainian-attributed escalatory behavior also follow a similar pattern over these expanded dates, though it also weaker than the main 2015–2018 time frame, as the Russian-initiated violence only shows a significant relationship with Russian protests when considering explosions only (Appendix Tables A-9, A-10). Ukrainian-initiated violence continues to show no relationship with protests in all specifications (Appendix Tables A-11, A-12).

## 5.3 Russian News Media Analysis

Analysis of Russian ceasefire violations thus provide strong suggestive evidence for Russian use of violence as a diversionary tactic, though it is difficult to confirm each link in the proposed mechanism given lack of access to the Russian leadership's discussions and decision rationale. One source which is available to strengthen the evidence for diversionary behavior, however, is the Russian media, both TV and print. According to the theory's expectations, we should see both the Russian government using media to divert attention from protests, complementing the use of violence, and the Russian government responding to unfavorable media coverage with violent diversions.

First, using protest-related articles identified with topic classification, we consider the relationship between protest articles and Russian ceasefire violations. Reports of protests within Russia reflect unfavorably on the government, and a theory of diversionary violence would expect to see greater escalatory violence after protest reporting. We find our theoretical expectation holds, with a positive, significant relationship between protest articles and Russian-initiated violations (Table 5).

We perform additional robustness checks using Russian residual explosions, as an even more extreme form of violation, and using only protest articles in news sources run by government opponents, which are likely to be the articles most concerning to the government. Both of these tests show an even stronger relationship between protest reporting and escalatory violence, as the theory would predict by focusing on higher-attention violence and competition with political opposition, respectively (Appendix Tables A-7, A-8).

Given the greater prevalence of TV media over print, we also test the theory's expectations using the TV and radio content identified through keyword analysis. Again aligned with the theory's expectations, we find an increase in usage of inflammatory, anti-Ukrainian terminology in Russian media following domestic protests, indicating the government may

Table 5: Russian Residual Non-Gunshot Ceasefire Violations by Protest Articles

April 2015 - October 2018

| | **Russian Non-Shot Violations** | |
|---|---|---|
| | (1) | (2) |
| | Ln Ru Non-Shot Viols | Ln Ru Non-Shot Viols |
| Prev 4 Days LnProtest Articles | 0.245** | |
| | (0.096) | |
| Prev 7 Days LnProtest Articles | | 0.355** |
| | | (0.147) |
| Observations | 1207 | 1186 |
| Clusters | 1207 | 1186 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Russian residual non-gunshot ceasefire violations (standardized residuals generated by regressing the natural log of Russian no-shot violations on the natural log of non-Russian no-shot violations) by Russian protest-related articles (obtained with BERTopic topic modeling).

be attempting to divert attention towards a common enemy in the Ukrainians (Table 6).[15]

# 6 Conclusion

These results are suggestive of a pattern of behavior in which protests in the Russian homeland may be linked to an increase in violence committed by pro-Russian troops acting as proxies. This supports the diversionary violence literature which suggests long-standing, territorial contests are often selected as the targets to heighten citizen responsiveness and approval of the conflict: given the persistent turmoil within Russia over the identity of Ukraine, many Russians and Ukrainians already had a heightened sensitivity to issues surrounding control of Ukraine. This sensitivity allowed for a stronger response to and support of the violence, in the eyes of the Russian leaders, and may have contributed to a calculation

[15] References to 'Nazis' and 'Nazism' were by far the most common among all the anti-Ukrainian terminology we tested, and results hold using only 'Nazi' references as the outcome variable.

Table 6: Anti-Ukrainian Inflammatory Language in TV/Radio by Protest Occurrence

April 2015 - October 2018

| | **Anti-Ukraine Content** | | | |
|---|---|---|---|---|
| | (1) Anti-Ukraine TV/Radio | (2) Anti-Ukraine TV/Radio | (3) Ln Anti-Ukraine TV/Radio | (4) Ln Anti-Ukraine TV/Radio |
| Previous 4 Days Protest | 3.677** | | | |
| | (1.659) | | | |
| Previous 7 Days Protest | | 3.313*** | | |
| | | (1.241) | | |
| Previous 4 Days Ln Protest | | | 0.057*** | |
| | | | (0.019) | |
| Previous 7 Days Ln Protest | | | | 0.074*** |
| | | | | (0.022) |
| Observations | 1297 | 1297 | 1297 | 1297 |
| Clusters | 1297 | 1297 | 1297 | 1297 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Absolute daily count of inflammatory, anti-Ukrainian terminology across Russian TV and Radio sources by the previous 4 and 7 days of protests within Russia.

that diversionary tactics might prove effective.

This paper further contributes to the debate in the literature on the kinds of turmoil which can lead to diversionary behavior, as it provides evidence that domestic unrest may be a driver of diversionary violence. While negative economic conditions and physical unrest are both commonly used as indicators of domestic discontent in the diversionary violence literature, our results support protests at least as potential diversionary triggers. In fact, our study provides these results at such a temporal granularity that economic conditions would not have the time to change, and thus would not be able to influence the diversionary patterns seen in our violations.

These results are suggestive of a pattern of diversionary behavior by proxy, a novel finding at the intersection of the diversionary violence and proxy warfare literature. Future research should perform thorough qualitative analyses, such as case studies, to track whether the increased ceasefire violations are truly the result of commands from Russian leadership designed to divert civilian attention. This could be expanded to examine which actors may be engaging in diversionary behavior; in particular, whether there is evidence of the hypothesized link in pro-Russian commanders, or whether it may arise from troop culture itself.

Future research should also include Russian television sources as an additional set of media data for text analysis, as television is an even more important mechanism for the Russian government to share information with – and shape the opinion of – Russian citizens.

While the data we develop and employ here consist of the period prior to Russia's February 2022 invasion of Ukraine, and not to the full-scale conflict itself, research using our developed ceasefire violations dataset has the possibility to provide numerous insights into motivations and features which may be hidden in macro-level datasets or unique to Russian circumstances. In particular, an examination of Russian behavior prior to its full-scale invasion of Ukraine may be able to provide much-needed insight into the factors which spurred Russia to escalate its proxy conflicts in the Donbas into a direct confrontation for the whole of Ukraine. We hope the novel ceasefire violations dataset presented here to provide a useful aid into this and other avenues of future research.

# References

Alrababa'h, Ala' and Lisa Blaydes. 2021. "Authoritarian media and diversionary threats: lessons from 30 years of Syrian state discourse." *Political Science Research and Methods* 9(4):693–708.

Angeletos, George-Marios, Christian Hellwig and Alessandro Pavan. 2006. "Signaling in a Global Game: Coordination and Policy Traps." *Journal of Political Economy* 114(3):452–484.

Angeletos, George-Marios, Christian Hellwig and Alessandro Pavan. 2007. "Dynamic Global Games of Regime Change: Learning, Multiplicity, and the Timing of Attacks." *Econometrica* 75(3):711–756.

Arel, Dominique and Jesse Driscoll. 2024. *Ukraine's Unnamed War: Before the Russian Invasion of 2022*. Cambridge University Press.

Baev, Pavel K. 2019. The Impacts of the Syrian Intervention on Russian Strategic Culture. Technical Report 033 George C. Marshall European Center for Security Studies.
**URL:** *http://www.marshallcenter.org/en/publications/security-insights/impacts-syrian-intervention-russian-strategic-culture-0*

Berman, Eli and David A. Lake, eds. 2019. *Proxy Wars: Suppressing Violence through Local Agents*. Cornell University Press.

Bohush, Denis, Oleksandra Baglai, Dace Kundrate, Vita Terauda, Linda Curika and Central and Eastern European Online Library. 2016. "Russian Propaganda Concerning Ukraine During the Syrian Campaign: an Innovative Approach to Assess Information Activities.".

Borshchevskaya, Anna. 2015. "Russia's Syria Propaganda." *Forbes* . Section: World Affairs.

Bueno de Mesquita, Ethan. 2010. "Regime Change and Revolutionary Entrepreneurs." *American Political Science Review* 104(3):446–466.

Bueno de Mesquita, Ethan and Mehdi Shadmehr. 2023. "Rebel Motivations and Repression." *American Political Science Review* 117(2):734–750.

Campello, Ricardo J. G. B., Davoud Moulavi and Joerg Sander. 2013. Density-Based Clustering Based on Hierarchical Density Estimates. In *Advances in Knowledge Discovery and Data Mining*, ed. Jian Pei, Vincent S. Tseng, Longbing Cao, Hiroshi Motoda and Guandong Xu. Lecture Notes in Computer Science Berlin, Heidelberg: Springer pp. 160–172.

Carlsson, Hans and Eric Van Damme. 1993. "Global games and equilibrium selection." *Econometrica* 61(5):989–1018.

Carter, Erin Baggot. 2020. "Diversionary cheap talk: economic conditions and US foreign policy rhetoric, 1945-2010." *International Interactions* 46(2):163–198.

Chatagnier, J Tyson. 2012. "The effect of trust in government on rallies' round the flag." *Journal of Peace Research* 49(5):631–645.

Chiozza, Giacomo and Henk E. Goemans. 2004. "Avoiding Diversionary Targets." *Journal of Peace Research* 41(4):423–443. Publisher: SAGE Publications Ltd.

Clark, David and Patrick Regan. 2016. "Mass Mobilization Protest Data.". Edition: V5 Section: 2016-01-13 12:06:37.262.

*Conflict in Ukraine's Donbas: A Visual Explainer*. 2022.
**URL:** *https://www.crisisgroup.org/content/conflict-ukraines-donbas-visual-explainer*

Correa, Sofía, Gaétan Nandong and Mehdi Shadmehr. 2024. "Grievance shocks and coordination in protest." *American Journal of Political Science* .

Durante, Ruben and Ekaterina Zhuravskaya. 2018. "Attack When the World Is Not Watching? US News and the Israeli-Palestinian Conflict." *Journal of Political Economy* 126(3):1085–1133.

Edmond, Chris. 2013. "Information manipulation, coordination, and regime change." *Review of Economic Studies* 80(4):1422–1458.

Egger, Roman and Joanne Yu. 2022. "A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts." *Frontiers in Sociology* 7:886498.

Filippov, Mikhail. 2009. "Diversionary Role of the Georgia-Russia Conflict: International Constraints and Domestic Appeal." *Europe-Asia Studies* 61(10):1825–1847. Publisher: Routledge.

Fravel, M. Taylor. 2010. "The Limits of Diversion: Rethinking Internal and External Conflict." *Security Studies* 19(2):307–341.

*GEC Special Report: Russia's Pillars of Disinformation and Propaganda*. 2020. Technical report US Department of State.
**URL:** *https://www.state.gov/russias-pillars-of-disinformation-and-propaganda-report/*

Gelpi, Christopher. 1997. "Democratic Diversions: Governmental Structure and the Externalization of Domestic Conflict." *Journal of Conflict Resolution* 41(2):255–282.

Gomza, Ivan. 2022. "The War in Ukraine: Putin's Inevitable Invasion." *Journal of Democracy* 33(3):23–30.

Gowa, Joanne. 1998. "Politics at the Water's Edge: Parties, Voters, and the Use of Force Abroad." *International Organization* 52(2):307–324.

Grootendorst, Maarten. 2022. "BERTopic: Neural topic modeling with a class-based TF-IDF procedure." *arXiv preprint arXiv:2203.05794* .

Götz, Elias. 2017. "Putin, the State, and War: The Causes of Russia's Near Abroad Assertion Revisited." *International Studies Review* 19(2):228–253.

Hale, Henry E. 2022. "Authoritarian Rallying as Reputational Cascade? Evidence from Putin's Popularity Surge after Crimea." *American Political Science Review* 116(2):580–594.

Haynes, Kyle. 2016. "Diversity and Diversion: How Ethnic Composition Affects Diversionary Conflict." *International Studies Quarterly* 60(2):258–271.

Haynes, Kyle. 2017. "Diversionary conflict: Demonizing enemies or demonstrating competence?" *Conflict Management and Peace Science* 34(4):337–358.

Jackson, Matthew O. and Massimo Morelli. 2007. "Political Bias and War." *American Economic Review* 97(4):1353–1373.

Kramer, Andrew E. 2022. "What Are the Minsk Accords, and Could They Defuse the Ukraine Crisis?" *The New York Times* .
**URL:** *https://www.nytimes.com/2022/02/08/world/russia-ukraine-minsk-accords.html*

Lankina, Tomila and Katerina Tertytchnaya. 2020. "Protest in electoral autocracies: a new dataset." *Post-Soviet Affairs* 36(1):20–36.

Lankina, Tomila V. 2018. "Lankina Russian protest event dataset.".

Leeds, Brett Ashley and David R. Davis. 1997. "Domestic Political Vulnerability and International Disputes." *Journal of Conflict Resolution* 41(6):814–834.

Levy, Jack. 1989. The Diversionary Theory of War: A Critique. In *Handbook of War Studies.* Boston: Unwin Hyman pp. 259–288.

Mankoff, Jeffrey. 2022. "Russia's War in Ukraine: Identity, History, and Conflict.".
**URL:** *https://www.csis.org/analysis/russias-war-ukraine-identity-history-and-conflict*

Marten, Kimberly. 2015. "Putin's Choices: Explaining Russian Foreign Policy and Intervention in Ukraine." *The Washington Quarterly* 38(2):189–204.

McInnes, Leland, John Healy and James Melville. 2020. "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction." *ArXiv e-prints* .

McLaughlin Mitchell, Sara and Clayton L. Thyne. 2010. "Contentious Issues as Opportunities for Diversionary Behavior." *Conflict Management and Peace Science* 27(5):461–485.

Mehrl, Marius and Ioannis Choulis. 2021. "Diversionary Politics and Territorial Disputes: Evidence from Turkish Airspace Incursions." *Foreign Policy Analysis* 17(3):orab007.

Mercer, Jonathan. 2005. "Rationality and Psychology in International Politics." *International Organization* 59(1):77–106.

Miller, Ross A. 1995. "Domestic Structures and the Diversionary Use of Force." *American Journal of Political Science* 39(3):760–785.

Mitchell, Sara McLaughlin and Brandon C. Prins. 2004. "Rivalry and Diversionary Uses of Force." *Journal of Conflict Resolution* 48(6):937–961.

Mitchell, Sara McLaughlin and Clayton L. Thyne. 2010. "Contentious Issues as Opportunities for Diversionary Behavior." *Conflict Management and Peace Science* 27(5):461–485.

Morris, Stephen and Hyun Song Shin. 1998. "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks." *American Economic Review* 88(3):587–597.

Mueller, John E. 1973. *War, Presidents, and Public Opinion.* John Wiley.

Nardelli, Alberto, Jennifer Rankin and George Arnett. 2015. "Vladimir Putin's approval rating at record levels." *The Guardian* .
**URL:** *https://www.theguardian.com/world/datablog/2015/jul/23/vladimir-putins-approval-rating-at-record-levels*

Oakes, Amy. 2006. "Diversionary War and Argentina's Invasion of the Falkland Islands." *Security Studies* 15(3):431–463.

Oberson, Frederic. 2021. OSCE Special Monitoring in Ukraine. In *Between Peace and Conflict in the East and the West: Studies on Transformation and Development in the OSCE Region*, ed. Anja Mihr. Cham: Springer International Publishing pp. 47–75.

*Permanent Council Decision No. 1117*. 2014.
**URL:** *https://www.osce.org/pc/116747*

Pifer, Steven. 2019. "Ukraine: Looking forward, five years after the Maidan Revolution.".
**URL:** *https://www.brookings.edu/blog/order-from-chaos/2019/02/22/ukraine-looking-forward-five-years-after-the-maidan-revolution/*

Pifer, Steven. 2020. "Crimea: Six years after illegal annexation.".
**URL:** *https://www.brookings.edu/blog/order-from-chaos/2020/03/17/crimea-six-years-after-illegal-annexation/*

Raleigh, Clionadh, Andrew Linke, Håvard Hegre and Joakim Karlsen. 2010. "Introducing ACLED: An Armed Conflict Location and Event Dataset: Special Data Feature." *Journal of Peace Research* 47(5):651–660.

Reimers, Nils and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing.* Association for Computational Linguistics.

Richards, Diana, T. Clifton Morgan, Rick K. Wilson, Valerie L. Schwebach and Garry D. Young. 1993. "Good Times, Bad Times, and the Diversionary Use of Force: A Tale of Some Not-So-Free Agents." *Journal of Conflict Resolution* 37(3):504–535.

Smith, Alastair. 1996. “Diversionary Foreign Policy in Democratic Systems.” *International Studies Quarterly* 40(1):133–153.

Stoner, Kathryn and Michael McFaul. 2015. “Who Lost Russia (This Time)? Vladimir Putin.” *The Washington Quarterly* 38(2):167–187.

Stukal, Denis, Sergey Sanovich, Richard Bonneau and Joshua A. Tucker. 2022. “Why Botter: How Pro-Government Bots Fight Opposition in Russia.” *American Political Science Review* 116(3):843–857.

Tarar, Ahmer. 2006. “Diversionary Incentives and the Bargaining Approach to War.” *International Studies Quarterly* 50(1):169–188.

Theiler, Tobias. 2018. “The Microfoundations of Diversionary Conflict.” *Security Studies* 27(2):318–343.

Tir, Jaroslav. 2010. “Territorial Diversion: Diversionary Theory of War and Territorial Conflict.” *Journal of Politics* 72(2):413–425.

Treisman, Daniel. 2011. “Presidential Popularity in a Hybrid Regime: Russia under Yeltsin and Putin.” *American Journal of Political Science* 55(3):590–609.

Tyson, Scott A and Alastair Smith. 2018. “Dual-Layered Coordination and Political Instability: Repression, Co-optation, and the Role of Information.” *The Journal of Politics* 80(1):44–58.

Weidmann, Nils B. and Espen Geelmuyden Rød. 2019. *The Internet and Political Protest in Autocracies*. Oxford University Press.

# SUPPLEMENTAL APPENDIX [FOR ONLINE ONLY]

Natalie Ayers, Christopher W. Blair, Joseph J. Ruggiero, Austin L. Wright, and Konstantin Sonin

## Contents



## List of Figures and Tables

Table/figure numbers are noted in left column. Page numbers are noted in the right column.

---

**Supplemental Figures**

---

## Supplemental Tables

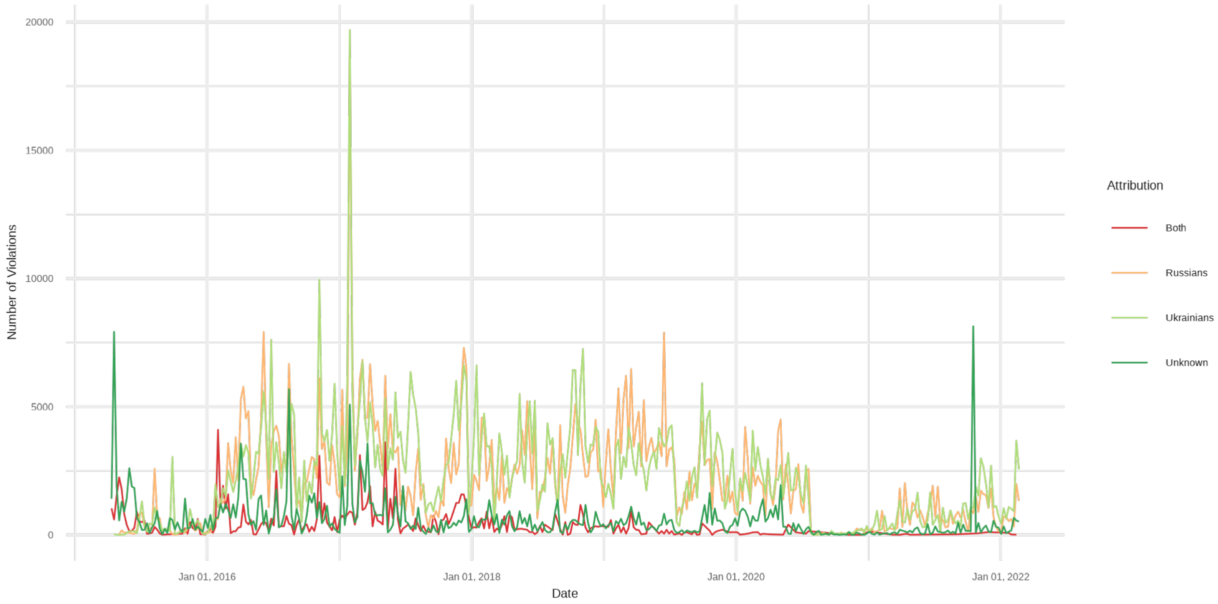
# A Empirical Appendix

In this brief empirical appendix, we introduce supplemental results.

## A.1 Descriptive Data Analysis

Figure A-1: Weekly Ceasefire Violations

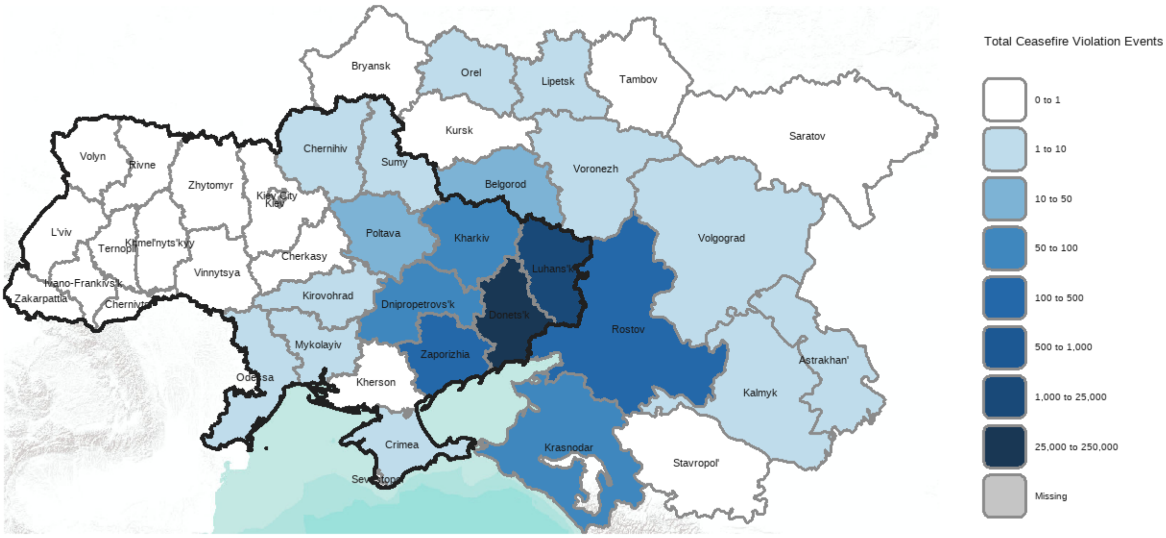
April 2015 - February 2022




*Note*: Count of weekly ceasefire violations for the entire dataset, from April 2015 – February 2022. Colored by group to which the violation is attributed.

Figure A-2: Total Recorded Ceasefire Violation Events

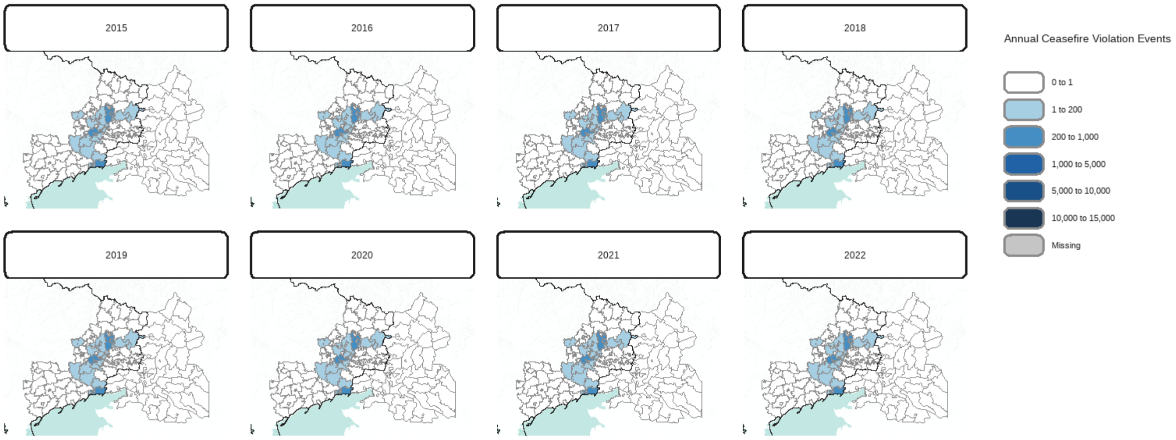
April 2015 - February 2022




*Note*: Count of total recorded ceasefire violation events in Ukraine and Russia by Admin 1 area. Events span from April 2015 – February 2022. The count represents the number of distinct violation events recorded, not the summed number of violations across all events (eg, a record identifying 10 bursts is counted as a single event for the purpose of this map).

Figure A-3: Annual Recorded Ceasefire Violation Events in Core Fighting Areas

April 2015 - February 2022




*Note*: Count of recorded ceasefire violation events annually by Admin 2 (Ukraine) or Admin 3 (Russia) area in the core zones of conflict, defined as admin 1 areas with 100+ ceasefire violations over the entire time period. Events span from April 2015 – February 2022. The count represents the number of distinct violation events recorded, not the summed number of violations across all events (eg, a record identifying 10 bursts is counted as a single event for the purpose of this map).

Table A-1: Total Yearly Ceasefire Violations by Type and Attribution

April 2015 - Feb 2022

| Year | Attribution | All Types | Explosion | Rocket | Projectile | Salvo | Round |
|---|---|---|---|---|---|---|---|
| 2015 | Both | 14328 | 11476 | 0 | 0 | 0 | 346 |
| | Russians | 10410 | 4747 | 19 | 1 | 0 | 1733 |
| | Ukrainians | 13309 | 4745 | 0 | 0 | 2 | 1162 |
| | Unknown | 32444 | 16487 | 0 | 0 | 26 | 1523 |
| 2016 | Both | 31811 | 12134 | 3 | 0 | 0 | 1616 |
| | Russians | 145780 | 47626 | 483 | 943 | 50 | 4835 |
| | Ukrainians | 151073 | 64023 | 867 | 1321 | 39 | 3075 |
| | Unknown | 49663 | 17668 | 16 | 4 | 2 | 2226 |
| 2017 | Both | 41264 | 6986 | 0 | 5 | 9 | 0 |
| | Russians | 178708 | 65106 | 674 | 13267 | 62 | 433 |
| | Ukrainians | 202678 | 77622 | 1183 | 11314 | 20 | 444 |
| | Unknown | 44903 | 20804 | 142 | 1361 | 58 | 321 |
| 2018 | Both | 16686 | 3284 | 0 | 0 | 0 | 0 |
| | Russians | 140490 | 36525 | 0 | 32448 | 0 | 1388 |
| | Ukrainians | 174347 | 35770 | 0 | 47156 | 1 | 1791 |
| | Unknown | 25507 | 7323 | 0 | 2114 | 6 | 25 |
| 2019 | Both | 6928 | 1177 | 0 | 14 | 0 | 0 |
| | Russians | 156582 | 29179 | 0 | 56770 | 0 | 1957 |
| | Ukrainians | 142297 | 39792 | 0 | 42327 | 0 | 1226 |
| | Unknown | 22798 | 5143 | 0 | 2618 | 0 | 0 |
| 2020 | Both | 1172 | 1366 | 0 | 33 | 0 | 0 |
| | Russians | 63245 | 18323 | 0 | 14442 | 0 | 48 |
| | Ukrainians | 65221 | 19981 | 0 | 20235 | 0 | 136 |
| | Unknown | 18829 | 5925 | 0 | 2021 | 0 | 0 |
| 2021 | Both | 439 | 439 | 0 | 41 | 0 | 0 |
| | Russians | 40350 | 10235 | 0 | 9833 | 0 | 110 |
| | Ukrainians | 45451 | 11729 | 0 | 13608 | 0 | 0 |
| | Unknown | 16285 | 2383 | 0 | 285 | 0 | 0 |
| 2022 | Both | 102 | 102 | 0 | 0 | 0 | 0 |
| | Russians | 6391 | 2973 | 0 | 658 | 0 | 0 |
| | Ukrainians | 11507 | 5603 | 0 | 2354 | 0 | 0 |
| | Unknown | 2501 | 1146 | 0 | 0 | 0 | 0 |

Full collection of ceasefire violations by violation type, aggregated by year and attribution. Full violation type breakdown continued in A-2.

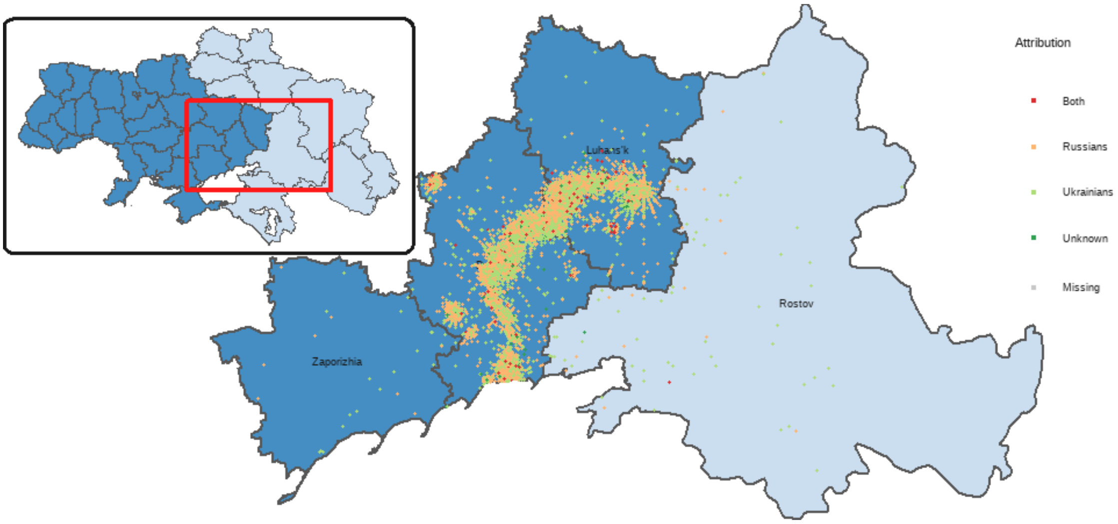
Table A-2: Total Yearly Ceasefire Violations by Type and Attribution, continued

April 2015 - Feb 2022

| Year | Attribution | Impact | Flare | Tracer | Shot | Burst | Flash |
|---|---|---|---|---|---|---|---|
| 2015 | Both | 28 | 0 | 0 | 234 | 1999 | 0 |
| | Russians | 46 | 0 | 0 | 742 | 2635 | 0 |
| | Ukrainians | 36 | 0 | 0 | 3093 | 2819 | 0 |
| | Unknown | 536 | 0 | 0 | 3525 | 4510 | 0 |
| 2016 | Both | 415 | 0 | 56 | 7111 | 10723 | 9 |
| | Russians | 998 | 36 | 4030 | 39555 | 43362 | 450 |
| | Ukrainians | 1750 | 120 | 3077 | 35201 | 35702 | 229 |
| | Unknown | 350 | 19 | 348 | 12589 | 15587 | 107 |
| 2017 | Both | 157 | 0 | 4 | 4285 | 28917 | 0 |
| | Russians | 13 | 154 | 24284 | 32306 | 39291 | 59 |
| | Ukrainians | 6 | 238 | 37246 | 28359 | 42317 | 23 |
| | Unknown | 8 | 17 | 2374 | 11347 | 12494 | 59 |
| 2018 | Both | 140 | 0 | 0 | 4382 | 8953 | 0 |
| | Russians | 6 | 14 | 490 | 25857 | 44356 | 797 |
| | Ukrainians | 69 | 32 | 513 | 35050 | 54926 | 961 |
| | Unknown | 0 | 0 | 3 | 7177 | 7161 | 46 |
| 2019 | Both | 0 | 0 | 0 | 1701 | 3700 | 0 |
| | Russians | 101 | 0 | 1 | 30271 | 39152 | 1213 |
| | Ukrainians | 51 | 0 | 17 | 22014 | 36935 | 1158 |
| | Unknown | 1 | 0 | 0 | 5831 | 5656 | 174 |
| 2020 | Both | 0 | 0 | 0 | 250 | 650 | 0 |
| | Russians | 0 | 0 | 0 | 14408 | 18114 | 922 |
| | Ukrainians | 4 | 0 | 0 | 7048 | 17431 | 993 |
| | Unknown | 1 | 0 | 0 | 5636 | 3246 | 64 |
| 2021 | Both | 0 | 0 | 0 | 0 | 0 | 0 |
| | Russians | 0 | 0 | 0 | 11829 | 8201 | 180 |
| | Ukrainians | 0 | 0 | 0 | 11742 | 7785 | 584 |
| | Unknown | 0 | 0 | 0 | 8037 | 6055 | 18 |
| 2022 | Both | 0 | 0 | 0 | 0 | 0 | 0 |
| | Russians | 0 | 0 | 0 | 1257 | 1322 | 180 |
| | Ukrainians | 0 | 0 | 0 | 1656 | 1829 | 65 |
| | Unknown | 0 | 0 | 0 | 468 | 780 | 7 |

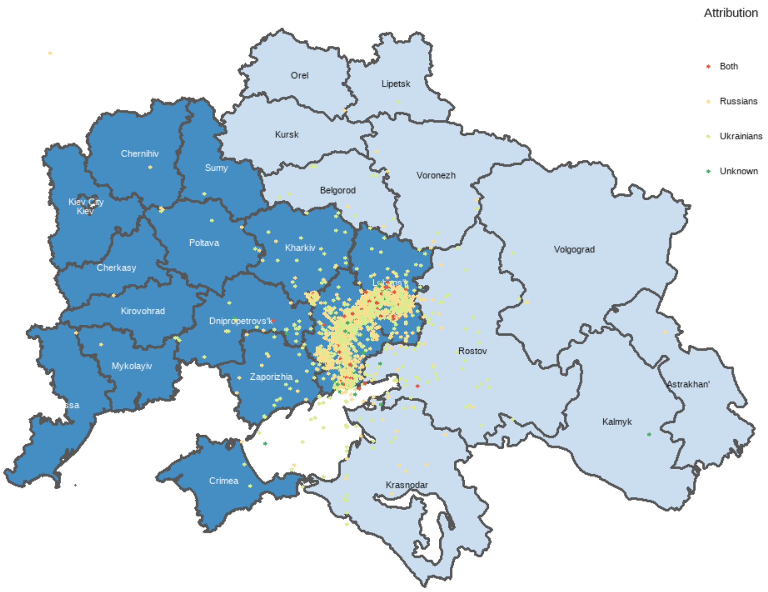
Full collection of ceasefire violations by violation type, aggregated by year and attribution, continued with remaining types from A-1

Figure A-4: Ceasefire Violations by Actor

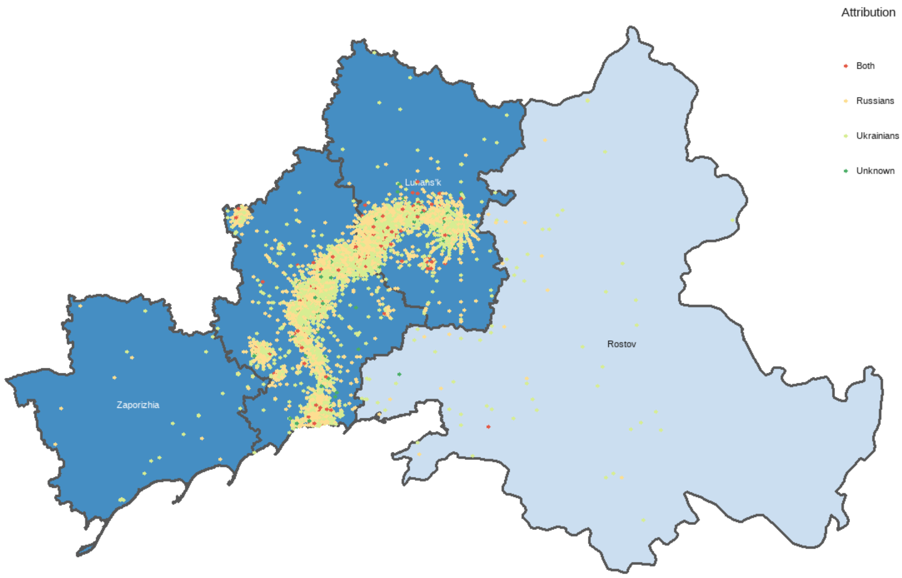



*Note*: Location of ceasefire violations occurring from April 2015 - February 2021. Color determined by perpetrator of violation. Primary map includes only core fighting areas, defined as admin 1 areas with 100+ ceasefire violations over the entire time period.

Figure A-5: Ceasefire Violations by Actor



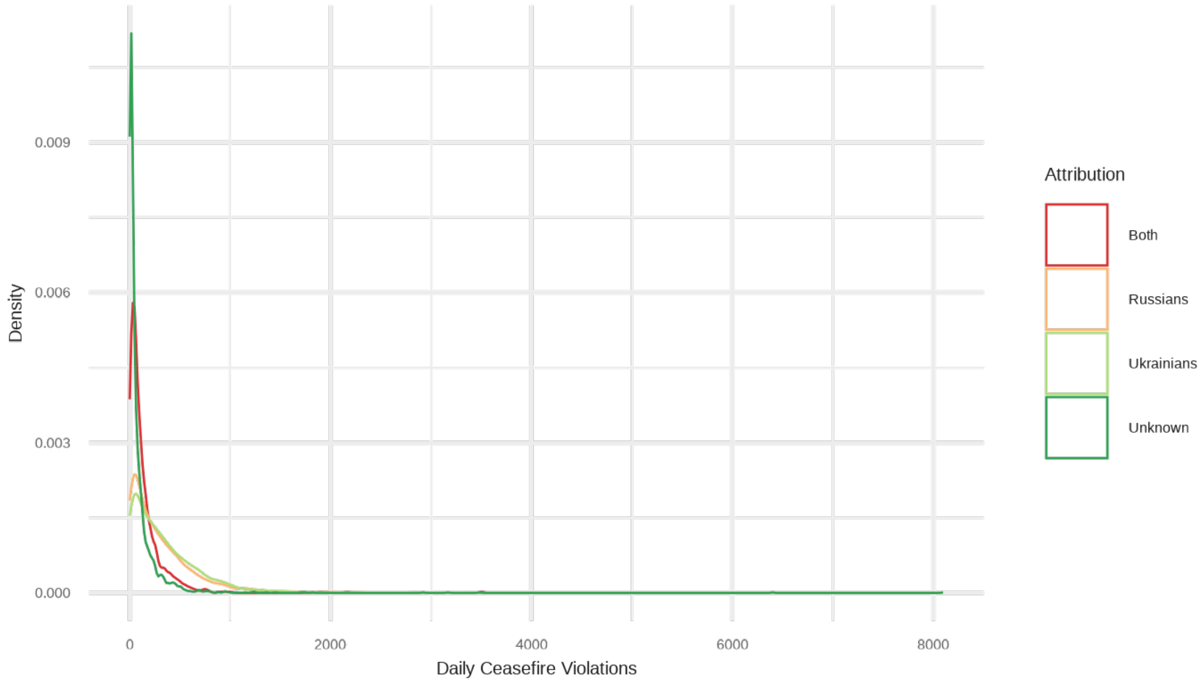

(a) Ceasefire Violations by Actor




(b) Ceasefire Violations by Actor in Core Fighting Areas

*Note*: Location of ceasefire violations occurring from April 2015 - February 2021. Color determined by perpetrator of violation. Top map (a) includes all violations reported, bottom map (b) focuses on core fighting areas, defined as admin 1 areas with 100+ ceasefire violations over the entire time period.

Figure A-6: Distribution of Daily Ceasefire Violations




*Note*: The distribution of daily counts of violations of all types. Results for the full time period, April 2015 - February 2022.

### A.1.1 Measuring Protest Activity

Other major sources of protest data include ACLED (Raleigh et al., 2010), the Mass Mobilization Protest Data (MM) (Clark and Regan, 2016), and the Mass Mobilization in Autocracies Dataset (MMAD) (Weidmann and Rød, 2019). While each are well-respected sources, for our purposes the reliance of both on international news media introduces the possibility of overlooking regional protests of interest. Further, while testing for diversionary behavior requires only considering protest activity targeting the government, ACLED includes both pro- and anti-government protests without a clear strategy for removing pro-government activity.[16] The MM and MMAD datasets, on the other hand, both implement participation minimums as a condition for inclusion. To ensure we have the most comprehensive information on protest events directed at the Russian government, we thus selected LAruPED for our primary analysis.

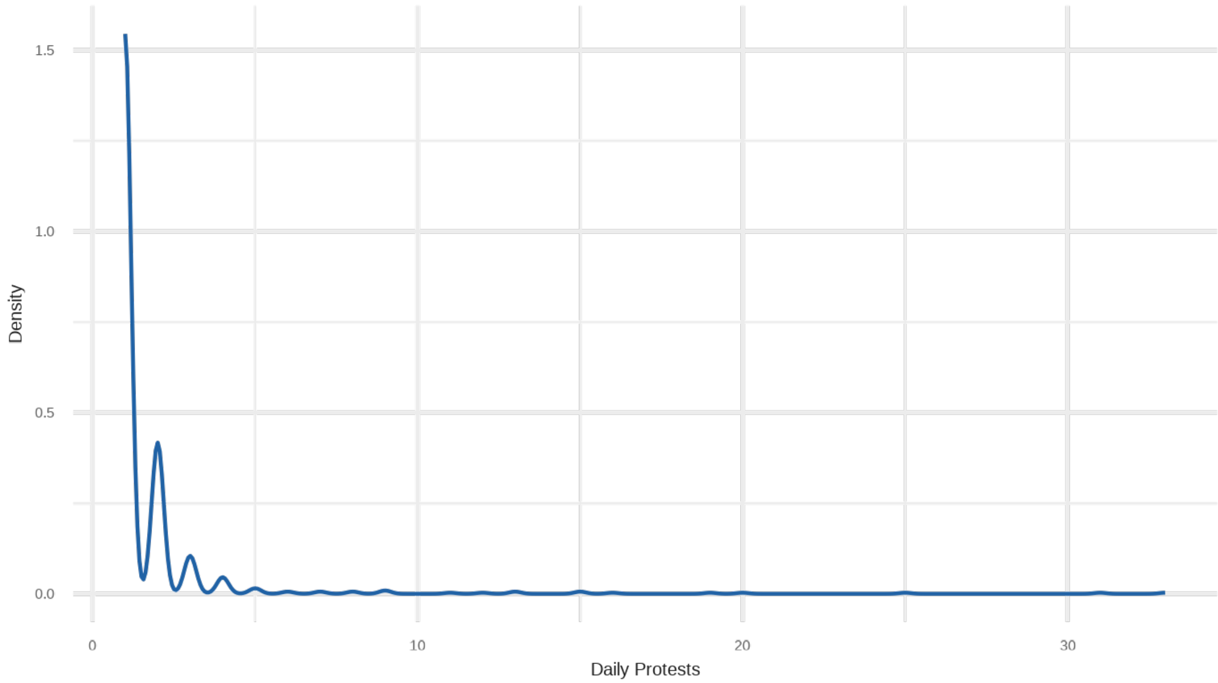
[16] There are 458 unique combinations of actors (coded assoc_actor_1) in the ACLED data for our time frame of interest. Many of these contain pro-Russian government groups, but differentiating them would require making many value judgments better left to Russian experts such as those running namarsh.ru.

Figure A-7: LAruPED Protests

April 2015 - March 2022


*Note*: Location of LAruPED and namash.ru recorded protests occurring from April 2015 - March 2022. Locations geocoded using the Mapquest Geocoding API from the LAruPED and namash.ru recorded locations.

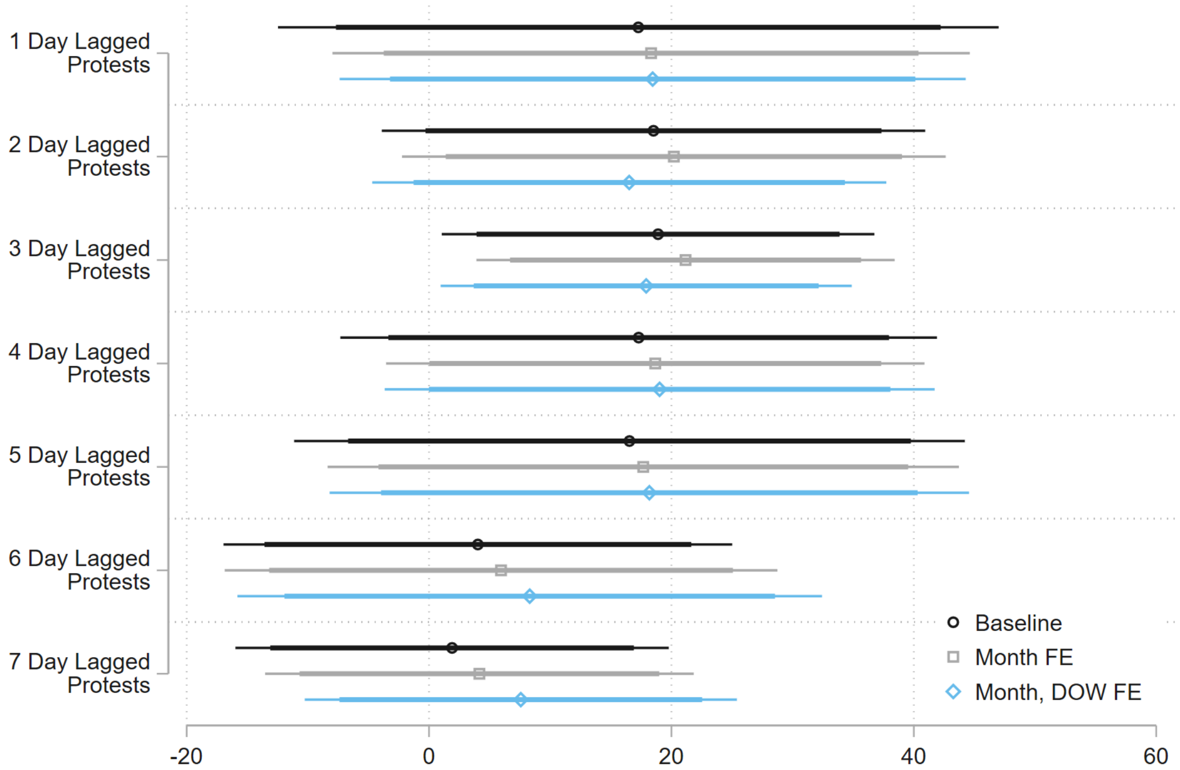
Figure A-8: Distribution of Daily Protests




*Note*: The distribution of daily counts of protests. Results for the time period April 2015 - March 2022.

### A.1.2 Additional 2015 - 2018 Ceasefire Violations Analysis

Figure A-9: Ceasefire Violations by Lagged Protests: 2015 - 2018




*Note*: The impact of seven, lagged days of protests on total violations, with robust, time-clustered standard errors. Results are shown at the baseline level with no fixed effects (circle), with month fixed effects (square), and with month and day of week fixed effects (diamond).

Table A-3: Russian Residual Explosions and Protests

April 2015 - October 2018

| | **Russian Residual Explosion Violations** | |
|---|---|---|
| | (1) Ln Violations | (2) Ln Violations |
| Previous 4 Days Ln Protests | 0.168*** | |
| | (0.063) | |
| Previous 7 Days Ln Protests | | 0.227*** |
| | | (0.074) |
| Observations | 1255 | 1255 |
| Clusters | 1255 | 1255 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Russian residual explosions (standardized residuals generated by regressing the natural log of Russian explosions on the natural log of non-Russian explosions) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests.

Table A-4: Ukrainian Residual Explosions and Protests

April 2015 - October 2018

| | **Ukrainian Residual Explosion Ceasefire Violations** | |
|---|---|---|
| | (1) Ln Violations | (2) Ln Violations |
| Previous 4 Days Ln Protests | 0.035 | |
| | (0.070) | |
| Previous 7 Days Ln Protests | | 0.086 |
| | | (0.090) |
| Observations | 1268 | 1268 |
| Clusters | 1268 | 1268 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Ukrainian residual explosions (standardized residuals generated by regressing the natural log of Ukrainian explosions on the natural log of non-Ukrainian explosions) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests.

### A.1.3 2015 - 2018 Russian News Analysis

Table A-5: Protests and Protest-Related Articles

April 2015 - October 2018

| | Ceasefire Violations | | | |
|---|---|---|---|---|
| | (1) Protest Articles | (2) Protest Articles | (3) Ln Protest Articles | (4) Ln Protest Articles |
| Previous 4 DaysProtests | 0.005 | | | |
| | (0.008) | | | |
| Previous 7 DaysProtests | | 0.006 | | |
| | | (0.005) | | |
| Previous 4 DaysLn Protests | | | 0.030 | |
| | | | (0.029) | |
| Previous 7 DaysLn Protests | | | | 0.034 |
| | | | | (0.030) |
| Observations | 1282 | 1282 | 676 | 676 |
| Clusters | 1282 | 1282 | 676 | 676 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Russian protest-related articles (obtained with BERTopic topic modeling) by LAruPED recorded protests in the preceding four (1) and seven (2) days, as well as the natural log of violations by the natural log of the preceding four (3) and seven (4) days of protests.

Table A-6: Ceasefire Violations by Protest Articles

April 2015 - October 2018

| | **Ceasefire Violations** | | | |
|---|---|---|---|---|
| | (1) Violations | (2) Violations | (3) Ln Violations | (4) Ln Violations |
| Previous 4 DaysArticles | 17.973 | | | |
| | (20.373) | | | |
| Previous 7 DaysArticles | | 15.086 | | |
| | | (17.689) | | |
| Previous 4 DaysProtest Articles | | | 0.251** | |
| | | | (0.116) | |
| Previous 7 DaysProtest Articles | | | | 0.345** |
| | | | | (0.156) |
| Observations | 1228 | 1207 | 1228 | 1207 |
| Clusters | 1228 | 1207 | 1228 | 1207 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Total ceasefire violations by Russian protest-related articles (obtained with BERTopic topic modeling).

Table A-7: Russian Residual Explosions by Protest Articles

April 2015 - October 2018

| | Russian Explosion Violations | |
|---|---|---|
| | (1) | (2) |
| | Ln Ru Explosions | Ln Ru Explosions |
| Prev 4 Days LnProtest Articles | 0.306*** | |
| | (0.092) | |
| Prev 7 Days LnProtest Articles | | 0.437*** |
| | | (0.145) |
| Observations | 1197 | 1176 |
| Clusters | 1197 | 1176 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications.Russian residual explosions (standardized residuals generated by regressing the natural log of Russian explosions on the natural log of non-Russian explosions) by Russian protest-related articles (obtained with BERTopic topic modeling).

Table A-8: Russian Residual Non-Gunshot Ceasefire Violations by Opponent-Media Protest Articles

April 2015 - October 2018

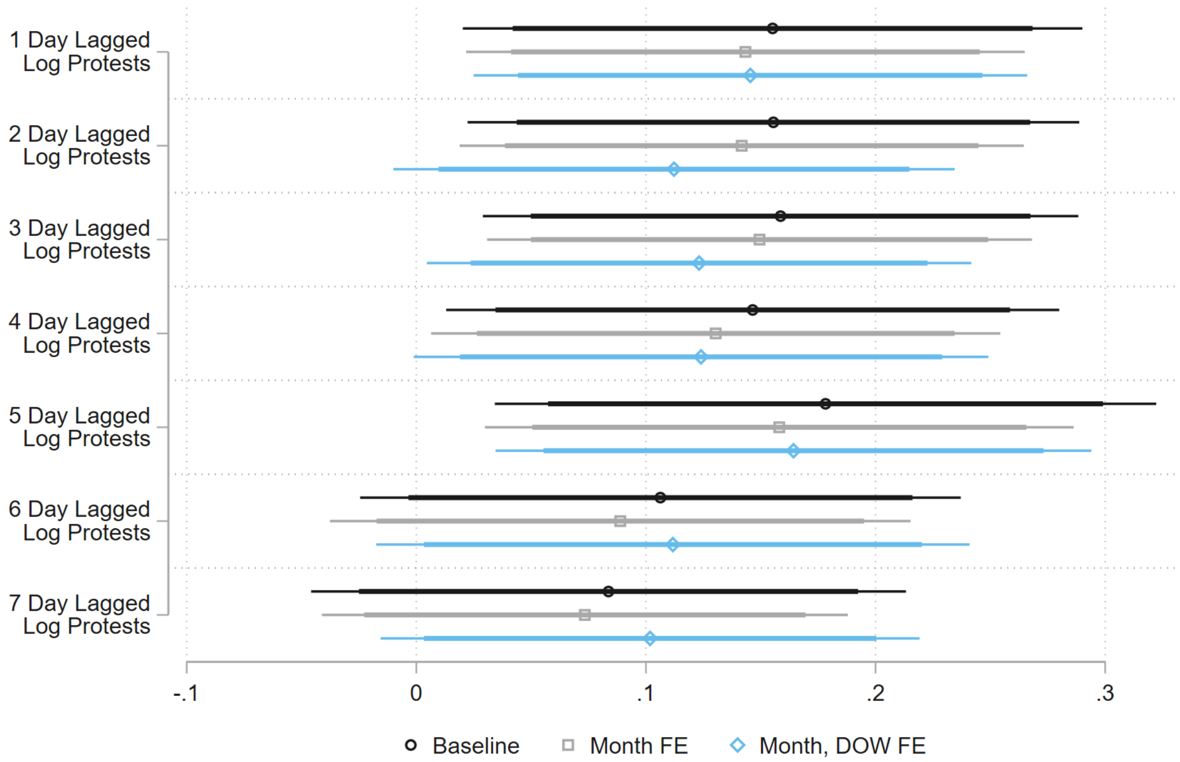
| | Russian Non-Shot Violations | |
|---|---|---|
| | (1) | (2) |
| | Ln Ru Non-Shot Viols | Ln Ru Non-Shot Viols |
| Prev 4 Days Opp LnProtest Articles | 0.516*** | |
| | (0.123) | |
| Prev 7 Days Opp LnProtest Articles | | 0.523*** |
| | | (0.121) |
| Observations | 1207 | 1186 |
| Clusters | 1207 | 1186 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications.Russian residual non-gunshot ceasefire violations (standardized residuals generated by regressing the natural log of Russian no-shot violations on the natural log of non-Russian no-shot violations) by Russian protest-related articles produced by media sources in opposition to the regime (obtained with BERTopic topic modeling).

## A.2 Extended Data Range: April 2015 - February 2021, Selected Dates

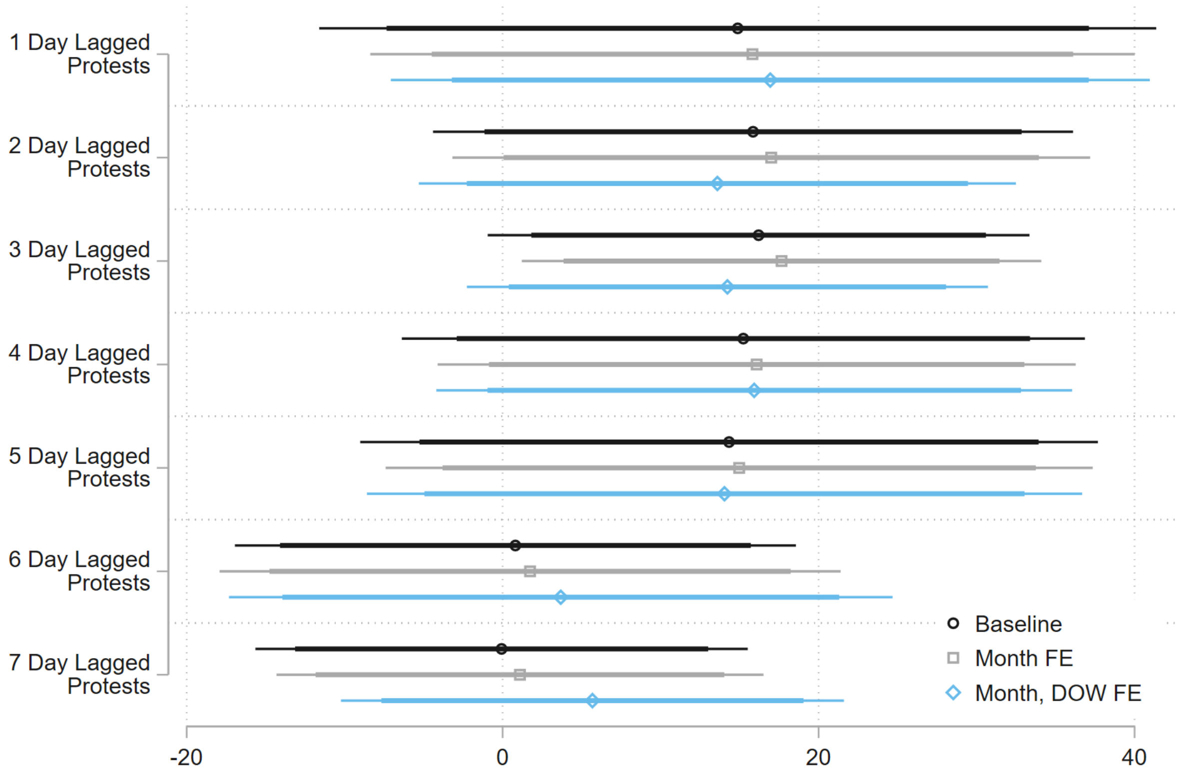
### A.2.1 Additional 2015 - 2021 Ceasefire Violations Analysis

Figure A-10: Ceasefire Violations by (Ln) Lagged Protests: 2015 - 2021




*Note*:The impact of seven, lagged days of log protests on log total violations, with robust, time-clustered standard errors. Results are shown at the baseline level with no fixed effects (circle), with month fixed effects (square), and with month and day of week fixed effects (diamond). Results given for the extended time period, which includes April 2015 - October 2018 with three additional time frames between 2018 and 2021 (see Figure 6).

Figure A-11: Ceasefire Violations by Lagged Protests: 2015 - 2021




*Note*:The impact of seven, lagged days of protests on total violations, with robust, time-clustered standard errors. Results are shown at the baseline level with no fixed effects (circle), with month fixed effects (square), and with month and day of week fixed effects (diamond). Results given for the extended time period, which includes April 2015 - October 2018 with three additional time frames between 2018 and 2021 (see Figure 6).

### A.2.2 Russian Escalatory Violence

Table A-9: Russian Residual Non-Gunshot Ceasefire Violations and Protests

April 2015 - February 2021, Selected Dates

| | **Russian Residual Non-Shot Ceasefire Violations** | |
|---|---|---|
| | (1) Ln Violations | (2) Ln Violations |
| Previous 4 Days Ln Protests | 0.064 | |
| | (0.049) | |
| Previous 7 Days Ln Protests | | 0.088 |
| | | (0.060) |
| Observations | 1652 | 1652 |
| Clusters | 1652 | 1652 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Russian residual non-gunshot ceasefire violations (standardized residuals generated by regressing the natural log of Russian non-gunshot violations on the natural log of non-Russian non-gunshot violations) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests. Results given for the extended time period, which includes April 2015 - October 2018 with three additional time frames between 2018 and 2021 (see Figure 6).

Table A-10: Russian Residual Explosions and Protests

April 2015 - February 2021, Selected Dates

| | **Russian Residual Explosions** | |
|---|---|---|
| | (1) Ln Explosions | (2) Ln Explosions |
| Previous 4 Days Ln Protests | 0.111** | |
| | (0.054) | |
| Previous 7 Days Ln Protests | | 0.134** |
| | | (0.063) |
| Observations | 1634 | 1634 |
| Clusters | 1634 | 1634 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Russian residual explosions (standardized residuals generated by regressing the natural log of Russian explosions on the natural log of non-Russian explosions) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests. Results given for the extended time period, which includes April 2015 - October 2018 with three additional time frames between 2018 and 2021 (see Figure 6).

### A.2.3 Ukrainian Escalatory Violence

Table A-11: Ukrainian Residual Non-Gunshot Ceasefire Violations and Protests

April 2015 - February 2021, Selected Dates

| | Ukrainian Residual Non-Shot Ceasefire Violations | |
|---|---|---|
| | (1)<br>Ln Violations | (2)<br>Ln Violations |
| Previous 4 Days Ln Protests | 0.011<br>(0.057) | |
| Previous 7 Days Ln Protests | | 0.041<br>(0.068) |
| Observations | 1652 | 1652 |
| Clusters | 1652 | 1652 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Ukrainian residual non-gunshot ceasefire violations (standardized residuals generated by regressing the natural log of Ukrainian non-gunshot violations on the natural log of non-Ukrainian non-gunshot violations) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests. Results given for the extended time period, which includes April 2015 - October 2018 with three additional time frames between 2018 and 2021 (see Figure 6).

Table A-12: Ukrainian Residual Explosions and Protests

April 2015 - February 2021, Selected Dates

| | Ukrainian Residual Explosions | |
|---|---|---|
| | (1) Ln Explosions | (2) Ln Explosions |
| Previous 4 Days Ln Protests | 0.016 | |
| | (0.056) | |
| Previous 7 Days Ln Protests | | 0.052 |
| | | (0.067) |
| Observations | 1643 | 1643 |
| Clusters | 1643 | 1643 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Ukrainian residual explosions (standardized residuals generated by regressing the natural log of Ukrainian explosions on the natural log of non-Ukrainian explosions) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests. Results given for the extended time period, which includes April 2015 - October 2018 with three additional time frames between 2018 and 2021 (see Figure 6).

## A.3 Original LAruPED Date Range, 2015 - 2016

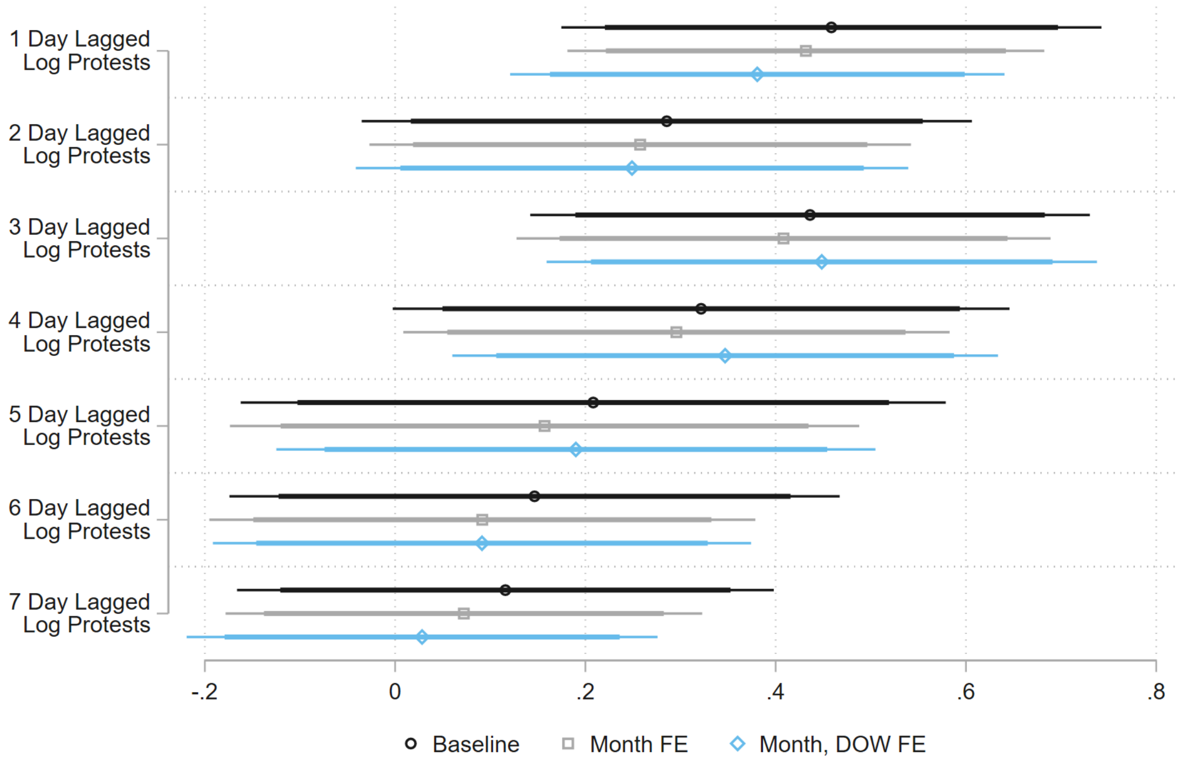
The data used in the primary analysis builds upon the original LAruPED dataset, which extends from 2015-2016. Given the availability of additional data from `namash.ru`, we felt it most appropriate to extend the time frame of our analysis to 2018, but we provide the original LAruPED-only analysis below. The results from this initial time period of 2015-2016 provide strong evidence for the use of diversionary violence spurred by Russian protests, with larger effects than in the primary 2015 - 2018 dataset.

Table A-13: Total Ceasefire Violations and Protests: 2015 - 2016

| | **Ceasefire Violations** | | | |
|---|---|---|---|---|
| | (1) Violations | (2) Violations | (3) Ln Violations | (4) Ln Violations |
| Previous 4 Days Protests | 86.639*** (24.352) | | | |
| Previous 7 Days Protests | | 63.340*** (19.040) | | |
| Previous 4 Days Ln Protests | | | 0.533*** (0.185) | |
| Previous 7 Days Ln Protests | | | | 0.578*** (0.223) |
| Observations | 621 | 621 | 621 | 621 |
| Clusters | 621 | 621 | 621 | 621 |

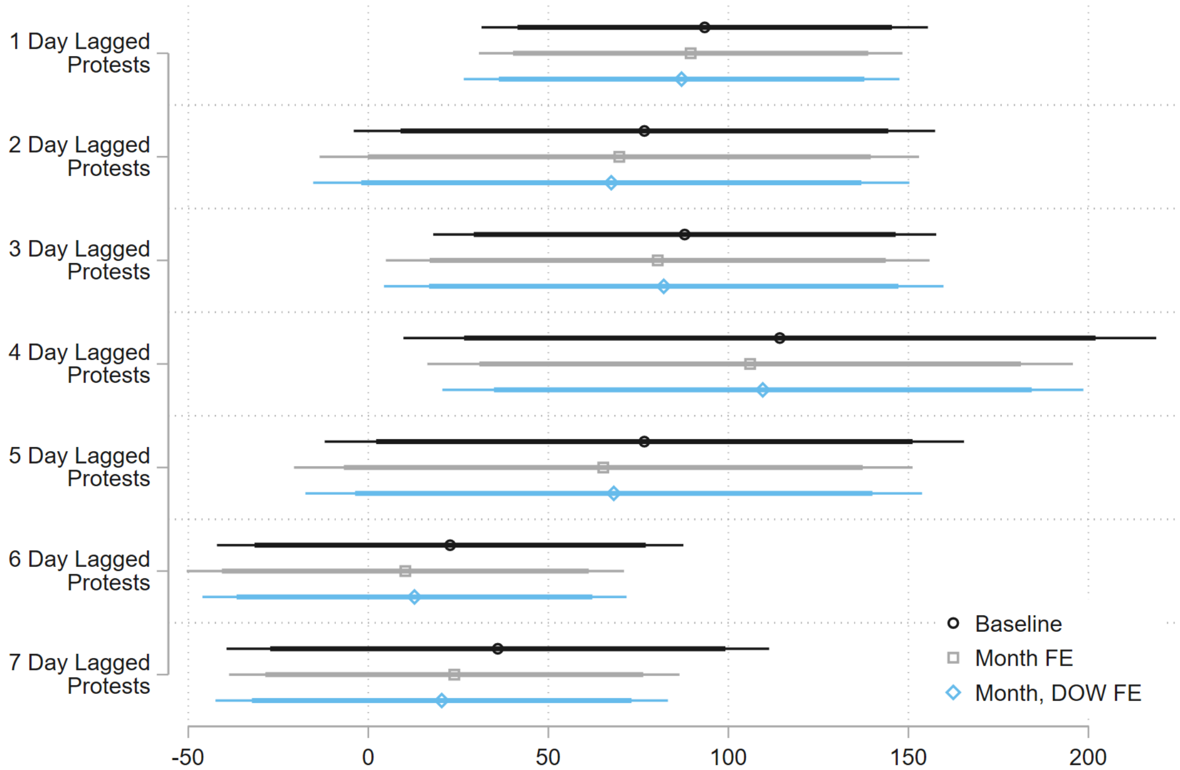
*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Combined ceasefire violations by LAruPED recorded protests in the preceding four (1) and seven (2) days, as well as the natural log of violations by the natural log of the preceding four (3) and seven (4) days of protests. Results given for the original LAruPED time period of April 2015 - December 2016.

Figure A-12: Ceasefire Violations by (Ln) Lagged Protests: 2015 - 2016




*Note*:The impact of seven, lagged days of log protests on log total violations, with robust, time-clustered standard errors. Results are shown at the baseline level with no fixed effects (circle), with month fixed effects (square), and with month and day of week fixed effects (diamond). Results given for the original LAruPED time period of April 2015 - December 2016.

Figure A-13: Ceasefire Violations by Lagged Protests: 2015 - 2016




*Note*:The impact of seven, lagged days of protests on total violations, with robust, time-clustered standard errors. Results are shown at the baseline level with no fixed effects (circle), with month fixed effects (square), and with month and day of week fixed effects (diamond). Results given for the original LAruPED time period of April 2015 - December 2016.

### A.3.1 Russian Escalatory Violence

Table A-14: Russian Residual Non-Gunshot Ceasefire Violations and Protests

April 2015 - December 2016

| | **Russian Residual Non-Shot Ceasefire Violations** | |
|---|---|---|
| | (1) Ln Violations | (2) Ln Violations |
| Previous 4 Days Ln Protests | 0.046 (0.125) | |
| Previous 7 Days Ln Protests | | 0.098 (0.165) |
| Observations | 598 | 598 |
| Clusters | 598 | 598 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Russian residual non-gunshot ceasefire violations (standardized residuals generated by regressing the natural log of Russian non-gunshot violations on the natural log of non-Russian non-gunshot violations) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests. Results given for the original LAruPED time period of April 2015 - December 2016.

Table A-15: Russian Residual Explosions and Protests

April 2015 - December 2016

| | **Russian Residual Explosions** | |
|---|---|---|
| | (1) Ln Explosions | (2) Ln Explosions |
| Previous 4 Days Ln Protests | 0.062 | |
| | (0.125) | |
| Previous 7 Days Ln Protests | | 0.102 |
| | | (0.156) |
| Observations | 590 | 590 |
| Clusters | 590 | 590 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Russian residual explosions (standardized residuals generated by regressing the natural log of Russian explosions on the natural log of non-Russian explosions) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests. Results given for the original LAruPED time period of April 2015 - December 2016.

### A.3.2 Ukrainian Escalatory Violence

Table A-16: Ukrainian Residual Non-Gunshot Ceasefire Violations and Protests

April 2015 - December 2016

| | Ukrainian Residual Non-Shot Ceasefire Violations | |
|---|---|---|
| | (1)<br>Ln Violations | (2)<br>Ln Violations |
| Previous 4 Days Ln Protests | -0.085<br>(0.142) | |
| Previous 7 Days Ln Protests | | -0.046<br>(0.181) |
| Observations | 601 | 601 |
| Clusters | 601 | 601 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Ukrainian residual non-gunshot ceasefire violations (standardized residuals generated by regressing the natural log of Ukrainian non-gunshot violations on the natural log of non-Ukrainian non-gunshot violations) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests. Results given for the original LAruPED time period of April 2015 - December 2016.

Table A-17: Ukrainian Residual Explosions and Protests

April 2015 - December 2016

| | **Ukrainian Residual Explosions** | |
|---|---|---|
| | (1) | (2) |
| | Ln Explosions | Ln Explosions |
| Previous 4 Days Ln Protests | -0.008 | |
| | (0.134) | |
| Previous 7 Days Ln Protests | | 0.101 |
| | | (0.177) |
| Observations | 599 | 599 |
| Clusters | 599 | 599 |

*Note*: * p <.10, ** p <.05, *** p <.01. Robust, time-clustered standard errors are in parentheses. Month and day of week fixed effects are included in all model specifications. Ukrainian residual explosions (standardized residuals generated by regressing the natural log of Ukrainian explosions on the natural log of non-Ukrainian explosions) by the natural log of the preceding four (1) and seven (2) days of LAruPED protests. Results given for the original LAruPED time period of April 2015 - December 2016.

# B Data Appendix

There are several sources that go into our main dataset.

## B.1

OSCE Strike Data

The data on location and timing of the strikes come from daily reports from the OSCE Special Monitoring Missing to Ukraine (henceforth, the "SMM"). These are published (almost) daily on the OSCE website [17]. Before April 15, 2015 they contained only word descriptions of ceasefire violations, but afterwards SMM started publishing accompanying PDF tables that the reports are based on. These PDF tables contain invaluable information on the location of the strikes, their time, type, weapon, etc.

## B.2 Liveuamap Data

War at the Donbass[18], an Een Story Map, is an exploration of the conflict in Ukraine that, among other things, attempts to visualize the areas held by the insurgents each month. They used data from Live Universal Awareness Map (Liveuamap)[19] which they downloaded for each day and displayed the polygons of the area held by the insurgents on the first day of every month. The data is available from May 2014 and we use it until December 2018.

Liveuamap is a mapped database of events related to the Ukrainian conflict that uses human-checked AI web crawlers to fill itself (although it is not entirely clear how the maps of the insurgent area are created).

## B.3 Geonames Database

We use the database of Ukrainian cities with their coordinates from `http://www.geonames.org/` (it contains 42009 unique cities). Fundamentally, it is a crowd sourced service and the quality of its sources may vary.

## B.4 Creating Main Dataset

### B.4.1 Reading PDFs

We scraped the search page of the OSCE Strike Data and downloaded each of the PDFs. Extracting data from them presented a challenge. They were created in Word, so the text was embedded in the files and could be easily extracted. However the PDF format does not preserve the cellular structure of the tables, so attributing text to a particular row or

[17] https://www.osce.org/ukraine-smm/reports/

[18] https://www.arcgis.com/apps/MapJournal/index.html?appid=53d7d6a0829b40ea8064204d83353585, accessed June 10, 2019.

[19] https://liveuamap.com, accessed June 10, 2019.

column was largely impossible [20]. Things were further complicated by the fact that column headers were only on the first page, and sometimes columns with merged rows would be entirely empty for the later pages (see Figures A-14 and A-15). This last problem rendered the tables unreadable by the software that was available to us, because of the failure to recognize these empty columns as columns.

Our solution was to utilize the fact that the tables were prepared in Word (i.e., were not scanned) and are, therefore, *perfectly* aligned. That meant that when looked at as a raster image, each column margin was simply a pixel-wide column of 1's. Collapsing the image into column sums and normalizing by the number of rows would produce something that is close to 1 for a margin, and something significantly less for a raster column that does not contain a margin (even if it contains text, the ratio of 1's to the number of rows would be nowhere close to one). A similar logic would hold with rows. This meant that we would have been able to identify the location of each cell and extract the text from that particular area cell by cell.

Here is our procedure in more detail.

1. Cconvert each PDF page into a PNG and import as a raster into R.
2. Use *raster::boundaries()* that detects the edges of objects.
3. Using *raster::clumps()*, we assign to each pixel of a connected object a unique value [21]. Figure A-16 gives examples for these steps. The largest connected object (in the sense of having the most pixels) are the table boundaries, so we can drop everything else and crop the raster (see Figure A-17).
4. In this margin structure each column can be easily identified by taking the column sums of the underlying matrix. For table columns, the ratio of column sums to the number of rows is close to 1 (we used a threshold of 0.96 [22]). Notice that occasionally some letters are "clumped" together with the margins [23], but that is way below the threshold.
5. Loop through each column, using similar logic to identify row margins. These steps allow us to obtain locations of each cell in the table.
6. Convert raster coordinates to PDF coordinates used by tabulizer package.

[20] PDF does preserve what the text looks like, so within one page it was theoretically possible to distinguish between columns based on the number of spaces between text. But different pages had different numbers of spaces between column (and sometimes no spaces at all), rendering the attempts to tease out information futile.

[21] For example, consider the beginning of the word "Explosion" in Figure A-16. All pixels in the letter "E" are assigned value 84. Letters "x" and "p" happen to have no space between them and are lumped together; their pixels have value 90.

[22] It might not be exactly one due to artifacts of compression.

[23] See Figure A-17 for an example. Notice the letters on the left that were placed so close to the margin that the function failed to distinguish them. This happens occasionally and is not a problem for our algorithm.

7. Extract the text sequentially from each cell with *tabulizer::extract_text()* by passing it the respective coordinates.

8. Lastly, simply collect the cells in the correct order into a table.

## B.5 Contents of Raw Data

The structure of the tables changed from time to time, but largely by way of splitting columns (for example, "Time/Date" into "Time" and "Date"). The contents remained constant and could be summarized as follows. For convenience, in explaining our cleaning efforts we will use the names of the variables suggested below, although they may slightly differ from the column headers in the PDFs.

- Position

  Position of the SMM that recorded a ceasefire violation. For example, "About 1km NW of the railway station in Yasynuvata (non-government-controlled, 16km NE of Donetsk)". It may be the location where a CCTV camera is mounted, or where the mission operatives were physically present at the time.

- Even Location

  An attempt to place the violation relative to the SMM position. For example, "1-2km SSW".

- Means

  How the violation was recorded (heard or seen by operatives or recorded by equipment).

- Observation

  Description of the nature of the violation, i.e., a shot, a burt, an explosion, a projectile, etc. Note that many tables combine Means, Number, and Observation into a single column titled "Observation".

- Number

  Number of shots, bursts, etc.

- Description

  Vital information about the observation: the direction of flight of a projectile (e.g., "SE to NW"), assessment of whether a shot came from the disengagement area, whether an explosion came from an impact, etc.

- Weapon

  Type of weapon used.

- DateTime

  Time and date of observation. For prolonged events, start and end times.

## B.6 Codebook

Here we describe the variables in the final dataset and briefly talk about their exact source. For the details on those that we have worked on extensively we refer the reader to Section B.7.

- TableID

  Unique ID associated with the report (table) found in the URL of the report (e.g., "https://www.osce.org/special-monitoring-mission-to-ukraine/*422084*").

- CleanDate

  Date of the violation. Taken from StartTime (see below).

- StartTime

  Date and time the violation started. The date without is parsed from DateTime. To get the year we needed to parse the date from the date mentioned in the initial HTML page with the link to the PDF table.

- EndTime

  Date and time the violation finished. Parsed similarly to StartTime. If the table does not give an interval, it is equal to the StartTime.

- StrikeLon

  Longitude of the strike. See Section B.7.

- StrikeLat

  Latitude of the strike. See Section B.7.

- Means

  How the violation was recorded (heard or seen by operatives or recorded by equipment).

- Weapon

  Type of weapon used.

- Type[type of violation]

  The number of particular types of violations observed, based on Observation and Number. Some observations record several types of violations at once (for example, "Heard 2 explosions and 5 rounds."), so we created a separate column for each type of violation: explosion, rocket, projectile, salvo, round, impact, flare, tracer, shot, burst, flash.

  Sometimes SMM used word approximations instead of numbers (e.g., "Heard an undetermined number of explosions."). We used the following approximations: uncountable or multiple uncountable – 50; undetermined number – 40; continuous – 50; multiple – 50; hundreds – 300; several – 6.

- TypeMinutes

  Number of minutes of fire, based on Observation. E.g., ”Heard an intense exchange of fire for 3 minutes”.

- TypeGeometry

  The geometry of the strike, based on Observation: direct or indirect. This is helpful in determining artillery strikes. For example, ”1 tracer of undetermined indirect fire from E-W”.

- TypeIntense

  Dummy for whether the observation was qualified as intensive, based on Observation. E.g., ”Heard an intense exchange of fire for 3 minutes”.

- TypeFire

  Dummy for whether Observation contains vague words like ”firing”, ”fire”, and ”shooting”.

- TypeImpactDummy

  Dummy for whether Observation contains the word ”impact”.

- TypeArtillery

  Dummy for whether Observation contains the word ”artillery”.

- TypeMuzzle

  Dummy for whether Observation contains the word ”muzzle”.

- AreaInside

  Dummy for whether the observation is qualified as being inside the disengagement area, based on Observation.

- AreaOutside

  Dummy for whether the observation is qualified as being outside the disengagement area, based on Observation.

- AreaUnclear

  Dummy for whether the SMM was unable to specify whether the violation occurred inside or outside the disengagement area, based on Observation.

- TypeTraining

  Dummy for whether the observation is qualified as being a training exercise, based on Observation. E.g., ”90-100 bursts (assessed as live-fire training exercise)”.

- Control

  Party controlling the position of the SMM: DPR, LPR, Government, and Non-Government. Parsed from Position, e.g., "SW edge of Avdiivka (government-controlled, 17km N of Donetsk)".

- WithinInsurgentArea

  Our estimation whether the coordinates of the strike are within the area controlled by the insurgents at the time of the strike.

- BorderDistanceKm

  Our estimation of the distance from coordinates of the strike to the closest point of the border of the area controlled by the insurgents at the time of the strike.

- Attribution

  Our estimation of the side responsible for the strike, Russians or Ukrainians

- CrossingDist

  For strikes within the insurgent area, the distance to the closest crossing point.

- CrossingDist_2

  For strikes within the insurgent area, the distance to the second closest crossing point.

- CrossingClosest

  For strikes within the insurgent area, the name of the closest crossing point.

- CrossingClosest_2

  For strikes within the insurgent area, the name of the second closest crossing point.

## B.7 Preparation of Selected Final Variables

### B.7.1 StrikeLat and StrikeLon

*General Approach*
The general approach to acquiring these coordinates was the following. Position contains the SMM location that we can parse and then find GPS coordinates to. We then look at Location that gives bearing and distance of the strike from the SMM position (e.g., "2-3km SSE"). Then, shifting the GPS coordinates associated with Position as prescribed by Location we get approximate latitude and longitude of the strike.

*Dealing with Merged Cells*
Sometimes the value of Position was split across two pages. We utilized the fact that it was almost always in the form of "[city] ([some additional information])" and glued the lines with improper parentheses, e.g., if the last line on a page is missing a closing parenthesis and the first line on the next page is missing an opening one.

If several cells were merged, this resulted in the same incomplete value being on several lines, but this was easy to deal with by replacing identical sequential lines with the glued value.

*Cleaning Position*

Position contains several important details. First of all, it is the position of the SMM from which we would be able to identify the location of the strike. It also often contains information about who controls the area (final variable Control), and also an approximation of the position of the SMM relative to one of the major cities [24]. Sometimes parentheses also contain information about the former name of the village/city.

We split Position on the opening parenthesis. Control, approximation, and former name are easy to parse from the second part, while the first part requires a little bit of cleaning. In addition to the village/city, there is often mention of SMM cameras, or observation points, or gas stations, etc., which we parse out trying to cut the string to only the name of the village/city and the position relative to it, e.g., "2.9km NNW of Lebedynske". Importantly, we remove indications like "South edge of" which, for larger cities, means that we do not get the position of the SMM accurately. We then extract the shift and the bearing as ObservationShift, and what remains we will call ObservationCity (to be used later but not to appear in the final dataset).

*Getting City Coordinates*

Now we need to assign coordinates to each village/city. For the some of the most often occurring locations we find the coordinates manually, for example, Stanytsia Luhanska Bridge or Donetsk Central Railway Station. The number of unique positions is too high, however, to do it manually, so we supplement the data using the following procedure.

We download the database of Ukrainian cities from http://www.geonames.org/ (it contains 42009 unique cities). A particular virtue of this database is the fact that for each city it contains numerous variations of spelling which is especially important for names transliterated from Cyrillics. For example, the city of Kadiyivka in the OSCE dataset is spelled Kadiivka – a spelling contained in the Geonames Dataset along with Kadievka, Kadiivka, Kadijewka, Kadijivka, Kadijiwka, Kadiyevka, Kadiyivka, etc. This, however, also becomes a vice, as many similar-sounding places may have forms of spelling that are identical. There are also villages that are in different regions but are named the same.

To tackle this, we employ the following procedure.

1. We load the Geonames Dataset, draw a bounding box in Ukraine that reliably contains all the activities associated with the war (top left corner at (49.41502, 36.31368)), and keep only those cities that fall within it. That allows us to disregard villages that are, for example, near Kiev.

2. We create a vector of unique names that we do not have coordinates for from three sources:

[24] For example, for Position "Popasna (government-controlled, 69km W of Luhansk)" we are given that it is 69 km West of Luhansk and therefore could check that the coordinates for Popasna that we obtained are correct.

- ObservationCity;
- Former name of the place if ever mentioned;
- ObservationCity variable stripped to only those words that begin with a capital letter (ObservationCityRough) [25].

3. For each of these places we run a loop that checks whether the name appears in the Geonames Dataset. If it appears only once, we assign coordinates to that place. If it appears more than once but there is an observation in Geonames Dataset that lists the given spelling as primary, we use that. Thus we create a data frame of unique places and their coordinates.

4. Lastly, we iteratively left join the places data frame to our dataset:
   - For observations missing city coordinates, we match places to ObservationCity;
   - For observations still missing city coordinates, we then match places to Formerly, which fills only those that we were unable to find matches to earlier;
   - Lastly, for observations still missing city coordinates, we match places to ObservationCityRough (see above).

*Getting Strike Locations*

Here is what a typical value of Location looks like: ”2.5-3.5km NNW”. Whenever the distance given as ”more than 1km” we use the value of 2, for ”more than 2” – 3. Whenever the distance is not known we approximate with 4. We then find all observations identifiable with RegEx request `"^[^( )]*m [A-Z]{1,3}$"`. We split each into StrikeDistance and StrikeDirection and convert meters into kilometers as needed.

Similarly, we split ObservationShift (the variable extracted from Position as described earlier) into ShiftDistance and ShiftDirection. We convert ShiftDirection and StrikeDirection into angle bearing, and whenever their -Distance counterparts give a range we replace it with the mean.

We then loop through our dataset and, for observations that we have respective data, calculate the coordinates of the strike. We begin by taking the initial coordinates of the city. If ShiftDistance and ShiftDirection are not NA (i.e., when the position of the SMM was relative to a particular city), we use *geosphere::destPoint()* that, given a starting point, initial bearding, and distance, computes the destination point traveling along the shortest path on the geodesic with flattening value $f = 1298.257223563$. We then use the same function to shift (further, if ShiftDirection was present) the point using StrikeDistance and StrikeDirection. This, finally, gives us StrikeLon and StrikeLat.

Attribution, WithinInsurgentArea, and BorderDistanceKm For each strike we want to know whether it falls/originates in the area controlled by the insurgents or not. To that end

[25] This is valuable for situations when our cleaning did not produce clean results. For example, if the word ”posiiton” is misspelled, then we do not remove it from the string (”SMM posiiton in Svitlodarsk”) and it ends up in the ObservationCity

we make use of the following two data sources. For the period before December 2018 we use Liveuamap Data. For the period starting December 2018 we utilize the OSCE daily reports. From that point onwards they start specifying the line of contact very accurately.

For the former source, we scrape the shapes for each month and georeference each of them separately in QGIS. For the latter, we draw the boundary in QGIS by hand.

For each observation point we then check whether its coordinates fall within the polygon for the respective month using *rgeos::gContains()* and calculate the distance to the closest point on the boundary using *geosphere::dist2Line()*. This gives us WithinInsurgentArea and BorderDistanceKm.

We then try to infer the attribution of each strike. If an incoming artillery strike flew from the North-East and landed in the area controlled by the insurgents it is reasonable to assume that it originated from the Ukrainian side. Alternatively, if the strike originated in the insurgent area we can argue that the Russians are responsible.

To help us with this task we use the information in the Observation column about the directionality of the strike. There are two kinds of phrases that may be useful to us. Sometimes, particularly when referring to explosions, SMM note whether they are incoming or outgoing. We form a variable Direction2 to be:

- "Outgoing" when Observation contains the word "outgoing";
- "Incoming" when Observation contains the words "incoming" or "impact"; and
- "Twoway" if:
  - It can be classified as both Outgoing and Incoming; or
  - Observation contains the word "overlapping" (as in "overlapping fire").

Other times the SMM notes the direction of the flight of the strike, for example, "Recorded 4 projectiles from NW to SE". We parse the directionality and create a variable Direction1.

If Direction1 is available, we calculate the point on the insurgent boundary that is the closest to the strike, calculate the bearing from the strike to that point, and compare Direction1 to that bearing. If the difference is less than 90 degrees we attribute the strike to the insurgents and to Ukrainians otherwise.

If Direction1 is not available, we use Direction2. If a strike is located within the insurgent area and Direction2 is "Outgoing", we attribute it to the insurgents. If the strike is located within the insurgent area and Direction2 is "Incoming", we attribute it to the Ukrainians. We flip this for the opposite case. We set Attribution to "Both" if the direction is "Twoway".

Lastly, if both Direction1 and Direction2 are missing, we consider the following types of observations to be *outgoing*: Rocket, Projectile, Salvo, Round, Flare, Tracer, Shot, Burst, Flash, Fire, Muzzle; and the following to be *incoming*: Explosion and Impact, and then employ the same procedure as for Direction2.

## B.8 The Life of a Strike

In this section we follow a couple of individual observations that may serve as a good example to help the reader navigate the steps in cleaning the data. Consider two observations that come from the Daily Report from October 4, 2017 (see Table A-18 below. Note that we replaced the original column names with cleaner versions) [26]. The resulting variables can be seen in Table A-19.

Table A-18: Example of Observations from the Daily Report from October 4, 2017

| DateTime | Position | Location | Observation | Weapon |
|---|---|---|---|---|
| 3 Oct, 13:23 | 2.9km NNW of Lebedynske (government-controlled, 16km NE of Mariupol) | 3km E | Heard 2 explosions (impacts) | N/K |
| 3 Oct, 13:23 | 2.9km NNW of Lebedynske (government-controlled, 16km NE of Mariupol) | 3km E | Heard 1 explosion (outgoing) | N/K |

Table A-19: Example of Observations from the Final Dataset

| TableID | StartTime | StrikeLon | StrikeLat | Means | TypeExplosion | Direction2 | Control | WithinInsurgentArea | BorderDistanceKm | Attribution |
|---|---|---|---|---|---|---|---|---|---|---|
| 347981 | 2017-10-03 13:23:00 | 37.779 | 47.151 | Heard | 2 | Incoming | Government | FALSE | 2.86 | Russkies |
| 347981 | 2017-10-03 13:23:00 | 37.779 | 47.151 | Heard | 1 | Outgoing | Government | FALSE | 2.86 | Ukrainskies |

TableID is the last number seen in the URL for the daily report. We parse DateTime to get the date of the violation and scrape the year from the html with the report.

For both observations in the example Position tells us that when the violation occurred the SMM was located near Lebedynske village whose coordinates we can easily find (47.127, 37.754). Shifting this by 2.9km NNW we get (47.15118, 37.73953). Then, from Location, we learn that the strike itself occurred 3 km to the East of the SMM position. Thus, we need to shift the coordinates 3 km further to get (47.15118, 37.77909). See Figure A-19 for visualization of locations.

From Position we also learn that that area (being the area where the SMM was located, not the strike itself) is under government control, hence the value of Control is "Government". From Observation we get that the SMM heard the violations (instead of, for example, seeing them) and that one was two incoming explosions while the other was 1 outgoing (this information goes into TypeExplosion and Direction2).

Given the coordinates, we can check whether the strikes fall within the territory currently controlled by the insurgents (producing the WithinInsurgentArea dummy) and the distance to the border in kilometers (BorderDistanceKm). Armed with that, we can utilize Direction2 to infer who was firing. The first observation describes incoming strikes landed in the Ukrainian territory, so we attribute them to the insurgents. The second observation is an outgoing explosion from the same spot, so we see it as Ukrainians firing onto the insurgents.

[26] https://www.osce.org/special-monitoring-mission-to-ukraine/347981

# C Proofs

**Lemma 1.** *Let the regime type that is indifferent between doing nothing and diversionary escalation be given by*

$$\tilde{\theta}(x^*) = x^* - \sigma\Phi^{-1}\left[\Phi\left(\frac{x^* - t}{\sigma}\right) + p\right]$$

*and let the regime that is indifferent between diversionary escalation and failing be given by* $\tilde{\tilde{\theta}}(x^*) = \Phi(\frac{1}{\sigma}(x^* - t)) + p$*, for any citizen threshold* $x^* \leq t + \sigma\Phi^{-1}(1-p)$*. Alternatively, for any* $x^* > t + \sigma\Phi^{-1}(1-p)$*, choose any* $\tilde{\tilde{\theta}}(x^*), \tilde{\theta}(x^*) \in \mathbb{R}$ *such that* $(\tilde{\tilde{\theta}}(x^*), \tilde{\theta}(x^*)) = \emptyset$*.*

*Then, in any symmetric monotone equilibrium defined by citizen threshold* $x^*$*, proxy war is preferred to no diversion if and only if* $\theta \in (\tilde{\tilde{\theta}}(x^*), \tilde{\theta}(x^*))$*.*

*Proof.* Suppose $x^* \leq t + \sigma\Phi^{-1}(1-p)$. By the normal cumulative distribution function, $\Phi(\frac{1}{\sigma}(x^* - \theta)) > \Phi(\frac{1}{\sigma}(x^* - \tilde{\theta}))$ for all $\theta < \tilde{\theta}$. Therefore,

$$\Phi\left(\frac{x^* - \theta}{\sigma}\right) - \Phi\left(\frac{x^* - t}{\sigma}\right) > \Phi\left(\frac{x^* - \tilde{\theta}(x^*)}{\sigma}\right) - \Phi\left(\frac{x^* - t}{\sigma}\right) = p$$

for all $\theta < \tilde{\theta}(x^*)$. Then, all types $\theta < \tilde{\theta}(x^*)$ prefer proxy war to doing nothing if and only if $\theta - \Phi(\frac{1}{\sigma}(x^* - \theta)) - p > 0$ or equivalently $\theta > \Phi(\frac{1}{\sigma}(x^* - \theta)) + p$. On the other hand, $p = \Phi(\frac{1}{\sigma}(x^* - \tilde{\theta}(x^*))) - \Phi(\frac{1}{\sigma}(x^* - t)) > \Phi(\frac{1}{\sigma}(x^* - \theta)) - \Phi(\frac{1}{\sigma}(x^* - t))$ for all $\theta > \tilde{\theta}(x^*)$.

Now suppose $x^* > t + \sigma\Phi^{-1}(1-p)$. Then, proxy war is never preferred to no diversion, $\theta - \Phi(\frac{1}{\sigma}(x^* - t)) - p < \theta - \Phi(\Phi^{-1}(1-p)) - p = \theta - 1 < \theta - \Phi(u)$ for all $u \in \mathbb{R}$. □

**Lemma 2.** *In any symmetric monotone equilibrium with threshold* $x^*$*, conventional war is preferred to no diversion if and only if* $\theta \in (w, x^* - \sigma\Phi^{-1}(w))$*.*

*Proof.* The result immediately follows $\theta - w > \max\{0, \theta - A(\theta, r = 0)\}$. □

**Lemma 3.** *Let* $\Theta_r := \{\theta : r(\theta) = r\}$ *be the set of types that choose regime action* $r$ *in any given equilibrium strategy profile. Then, if citizens play a threshold strategy given by* $x^*$ *and diversion occurs on the path of play* $\Theta_0 \neq \mathbb{R}$*,* $\Theta_1 \neq \emptyset$ *if and only if* $\Theta_2 = \emptyset$*.*

*Proof.* By Lemma 1, the set of types that engage in diversionary escalation will be given by the set $\Theta_1 = (\tilde{\tilde{\theta}}(x^*), \tilde{\theta}(x^*))$ if proxy war occurs in equilibrium or $\Theta_1 = \emptyset$ if not.

First, suppose $\Theta_1 = \emptyset$. Then, by Lemma 2, either the set $\Theta_2 = (w, x^* - \sigma\Phi^{-1}(w))$ is an open interval or is empty $\Theta_2 = \emptyset$, and $\Theta_0 = \Theta_2^c$. $\Theta_0 \neq \mathbb{R}$ then implies that $\Theta_2 \neq \emptyset$.

Alternatively, consider $\Theta_1 \neq \emptyset$. A regime of type $\theta$ will choose conventional war $r(\theta) = 2$ if and only if $w < \min\{\theta, \Phi(\frac{1}{\sigma}(x^* - \theta)), \Phi(\frac{1}{\sigma}(x^* - t)) + p\}$. For this to occur, it is required that $w < \Phi(\frac{1}{\sigma}(x^* - t)) + p$. Because both sides of this inequality are fixed in $\theta$, however, the condition either always or never holds. If never, then $\Theta_2 = \emptyset$ and therefore $\Theta_0 = \Theta_1^c$. If always, then $\Theta_1 = \emptyset$, a contradiction. □

**Lemma 4.** *For any $\gamma \in [0,1]$, let*

$$f(\theta;\tau,\sigma,z,t,\gamma) := \gamma\Phi\left(\frac{\theta-\mu(\theta+\sigma\Phi^{-1}(\theta))}{\nu}\right) + (1-\gamma)\Phi\left(\frac{t-\mu(\theta+\sigma\Phi^{-1}(\theta))}{\nu}\right) - c \quad \text{(A1)}$$

*taking the definitions of $\mu(x) := \frac{\sigma^2 z+\tau^2 x}{\tau^2+\sigma^2}$ and $\nu^2 := \frac{\tau^2\sigma^2}{\tau^2+\sigma^2}$. Then, $f(\theta;\cdot)$ is monotonic in $\theta$ if $\sigma < \tau^2\sqrt{2\pi}$.*

*Proof.* By definition of $\mu(\cdot)$ and $\nu$, we can rewrite $f(\theta;\cdot)$ in terms of $\sigma$, $\tau$, and $z$. Plugging in and simplifying, we recover

$$f(\theta;\cdot) = \gamma\Phi\left(\frac{\sigma(\theta-z)-\tau^2\Phi^{-1}(\theta)}{\tau\sqrt{\sigma^2+\tau^2}}\right) + (1-\gamma)\Phi\left(\frac{\sigma^2(t-z)+\tau^2(t-\theta)-\tau^2\sigma\Phi^{-1}(\theta)}{\sigma\tau\sqrt{\sigma^2+\tau^2}}\right) - c.$$

Taking the partial derivative with respect to type, we have

$$\begin{aligned}\frac{\partial f(\theta;\cdot)}{\partial\theta} = \frac{1}{\sigma\tau\sqrt{\sigma^2+\tau^2}}\Bigg[&\gamma\phi\left(\frac{\sigma(\theta-z)-\tau^2\Phi^{-1}(\theta)}{\tau\sqrt{\sigma^2+\tau^2}}\right)\left(\sigma^2-\frac{\tau^2\sigma}{\phi\left(\Phi^{-1}(\theta)\right)}\right)\\ &-(1-\gamma)\phi\left(\frac{\sigma^2(t-z)+\tau^2(t-\theta)-\tau^2\sigma\Phi^{-1}(\theta)}{\sigma\tau\sqrt{\sigma^2+\tau^2}}\right)\left(\tau^2+\frac{\tau^2\sigma}{\phi\left(\Phi^{-1}(\theta)\right)}\right)\Bigg].\end{aligned}$$

Leveraging the fact that $\max_\theta \phi(\Phi^{-1}(\theta)) = \frac{1}{\sqrt{2\pi}}$, the condition that $\sigma < \tau^2\sqrt{2\pi}$ is sufficient for $f(\theta;\cdot)$ to be monotonic in $\theta$. □

*Proof of Proposition 2.* Suppose that $\theta^* \notin (\tilde{\tilde{\theta}}(x^*), \tilde{\theta}(x^*)) \neq \emptyset$. By definition of $\tilde{\tilde{\theta}}(x^*)$ and $\tilde{\theta}(x^*)$, and noting $x^* = \theta^* + \sigma\Phi^{-1}(\theta^*)$, we can recover the expression

$$\tilde{\theta}(x^*) = \theta^* + \sigma\Big[\Phi^{-1}(\theta^*) - \Phi^{-1}(\tilde{\tilde{\theta}}(x^*))\Big].$$

If $\tilde{\tilde{\theta}}(x^*) \geq \theta^*$, then $\tilde{\theta}(x^*) \leq \theta^*$, a contradiction. Therefore, we have $\tilde{\tilde{\theta}}(x^*) < \theta^*$. However, it is then implied that $\tilde{\theta}(x^*) > \theta^*$, a contradiction. □

*Proof of Proposition 3.* For contradiction, suppose diversionary escalation occurs in symmetric monotone equilibrium for a measurable set of types as $\sigma \to 0^+$. Let equilibrium thresholds $x^*(\sigma) \to x^*$ and $\theta^*(\sigma) \to \theta^*$ as $\sigma \to 0^+$.

First, suppose $\theta^* \geq t$. Then, by $\Phi(\frac{1}{\sigma}(x^*-\theta^*)) < \Phi(\frac{1}{\sigma}(x^*-t)) + p$, there is no diversionary escalation. Second, consider $\theta^* < t$. For any thresholds $(x^*(\sigma), \theta^*(\sigma))$ and signal quality $\sigma > 0$, the ex-ante probability of diversionary war can be expressed in terms of $\theta^*(\sigma)$ by the definition of $\tilde{\tilde{\theta}}(x^*(\sigma))$ and $\tilde{\theta}(x^*(\sigma))$ and the fact that $x^*(\sigma) = \theta^*(\sigma) + \sigma\Phi^{-1}(\theta^*(\sigma))$,

$$\begin{aligned}\gamma(\theta^*(\sigma);\cdot) = 1 + \Phi&\left(\frac{1}{\tau}\left[\Phi\left(\frac{\theta^*(\sigma)-t}{\sigma}+\Phi^{-1}(\theta^*(\sigma))\right)+p-z\right]\right)\\ &-\Phi\left(\frac{1}{\tau}\left[\theta^*(\sigma)+\sigma\Phi^{-1}(\theta^*(\sigma))-\sigma\Phi^{-1}\left(\Phi\left(\frac{\theta^*(\sigma)-t}{\sigma}+\Phi^{-1}(\theta^*(\sigma))\right)+p\right)-z\right]\right).\end{aligned}$$

Plugging into equation (A1), we recover the expression

$$f(\theta;\tau,\sigma,z,t,p) = \gamma(\theta;\cdot)\Phi\left(\frac{\theta-\mu(\theta+\sigma\Phi^{-1}(\theta))}{\nu}\right) + (1-\gamma(\theta;\cdot))\Phi\left(\frac{t-\mu(\theta+\sigma\Phi^{-1}(\theta))}{\nu}\right) - c.$$

For any $\sigma > 0$ and corresponding equilibrium threshold $\theta^*(\sigma)$, $f(\theta^*(\sigma);\cdot) = 0$. Noting that $\theta^* < t$, we can recover $\lim_{\sigma\to 0^+} f(\theta^*(\sigma);\cdot) = \Phi(t-\theta^*) - c$, implying $\theta^* = t - \Phi^{-1}(c)$.

First, if $c \leq \frac{1}{2}$, then $\theta^* = t - \Phi^{-1}(c) \geq t$, a contradiction. Second, if $c > \frac{1}{2}$, then

$$\lim_{\sigma\to 0^+} \tilde{\tilde{\theta}}(x^*(\sigma)) = \lim_{\sigma\to 0^+} \Phi\left(\frac{1}{\sigma}\Phi^{-1}(c) + \Phi^{-1}(t-\Phi^{-1}(c))\right) + p = 1 + p > 1.$$

However, by Proposition 2, diversionary escalation requires $\theta^* \in (\tilde{\tilde{\theta}}(x^*), \tilde{\theta}(x^*))$. If $\theta^* > 1$, then $A(\theta^*, r = 0) = 0$, a contradiction. $\square$

Figure A-14: Example of the first page of a table. The rows several columns are merged where they contain the same values.

**Table of ceasefire violations as of 15 April 2018[1]**

| SMM position | Distance | Direction | Means | No. | Observation | Description | Weapon | Date | Time |
|---|---|---|---|---|---|---|---|---|---|
| SMM camera at Donetsk Filtration Station (15km N of Donetsk) | 1-3km | S | Recorded | 1 | Explosion | Undetermined | N/K | 13-Apr | 19:39 |
| | | | | 1 | Explosion | Undetermined | | | 19:41 |
| | | | | 1 | Projectile | W to E | | | 19:54 |
| | | | | 5 | Projectile | E to W | | | 21:42 |
| | | | | 2 | Projectile | E to W | | | 21:43 |
| | | | | 1 | Explosion | Undetermined | | | 22:34 |
| | | | | 2 | Projectile | W to E | | 14-Apr | 01:09 |
| | | | | 2 | Projectile | W to E | | | 01:17 |
| | | | | 1 | Explosion | Undetermined | | | 01:47 |
| | | | | 1 | Projectile | W to E | | | 02:06 |
| | | | | 4 | Projectile | W to E | | | 02:25 |
| | | | | 3 | Projectile | W to E | | | 02:28 |
| | | | | 1 | Projectile | W to E | | | 02:29 |
| | | | | 1 | Projectile | E to W | | | 02:30 |
| | | | | 1 | Projectile | W to E | | | |
| | | | | 1 | Projectile | E to W | | | 02:46 |
| | | | | 1 | Projectile | E to W | | | 02:48 |
| | | | | 2 | Projectile | W to E | | | |
| | | | | 2 | Projectile | E to W | | | 02:49 |
| | | | | 1 | Projectile | W to E | | | |
| | | | | 2 | Projectile | E to W | | | 02:50 |
| | | | | 2 | Projectile | E to W | | | 02:51 |
| | | | | 1 | Projectile | W to E | | | 03:22 |
| | | | | 2 | Explosion | Undetermined | | | 04:29 |

[1] The table only includes ceasefire violations directly observed by the SMM and it may include those also assessed to be live-fire exercises, controlled detonations, etc. Details provided – in terms of distance, direction, weapons-type, etc. – are based on assessments provided by monitors on the ground, and are not always necessarily precise. When information is not known (indicated with an “N/K”), the SMM was unable to ascertain such information due to distance, weather conditions and/or other considerations. Ceasefire violations recorded by more than one patrol and assessed to be the same are entered only once.

14856/18mf

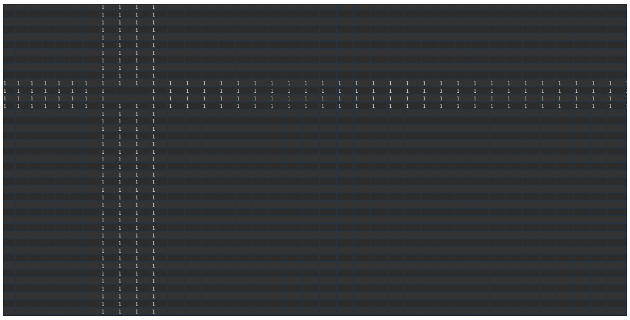
Figure A-15: Example of a subsequent page of a table. Several columns are empty, because the rows are merged with the ones of the previous page (see Figure A-14).

2

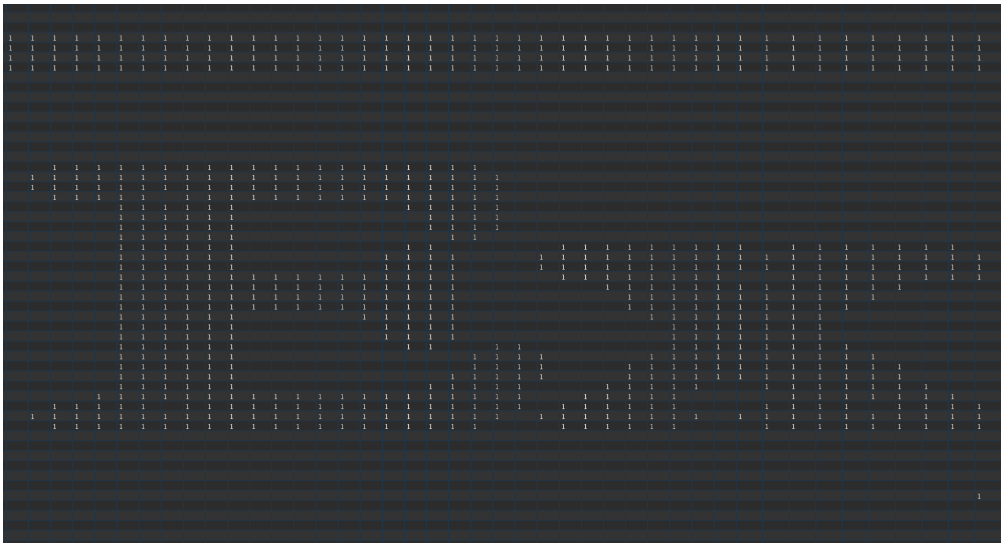
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | 2 | Explosion | Undetermined | | | 04:43 |
| | | | | 6 | Projectile | W to E | N/K | | 20:22 |
| | | | | 3 | Projectile | W to E | | | 20:23 |
| | | | | 6 | Projectile | W to E | | | 20:31 |
| | | | | 3 | Explosion | Undetermined | | | 20:37 |
| | | | | 1 | Explosion | Undetermined | | | 20:38 |
| | | | | 12 | Projectile | W to E | | | 20:56 |
| | | | | 1 | Projectile | W to E | | | 20:57 |
| | | | | 3 | Projectile | W to E | | | 21:01 |
| | | | | 3 | Projectile | W to E | | | 21:03 |
| | | | | 1 | Explosion | Undetermined | | | 21:04 |
| | | | | 5 | Projectile | W to E | | | 21:07 |
| | | | | 1 | Projectile | W to E | | | 21:10 |
| | | | | 2 | Projectile | W to E | | | 21:13 |
| | | | | 1 | Projectile | W to E | | | 21:15 |
| | | | | 1 | Explosion | Undetermined | | | 21:16 |
| | | | | 2 | Projectile | W to E | | | 21:19 |
| | | | | 1 | Explosion | Undetermined | | | 21:22 |
| | | | | 1 | Explosion | Undetermined | | | 21:41 |
| | | | | 1 | Projectile | E to W | | | 21:44 |
| | | | | 1 | Projectile | W to E | | | 21:51 |
| | | | | 1 | Explosion | Undetermined | | | 21:59 |
| | | | | 2 | Explosion | Undetermined | | | 22:07 |
| | | | | 1 | Explosion | Undetermined | | | 22:16 |
| | | | | 1 | Explosion | Undetermined | | | 22:49 |
| | | | | 3 | Projectile | E to W | | | 22:56 |
| | | | | 1 | Projectile | W to E | | | 22:59 |
| | | | | 2 | Projectile | W to E | | | 23:06 |
| | | | | 2 | Projectile | E to W | | | 23:07 |
| | | | | 1 | Explosion | Undetermined | | | |

Figure A-16: Example of what table margins and letters look like under *raster::boundaries()* and *raster::clumps()*.

(a) Boundaries of a margin intersection.



(b) Boundaries of letters "Ex" with a cell margin above.



(c) Clumps of letters "Ex" with a cell margin above.

Figure A-17: Margin structure of page 1 (see Figure A-14).

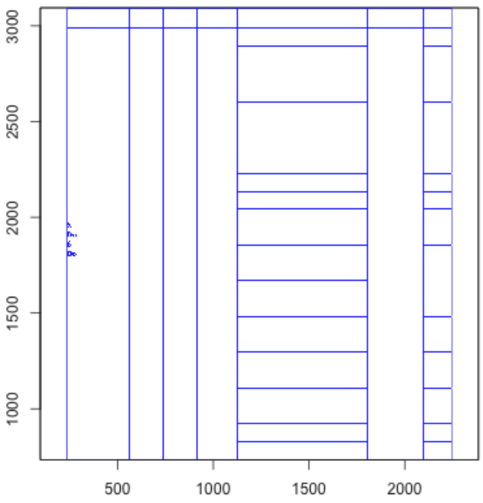



Figure A-18: Example of column 6 of page 1 (see Figure A-14).

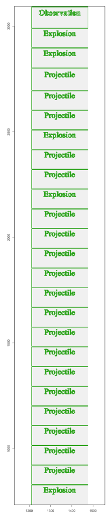


Figure A-19: Visualization of the example with the strike from TableID 347981. SMM position is 2.9 km NNW of Lebedynske, and the strike itself was observed 3 km East of that. The aerial image obtained from Google Maps.