Social media and Russian territorial irredentism: some facts and a conjecture

Jesse Driscoll & Zachary C. Steinert-Threlkeld

To cite this article: Jesse Driscoll & Zachary C. Steinert-Threlkeld (2020) Social media and Russian territorial irredentism: some facts and a conjecture, Post-Soviet Affairs, 36:2, 101-121, DOI: 10.1080/1060586X.2019.1701879

To link to this article: https://doi.org/10.1080/1060586X.2019.1701879

Published online: 12 Dec 2019.
Social media and Russian territorial irredentism: some facts and a conjecture

Jesse Driscoll and Zachary C. Steinert-Threlkeld

ABSTRACT
After Kremlin policymakers decided to incorporate the territory of Crimea into Russia, updates on public attitudes in Russian-speaking communities elsewhere in Ukraine would have been in high demand. Because social media users produce content in order to communicate ideas to their social networks, online political discourse can provide important clues about the political dispositions of communities. We map the evolution of Russian-speakers’ attitudes, expressed on social media, across the course of the conflict as Russian analysts might have observed them at the time. Results suggest that the Russian-Ukrainian interstate border only moved as far as their military could have advanced while incurring no occupation costs – Crimea, and no further.

ARTICLE HISTORY
Received 6 February 2019
Accepted 20 October 2019

KEYWORDS
Ukraine–Russia war; social media; Novorossiya; Crimea; russkoyazichnoe naselenie

Post-Soviet Affairs
2020, Vol. 36, No. 2, 101–121
https://doi.org/10.1080/1060586X.2019.1701879

© 2019 Informa UK Limited, trading as Taylor & Francis Group

I would like to remind you that what was called Novorossiya, back in the tsarist days—Kharkov, Lugansk, Donetsk, Kherson, Nikolayev, and Odessa—were not part of Ukraine back then. The territories were given to Ukraine in the 1920s by the Soviet government. . . . Why? God only knows! . . .

–Vladimir Putin, 2014

Introduction: some facts

On 21 November 2013, protests in Kyiv started against President Victor Yanukovych. On the night of 21 February 2014, Yanukovych fled Ukraine. On or around 27 February 2014, Russian special forces entered Crimea. Russia’s interstate borders seemed to be expanding, and Ukraine’s contracting, under a “regime phase” of territorial realignment (Lustick 1993, 123). With the map already re-drawn, and no Ukrainian authority to make arrests, newly formed pro- and anti-government militias acted on their own accord. Throughout 2014 these militias clashed with each other and brutalized civilians. Few of the anti-government militias had much success seizing or holding government buildings or other symbols of power. The exception to this general rule was the eastern Donbas region of Ukraine, where indigenous insurgents captured the regional apparatus of the state in two regions (oblasts), Donets’k and Luhans’k.

Militias clashed, then consolidated, and, eventually, formed stable coalitions with hierarchical chains of command. Those coalitions are today referred to as “the Ukrainian army” and “the secessionist rebels” in both academic and policy shorthand. In the first two years of fighting, approximately 3,000 civilians were killed (OHCHR (Office of the United Nations High Commissioner for Human Rights) 2019, 7). More than one million individuals fled their homes as refugees or internally displaced persons (IDPs). Property and industry damage is estimated in the tens of billions of dollars. Zones of fighting calcified into stable front lines in the winter of 2015 after Russia sent
regular troops to tip the scales at two critical junctures, the battles of Ilovaisk and Debaltseve. Territory has not changed hands significantly since those battles. As the war conventionalized along a territorially fixed line of contact, the brutalization of civilians slowed.

We are now five years in. What has emerged is not the ethnic bloodbath some experts feared. There has been no Srebrenica. Ethnic identities have hardened, perhaps, but not in simple ways pitting ethnicized Russians against ethicized Ukrainians. Civilizational and religious fault lines have not been weaponized as some pessimists predicted. Most Russian-speakers living in Ukraine in 2014 rejected calls for rebellion against the post-Yanukovych government.

Many born in historical Novorossiya went further, actively performing their Ukrainian patriotism by volunteering to fight off an invasion. Consider Figure 1. It is a map of Ukrainian martyrs per capita by rayon. Each dot represents the fraction of an oblast’s population that have died on the front lines, usually victims of shells fired from the territory of the self-declared “Donets’k People’s Republic” (DNR) or “Luhans’k People’s Republic” (LNR). It has been well documented that many pro- and anti-Putin fighters on both sides have pilgrimaged from distant lands, but it is also well documented that many fighters are locally recruited. Russian is the lingua franca on both sides of the line of contact. Many soldiers imagine themselves to be fighting for their homes. Ukraine’s war pits Russian-speakers who accept the premises of the Russian state narrative against Russian-speakers who are inoculated against that narrative.

**Figure 1.** Military deaths as a fraction of total rayon population on the Ukrainian side since the spring of 2014. Data on the birthplace of the deceased are from Ukrainian Memorial. Data on rayon populations are from the 2001 Ukrainian census. A thick line surrounds the territory of historical Novorossiya, which in defiance of early predictions is producing anti-invasion/anti-Putin martyrs at a consummate with other parts of Ukraine. The largest dot represents the city of Dnipro. Crimea is shown on this map and others in this paper as part of Ukraine in order to reflect the plasticity of interstate borders that would have been felt at the time.
This paper uses social media data to reconstruct how the Russian-state narrative was received by Russian-speakers living in Ukraine during the critical period between February 2014 (when Yanukovych fled Kyiv) and the Battle of Ilovaisk (when the Russian military intervened directly and froze the territorial front lines). Our conjecture is that during that time, policy elites in Moscow would have been considering using their conventional military to move the undeclared front lines of the war farther West. These planners would have been hungry for information on the social attitudes of Ukrainian Russian-speakers (*russkoyazychnoe naselenie*). Russian planners would have wanted to know if they were interested in opting out of the Ukrainian polity.

We show that Russian-language social media traffic could have been one new source of military intelligence. Since the prevalence of overtly political behaviors on social media provides important clues about the political dispositions within communities, a growing body of scholarship has taken advantage of these data to understand contentious action in Ukraine (Metzger, Nagler, and Tucker 2015; Onuch 2015; Metzger and Tucker 2017; Wilson 2017). Our departure from previous studies is emphasizing the potential for these data to be repurposed for crisis decision-making. As proof of concept, we reconstruct a number of different maps of social attitudes shared by Russian-speakers active on social media. Our dataset contains approximately 7 million online user entries (tweets), all generated within the territorial borders of Ukraine. Aggregated patterns in the data we analyze provide a measure – noisy, but informative – of how many self-identified Russians living in Ukraine would have favored border revision. Most did not.

Our supposition is that if Russian strategists were considering expansion beyond Crimea, they would have been able to use social media information to assess, with a great deal of precision and in real time, the reception that they would likely receive. Since interstate border changes are rare events, the re-purposing of public data for military reconnaissance has not yet been considered despite excellent studies of how polarized media bubbles allow conflicting coverage of the same events (Warren 2014; Baum and Zhukov 2015; Peisakhin and Rozenas 2018), how internet connectivity enables cyber-operations (Gartzke 2013; Kostyuk and Zhukov 2017), and the advantages that some states seek by deliberately muddying the historical record (Beissinger 2015b; Laurelle 2015; Szostek 2017; Hopf 2016; Hale, Shevel, and Onuch 2018; Snyder 2018).

**Background: divergent narratives**

Violence between self-identified Russians living in Eastern Ukraine and their self-identified Ukrainian neighbors was not an issue after the disintegration of the Soviet Union. In a study comparing the characteristics of four Russian-speaking “beached diasporas” – communities that found themselves living on parts of what they construed as their homeland, but divided among new post-Soviet states – Laitin (1998) attributes peaceful interethnic relations in Ukraine to a combination of deterrence and the ambiguity of political identity boundaries:

> The major mechanism holding back interethnic violence in Ukraine … is the feeling by Russians … that if they were ever terrorized (*qua* Russians) by the [Ukrainians], the Russian Federation would come to their aid. … But another mechanism reducing the likelihood of interethnic violence … is the embarrassing fact (for both sides) that the boundaries of opposition are not at all clear.¹

The breakdown of this peaceful equilibrium began in the fall of 2013. On 21 November 2013, Ukrainian President Viktor Yanukovych declined to sign an association agreement with the European Union (EU) in order to explore membership in Russia’s Eurasian Economic Union (EAEU). This reversal was seen as the culmination of years of friction and competition between Russia and Western Europe (Colton 2016; Charap and Colton 2017). Social forces mobilized. Maidan Square in central Kyiv became a focal point for “Euromaidan” protests that, as the weeks passed, took on an all-or-nothing anti-regime flavor. Clashes between state security forces and armed protesters gradually produced martyrs. Ukraine’s government imploded on 21 February 2014. Yanukovych fled Kyiv that night.
Russian special forces seized Crimea a few days later. The popular understanding within Russia was that its military was “coming to the defense” of ethnic Russians at risk. Over the next few months, military drills along the border provided cover for 50,000 Russian soldiers to mass, signaling that Russia could invade “at a moment’s notice” (Charap and Colton 2017, 132). The areas that eventually became the front lines of the conventional war – parts of Donet’sk and Luhans’k – were areas directly adjacent to Russian territory, where secessionist militias could anticipate the possibility of easy military resupply. Between February and August, on the Western side of the gradually solidifying conventional front lines, there were sporadic attempts by provocateurs sympathetic to the Russian cause to provoke general uprising in Eastern Ukraine (historical Novorossiya). Most failed. The Russian military did not send aid to militias outside of Crimea or the Donbas. Though widespread speculation of clandestine Russian assistance persists, and is made plausible by a few prominent pro-Kremlin volunteers – and more facts may yet come to light – direct Russian intervention did not occur until July (four months after Crimea and 10 weeks after the Ukrainian government began its “Anti-Terrorist Operation” (ATO) to forcibly re-incorporate the East). In the end, Russia only sent conventional ground forces to assist secessionist militias in the areas of the Donbas that had already demonstrated capacity to hold government buildings for months.

Though Russia did not engage in overt kinetic military activity outside of Crimea, Russian-language media broadcasts during the time represent an exemplary information warfare campaign. One goal was to solidify Russian domestic opinion. Another was to encourage Russian-speakers within Ukraine to take advantage of the temporary window of Ukrainian state incapacitation and rise up. Petersen (2001) identifies three analytically distinct triggering mechanisms that can impel leaderless resistance: (1) the amplification of emotions of resentment, especially caused by prospective status reversals for one’s ethnic group and subordination to another ethnic group (especially a hated one); (2) co-ordination on a few focal points and infusing them with special symbolism; and (3) valorization of heroic resistance, assuring citizens that incurring small risks of martyrdom will be accompanied by large community status rewards.

All three triggering mechanisms were prominent in the content of Russian television coverage of post-Maidan Ukraine. Emphasis on status reversals for Russians was overt. A constant barrage of news stories – including fabricated stories about Russian boys being crucified by Ukrainian far-right groups and staged photographs of soldiers proudly displaying flags of the Azov paramilitary group alongside the NATO flag and a Nazi flag – left no doubt that Russians, stranded in Ukraine, were potential hostages and under imminent threat. The subordination of Russians in a new status hierarchy below Ukrainians was a reoccurring theme. Valorization of heroic resistance to Ukrainian fascism was accompanied by promises of status rewards to patriotic volunteers from across the former Soviet Union. The reciprocal decision by the post-Maidan Ukrainian government to respond to uprisings with a national counterinsurgency policy called the “Anti-Terrorist Operation” (ATO) was also obviously strategic messaging, meant to resonate in NATO capitals and in the imaginations of Ukrainian patriots (Scholz 2016; Boyd-Barrett 2017).

This gloss is not meant as a comprehensive history, simply an amuse-bouche to whet the appetite for empirical exposition. After an unexpected regime change in Kyiv, Russian-speakers were provided two competing narratives to make sense of the tectonic political shift. “Different sets of anchoring” keywords – one promulgated by the Kremlin and the other promulgated by the new government in Kyiv – resulted in bifurcated narratives. These narratives contain well-understood focal points (coup, fascist, terrorist, invasion, etc.) calibrated to exile one’s political enemies from respectable coalition politics. Table 1 summarizes the competing narratives.

Methods: mapping divergent narratives

Starting on 26 August 2013, we connected to Twitter’s streaming application programming interface (API), requesting only tweets with GPS coordinates. We first filtered for time, focusing on the 188 days from 22 February 2014 to 28 August 2014. This filter generated a sample of roughly 940,000,000
Table 1. Common components of competing narratives.

<table>
<thead>
<tr>
<th>Public diplomacy “keywords and focal points” points for Ukraine</th>
<th>Pro-Kremlin narrative</th>
<th>Anti-Kremlin narrative</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maidan Events</strong>: Coup, Unconstitutional, Fascists</td>
<td><strong>Maidan Events</strong>: Revolution of Dignity</td>
<td></td>
</tr>
<tr>
<td><strong>Post-Maidan Violence</strong>: Civil War</td>
<td><strong>Post-Maidan Violence</strong>: Invasion, Aggression</td>
<td></td>
</tr>
<tr>
<td><strong>Crimea</strong>: Self-Determination, Homecoming</td>
<td><strong>Crimea</strong>: Invasion, Occupation</td>
<td></td>
</tr>
<tr>
<td><strong>The appropriate Russia-Ukraine relationship, taking the relevant historical facts into account, ought to be one of:</strong> … “</td>
<td>… <strong>natural hierarchy. Post-Soviet borders are historically contingent in any case.</strong></td>
<td></td>
</tr>
<tr>
<td>“Above all, we should acknowledge that the collapse of the Soviet Union was a major geopolitical disaster of the century … Tens of millions of our co-citizens and compatriots found themselves outside Russian territory.” (Putin, Annual Assembly Address, 2015; <a href="http://en.kremlin.ru/events/president/transcripts/22931">http://en.kremlin.ru/events/president/transcripts/22931</a>)</td>
<td>“All Members shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations.” (UN Charter, Article 2, Section 4)</td>
<td></td>
</tr>
<tr>
<td><strong>Future historians, writing about the Maidan events, will describe them as:</strong> … “</td>
<td>… <strong>a coup by far-right social forces, emboldened by material and moral support of the NATO alliance and Western intelligence agencies.</strong></td>
<td></td>
</tr>
<tr>
<td>“… from all appearances [the protesters] are controlled not so much by the oligarchic groups [in Ukraine] but to a significant extent by Polish and British intelligence services. … Given this, Russia is obliged to interfere in the geopolitical intrigue of the European Union, aimed against the territorial integrity of Ukraine.” (Novaya Gazeta 2015 attributed to Malofeev; Toal 2017, 248; <a href="https://www.unian.info/politics/1048525-novaya-gazetas-kremlin-papers-article-full-text-in-english.html">https://www.unian.info/politics/1048525-novaya-gazetas-kremlin-papers-article-full-text-in-english.html</a>)</td>
<td>“The problem of ‘prodazhnost’ [corruption] was also that in a world where everyone could be bought, there was no trust among people. Trust was something rare and precious, given and received only among close family and friends … The Maidan was all the more a miracle in such a society.” (Shore 2018, 267)</td>
<td></td>
</tr>
<tr>
<td><strong>The proximate cause of the violence in East Ukraine is:</strong> … “</td>
<td>… <strong>the CIA coup which brought fascists to power.</strong></td>
<td></td>
</tr>
<tr>
<td>“Would it be acceptable for Russia, considering its international standing, to keep mum and recognize the coup in Ukraine, and to leave Russians and Russian speakers in Ukraine in the lurch after the first order issued by the organizers of the anti-constitutional armed revolt, which was supported by their foreign sponsors, banned many things that were connected with the Russian language? … Had we not done what we did, we would have betrayed our civilization.” (Lavrov 2017 as quoted in Freeman 2019, 82)</td>
<td>“[T]here is considerable evidence to indicate that Russian state security structures worked in partnership with ostensibly private but functionally extended state networks of influence – oligarchic groups, veterans organizations, nationalist movements, biker gangs, and organized criminal networks – to encourage, support, and sustain separatist rebellion in eastern Ukraine from the very outset.” (Toal 2017).</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th></th>
<th>Pro-Kremlin narrative</th>
<th>Anti-Kremlin narrative</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Any account of the violence in East Ukraine is incomplete if it does not reference deeper structural causes, such as: …”</td>
<td><strong>… decades of Western policies to encircle Russia, expanding NATO and aggressively pushing regime change in post-Soviet states under the aegis of democracy promotion.</strong></td>
<td><strong>… a basic incompatibility of values between Putin’s regime and the Western security community.</strong></td>
</tr>
<tr>
<td>“Soldiers fighting to secede from Eastern Ukraine are best described as: …”</td>
<td><strong>… Russian patriots.</strong></td>
<td><strong>… terrorist insurgents.</strong></td>
</tr>
<tr>
<td><strong>“One state and, of course, first and foremost the United States, has overstepped its national borders in every way. This is visible in the economic, political, cultural and educational policies it imposes on other nations.”</strong> (Putin, Munich Conference, 2007; <a href="http://en.kremlin.ru/events/president/transcripts/24034">http://en.kremlin.ru/events/president/transcripts/24034</a>)</td>
<td><strong>“Now if this sounds familiar, it’s what Hitler did back in the 30s… Hitler kept saying: ‘They’re not being treated right. I must go and protect my people.’ And that’s what’s gotten everybody so nervous.”</strong> (Hillary Clinton; <a href="https://www.reuters.com/article/us-usa-politics-clinton/hillary-clinton-tries-to-fix-putin-hitler-comparison-idUSBREA242H520140305">https://www.reuters.com/article/us-usa-politics-clinton/hillary-clinton-tries-to-fix-putin-hitler-comparison-idUSBREA242H520140305</a>)</td>
<td><strong>“… terrorist organization: a stable association of three or more persons established to terrorist activity, in which made the distribution functions have certain rules of conduct required of persons while preparing and committing terrorist acts.”</strong> (Law of Ukraine – On the Fight Against Terrorism; <a href="https://bit.ly/2SAaQjL">https://bit.ly/2SAaQjL</a>)</td>
</tr>
<tr>
<td><strong>“The actions of the people of Crimea and Sevastopol remind me of the actions of Red Army soldiers during the first tragic months after the breakout of the Great Patriotic War, when they tried to battle through to join their comrades and carried their field flags close to their hearts.”</strong> (Putin, Crimean Spring, 2019; <a href="http://en.kremlin.ru/events/president/transcripts/speeches/60096/print">http://en.kremlin.ru/events/president/transcripts/speeches/60096/print</a>).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
geotagged tweets, which we then reduced to 6,880,623 tweets originating within the territorial borders of Ukraine (Crimea inclusive).

We divide the sample into three periods: Crimea (dated from the flight of Viktor Yanukovych on 22 February until the 15 March voting referendum in Crimea); the post-Crimea period (in which local forces organized for secession knowing that the Ukrainian-Russian interstate border was in flux, which ended with the election of Petro Poroshenko via a mass-participation nonviolent voting exercise on 26 May); and the subsequent conventionalized artillery war in the Donbas region (27 May 2027 until 28 August). By the time By the time Russian armor intervened directly in late August at the battle of Ilovaisk, it was clear that the ease of seizing Crimea would not be repeated.

Two primary considerations caused us to prefer Twitter to Facebook or VKontakte. One concern was minimizing platform bias. Platform choice was itself a signal of political preferences: Facebook had been an organizing platform for pro-Maidan activists, so Vkontakte had a reputation as an “anti-Facebook” meeting place for anti-Maidan users (Gruzd and Tsyganova 2015). Twitter, by contrast, was new enough at the time that it had little reputation beyond being a popular social media platform that did not censor. It contained a sufficient population of both pro-Kremlin and anti-Kremlin users for analysis.

Multiple studies have shown that Twitter contained, and perhaps still does, substantial numbers of pro-Kremlin and anti-Kremlin Russian-language accounts and tweets. An analysis of Russian Twitter users from 2010–2011 shows that many users focused on Ukraine, including a sizable group with a positive attitude towards Russia (Kelly et al. 2012). Pro-Putin attitudes were popular on Twitter in Russia around the 2011 Duma and 2012 presidential elections and continued to track offline events (Spaiser et al. 2017).

Second, it was more practical for our team to acquire and work with large quantities of Twitter data than VKontakte or Facebook data in 2014. VKontakte’s API is neither well documented in English nor reliably uncensored. Using Facebook profile data requires working with internal researchers, providing the company veto power at every stage of research.

To make responsible inferences about public opinion expressed on social media, it is necessary to contrast earnest reproduction of keywords that signal support for the Kremlin narrative against the prevalence of users overtly rejecting those arguments. For this, we applied language filters. Metzger, Nagler, and Tucker (2015) demonstrate that many multilingual Twitter users performed solidarity with Maidan protesters by activating their Ukrainian identity by communicating on Twitter in Ukrainian, but they switched back into Russian after the success of Maidan. An interpretation is that they switched back in order to participate in online information warfare – the replication (or self-production) of pro- or anti-Kremlin propaganda. Our study restricts the investigation to Russian-language content since the Russian-speaking populations (russkoyazychnoe naselenie) are the subsample of Ukrainian citizens whose beliefs would have been most salient to Kremlin strategists.

We employ a dictionary of keywords for parsimony and interpretability. The six months in this study were marked by a series of dramatic, contentious events with their own vocabulary. A few weeks after the voting exercise in Crimea, coordinated protesters occupied government buildings in the eastern Ukrainian city of Kharkiv in April, demanding a referendum on independence. There was an attempt to storm a police station in Mauripol to seize heavy weapons. There was a military siege on the city of Sloviansk that would last through early July (live-webcast, with constant YouTube updates). In early May, violent clashes in Odessa left 42 people dead when a building caught fire. The framing language of “fascism” and “terrorism” was prevalent in descriptions of all these events. Table 2 presents the dictionaries.

Two steps code the tweets. First, a Python script filters all tweets from Ukraine so that each tweet in the sample contains at least one keyword from a narrative’s dictionary. We considered complicating our bag-of-words approach with cases and declensions but, based on initial visual inspection of the sample, opted for a dictionary including only nouns and adjectives in the nominative case for initial filtering. The irregular use of declensions (and irregular spellings generally) on Twitter may be a confound, but we opt for clear coding of meanings in the initial dictionary (knowing we can use
these accounts to build a supervised model that will pull other important variants, including the same words with different declensions – see below). Given the much longer and larger pro-Kremlin dictionary, we have no reason to believe the decision to search only in the nominative biases inferences systematically.

After manually screening for automated accounts (“bots”), this process yielded 5,328 tweets from 1,339 individual accounts. Second, teams of Russian-speakers – four native Ukrainians and three fluent Russian-speaking residents of North America – read each tweet and coded it as pro-Kremlin or anti-Kremlin. This second step is necessary both because irony confuses unsupervised computer classifiers and sometimes has poor inter-coder reliability and because visual inspection was the most reliable way to spot automated accounts. To understand the demographics and professions of these users, we searched Google, Facebook, and VKontakte for each user in our sample. Tentative results suggest that the sample skews slightly young (16–36) and male, but with a bulk of accounts unidentifiable on these characteristics.8

Having identified the 5,328 tweets containing at least one of the keywords from Table 2, we built two separate supervised models to identify pro-Kremlin and anti-Kremlin tweets that our dictionaries might have missed. Hand-coded tweets were used as a training set on which we built each model. Through processes described in our online Supplementary Materials, we stemmed all words, removed stopwords, and dropped all Foursquare-account generated data.9 The resulting classifiers identify 58,689 tweets as pro-Kremlin and 107,041 as anti-Kremlin. Training the supervised models provided us a higher proportion of pro-Russia narrative than the dictionary, bringing our results more in line with other studies of the Ukrainian and Russian Twittersphere (Kelly et al. 2012; Spaiser et al. 2017; Wilson 2017).

Results

Results and interpretation

The keyword filter finds a large majority of the tweets (85%) to be anti-Kremlin. This result surprises, given that the pro-Kremlin selection dictionary is much larger than the anti-Kremlin dictionary.10 Figure 2 is a time plot of the raw data, organized by narrative track. The dotted lines in this figure divide the sample into the three periods discussed above. Until mid-May, the two narratives peak on the same days, suggesting an online clash of narratives as locals used their
accounts to narrate the same offline events competitively (signaling solidarity by performing the “pro-Kremlin or anti-Maidan” line or performing “anti-Kremlin or pro-Maidan” line for an audience in the social network). 11

The greatest density of pro-Kremlin tweets occurred in April and May. During this period, the Russian military had consolidated control of Crimea, but it was unclear whether the Kremlin would come to the assistance of militias who had seized territory and advertised their desire to secede. The anti-Kremlin narrative did not emerge as dominant until the government’s dedicated counterinsurgency policy (the ATO, “Anti-Terrorist Operation”) is initiated in May. The visible outlier in anti-Russian Twitter activity on 17 July is descriptions of the downing of Malaysian Airlines Flight 17 as terrorism.

Figure 2. Vertical dotted lines divide the sample into three periods: Crimea (dated from the flight of Viktor Yanukovych until the 15 March referendum); the post-Crimea period in which local forces organized for secession (16 March to 26 May); and the subsequent conventionalized artillery war in the Donbas region (27 May until 28 August). Cauterized uprisings by Russian-speakers in various parts of Novorossiya occur in late April and early May. The Ukrainian Government’s ATO (“Anti-Terrorist Operation”) is initiated in May. The visible outlier in anti-Russian Twitter activity on 17 July is descriptions of the downing of Malaysian Airlines Flight 17 as terrorism.

accounts to narrate the same offline events competitively (signaling solidarity by performing the “pro-Kremlin or anti-Maidan” line or performing “anti-Kremlin or pro-Maidan” line for an audience in the social network). 11

Just as in all highly polarized news coverage, in some cases these competitive narratives “latched on” to the same offline events, providing different frames for their description, and in other cases the narratives “change the subject” and simply focus on different events. The Twitter Streaming API provides only 1% of tweets and thus prevents the reconstruction of full threads. We must make inferences about offline events based on what amount to conversation snippets. It is still straightforward to observe narratives in tandem. In the immediate aftermath of the downing of Malaysian Airlines Flight 17 (MH17) on 17 July 2014, Pro-Kremlin tweets promulgated the narrative that the Ukrainian military had shot down its own plane. Anti-Kremlin tweets promulgated the more standard
narrative outside of Russia (that a BUK missile had been fired by separatists and that Russia was a state supporter of terrorism). We manually analyzed the subsample of 567 tweets from our human-coded sample from the day before to the week after and found that even in this week, when the airline crash provided a clear focal point for news coverage, a plurality of tweets in our sample referred to other aspects of the conflict (e.g. movement of weapons across the Russian border, POWs captured and exchanged, Ukrainian state security (SBU) operations, stockpiled weapons, battlefronts in Luhans’k, individuals tweeting the location of separatists to authorities, etc.) or employed generic name-calling. Approximately twice as many anti-Kremlin tweets as pro-Kremlin tweets (41% compared to 20%) in our sample referenced the event, suggesting it was more common for the pro-Russian narrative to dwell on other themes.

The front lines, where Russian military intervention would have been operationally feasible at low cost, are of special interest. Figure 3 replicates Figure 2, but using the larger machine-learning dataset and only examining geotagged tweets generated in Novorossiya. This region is where uprisings of ethnic Russians never occurred, but if they had, might have been joined with logistical support from the Russian military.\textsuperscript{12} Note that the two narratives track each other relatively evenly for the first two periods, suggesting an ongoing battle on social media for hearts and minds in Russian-speaking communities in Novorossiya in the early months of the conflict, followed by dominance of anti-Kremlin messaging in the third period once conventional front lines solidified.

The Supplementary Materials show the same graph using data from all of Ukraine, and, not surprisingly, the Anti-Kremlin discourse clearly dominates the entire period. We also undertake two additional robustness checks to ensure that automated accounts (“bots”) do not drive the results from the machine-learning models. First, we submit every user to Botometer, a service that produces a probability estimate that an account is a “bot,” and drop tweets from any account with a probability of being a bot greater than or equal to .4 (Varol et al. 2017). The Anti-Kremlin narrative dominates the Pro-Kremlin one even more once these tweets are removed. Second, we drop tweets from any

\textbf{Figure 3.} The subset of the raw data from the machine-learning dataset ($N = 166,454$) after dropping tweets, replicating Figure 2 using only data generated from oblasts in historical Novorossiya. In March and early April, the pro-Kremlin narrative ebbed and flowed, but was generally dominant in the Novorossiya sample. As the conventional warfare phase gets underway, the anti-Kremlin narrative begins to dominate.
account at or above the 95th percentile of the tweet frequency or friend:follower distribution, as these behaviors are common features of bots (Bessi and Ferrara 2016). Dropping tweets on these criteria does not substantively alter our inferences.

To enable spatial comparisons across Ukraine, we exploit the variation in the prevalence of each narrative across oblasts as a fraction of overall Russian Twitter behavior. We first calculate the percentage of all Russian-language geotagged tweets originating within an oblast. This number is the denominator. To calculate the numerator, we repeat the oblast-level calculation for the population of tweets that contain keywords from either narrative track. The percentage of all tweets in Russian from each oblast (A_i), the percent of all pro-Kremlin tweets in Russian from each oblast (B_i), and the percent of all anti-Kremlin tweets from each oblast (C_i) can be used to compare the percentage difference between B_i and A_i, and the percentage difference between C_i and A_i. This quantity measures the over- or under-production of pro-Kremlin (B_i) or anti-Kremlin (C_i) tweets, against an oblast-specific production baseline. So long as we observe some pro-Kremlin behavior in every oblast there are enough data to make comparisons across space.

The non-parametric nature of this operationalization generates two advantages. First, it mechanically controls for an oblast's Russian-speaking Twitter population and any omitted demographic, social, economic, technological, or political variables that might correlate with the overall percentage of Russian-language Twitter users (since the denominator of each oblast is the total number of Russian-language geotagged tweets). Second, it avoids mechanically capturing the East–West (Russian-Ukrainian) cleavage since it incorporates oblast-specific amounts of Russian production. For example, it is not surprising that Crimea would produce a lot of pro-Russia tweets. What is surprising is that it produces more than would be expected given its baseline production of tweets in Russian. This measure therefore captures the residual “cultural package” of pro-Kremlin attitudes that outlived the institutional implosion of the Party of Regions.

Figure 4. Relative over- or underproduction of narrative track keywords, by oblast, using the full sample of hand-coded data. Oblasts are arrayed roughly from west to east. The pro-Kremlin outliers are Crimea and its capital city, Sevastopol. Overproduction of anti-Kremlin narratives takes place in the west, in the center Kyiv, and on the military front line of the Donbas.
Figure 4 displays relative over- or underproduction of narrative track keywords by oblast using the full sample of our hand-coded data (N = 5,328). Oblasts are “arrayed roughly from” west to east. The clear outlier is Crimea and its capital Sevastopol. The other oblasts where social media users reproduced the pro-Russia narrative were Transcarpathia, Khmel’nyts’kyi, Zhytomyr, Odessa, Dnipropetrovs’k, Kharkiv, Zaporizhzhia, and Donets’k. Overproduction of the anti-Kremlin narrative is more common in the West, but also, crucially, in Mykolayiv, Donets’k, and Luhans’k. These areas are precisely where it would have been easiest for Russia to expand if it wanted to. The Spearman rank correlation between this relative production measure and the proportion of pro-Kremlin tweets by oblast is .58. That many people in these supposedly pro-Russia oblasts were against Russia may have given pause to military planners.

Figure 5 replicates 4 with the machine-coded dataset. Broad trends are similar, but two differences deserve notice. First, the Crimean oblasts are no longer extreme outliers. Second, the East-West dimension of the data is now much more pronounced. Russian-language Twitter behaviors in the seven oblasts that were formerly part of the Habsburg Empire systematically underproduce the focal point keywords in the pro-Kremlin dictionary and overproduce keywords from the anti-Kremlin dictionary. The opposite trend occurs in eight of the nine easternmost oblasts, and also in Odessa. The most active “front lines” of the social media conflict were Mykolayiv and Luhans’k. The Spearman rank correlation between this relative production measure and the proportion of pro-Kremlin tweets

**Figure 5.** Relative over- or underproduction of narrative track keywords, by oblast, using the full sample of machine-coded tweets. Oblasts are arrayed roughly from west to east. Historical Novorossiya overproduces pro-Kremlin narratives and underproduces anti-Kremlin narratives. Mykolayiv, Kyiv, and Luhans’k are notable for the relative overproduction of both narratives.
Interpretation

Our supposition is that spatiotemporal trends in these online social behaviors would have correlated with the offline social behaviors that would have been easily visible to civilians, journalists, or embedded observers reporting to Russian intelligence. The behaviors described spatially in Figure 1, and the main results in Figures 2 and 3 reveal the limits of the Kremlin’s capacity to compete with the West in soft power projection. The Kremlin’s narrative of events seems to have found limited reception in the Russkii Mir, even in Russia’s historical sphere of influence, even for a population historically sympathetic to its message (such as those living in

Figure 6. Relative over- or underproduction of narrative track keywords, by oblast, using the full sample of machine-coded tweets. To visualize, we bin the results: whether the oblast produced 50–100% fewer tweets than expected, 0–50% fewer, no change, 0–50% more, 50–100% more, or more than double the number of tweets expected. Darker colors indicate a relative surge in tweets containing target keywords relative to overall Twitter traffic in the district.

by oblast is .62. Figure 6 includes two paired visual maps of the data in Figure 5, with oblasts shaded corresponding to their relative narrative production.
parts of Novorossiya that had reliably delivered votes to the Party of Regions), even when Kremlin-influenced producers monopolized the airwaves (Peisakhin and Rozenas 2018). All factors suggested, a priori, hegemonic dominance of the pro-Kremlin narrative. Ex-post analysis of outcomes reveals a decidedly mixed picture.

These data suggest that a key point of failure for a Russian “social tip” towards widespread pro-Kremlin sedition against the post-Maidan Ukrainian political regime occurred in Luhans’k. Eventually this oblast emerged as the front line of conventional warfare. In Figure 6, while the patterns of production in Luhans’k are not strongly differentiated from the rest of Novorossiya in terms of pro-Kremlin production, this oblast was a focal point for anti-Kremlin social media activity. The conventional front line was a front line in a war of ideas first: the influx of military activity brought volunteer journalists with Twitter accounts. This altered the sample to reflect a different set of social dynamics (and social media dynamics) than elsewhere.\(^\text{16}\)

Figure 7(a,b,c) allows visual inspection of the 166,454 tweets in a rough time series. The progression in the variation in attitudes by Russian-speakers in Ukrainian oblasts emerges in these snapshots, as a Russian intelligence analyst might have seen them. Data from the first period, 22 February

\[\text{Figure 7. Visual inspection of the 166,454 tweets in a rough time series. (a) (top) 02.22.2014–03.15.2014. (b) (middle) 03.16.- 05.26.2014. (c) (bottom) 05.27.2014–08.28.2014.}\]
through 15 March (Figure 7(a)), suggest there was support for the Russian narrative, and a relative dearth of anti-Kremlin pushback, in territories near Crimea. Occupied Sevastopol’ was, to our surprise, a site of contestation according to these data. Kharkiv especially, but also Zaporizhzhia, Dnipropetrovs’k, and Odessa, might have been tempting targets for annexation in March and April. Between mid-March and late May (Figure 7(b)), there were a few attempts by pro-Russia forces to engineer uprisings. The anti-Kremlin tweets in Mykolayiv in period 2 (Figure 7(b)) probably reflect sentiments by residents, after uprisings in Odessa, expressing fears that Russian planners might be tempted to create a land bridge linking Crimea to Transnistria. After the election of Petro Poroshenko, the consolidation of the post-Maidan Ukrainian state, and the intensification of an artillery war (Figure 7(c)), Russian military intervention would have been more difficult. Outside of Mykolayiv and Luhans’k Russian-speaking, social media users in Novorossiya continued to be relatively receptive to the Kremlin’s point of view in this period. This interpretative exercise is not meant to be the final word on public sentiment by Russian-speaking communities in Novorossiya – just an exposition of how these tools could have been used by a computer-literate observer monitoring social media trends.

**Caveats and speculations**

Twitter was essentially an anti-Kremlin platform in 2014 during our study period everywhere in Ukraine except Crimea. Our supposition is that spatiotemporal trends in online behaviors, which we can measure from a distance, correlate with the offline signaling behaviors that would have manifested in Russian-speaking communities, but it is worth re-emphasizing that extrapolating wider trends in public opinion from these data is hazardous. There may well have been a hidden density of pro-Kremlin Russian-speakers that decamped from Twitter and continued to communicate on Tor, in chat rooms, on forums, or using platforms beyond the scope of our analysis. That said, social media communications can be used to estimate levels of support for seditious political attitudes. Since border revisions are very rare, and since social media are very new, inferential limitations should be made explicit.

Our decisions targeted the opinions of a diverse population (Russian-speaking Twitter users residing within Ukraine) during a complicated period of institutional collapse and state weakness (February–August 2014). Context-specific variables matter. We make no claim to external validity. Twitter is not a perfect substitute for representative public opinion sampling. Since people employ social media for many different reasons and the technology itself is new, the profiles of non-users and users are likely to be systematically different. Even for those that enthusiastically “opt-in” to online politics, every 140-character tweet (or status on Facebook) is not an authentic political act. There is exciting behavioral work to be done, but until there is academic convergence on best practices for how to interpret spontaneous performances on social media, frontier-mapping exercises like ours should be treated with caution.

Three assumptions must hold to justify online observation as measurement for offline behavior. First, it must be the case that individuals in our sample do not maintain a performance identity on social media that promulgates information contrary to offline beliefs. The same people should share the same ideas on Twitter as on other platforms – and, more importantly, as around kitchen tables or on soccer fields. This assumption is plausible but contestable. Context-specific research is needed to sort extremist cheap talk on social media from sincerely held extremist beliefs, especially for populations flirting with radicalization in active war zones.

Second, any study of community signaling behaviors tacitly assumes that the imagined audiences for tweets are local friend and family networks. Spatial comparisons of production patterns across oblasts may provide a window into public sentiment if density of production is related to latent characteristics of the communities Twitter users are trying to influence (and have private information on, as community members). This assumption is plausible, and consistent with academic understandings (McGee, Caverlee, and Cheng 2011), but more site-specific research is needed.
Third, the native population of an area must produce the bulk of the data in a sample. It would be a huge problem if the majority of data emanated from non-community members such as journalists or mercenaries. Strategic efforts to create a false impression of local support through the use of bots or clandestine operatives (which is why we laboriously coded user characteristics) could also contaminate inferences. Manual inspection of the 5,328 tweets and 1,339 accounts convinced us that this study contains relatively few accounts originating from outside an oblast, but we admit caution on this point. A sophisticated information operation, if prepared years in advance, could foil visual inspection of the sort that we employ in this paper. Our inferences assume we are measuring foreign-directed grassroots activism, not astroturf.

Weighed against these concerns are certain advantages of analyzing data from social media platforms. Unstructured data from populations that would be otherwise impossible to reach (in this case Crimea or behind the lines of control in the Donbas) can be analyzed. Unlike surveys, there is no attempt to claim population representativeness, so neither social desirability bias nor strategic non-response confounds inferences. Unlike ethnographic observation, which is limited by the range of the researcher’s own sensory equipment, research designs that employ social media data can compare patterns of production that occur at the same time in many places.

Perhaps the most salient objection to these results is that they are not novel. The East-West split has defined Ukrainian politics since independence (Barrington and Herron 2004; Darden and Grzymala-Busse 2006; Clem and Craumer 2008; Constant, Kahanec, and Zimmerman 2011, 2012; Frye 2015; Zhukov 2016). Using new social media data to draw costly maps that reproduce old maps (such as the second map in Figure 6) may be criticized as old wine in new bottles. There are three reasons not to dismiss this paper’s methodology or results so quickly.

First, unlike a cross-sectional survey, these data mirror the series of updates that would have arrived, in real time, to Russian military personnel during a period of crisis bargaining. New information would have been at a premium for Kremlin policymakers. Maidan, the implosion of the Party of Regions, and Russia’s seizure of Crimea were major events. Old understandings of public opinion would have been held up under close scrutiny. In that moment of crisis, no party, academic or military, would have had time to collect or analyze survey data.

Second, this research has generated new knowledge. The outline of Novorossiya is included in “Figure 1” as a reminder that old maps are not all useful guides to high-stakes behavior by Russian-speaking communities. There were many surprises among our research team as we conducted this study. The dominant narrative used by political elites in Kyiv describing this period is one of the Russian agents sowing discord (which complements the dominant narrative in the United States stating that Russia is an innovator in the information warfare domain). As such, we anticipated finding widespread geographic support for Russia (expressed in the “pro-Kremlin” narrative) and extensive evidence of astroturfing (bots or dubious accounts reproducing Russian talking points). Neither appeared. Only 24 accounts, responsible for 196 tweets, were from bots. Outside of occupied Crimea, most Russian speakers did not use Twitter as a forum to voice support for Russia. The facts were surprising to our team but stubbornly clear.

Our supposition is that the failure of the pro-Kremlin narrative to catch on would have been an important source of military intelligence for Russian planners in 2014. Recall that having begun the process of redrawing the post-Soviet territorial map, it was not clear where Russia would define the natural end-point to its irredentism. Russian mechanized units could have moved quickly to establish facts on the ground if they had expected to find a population ready to greet them as liberators. The front lines of Ukraine’s conflict could easily be many kilometers farther west. Some claim that Russia did not go farther because its leadership feared international censure, but Russian diplomats could have easily justified the action, much as they justified Crimea, by invoking familiar “responsibility to protect” and “self-determination” arguments. That works if and only if many Russians call for help, however. The information that military planners needed for a more ambitious policy, but did not have, is whether they were likely to encounter resistance. If the Kremlin had access to data like ours, they would have known that they were unlikely to be greeted as liberators by many Russian-speaking communities –
even in the Eastern Donbas. The Russian-Ukrainian interstate border moved only as far as Russian forces could advance while incurring no occupation costs – Crimea, and no further.

Third, the question of whether new kinds of technologies – in this case, social media – enable irredentist mobilization is intrinsically worthy of study. If we are correct, social media have under-appreciated implications for revisionist powers trying to assess occupation costs prospectively. “This application” is analytically separate from other well-analyzed applications of social media (e.g. lowering the costs of collective action, lowering the costs for state actors for surveillance of dissident networks, real-time source-checking of “fake news,” “and so on”). When war weaponized radio and film, states had a comparative advantage in what might be called “memetic supply” (the production and dissemination of narrative embedded in memorable slogans, catchy songs, and viral images). Until the recent proliferation of inexpensive smartphones, states did not have the capability to reliably and systematically measure “memetic demand” in real-time. Our empirical results suggest that this capability probably already exists.

Conclusion

Though the master cleavage of this war (Kalyvas 2006: 14, 389) is the East–West (Russophone-Ukrainophone) division, our design is calibrated to illuminate intra-Russian-speaking discordant politics. Our tests assume that Russian-speaking civilians were purposive agents competing with each other to explain the many unexpected offline political upheavals that took place within Ukraine during the study period. Social media behaviors are public signals analogous to scrawling graffiti, whistling a patriotic tune on a bus, talking loudly about politics in a public setting, or flying a flag. Since social media users add content to platforms in order to communicate ideas to their social network, the prevalence of overtly political behaviors can provide important clues about the political dispositions of the community that is the imagined audience for those messages.

We present no evidence supporting the claim that Russian military actions in 2014 were altered as a result of re-purposing social media trends for military intelligence – merely a variety of evidence consistent with our conjecture that such re-purposing is now possible. Social media data are straightforward to analyze systematically and can be collected at a relatively low cost. Following Kostyuk and Zhukov (2017, 3), we favor the analogy between information warfare techniques and airplanes at the start of the First World War. Recall that planes were used primarily for reconnaissance before they were used to drop bombs. Conventional militaries are just beginning to explore the ways that emergent information technologies can shape battlefields. As techniques for real-time data mining become commodified, they will be integrated into best practices for counterinsurgency (Berman, Felter, and Shapiro 2018) and, more generally, into military planning. This paper has shown one way in which they could have been useful.

Notes

1. Laitin (1998, 185), emphasis added.
2. The entire local government institutional apparatus, which represented remnants of the suddenly defunct Party of Regions, accepted Russian rule almost immediately. That said, how much popular support there was for Russian military actions within the permanent population of Crimea will continue to be disputed. See, for example, Suslov (2015) and Faizullaev and Cornut (2017).
4. See Peisakhin and Rozenas (2018). For evidence of saturation of Ukraine-related stories on Russia’s Channel One News, see especially their Figure 1.
5. Russia’s narrative reinforced analogies to World War II (e.g. by “NATO” with “Nazis” using consonant repetition, substituting “Germany” for “the EU,” the explicit claim that Maidan was a CIA coup, etc.) and rabid anti-Americanism. See, for example, Cottiero et al. (2015).
6. We foresee two distinct potential methodological objections to this methodology: (a) that Twitter users are systematically different from non-users; and (b) that Twitter users who geotag tweets are different than other Twitter users. The first concern will be addressed in the text below. Although we do not have an empirical strategy to address the second concern, other research teams employ geotagged Tweets in studies of behavior in Ukraine and find patterns consistent with behavioral expectations (Wilson 2017). More research on geotagged tweet bias, however, is needed. In the United States, geo-located accounts are more likely to be from smartphones, residents of cities, certain minorities, and higher income US census tracts (Malik et al. 2015).

7. Though complex syntax and subtle reasoning can stymie dictionary classifiers (Schwartz and Ungar 2015), we chose a “bag-of-words” approach because 140-character tweets are direct and short. A similar “bag-of-words” method has been used in other event-based studies relying on tweets (Ritter, Etzioni, and Clark 2012; Ramakrishnan et al. 2014).

8. Only 24 accounts, and 196 tweets, are from bots, based on manual inspection. All hand-coded analysis in the paper’s main results excludes them.

9. The first output of the classifier identifies 204,189 tweets as concerning either narrative, 144,776 of which are anti-Kremlin. Validating the output showed many were of the form of “I’m at [place],” indicating that the tweet was created on the app Foursquare. Since no Foursquare tweets had appeared in the hand-coded data, this confound was not discovered until the referee process. Simply removing all tweets that start with “I’m at” (a total of 37,735) from the study did not change results substantively, but we rely on the smaller dataset excluding Foursquare data for all analyses in this paper. See the online Supplementary Materials for more details.

10. The proportion in our sample may be higher than 85%. Hand-coded tweets from the dictionary method contain 273 clearly pro-Kremlin, 4,338 clearly anti-Kremlin, and another 543 with disputable content (e.g. with intercoder variation across the options of “neither” or “both”).

11. The Supplementary Materials contain evidence of spatial and temporal correlations between social media behaviors that earnestly reproduce the anti-Ukraine (“fascist”) narrative and ironic “trolling” behaviors (using the fascist keywords, but intending to mock that position).

12. The quote from Vladimir Putin in the epigraph, openly questioning the legitimacy of the border between Ukraine and Russia, was delivered with scripted sincerity and references these oblasts.

13. We did not aggregate to a lower geographic level, such as city or raion, because of lack of tweets available in our hand-coded sample, especially if we also wanted to subdivide the data into smaller bins by time period.

14. For example, in one period, Crimea produced 4.3% of all Russian-language tweets, but it produced 10.93% of all Russian-language tweets reproducing the pro-Kremlin narrative. Our method would calculate that it produced 154.19% (10.93–4.30)/(4.30) more pro-Kremlin tweets than the baseline expectation. Results are robust to an alternative model specification in which the denominator is the percentage of all geotagged tweets originating from within each oblast regardless of language.

15. We speculate the relative overproduction of pro-Russia discourse in Crimea is explained jointly by a few causal processes: (a) overproduction as a reflection of an authentic broad-based outpouring of support for rejoining the homeland; (b) information operations conducted by pro-Kremlin agents who were not indigenous citizens of the peninsula; (c) the Russian military presence deterred the production of the “anti-Kremlin” narrative by indigenous citizens, creating strategic self-censorship by citizens who left the Twitter-sphere once military occupation was a fait accompli; and (d) the linked claim that in other parts of Ukraine, post-Maidan residual state capacity deterred the irredentist Russian narrative, but those constraints ceased to be present in Crimea very early in our study. Our design cannot disentangle these mechanisms. Because Twitter users are anyway not representative of the entire population, these trends in our data should not be interpreted as evidence that Russian annexation was overwhelmingly popular with the population, an “authentic” victory for national self-determination, or anything of the sort.

16. Luhans’k is not the only oblast that overproduces both narratives. Mykolayiv, Kyiv, Kyiv City, and occupied Sevastopol’ are also sites of contestation.

17. Yet we are aware of no public polling during this period in Crimea, Luhans’k, or Donets’k. The shortcomings of social media should be weighed against the ability to conduct studies that danger would otherwise forbid.

18. See Hill, Zhang, and Rothschild (2016) and Malik et al. (2015). Because behaviors on Twitter replicate known offline phenomena such as Dunbar’s Number (Dunbar et al. 2015) and diurnal patterns of activity (Goldar and Macy 2011), we are cautiously comfortable with this assumption. For a more thorough defense of measuring offline data with online sources, see Steinert-Threlkeld (2018).

Acknowledgement

For excellent research assistance at various stages of the project, we thank Ganna Grebennikova, Aleksandr Karpekov, Yaroslav Makarov, Anton Prokopyev, and Iulija Puchko Wilson. For generosity of critiques, ideas, and time, we thank Dominique Arel, Rex Douglass, James Fearon, James Fowler, Erik Gartzke, Steve Hanson, Jon Lindsay, Margaret Roberts, Phil Roeder, Jacob Shapiro, Branislav Slantchev, Barbara Walter, and Jeremy Weinstein. The work has benefitted from critiques from participants at the UCSD Human Nature Group, the Citizen Lab at the University of Toronto, the Danyliw
Seminar at the University of Ottawa, the Association for Nationalities, the Central Eurasian Studies Society, and the comparative politics groups at UCSD, the University of Michigan, and the College of William and Mary. The editor and two anonymous reviewers helped improve the paper at multiple steps. Translation and transliteration choices for proper nouns from Russian or Ukrainian into English in the main text are somewhat haphazard and ought not be decoded as taking sides in philology wars. All remaining errors are accidental.

ORCID

Jesse Driscoll [http://orcid.org/0000-0002-6880-6268]
Zachary C. Steinert-Threlkeld [http://orcid.org/0000-0002-5122-9660]

References


