Reporting Bias and Information Warfare

Yuri M. Zhukov* and Matthew A. Baum†

International Studies Association Annual Convention
Atlanta, GA, 16-19 March 2016

Abstract

By systematically under- or over-reporting violence by different actors, media organizations convey potentially contradictory information about how a conflict is likely to unfold, and whether outside intervention is necessary to stop it. These reporting biases affect not only statistical inference, but also public knowledge and policy preferences. Using new event data on the ongoing armed conflict in Eastern Ukraine, we perform parallel analyses of data from pro-government, pro-rebel and third party sources. We show that actor-specific reporting bias can yield estimates with vastly different implications for conflict resolution: Ukrainian sources predict an equilibrium with frequent unilateral escalation by rebels, rebel sources predict unilateral escalation by government troops, while Russian and international sources predict that transgressions by either side should be quite rare. Experimental evidence suggests that news consumers tend to support intervention against whichever side is shown to be committing the violence, while news reports without information about actors and tactics reduce support for intervention. We argue that these kinds of reporting biases could potentially make the conflict more difficult to resolve – hardening local attitudes against negotiated settlement, while reducing outside support for third-party intervention.

*zhukov@umich.edu, Assistant Professor of Political Science, University of Michigan
†Matthew_Baum@harvard.edu, Marvin Kalb Professor of Global Communications, Harvard Kennedy School
How we respond to a civil conflict depends on what we know about it. That, in turn, depends on where we get our information. Not every event is observable, and not every observed event is publicly reported. Information providers diverge in the events and actors that attract their attention. One source may focus disproportionately on violence by rebels, another may emphasize government operations, while a third may not attribute violence to any armed group at all. They may do so for commercial or partisan reasons, or because the government controls the press and requires it. Each source offers a different perspective on the conflict, and how violence begins, perpetuates and stops. This variation constitutes reporting bias – the systematic under- or over-reporting of events.

Conflict scholars have recently sought to explain the directions and magnitudes of these biases (Davenport and Ball, 2002), including how the intensity (Weidmann, 2014), location and timing of violence (Hammond and Weidmann, 2014) affect its visibility in the press, and its inclusion in social science datasets (Eck, 2012). These studies have uncovered systematic differences between media- and government-generated conflict data (Weidmann, 2014), between different types of media sources (Earl et al., 2004), and countries of publication (Drakos and Gofas, 2006; Baum and Zhukov, 2015). Such research has sought, thus far with mixed success, to resolve the myriad problems reporting bias can produce for statistical inference.

As important as this research program is for how scholars study conflicts, reporting bias is more than a nuisance for social scientists. It is information warfare. Even if such bias does not directly stem from the influence of the warring parties, it can lead the public to misunderstand the nature of a conflict. Research on effects of selective exposure to partisan media in the United States (Stroud, 2011; Arceneaux, Johnson and Murphy, 2012; Iyengar and Hahn, 2009) shows that exposure to one-sided arguments can drive opinions apart and make compromise more difficult. This tendency should be especially powerful in coverage of war, where consumers typically have little direct personal knowledge beyond what they consume in the press (Baum and Groeling, 2009; DellaVigna et al., 2011). Despite this evidence, conflict research has so far failed to adequately measure the substantive magnitude of reporting bias, how it shapes public understanding of conflicts and resulting support for, or opposition to, taking action to end them.

This study makes several unique contributions to the emerging literature on reporting bias during war. First, while previous studies have focused on the aggregate level (i.e. reporting bias in overall coverage of a conflict) we focus on actor-specific and tactic-specific reporting bias – that is, systematic differences in media coverage of rebels vs. government forces, and of indiscriminate vs. tar-
geted acts of violence. It is relatively straightforward to identify a unique conflict event: actors, casualties, locations and dates. Yet sources differ in which of these ‘hard facts’ they report. For instance, a news outlet may describe violence by actor A simply as ‘violence,’ while describing violence by actor B as ‘violence by actor B.’ News consumers will then, perhaps wrongly, perceive actor B as more responsible for violence than actor A.

Second, in light of this potential for misperception, we explore the implications of these biases not only for statistical inference, but also for public knowledge and public policy. Differential coverage of rebel and government attacks reveals information about how violence by one actor affects violence by the other, which side is more likely to cooperate in equilibrium, which side is more likely to unilaterally escalate, and which is systematically more restrained. This information, in turn, shapes expectations about how a conflict is likely to unfold, and whether and what type of outside intervention is necessary or appropriate to stop it. In the above example, if the public receives a biased message of “more violence by actor B,” it may conclude that actor B is principally to blame for the conflict and support policies aimed at defeating actor B as a means to resolve the conflict. That same conclusion of “more violence by B” will also enter the datasets social scientists use, which may compromise statistical models aimed at explaining the causes and effects of civil conflict.

Using new event data from the ongoing armed conflict in Eastern Ukraine, we perform parallel analyses of media-generated event data from pro-government, pro-rebel and third party sources, to examine how reporting bias affects the empirical study of armed conflict. We investigate the extent to which different sources suggest different patterns of strategic interaction between warring sides, and advance different conclusions about the causes, location and timing of violence.

The Ukrainian conflict presents an opportune test case for several reasons. Due to its location in an economically developed and densely populated part of Eastern Europe, the conflict has received extensive coverage in local and foreign press, producing an abundant supply of event data. It is also a conflict where reporting biases are likely to have significant effects on public knowledge. Political authorities in Ukraine, Russia and rebel-held territories have imposed tight restrictions on news coverage, while limiting alternative sources of information for media consumers. As we show, this highly politicized media environment has helped determine which audiences receive information from which sources, about which actors. Given the highly politicized nature of the conflict’s coverage, scholarly efforts to explain its causes and predict its future course will depend on our ability to understand and account for these biases.
We find that actor-specific reporting bias can profoundly affect both statistical inference and public opinion. According to data from Ukrainian sources, rebels are more likely than the government to unilaterally escalate violence. According to rebel sources, the opposite is true. Both Ukrainian and rebel sources predict more violence in equilibrium than do Russian and international sources. Each perspective has its own implications for how different actors behave in war, the need for third-party intervention, and whether such intervention should be neutral or one-sided.

To investigate the effects of actor-specific reporting biases on policy preferences, we conduct a survey experiment, drawing participants from two national populations (United States and India). We find that respondents in both countries tend to support intervention against whichever side is shown to be committing the violence. By contrast, event reports without information about specific actors and tactics reduce support for intervention, and increase support for a limited, impartial response. These findings suggest that, in addition to confounding the statistical analysis of conflict, reporting bias can mobilize support for and opposition to specific armed groups.

The broader implication of our research is that reporting biases could potentially make conflicts more difficult to resolve. For audiences inside a conflict zone, selective exposure and systematic over-reporting of unilateral violence by the ‘other side’ is likely to harden attitudes against negotiated settlement. For audiences outside a conflict zone, more balanced and neutral coverage in mainstream news reporting – as well as in the event reports of international governmental organizations – may reduce support for external intervention to end a conflict.

In addition to the rebel-government interaction and its implications for conflict resolution, we find that actor-specific reporting bias can affect inferences about when and where violence is most likely to occur. Ukrainian and rebel sources disagree over how territorial control shapes the distribution of violence (Kalyvas 2006), but all sources show that economic vulnerability has been a stronger predictor of violence in Ukraine than language or ethnicity (Zhukov 2016).

Our findings contribute to political science and communications research on media bias (Davenport and Ball 2002; Weidmann 2016), and particularly to the growing literature on information and communications technology (ICT) and violence (Dafoe and Lyall 2015). Due to the pervasive nature of reporting bias, these findings have broader implications for all empirical research on conflict that relies on event data – whether the sources of the data are media firms (Raleigh et al. 2010), NGOs (Lyall 2010) or government archives (Berman, Shapiro and Felter 2011; Sullivan 2014).
This paper proceeds as follows. The first section provides an overview of recent research on media bias in the study of armed conflict. The second offers background on the Ukrainian case, and summarizes main differences in coverage between Ukrainian, rebel, Russian, and international information providers. The third section considers the consequences of these biases for data analysis and theory testing. The fourth section considers the consequences for public opinion, using a survey experiment. The fifth section summarizes our results.

1 Reporting bias and the study of armed conflict

Information providers differ in how they describe an event, and whether they choose to report it at all. Due to the growing reliance of armed conflict research on event data, these differences in reporting have never been more important to the study of violence than they are today. Recent years have seen several notable efforts by scholars of armed conflict to identify the sources of reporting bias, and their consequences for statistical inference.

1.1 Causes of reporting bias

For an event to become news, someone must observe and report it, and an information provider (e.g. media firm, government agency, or non-governmental organization) must publish the relevant details (i.e. actors, casualties, location, date). For news to become data, social scientists need to detect the event report, classify it into a distinct category (e.g. rebel attack, government operation), and convert it into a format suitable for more rigorous analysis. Although selection issues abound in both processes, our focus in this paper is on the first requirement – why some events become news but others do not. In particular, we examine the ‘whodunit’ problem: why information providers report events perpetrated by some actors more than others, and how the actor-specific reporting biases shape what citizens and scholars know about conflict.

One of the most basic sources of reporting bias is lack of information: not all events are equally visible to observers. Events in densely populated urban areas tend to have more eyewitnesses than events in rural areas (Danzger [1975]). The likelihood that eyewitnesses report the observed event may depend on the proximity of event locations to reporting agencies (Moeller [1999] Gans [1980]), or the availability of communications infrastructure, like cell phone towers (Weidmann [2016]). Beyond remoteness, some event locations – like those with ongoing battles
– may simply be too dangerous for reporters to access (Weidmann, 2014).

Once information providers learn of an event, they decide whether to report it – internally or publicly. Here, the incentives of reporting agencies vary greatly. Profit-oriented media firms tend to publish information that maximizes their audience. ‘Newsworthy’ events tend to be large-scale (Woolley, 2000), rare (Snyder and Kelly, 1977), new (Davenport and Stam, 2006), or otherwise salient to the intended audience (Galtung and Ruge, 1965). Where the opportunity costs of event coverage are high, as in print journalism or other media with limited space or time to feature competing stories (Snyder and Kelly, 1977; Davenport and Ball, 2002), the relative ‘newsworthiness’ of an event is a far stronger predictor of coverage than it is for less physically constrained digital media, like newswires, blogs or social media platforms (Wu, 1998; Shoemaker and Cohen, 2012). These market incentives may compound or offset other potential sources of media bias, like ownership structure (Djankov et al., 2001; Gehlbach and Sonin, 2014) and editorial political ideology (Davenport, 2009).

Ironically, another newsworthiness criteria, the norm of providing balanced reporting (Baum and Groeling, 2009) – that is, including the perspectives of both sides – may ultimately be at least as consequential for public opinion as any of these other factors. Balance implies neutrality. Neutrality in a conflict where one side bears the bulk of responsibility, in turn, may be quite different from “truth” or “objectivity.” Borrowing an example from American politics, for a number of years in the mid-2000s, much of the mainstream American media pursued a policy of balance, or neutrality, in its coverage of climate change. This meant that when a scientist appeared on a news outlet, like CNN, discussing the scientific evidence in support of human caused climate change, the network would feature a climate skeptic arguing the other side. This is despite the fact that there existed a virtual consensus in the global scientific community that human caused climate change was both real and a threat demanding a policy response. By treating the views of the skeptic, who represented at most a tiny fraction of scientific opinion, as of equal consequence, the network created a false equivalence between the two (Mayer, 2012). This made it difficult for viewers to understand that one side did, in fact, represent the overwhelming dominant scientific view. In the context of civil conflict, if a media outlet provides equal time to the perspectives of both sides regarding some violent act or fails to attribute blame, even when one side is primarily responsible, consumers will lack the information they need to appropriately attribute responsibility. This, in turn, may depress or misdirect support for external intervention.

Government and NGO sources face somewhat different, yet also in some ways
overlapping incentives. They are different in that government records are not con-
strained by market pressures, and tend to report a higher proportion of observed
events than media sources (McCarthy, McPhail and Smith [1996], Weidmann
2014). However, the specific mission of the government agency (e.g. internal vs.
public reporting), secrecy, and lag time to archival declassification can still
limit the scope of this reporting. NGOs face similar problems of specialization
– focusing, for instance, on particularly egregious human rights violations, rather
than the day-by-day dynamics of armed conflict (Davenport and Ball [2002]).

Conversely, like their news media counterparts, governmental and nongovern-
mental agencies at the state and supra-national level frequently hold neutrality and
balance as core values. As previously discussed, conflating balance, or neutrality,
with objectivity or “truth” can lead to the creation of false equivalency, in which
the culpability of the warring parties is wrongly treated as equivalent. This form
of reporting bias holds potentially profound implications for public attitudes as
well as for public policy.

Beyond source-specific variation in coverage, recent research has highlighted
the importance of the political environment in which information providers are
based. The extent to which media firms are able to act in accordance with ‘news-
worthiness’ considerations depends on the level of press freedom in their me-
dia market (Baum and Zhukov [2015]). Even where they do not directly own
the media, ruling regimes can impose regulations on what media can and cannot
report (Whitten-Woodring and James [2012]) or create norms of self-censorship
(Djankov et al. [2001]), producing cross-national variation in coverage of certain
categories of events (Drakos and Gofas [2006]). Even democratic regimes may im-
pose wartime restrictions on coverage of sensitive topics, particularly those that
may compromise ongoing operations or discredit government policy (Sweeney,
2001; Norris, Kern and Just [2003]; Allan and Zelizer [2004]; Hallin [1989]).

Any one of these potential sources of bias may affect the relative likelihood
that government or rebel violence will receive coverage. On the supply side, re-
porters may have more difficulty accessing rebel-held areas than government-held
ones. On the demand side, commercial and political incentives may discourage
reporting of government violence, and encourage reporting of rebel attacks. The
direction of these biases may differ across information providers and countries.
Although existing research has examined reporting biases with respect to other
‘hard facts’ of conflict events – like casualties (Gohdes and Price [2012]), location
and timing (Hammond and Weidmann [2014], Weidmann [2014]) – actor-specific
reporting bias has mostly eluded rigorous study.
1.2 Consequences of reporting bias

If violent events by some actors are more likely to receive coverage than violence by others, what impact will these biases have on public knowledge and public policy, or on the empirical study of conflict? Research on the consequences of reporting bias is less voluminous than research on its causes, but several findings have emerged. Beginning with the consequences for empirical research, the impact of reporting bias depends on two primary considerations: whether the direction and magnitude of the bias is correlated with the explanatory variable of theoretical interest, and whether the bias is common to all sources.

If reporting bias is uncorrelated with the explanatory variable, then selective reporting affects results as a source of measurement error rather than selection bias (Weidmann, 2014). An example might be the non-detection of events due to random disruptions in a communication network, or differential reporting due to an observable factor other than, and orthogonal to the one under theoretical investigation. The chief concern in these cases is that a potentially large number of false negatives in the data will favor models that under-predict the occurrence of violence, and under-estimate the strength of causal relationships (Type II error).

A different sort of problem emerges if reporting bias is correlated with the explanatory variable of theoretical interest. Although the impact on empirical estimates will depend on the direction of this selection bias, recent studies on such cases have demonstrated a heightened risk of detecting an empirical relationship where none may exist (Type I Error). For example, if there is more reporting in places with more cell phone towers, then the size of the ‘cell tower effect’ on violence will be biased upward (Weidmann, 2016).

A key consideration in how to account for these problems is whether the reporting bias is common to all sources. If the direction and magnitude of bias vary across sources, then the problem is, in one sense, easier to empirically address. Recent research has explored various methods to offset gaps in coverage with information from other sources, including multiple systems estimation (Ball et al., 2003), capture-recapture techniques (Nichols, 1992; Hendrix and Salehyan, 2015), and the pooling of multiple sources with a one-a-day filter to prevent double counts (Leetaru and Schrodt, 2013).

Turning to the consequences for citizens and public policy, while social scientists have tools to correct or at least detect reporting bias, news consumers generally do not. Most consumers have neither the time nor interest to seek out multiple alternative sources of information (Popkin, 1994), and tend to consume media that already aligns with their worldview (Stroud, 2011; Iyengar and Hahn, 2009).
Although political communication scholars have explored the effect of news coverage on public opinion and knowledge (Zaller 1992; Prior 2007; Baum and Kernell, 1999), conflict scholars have generally avoided this topic. This gap is surprising, since much actor-specific reporting bias results from a conscious effort to influence and mobilize public opinion. In the context of an ongoing armed conflict, such bias reflects the proactive policies of conflict actors seeking to alter the external environment to their own advantage. That is, it is information warfare. Governments and activists recognize that selective news coverage can be a powerful tool of propaganda, and tend to promote the kind of reporting that is favorable to their cause, and restrict information that could be damaging (Kumar 2006; Rampton and Stauber 2003; Taylor 1992). The U.S. government, for instance, does not report civilian casualties resulting from drone strikes or counterinsurgency operations. Protest movements tend to deny or underemphasize violent elements within their own ranks, while calling attention to the brutality of the police response. Similar competitions between governments and protesters ensue in estimates of crowd-size at protest rallies (McPhail and McCarthy, 2004). Such biases are particularly acute for information providers whose audience has a direct stake in the conflict – like government agencies and media outlets located in close proximity to a conflict zone. Depending on one’s source of information, a news consumer will likely see only one side of a multifaceted story. One-sided information streams, in turn, can have important effects on public attitudes, including polarizing attitudes of individuals exposed to conflicting one-sided streams (Pariser 2012; Stroud 2011; Levendusky, 2013). This, in turn, makes compromise, and hence peaceful resolution, more difficult and may thus serve to perpetuate conflict.

2 Ukraine’s information war

One of the defining features of the ongoing armed conflict in Ukraine has been an ‘absence of transparent, agreed-upon truth’ (Darden, 2014). After the Euromaidan protest movement swept President Viktor Yanukovych from power in February 2014 – and Russia annexed the Crimean peninsula – residents of Ukraine’s eastern and southern provinces launched a series of demonstrations against the new authorities in Kyiv. These demonstrations escalated into a Russian-backed separatist rebellion in Ukraine’s eastern Donbas region, comprising the heavily-industrialized and densely populated provinces of Donetsk and Luhansk.

Before the revolution in Kyiv, Russian media had a heavy presence in Ukraine,
particularly in Crimea and other parts of the country’s south-east (Broadcasting Board of Governors, 2014). In contrast to Western media portrayals of the Euromaidan as a largely peaceful protest movement confronting riot police and hired thugs, mainstream Russian media devoted their coverage to nationalist militants storming parliament and hurling Molotov cocktails. Both images were in a narrow sense true, but neither represented a complete picture of events. The Russian perspective on events seemed to leave an impression on angry crowds in Crimea and the Donbas, who condemned the Euromaidan movement as a ‘Western-backed coup’ and ‘fascist junta.’

Concerned over the mobilizational potential of Russian media, Ukraine’s post-revolutionary authorities took a series of steps to create an ‘hermetically sealed information environment’ (Vikhrov, 2014). In March 2014, before the first shots were fired in east, Kyiv banned Russian federal broadcasters from Ukrainian cable TV, followed several months later by bans on some Russian films and serials, and travel bans on Russian journalists. In December, Ukraine established a Ministry of Information Policy to protect Ukrainians from ‘unreliable information,’ register media outlets and define professional journalistic standards. To spread government-approved content in social media, the Ministry launched an ‘Information Army’ of patriotic volunteers.

Ukrainian authorities also exerted direct pressure on some information providers. In September 2014, Ukraine’s Security Service (SBU) raided the offices of the newspaper Vesti, accusing it of violating Ukraine’s territorial integrity through its coverage of the Donbas conflict. In February 2015, Ukrainian authorities arrested a blogger on charges of treason, for posting a YouTube video criticizing the government’s military mobilization campaign. The same month, Ukraine’s Television and Radio Council accused popular TV host Savik Shuster of violating a law on ‘incitement of hatred’ after a Russian journalist criticized the government’s military operations on his show.

In the rebel-held territories of the Donbas, separatists moved to create a similar zone of ‘informational sovereigny’ (Pomerantsev, 2014). After seizing the Donetsk regional administration building in April 2014, one of the rebels’ next steps was to take control of TV towers in the region, take Ukrainian channels off the air, and put Russian ones back on. Later that year, the self-proclaimed Donetsk People’s Republic established an official News Agency (DAN), while multiple privately-owned pro-rebel outlets emerged to fill the regional media vacuum. Suspicious of journalists from outside Russia and the region, rebels detained several reporters on suspicions of espionage, including an American journalist with Vice News.
In 2014, across rebel- and government-controlled territories of Ukraine, there were 7 documented killings of journalists, 286 physical assaults, 78 abductions, multiple physical attacks on offices and cyberattacks on websites (Freedom House, 2015). Many of these developments have predictably raised concerns over freedom of speech (Gorodnichenko and Mylovanov, 2015). Some analysts have worried that an informational firewall between dueling and contradictory media narratives will only deepen existing divisions (Darden, 2014).

How has Ukraine’s information war affected public attitudes toward the conflict? Survey evidence suggests that very few Ukrainians outside of the Donbas see Russian state media as a reliable or truthful source – which may be evidence either of the success of Ukraine’s counter-propaganda efforts, or ineffectiveness on Russia’s part (Snegovaya, 2015). Residents of rebel-held areas appear to have a similarly skeptical view of Ukrainian media, particularly due to the latter’s unwillingness to report on civilians killed by pro-government troops – incidents which Kyiv routinely denies (The Economist, 2015).

Despite much anecdotal speculation over who is ‘winning’ Ukraine’s information war, there have been no systematic empirical studies on variation in coverage across information providers, or the impact of this variation on statistical results and public opinion.

### 2.1 Quantifying the conflict in Ukraine

To take stock of reporting biases in the Ukrainian conflict, we examine new violent event data based on human-assisted machine coding of news reports, press releases and blog posts from Ukrainian, rebel, Russian and international sources. These sources include official newswires, television channels, internet news sites, blogs. We also included the Russian-language edition of Wikipedia, and daily briefings from the OSCE Special Monitoring Mission to Ukraine. For each data source, we created a separate electronic text corpus that contained all incident reports published on the Donbas since between February 2014 and August 2015. Altogether, our data include 40,741 violent events reported by 17 information providers, between February 23, 2014 and July 31, 2015.

To determine the geographic locations of events mentioned in the reports, we ran an automated geocoding script that identified populated place names referenced in the text, and matched them against the U.S. National Geospatial Intelli-
gence Agency’s GeoNames database. Table 1 shows the resulting spatial distribution of events, along with a description of each source.

We used a supervised learning algorithm (Support Vector Machine) to classify each event into a series of pre-defined categories, by event type, initiator, target, tactic, and casualties. The events of primary interest are rebel violence and government violence. We define incidents of rebel violence as specific acts of organized violence initiated by any anti-Kyiv armed group or regular Russian Armed Forces. Incidents of government violence involve organized violence by any pro-Kyiv armed group. For each dataset shown in Table 1, the authors and a team of research assistants read and classified a randomly-selected training set of 130-600 reports (depending on the size of the corpus), in Russian, Ukrainian and/or English. We used these manually-coded training data to train the SVM classifier to construct 17 separate datasets of violence in eastern Ukraine, one for each information provider.

2.2 Actor-specific reporting bias in Ukraine

How do the sources in Table 1 differ in their coverage of the Donbas conflict? Do all sources report the same kinds of events by the same actors, or do they focus

---

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

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

3 Pro-Kyiv groups include Ukraine’s regular Army, Air Force, Airborne troops, Marines, Border Guard, SBU, Interior Ministry, local police, National Guard or any of 46 volunteer battalions (e.g. Azov, Aydar, Dnipro-1, Donbas) and independent right-wing militias (e.g. Right Sector).

4 To account for potential disagreement between coders, at least two sets of eyes read each training set document, including the author and another member of the research team. Inter-coder reliability statistics, reported in the online appendix, indicate a high and statistically significant level of agreement between coders on the relevant categories, including where coders read the same documents in different languages.
on one armed group more than another? To answer these questions we report the relative propensity of information providers to attribute an event to each warring side, conditional on reporting the event in the first place.\footnote{We estimate these quantities from the joint predicted probabilities of a bivariate probit model, regressing actor attribution (rebel, government, both, neither) on the country affiliation of the information provider, and a series of other event attributes that may affect the likelihood of coverage (e.g. local population density, terrain, proximity to road infrastructure, proximity to the front line).}

Figure 1 reveals large systematic differences in the armed actors who receive coverage in Ukrainian, rebel, Russian and international sources. On average, Ukrainian sources devote more news coverage to rebel violence (.43, 95% CI: .42, .44) and less coverage to government operations (.16, 95% CI: .15, .16) than any other group of sources. Rebel media show an even more extreme imbalance, in the opposite direction, with over half their reporting on government violence (.54, 95% CI: .53, .56) and less than ten percent on rebel violence (.08, 95% CI: .08, .09). Russian sources have the same direction of bias as rebel sources, but with lower magnitude – 19 percent of events are attributed to rebels (95% CI: .18, .20) and 36 percent to the government (95% CI: .34, .37).

A very different picture appears in international sources, like OSCE reports and Wikipedia. These data are more ‘neutral,’ in the sense that they are unlikely to attribute violence to any armed group at all. The language in these reports tends to be more passive and non-specific (e.g. ‘shelling was reported near village X’) than language in local media. For the OSCE, this finding is consistent with anecdotal reports that – because it must maintain working relations with all sides – the monitoring organization is exceedingly cautious about attributing violent events to specific initiators. For Wikipedia, this pattern may reflect the crowd-sourced nature of the data – users can flag entries as lacking neutrality, and remove offending information until a stripped-down ‘compromise’ is reached.

The extent of actor-specific reporting bias varies by source. Figure 3 disaggregates the country groups, and shows the proportion of reported rebel and government attacks by individual information provider. Sources on the left side report almost exclusively on government violence, and those on the right report mainly rebel violence. The sources with the most extreme biases, in each direction, are NewsFront and DAN on the rebel side and Information Resistance (Sprotyv) on the Ukrainian side. This distribution is not surprising, since DAN is the mouthpiece of rebel authorities in Donetsk, and Sprotyv is a Ukrainian military blog established for the explicit purpose of countering Russian propaganda with ‘objective’ information from the front. One Russian media outlet – the independent, opposition-oriented Dozhd TV channel – occupies a space on this spectrum closer
to the median Ukrainian source. Between rebel and Ukrainian media, however, there is a much clearer separation – the ‘left-most’ Ukrainian outlet is still to the right of the ‘right-most’ rebel outlet.

Figure 3 shows that Ukrainian and rebel media not only tend to report disproportionately on violence by the ‘other side,’ but they disproportionately focus on indiscriminate violence by the ‘other side’ – events like artillery and rocket shelling, and the use of heavy armor. Two thirds of Ukrainian media reports of rebel violence describe indiscriminate uses of force (95% CI: .63, .70), compared to 41% in rebel media (95% CI: .36, .46). Coverage of indiscriminate government violence is a near-mirror image: 41 percent of the government violence reported by Ukrainian sources is indiscriminate (95% CI: .37, .45), compared to 65 percent for rebel sources (95% CI: .62, .69). Russian and international sources, as before, lie somewhere in between.

By way of an illustration, Figure 4 shows density plots with examples cases of selective reporting for each group of sources. Almost all rebel uses of force reported by Donetsk News Agency (DAN) are selective (e.g. arrests, detentions, assassinations, small arms firesights and sniper fire). Interfax-Ukraine shows the opposite: rebel violence is usually indiscriminate, and government violence is almost always selective. Ironically, Interfax-Ukraine is a Russian-owned newswire, with editorial offices in Kyiv. The popular Russian online news outlet Lenta.ru offers near-equal coverage of indiscriminate violence by each actor. OSCE monitoring reports – a very small proportion of which actually include actor attribution – suggest that both sides are indiscriminate, with rebels claiming a slight edge.

An information provider’s county or group affiliation is, of course, not the only determinant of actor attribution. Some reports may not specify the initiator simply because such information is not available. In line with existing evidence on the ‘supply-side’ causes of selective reporting (Weidmann, 2016), we find that – for all sources – there is significantly more attribution in places with more witnesses (high population density), more visibility (open terrain), and more accessibility (proximate to a major road).

Beyond patterns of selective reporting like these, which are common to all sources, our evidence suggests that there also exist systematic differences in the actors whose violence individual sources cover. The direction of these biases – along with the disproportionate emphasis on indiscriminate violence – aligns with what one might expect if information providers were active participants in an information war intended to discredit the opposing side and mobilize public opinion against it. Whether these biases can actually produce this effect is an open empirical question, which we address in the next two sections.
<table>
<thead>
<tr>
<th>Name</th>
<th>Map</th>
<th>Info</th>
<th>Name</th>
<th>Map</th>
<th>Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>112 (Ukraine)</td>
<td>TV</td>
<td>Rus/Ukr-language</td>
<td>Lenta.ru (Russia)</td>
<td>Online</td>
<td>Rus-language</td>
</tr>
<tr>
<td>Channel 5 (Ukraine)</td>
<td>TV</td>
<td>Ukr-language</td>
<td>NewsFront (rebel)</td>
<td>Online</td>
<td>Rus-language</td>
</tr>
<tr>
<td>BFM (Russia)</td>
<td>Online</td>
<td>Rus-language</td>
<td>OSCE (international)</td>
<td>Online</td>
<td>English-language</td>
</tr>
<tr>
<td>DAN (rebel)</td>
<td>News agency</td>
<td>Rus-language</td>
<td>RusVesna (rebel)</td>
<td>Online</td>
<td>Rus-language</td>
</tr>
<tr>
<td>Dozhd (Russia)</td>
<td>TV</td>
<td>Rus-language</td>
<td>Sprotyv (Ukraine)</td>
<td>Online</td>
<td>Rus-language</td>
</tr>
<tr>
<td>Espreso (Ukraine)</td>
<td>TV</td>
<td>Ukr-language</td>
<td>Ukrinform (Ukraine)</td>
<td>News agency</td>
<td>Rus/Ukr-language</td>
</tr>
<tr>
<td>Gazeta.ru (Russia)</td>
<td>Online</td>
<td>Rus-language</td>
<td>Vesti (Russia)</td>
<td>TV</td>
<td>Rus-language</td>
</tr>
<tr>
<td>Interfax.ru (Russia)</td>
<td>News agency</td>
<td>Rus-language</td>
<td>Wikipedia (international)</td>
<td>Online</td>
<td>Rus-language</td>
</tr>
<tr>
<td>Interfax.ua (Ukraine)</td>
<td>News agency</td>
<td>Rus/Ukr-language</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: Actor-specific reporting propensities.

Figure 2: Actor-specific reporting propensities, by source.
Figure 3: Reports of indiscriminate violence by actor.

(a) Rebel

(b) Government

\[ P(\text{indiscriminate} \mid \text{rebel violence}) \]

\[ P(\text{indiscriminate} \mid \text{govt violence}) \]
Figure 4: Reports of indiscriminate violence by actor, examples. Densities represent proportion of events per municipality classified as indiscriminate, by actor.

(a) Ukraine: Interfax.ua  
(b) Rebel: DAN  
(c) Russian: Lenta  
(d) International: OSCE
3 Implications for theory-testing and prediction

The previous discussion revealed two things. First, political authorities on both sides of the Ukrainian conflict have sought to control media coverage of the war, and limit consumers’ access to alternative sources of information. Second, this ‘information war’ has yielded significant actor-specific reporting biases in media coverage of the Ukrainian conflict, with Ukrainian sources covering mostly rebel violence, rebel sources covering Ukrainian government violence, and foreign sources somewhere in the middle.

In this section, we report the results of statistical analyses that rely exclusively on each set of sources – Ukrainian, rebel, Russian and international – and evaluate the consequences of actor-specific reporting bias for statistical inference. Our primary interest is in the strategic interaction between rebels and the government – how the choices of one actor influences the choices of the other, and whether one side is more likely to cooperate or escalate than the other. We also examine the political geography and political economy of conflict – where and when violence by each actor is most likely to occur, and how factors like territorial control, ethnicity, and economic vulnerability influence its relative intensity.

3.1 Propensity for escalation

What do different data sources tell us about rebels’ and government’s relative propensity for escalation? Each source offers its own perspective on how fighting between these armed groups is likely to unfold, and what sort of equilibrium, or steady state, may emerge in the absence of outside intervention. This equilibrium may be a largely peaceful one, where violence diminishes organically and neither side is likely to unilaterally escalate. In this case, a news consumer or policymaker may conclude that outside intervention is not necessary to reduce violence. If the equilibrium is a more violent one, where transgressions by one or both actors are common, the implications for conflict resolution are quite different: for violence to decline, third-party enforcement should be directed at whichever side is more prone to unilaterally escalate.

To evaluate the impact of actor-specific reporting biases on these conflict dynamics, we model the strategic interaction between rebels and government as a stochastic process. Consider a repeated game played by two actors, indexed \( k = 1, 2 \) (e.g. rebels and government). In every round, both actors choose between a low and high level of violence \( (L, H) \). The strategy space yields a system with four possible states: \( LL, LH, HL, HH \), which we index \( i = 1, 2, 3, 4 \). Although
event data cannot tell us about the payoffs actors attach to each state, they do reveal the strategies of the two actors over time – the probability that each actor plays \( L \) or \( H \) in the current round, given the state of the system in the previous round.

A transition matrix \( \mathbf{M} = [m_{ij}] \) governs the tendency of the system to move between the four states, with \( m_{ij} \) representing the probability that the game will be in state \( i \) at \( t \) after it was in state \( j \) at \( t - 1 \). With two players, the transition matrix has 16 elements:

\[
\mathbf{M} = \begin{pmatrix}
    LL & LH & HL & HH \\
    m_{1,1} & m_{1,2} & m_{1,3} & m_{1,4} \\
    m_{2,1} & m_{2,2} & m_{2,3} & m_{2,4} \\
    m_{3,1} & m_{3,2} & m_{3,3} & m_{3,4} \\
    m_{4,1} & m_{4,2} & m_{4,3} & m_{4,4}
\end{pmatrix}
\] (1)

Over time, the system will converge to a stationary distribution \( \pi = \pi \mathbf{M} \),

\[
\pi = (\pi_1, \pi_2, \pi_3, \pi_4)
\]

\[= (\pi_{LL}, \pi_{LH}, \pi_{HL}, \pi_{HH})\]

The stationary distribution represents the steady state or equilibrium of the repeated game. Over the long run, whatever the starting state, the chain will spend approximately \( \pi_i \) of its time in state \( i \), with \( 0 \leq \pi_i \leq 1 \) and \( \sum_i \pi_i = 1 \).

For instance, \( \pi = (0.99, 0, 0, 0.01) \) suggests that the system will converge toward a relatively peaceful equilibrium, spending 99 percent of its time in state \( LL \) and 1 percent in \( HH \). By contrast, \( \pi = (0, 0, 0.5, 0.5) \) suggests that the system will remain quite violent, spending half its time in a state of one-sided government violence (\( HL \)) and half under two-sided violence (\( HH \)).

Our quantity of interest is the relative propensity for unilateral escalation,

\[ \omega = \pi_2 - \pi_3 \] (3)

If \( \omega < 0 \), unilateral escalation by the rebels (\( HL \)) is more common in equilibrium than unilateral escalation by the government (\( LH \)). If \( \omega > 0 \) the rebels are more likely than the government to show unilateral restraint. The appendix provides a more detailed derivation and discussion of this statistic.

We can obtain an empirical estimate of \( \omega \) from the predicted probabilities of a bivariate probit model, in which the outcomes of interest are pairs of interdependent decisions by rebels and the government – indexed \( k = 1, 2 \), respectively.
– to use \((H)\) or not use violence \((L)\) in a particular time and place. We define two binary dependent variables, \(Y_{1jt}\) and \(Y_{2jt}\), which take the value of 1 if the actor plays \(H\) or 0 if they play \(L\). The unit of analysis is a district \((j)\) - week \((t)\), but we drop the spatial index \(j\) from the notation below. As in a standard probit model, we assume that the observed dependent variables, \(Y_k\), are dichotomized versions of latent continuous variables, \(Y_k^*\), with the observation mechanism

\[
Y_k = \begin{cases} 
1 & \text{if } Y_k^* \geq 0 \\
0 & \text{if } Y_k^* < 0 
\end{cases}
\]

If the two actors make their choices independently, we can model the joint outcome using two separate equations

\[
\begin{align*}
Y_1^* &= \mu_1 + u_1 \\
Y_2^* &= \mu_2 + u_2 \\
u_1 &\sim N(0, 1) \\
u_2 &\sim N(0, 1)
\end{align*}
\]

where \(\mu_k\) are means for \(Y_k^*\) and \(u_k\) are i.i.d. error terms.

The independence restriction would be violated, however, if the unobserved components of the two models \((u_1, u_2)\) were correlated:

\[
\begin{align*}
u_1 &= \eta + \varepsilon_1 \\
u_2 &= \eta + \varepsilon_2
\end{align*}
\]

where one error component \((\varepsilon_k)\) is unique to each actor, but another \((\eta)\) is shared across both. Where this is the case, the joint outcome \((Y_1, Y_2)\) can be modeled using the marginal probabilities for each dependent variable and a correlation parameter \(\rho\), which describes how the two dependent variables are related. In this case, the latent variables \((Y_k^*)\) follow a bivariate Normal distribution

\[
\begin{bmatrix} Y_1^* \\
Y_2^* \end{bmatrix} \sim N_2 \left( \begin{bmatrix} \mu_1 \\
\mu_2 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\
\rho & 1 \end{bmatrix} \right)
\]

where \(\rho\) is a scalar correlation parameter. If \(\rho = 0\), the two variables are independent and the above expression reduces to two standard Normal distributions, permitting us to model the joint outcome using two separate probit models. If \(\rho \neq 0\), a bivariate probit specification is more appropriate.

We model the means of these distributions as follows\(^6\)

\[
\begin{bmatrix} \mu_{1,t} \\
\mu_{2,t} \end{bmatrix} = \begin{bmatrix} y_{2,t-1} \xi_1 + y_{1,t-1} \alpha_1 + y_{2,t-1} y_{1,t-1} \gamma_1 + x_{1,t} \beta_1 + W_{y_{1,t-1}} \rho_1 \\
y_{1,t-1} \xi_2 + y_{2,t-1} \alpha_2 + y_{1,t-1} y_{2,t-1} \gamma_2 + x_{2,t} \beta_2 + W_{y_{2,t-1}} \rho_2 \end{bmatrix}
\]

\(^6\)In the following, we suppress the observation index \(j\).
where $x_{k,t}$ is a vector of covariates for actor $k$ at time $t$, $y_{k,t-1}$ is a time-lagged dependent variable for actor $k$, and $\alpha_k, \beta_k, \zeta_k$ and $\gamma_k$ are regression coefficients. To account for spillovers of violence from neighboring districts, we include spatio-temporal lags of the dependent variable, where $W$ is a row-normalized spatial weights matrix, and $\rho_k$ is the autoregressive parameter.

The model assumes that the influence of an actor's own past actions on current behavior is conditional on the previous actions of the adversary. If actor 2 played $L$ in the previous round ($y_{2,t-1} = 0$), the effect of $y_{1,t-1}$ on $y^*_1$ is $\alpha_1$; if actor 2 played $H$ ($y_{2,t-1} = 1$), the effect is $\alpha_1 + \gamma_1$.

Tables 2 reports empirical estimates of the stationary distribution $\pi$, for each group of sources. The results indicate strong disagreement over the type of equilibrium that is most likely to unfold, in the absence of outside intervention. Russian and international sources suggest relatively peaceful equilibria, with $\hat{\pi}_1 = .67$ and $\hat{\pi}_1 = .71$, respectively. In other words, according to these information providers from outside Ukraine and rebel territories, rebel-government interactions will be non-violent at least two thirds of the time, once the system reaches equilibrium.

Local sources paint a more ominous picture. If the conflict continues to play out as reported in Ukrainian media, the two sides will be at peace just 41 percent of the time ($\hat{\pi}_1$), and will experience one- or two-sided violence during the remaining 59 percent. Rebel sources are even more pessimistic, with the system staying non-violent just 20 percent of the time.

Which actor is most likely to break the peace, according to each set of sources? Table 3 reports empirical estimates of the relative propensity for escalation $\omega$. As one might expect, the greatest disparity here is between Ukrainian and rebel sources. Ukrainian sources predict that rebels are much more likely to unilaterally escalate than government troops ($\hat{\omega} = -.17$). Rebel sources predict an even stronger pattern in the opposite direction, with government troops far more likely to unilaterally escalate than the rebels ($\hat{\omega} = .30$). Russian sources predict that the two sides are almost equally likely to escalate ($\hat{\omega} = .01$), while international sources suggest slightly more escalation on the government side ($\hat{\omega} = .07$).

### 3.2 Political geography of violence

Although the above results tell us about how equilibrium levels of violence might look in an average Donbas district, we may expect the stationary distribution to vary over space. What can our conflicting data sources tell us about when and where violence by each actor is likely to occur? Will we see the same stationary
Table 2: Stationary distribution of violence. $\hat{\pi}_1$: neither (LL), $\hat{\pi}_2$: government only (LH), $\hat{\pi}_3$: rebel only (HL), $\hat{\pi}_4$: both (HH). 95% confidence intervals in parentheses.

<table>
<thead>
<tr>
<th>Sources</th>
<th>$\hat{\pi}_1$ (LL)</th>
<th>$\hat{\pi}_2$ (LH)</th>
<th>$\hat{\pi}_3$ (HL)</th>
<th>$\hat{\pi}_4$ (HH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ukraine</td>
<td>0.410 (0.393,0.43)</td>
<td>0.052 (0.043,0.061)</td>
<td>0.221 (0.213,0.228)</td>
<td>0.317 (0.315,0.318)</td>
</tr>
<tr>
<td>Rebel</td>
<td>0.201 (0.19,0.21)</td>
<td>0.348 (0.346,0.351)</td>
<td>0.049 (0.037,0.063)</td>
<td>0.402 (0.382,0.422)</td>
</tr>
<tr>
<td>Russia</td>
<td>0.666 (0.61,0.729)</td>
<td>0.087 (0.069,0.104)</td>
<td>0.081 (0.065,0.098)</td>
<td>0.165 (0.136,0.188)</td>
</tr>
<tr>
<td>Int’l</td>
<td>0.711 (0.653,0.77)</td>
<td>0.128 (0.106,0.148)</td>
<td>0.059 (0.045,0.074)</td>
<td>0.101 (0.079,0.124)</td>
</tr>
</tbody>
</table>

Table 3: Relative propensity for unilateral escalation. $\omega < 0$: more rebel unilateral escalation, $\omega > 0$: more government unilateral escalation. 95% confidence intervals in parentheses.

<table>
<thead>
<tr>
<th>Sources</th>
<th>$\hat{\omega}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ukraine</td>
<td>-0.169 (-0.170,-0.167)</td>
</tr>
<tr>
<td>Rebel</td>
<td>0.299 (0.309,0.288)</td>
</tr>
<tr>
<td>Russia</td>
<td>0.006 (0.004,0.006)</td>
</tr>
<tr>
<td>Int’l</td>
<td>0.069 (0.061,0.074)</td>
</tr>
</tbody>
</table>

distribution in rebel-controlled and government-controlled territories? How might other potential drivers of conflict – like the local ethno-linguistic balance or local economic vulnerability – shape these dynamics in the long run?

Figure 5 shows the stationary distribution of violence, conditional on the local balance of territorial control between rebels and the government. We ran these simulations using parameter estimates from (8), with the black line representing $\hat{\pi}_1$ (LL, no violence), the blue line $\hat{\pi}_2$ (LH, one-sided government violence), the red line $\hat{\pi}_3$ (HL, one-sided rebel violence), and the purple line representing $\hat{\pi}_4$ (HH, two-sided violence).

---

7For information on territorial control, we drew on three sets of sources: (1) official daily situation maps publicly released by Ukraine’s National Security and Defense Council, (2) daily maps assembled by the pro-rebel bloggers ‘dragon_first_1’ and ‘kot_ivanov,’ and (3) Facebook posts on rebel checkpoint locations. For the first two sources, we georeferenced and vectorized daily map updates into spatial polygons. To construct polygons from user-reported checkpoint locations, we used the geographic convex hull of checkpoint coordinates. We overlaid these polygons with locations of municipalities and districts, to calculate the proportion under rebel control each day.
As before, the most telling differences in Figure 5 are between Ukrainian and rebel sources. Both predict that peace (LL) is more difficult to sustain in areas of divided control than in areas with complete rebel or government control. Yet the two sets of sources disagree over which actor is most likely to escalate, and where. According to Ukrainian sources, divided territorial control increases the probability of two-sided violence (HH), but does not as strongly affect one-sided violence by either of the armed groups. At every level of territorial control, unilateral escalation by the government (LH) is the least likely of the four scenarios, followed by one-sided rebel violence (HL). Although there is a substantial increase in two-sided violence in areas of divided control, the purple curve does not cross the curve for LL, suggesting that – even in the most contested areas – peace will be the modal outcome.

Rebel sources predict a more diverse set of equilibria. In locations of complete government control (far left), peace is the most common state of the system, followed by one-sided government violence. In areas of divided control (middle), the dominant state is two-sided violence, followed by one-sided government violence. Yet in areas of incomplete rebel and government control (about .25 and .75 on the horizontal axis), unilateral government escalation is the most common state. Unsurprisingly, one-sided rebel violence is consistently the least likely state.

All four groups of sources agree on a more general point: there is more violence in areas of divided control. The black line representing LL is U-shaped in each set of simulations. This finding underscores the symmetric and conventional nature of the warfare in Ukraine. Rather than focusing most of their efforts on guerrilla-style hit-and-run attacks in government-controlled territory, rebels have concentrated on taking, holding and consolidating control of physical territory. Likewise, the government has sought to defend against and push back these rebel territorial advances. Although each side has periodically conducted raids across enemy lines, along with the policing of territories under their own control, the bulk of the fighting has been in places where the areas of control overlap. This is a very different pattern of fighting from that seen in more irregular insurgency-counterinsurgency conflicts (Kalyvas 2006).

Figure 6 show analogous simulations for the relationship between violence and the proportion of a district’s population that claims Russian as its native language, according to the 2001 Ukrainian census. Public debate around the Ukrainian conflict has tended to emphasize Russian nationalism as a driver of the unrest. The civil conflict literature has identified several mechanisms by which geographically-concentrated minorities may have an advantage in mobilizing a rebellion, including the ability to monitor and punish defectors, attract diaspora
support, and overcome various collective action problems (Bates, 1983; Fearon, 2006; Weidmann, 2009). The Donbas would then appear to be an easy case for the ‘ethnic effect’ – 38 percent of Donetsk’s population is ethnically Russian, and 75 percent claims Russian as their native language. Russian nationalists had a visible presence during the protests in early 2014 and local residents expressed concern over the future status of the Russian language after the revolution in Kyiv.

The data do not support the view that Russian-speaking areas are more prone to violence – by rebels or forces loyal to the Ukrainian government. All four sets of information providers suggest the opposite – that Russian language either had no effect on violence at all, or a weakly negative impact. Not even rebel sources – which one might expect to over-report government violence in Russian-speaking areas – indicate the existence of an ‘ethnic effect.’ Indeed, the probability of peace ($LL$) slightly increases with the share of the Russian-speaking population.

Figure 7 explores a third set of explanations for violence: vulnerability to economic shocks. Following Zhukov (2016), we measure economic vulnerability with the proportion of the local population employed in the machine-building industry prior to the war. To a greater extent than any other industry in Ukraine, machine-building was exposed to a series of negative shocks due to Ukraine’s Comprehensive Free Trade Agreement (CFTA) with the EU, and retaliatory trade barriers with Russia. Some 45 percent of Ukraine’s machinery industry was concentrated in the Donbas before the war, with exports to Russia accounting for up to 90 percent of local production in some towns. Traditionally protected from competition, with production highly specialized for the Soviet and then Russian markets, the machinery industry was among the chief ‘losers’ of the Ukraine-EU CFTA deal. Indeed, Yanukovych’s decision not to sign this agreement, under pressure from Russia and local businesses in the Donbas, sparked the Euromaidan protests in the first place.

As Figure 7 shows, economic vulnerability has a consistently deleterious effect on violence, across all sets of sources. Consistent with previous empirical findings on this conflict (Zhukov, 2016), all groups of information providers – Ukrainian, rebel, Russian and international – suggest that two-sided violence ($HH$) is more likely and peace ($LL$) is less likely in districts where a high proportion of the local population was employed in the machinery industry prior to the war. When compared with the flat lines for Russian language in Figure 5, this result shows that Ukraine’s pro-Russian rebels were ‘pro-Russian’ not because they spoke Russian at home, but because their economic livelihood had depended on close trade and economic ties with Russia.
Figure 5: Relationship between violence and territorial control depends on source. Quantities reported are $E[\hat{\pi}|X]$, where $\hat{\pi}$ is the estimated stationary distribution of violence, and $X$ is rebel territorial control. $\hat{\pi}$: neither (LL), government only (LH), rebel only (LH), both (HH). Dotted lines are 95% confidence intervals.
Figure 6: No relationship between violence and Russian speakers. Quantities reported are $E[\hat{\pi}|X]$, where $\hat{\pi}$ is the estimated stationary distribution of violence, and $X$ is the proportion of the local population that claims Russian as a native language. $\hat{\pi}$: neither (LL), government only (LH), rebel only (HL), both (HH). Dotted lines are 95% confidence intervals.

(a) Ukrainian sources
(b) Rebel sources
(c) Russian sources
(d) International sources
Figure 7: Violence increasing in economic vulnerability. Quantities reported are $E[\hat{\pi}|X]$, where $\hat{\pi}$ is the estimated stationary distribution of violence, and $X$ is the proportion of the local population employed in the machine-building industry prior to the war. $\hat{\pi}$: neither (LL), government only (LH), rebel only (HL), both (HH). Dotted lines are 95% confidence intervals.

(a) Ukrainian sources  
(b) Rebel sources  
(c) Russian sources  
(d) International sources
4 Implications for public opinion

The previous section demonstrated that actor-specific reporting bias has a substantive impact on predictions of how conflict is likely to unfold, and how sustainable a negotiated settlement may be. Ukrainian sources predict an equilibrium in which unilateral escalation by rebels is much more common than unilateral escalation by government troops. Rebel sources predict an opposite equilibrium, where unilateral escalation by government troops is pervasive. Russian and international sources predict equilibria in which violence is generally less likely, and unilateral transgressions are more rare.

These predictions hold vastly different implications for conflict resolution. Internally, information consumers may doubt that an actor inclined to use unilateral violence can stick to the terms of a negotiated agreement. Externally, the relative propensity for escalation can shape perceptions over how intractable a conflict is likely to be, whether third-party enforcement is necessary to stop it, and whether that response should be impartial or directed at one side.

To more directly explore the impact of actor-specific bias on policy preferences, we ran a survey experiment, in which we exposed subjects to news stories about a generic civil conflict, and asked them about their attitudes toward third-party intervention in that conflict. Our subject pool included 1,596 respondents in the United States and 1,366 in India. We conducted the survey in April-May 2014, using Amazon Mechanical Turk to recruit survey subjects. Due to space constraints, the results reported below are based on the sub-sample of U.S. respondents only; however, the Indian sample produced results of near-identical magnitude, with the same direction and statistical significance (see appendix).

We employed a 2-by-2 factorial design survey to test the effect of both actor attribution (i.e. rebels or government) and war-fighting tactics (i.e. selective and indiscriminate) – the same two types of biases we see in our event data. The survey for each group consisted of a short news story about violence in a generic, non-specified civil conflict, followed by a battery of questions designed to elucidate the effect of the stories on public opinion. For each version of the treatment, we kept the bulk of the text intact, modifying only the actors committing the violence, and their reported tactics. We randomly assigned participants to one of the four treatment regimes, or to a control group, which received the same text without any information about actor or targets. To help respondents comprehend the relative destructiveness of selective and indiscriminate tactics, we accompanied each text with a photo – showing either an individual arrest (selective), a destroyed house (indiscriminate), or an armed group riding on an armored personnel carrier.
These photos were the same for rebel- and government-initiated events. Table 4 summarizes the four treatments, and the corresponding versions of the text passage we asked participants to read.

In the subsequent battery of questions, we asked participants about the type of response – by the international community and their home country – they felt was appropriate, with options including ‘No response,’ ‘Economic sanctions,’ ‘Military aid,’ and ‘Military intervention’ (see appendix for full survey). In addition, we asked whether the response should be impartial, or directed against one of the two sides. To account for other potential drivers of policy preferences, we concluded the survey with a battery of general questions about participant demographics, political ideology, military background and political knowledge. We also included several ‘attention filter’ questions, to exclude respondents who either did not read the news story or were clicking through the questionnaire at random.

The survey yields three main findings. First, news stories that disclose more information about actors and tactics increase support for intervention and reduce support for impartiality. Figure 8 reports model-based estimates of average survey responses for the two questions (support intervention, support taking sides), conditional on assignment into the control group (non-specific coverage) or any of the four treatment groups (actor/tactic-specific) listed in Table 4. In addition, the models control for participant demographics and political ideology.

Relatively few respondents favored a completely ‘hands-off’ response to reports of violence, with more than 90 percent of subjects favoring some form of diplomatic or military response. Yet among participants who read a story with information on actors involved and tactics used, overall support for intervention was 2.5 percent higher on average (95% CI: 1.3, 4.7), rising from 95.5 to 98 percent. The effect on respondent impartiality was stronger. Although only a quarter of respondents who read non-specific coverage favored taking sides in the conflict, one third of those in the four treatment groups did the same – an increase of 32 percent relative to the control condition (95% CI: 18.4, 48.6).

Our second finding is that actor attribution is more important for public opinion than information about tactics. Respondents tend to favor intervention against whichever actor is reported to have initiated the violence. Whether that violence was selective or indiscriminate matters less than the group responsible for the action. Figure 9 reports average predicted levels of support for anti-rebel and anti-government treatment groups, controlling for participant demographics and political ideology.

---

8The predicted probabilities in Figure 8 are based on the parameters of a logit model, regressing individual responses on treatment status, age, gender, education, social conservatism, and past military service.
anti-government interventions for the four treatment groups exposed to actor- and tactic-specific coverage. Support for anti-rebel intervention is about equally high (about 23 percent) among subjects who read reports of selective and indiscriminate rebel violence, and equally low (13-15 percent) among subjects who read about either type of government violence.

Attitudes toward anti-government intervention look slightly different. Although selectivity of tactics does not affect responses among subjects who read about government violence (18-19 percent support), reports of indiscriminate rebel violence appeared to generate a slight increase in support for anti-government intervention. This result is counterintuitive, since the government is not the side committing the violence – although rebel indiscriminate violence may signal the government’s inability to protect its citizens. Yet the effect in question is also quite uncertain, with average support for anti-government intervention following rebel indiscriminate violence lying within the 95 percent confidence interval for selective rebel violence. Actor attribution also has a substantively larger effect on participant responses. Among participants who read stories of selective violence, support for anti-government intervention was 107 percent (95% CI: 80, 138.1) higher if government troops, rather than rebels, were reported to have committed that violence. This difference is smaller for participants who read about indiscriminate violence, but in the same direction (54 percent, 95% CI: 43.2, 66.4).

The relatively weak effect of tactic-specific reporting bias is surprising, since the photos we displayed in our survey instrument highlighted specifically differences between selective and indiscriminate tactics, rather than between rebel and government violence (Table 4). In theory, this emphasis should have inflated rather than attenuated the ‘tactic effect.’ The fact that indiscriminate violence did not have a stronger impact on respondents’ policy choices suggests that the differential reporting of opponents’ tactics we saw in Ukrainian and rebel sources (Figure 4) may not be the most compelling way to affect public attitudes.

Our third experimental finding is that the increase in support for intervention holds for non-military and military policy options. Figure 10 reports the percent increase in support for three types of intervention against each side (economic sanctions, aid to opponents, and direct military intervention), following reports of rebel (left) and government (right) violence. Respondents who read stories of rebel violence were 71 percent more likely (95% CI: 42, 106) to support anti-rebel economic sanctions and 90 percent more likely (95% CI: 69, 115) to support anti-rebel military intervention. Participants who read about government violence showed a greater increase in support for sanctions than military intervention – 178 percent (95% CI: 85, 319) to 114 percent (95% CI: 75, 172) – although the differ-
ence between these intervention types was still within the margin of error.

If more detailed news reports generate stronger support for intervention, and lead news consumers to pick sides, then we should expect Ukrainian and rebel sources to have a greater impact on public opinion than Russian or international sources. The relatively low-information content of OSCE reports and other information providers from outside the conflict zone should dampen pro-intervention preferences. We should also expect Ukrainian and rebel sources to generate opposing preferences with respect to the direction and partiality of this intervention. The heavy focus on rebel violence in Ukrainian media is likely to generate support for anti-rebel intervention; rebel media emphasis on government violence should generate support for anti-government intervention. The scope of this intervention may be limited (e.g. sanctions) or extensive (e.g. ground invasion), but the direction of the increased support will be the same.

Our experimental results are not evidence that Ukrainian and rebel media consciously manipulate news coverage – either to attract external support or to undermine local confidence in the opponent’s credibility as a negotiating partner. Yet if such an intent did exist, we would expect information providers to adopt the exact types of actor-specific reporting biases that we have seen in the Ukrainian case.
Table 4: Survey instrument.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Photo</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 rebel violence, selective</td>
<td><img src="image1.png" alt="Photo" /></td>
<td>‘Hundreds are missing in the war-torn country this week, as rebel forces escalated their operations against suspected government supporters. International monitors reported that 50 individuals died and at least 200 are missing after a series of assassinations and detentions by pro-rebel forces in provincial towns.’</td>
</tr>
<tr>
<td>T2 rebel violence, indiscriminate</td>
<td><img src="image2.png" alt="Photo" /></td>
<td>‘Hundreds are missing in the war-torn country this week, as rebel forces escalated their operations against suspected government supporters. International monitors reported that 50 individuals died and at least 200 are missing after a barrage of heavy artillery shelling by pro-rebel forces in provincial towns.’</td>
</tr>
<tr>
<td>T3 govt violence, selective</td>
<td><img src="image3.png" alt="Photo" /></td>
<td>‘Hundreds are missing in the war-torn country this week, as government forces escalated their operations against suspected rebel supporters. International monitors reported that 50 individuals died and at least 200 are missing after a series of assassinations and detentions by pro-regime forces in provincial towns.’</td>
</tr>
<tr>
<td>T4 govt violence, indiscriminate</td>
<td><img src="image4.png" alt="Photo" /></td>
<td>‘Hundreds are missing in the war-torn country this week, as government forces escalated their operations against suspected rebel supporters. International monitors reported that 50 individuals died and at least 200 are missing after a barrage of heavy artillery shelling by pro-regime forces in provincial towns.’</td>
</tr>
<tr>
<td>C no info on actors, tactics</td>
<td><img src="image5.png" alt="Photo" /></td>
<td>‘Hundreds are missing in the war-torn country this week, as rebel and government forces escalated their campaigns. International monitors reported that 50 individuals died and at least 200 are missing after the two sides clashed in provincial towns.’</td>
</tr>
</tbody>
</table>
Figure 8: Detailed reports increase support for intervention.

Q: “Based on what you have read, what type of international response do you consider appropriate?”

A: None; Intervention (any)

Q: “As a general matter, should the international response be neutral, or support one of the two sides?”

A: Neutral; Take Sides
Figure 9: Actor attribution matters more than tactical descriptions.

Q: “Against whom should the international response be directed?”

A: Government, Rebels

Figure 10: Effect holds across different scales of intervention.

Q: “Based on what you have read, what type of international response do you consider appropriate?”

A: None (baseline); Economic sanctions; Aid to opponent; Military intervention
5 Conclusion

This study set out to advance the nascent research program on reporting bias in civil conflict, by taking a more direct look at the consequences of selective news coverage for academic and public knowledge. In so doing, we focused on the empirically common, but relatively understudied actor-specific reporting bias – the tendency of information providers to report violence by some actors more than violence by others. Unlike prior research, we treated reporting bias not merely as an inside-the-church problem for social scientists seeking to understand civil conflicts after the fact, but also as a source of influence on consumers of biased news reports: citizens and governments. To that end, we conducted a survey experiment to explore how the sorts of biases revealed in our event data analysis might influence public support for, or opposition to, intervening in civil conflicts.

Our results show that – by casting one actor as ‘more violent’ than the other – actor-specific reporting bias can have a profound impact on both statistical inference and public opinion. Data from one set of sources may predict a relatively peaceful equilibrium, where neither side is likely to unilaterally escalate the level of violence. Another source may predict a more violent equilibrium, in which violations are common, and one side is disproportionately more likely to attack than the other. These opposing perspectives on the conflict carry different implications for policy, particularly as regards the utility of outside intervention, and the actors’ relative ability to honor a negotiated agreement. These findings demonstrate that – beyond sources of reporting bias already well-established in the literature, like access on the supply side and ‘newsworthiness’ on the demand side – selective coverage may be used as a form of information warfare, with the goal of shaping public opinion.

In our analysis of event data compiled from multiple information providers, we found that Ukranian news sources disproportionately emphasize violence by rebels, while rebel sources emphasize the opposite: violence by the Ukranian government forces. Both Ukranian and rebel sources, in turn, frame their coverage of the other side’s violence as overwhelmingly indiscriminate – using heavy weapons and indirect fire methods that carry a high risk of non-combatant civilians. For sources outside the conflict zone – like media outlets in Russia, and international organizations like the OSCE – we found a subtler form of bias: a tendency not to attribute responsibility for violence to either side, and frame both sides’ violence as about equally indiscriminate.

Our experimental findings show that respondents were more supportive of intervening in a conflict against the side characterized as perpetrating the violence.
They were less supportive of intervening when confronted with neutral coverage of the conflict. Interestingly, characterizing the violence as indiscriminate or directed against military combatants did not significantly influence participants’ support for intervention. Our respondents supported intervening against the party responsible for violence, regardless of the nature of that violence.

Each of these findings was robust across both American and Indian survey participants, suggesting that these patterns are not artifacts of either the uniquely dominant U.S. position in the global system or of India’s prominent role in numerous international peacekeeping missions.

The absence of any effect of characterizing violence as indiscriminate, also robust across Americans and Indians, surprised us. The implication is that attributions of responsibility for violence loom larger in public attitudes than do the details concerning the nature of the violence which one or the other side commits.

The relative direction and magnitude of actor-specific reporting biases in Ukraine represent the exact opposite of what would be needed to quickly resolve the conflict. The fact that neutral coverage – which prevails in international reporting of events in Ukraine – suppresses support for intervention raises troubling questions concerning the capacity of world leaders to generate the necessary public support for peacekeeping missions. When information providers ‘play it down the middle’ – whether due to the journalistic value of balanced, neutral coverage, or due to more cynical ‘false equivalency’ – consumers have more difficulty ascribing responsibility, and are less likely to support outside intervention. This dulling effect of neutral coverage on external public opinion stands in stark contrast to the polarizing effect of ‘one-sided’ local news coverage. The net effect of these reporting biases is that domestic audiences may become less interested in striking a bargain with the opposing side, while outside audiences will become less interested in intervention.

Reversing these two sets of biases is, of course, easier said than done. In the absence of attributions of responsibility for violence, leaders and activists interested in conflict resolution will need to better inform journalists about the details of specific incidents. Where attribution exists, governments and NGOs will need to ensure that audiences have access to more than one source of information.

Our study suggests that reporting bias can have a potentially significant impact on public attitudes toward conflict resolution, one that to date has gone largely unrecognized by scholars and practitioners. Future research should thus extend this analysis to additional civil conflicts, actors, and media outlets to determine whether these findings generalize beyond Ukraine, and in doing so to further refine our estimates of the nature, extent, and consequences of reporting bias.
Appendix

Derivation of \( \omega \)

The transition matrix in (1) represents an unrestricted case, allowing for unlimited dependence in the decisions of the two combatants. Let us now introduce the restriction that strategy choices are made independently – a reasonable assumption where military decisions are heavily driven by internal factors, such as doctrinal rigidity and centralized command and control, rather than adaptation to the dynamic interactions of conflict (Smith, Sola and Spagnolo [2000]). Let \( p_k \) be the probability that actor \( k \) plays \( L \) after its opponent plays \( L \). Let \( q_k \) be the probability that actor \( k \) plays \( L \) after the opponent plays \( H \). By way of illustration, a tit-for-tat strategy would dictate that \( p = 1 \) and \( q = 0 \), and this reactive strategy holds every round, irrespective of the current state of the system. In a two-player game, the rebels’ reactive strategy is \( S_1(p_1,q_1) \) and the government’s strategy is \( S_2(p_2,q_2) \). This restriction produces the following transition matrix:

\[
M = \begin{pmatrix}
LL & LH & HL & HH \\
LL & p_1p_2 & q_1p_2 & p_1q_2 & q_1q_2 \\
LH & p_1(1-p_2) & q_1(1-p_2) & p_1(1-q_2) & q_1(1-q_2) \\
HL & (1-p_1)p_2 & (1-q_1)p_2 & (1-p_1)q_2 & (1-q_1)q_2 \\
HH & (1-p_1)(1-p_2) & (1-q_1)(1-p_2) & (1-p_1)(1-q_2) & (1-q_1)(1-q_2)
\end{pmatrix}
\]

By way of an example, the probability that the conflict transitions from state \( HL \) (high rebel violence, low government violence) to \( LH \) (low rebel violence, high government violence) is \( p_1(1-q_2) \), or the probability that rebels play \( L \) after the government plays \( L \) and the government plays \( H \) after rebels play \( H \).

Following (Nowak [1990]), we define the quantities \( r_k = p_k - q_k \), \( k = 1, 2 \) and \( s_1 = \frac{q_2r_1+q_1}{1-r_1r_2}, s_2 = \frac{q_1r_2+q_2}{1-r_1r_2} \). Let \( \bar{x}_t \) be the probability distribution of the game after \( t \) rounds. Assuming that the matrix \( M \) is regular, implying \( |r_1r_2| < 1 \), we can derive its stationary distribution.

If \( M \) has a unique largest eigenvalue \( \lambda_0 = 1 \) (by the Perron-Frobenius theorem, this is true if the chain is irreducible and aperiodic) then the stationary distribution \( \bar{x} \) will be the eigenvector \( \bar{v}_0 \) associated with the dominant eigenvalue \( \lambda_0 \), normalized to sum to one.
The eigenvector \( \nu_0 \) associated with the dominant eigenvalue of \( M \) is

\[
\vec{x} = \begin{bmatrix}
\frac{(q_1+q_2 r_1)(q_2+q_1 r_2)}{(1-r_1 r_2-q_1-q_2 r_1)(1-r_1 r_2-q_2-q_1 r_2)} \\
\frac{1-r_1 r_2-q_1-q_2 r_1}{q_1+q_2 r_1} \\
\frac{1-r_1 r_2-q_2-q_1 r_2}{q_2+q_1 r_2} \\
1
\end{bmatrix}
\]

which simplifies to

\[
\vec{x} = \begin{bmatrix}
s_1 s_2 \\
\frac{s_1}{1-s_1} \\
\frac{s_2}{1-s_2} \\
1
\end{bmatrix}
\]

Normalized by \((1-s_1)(1-s_2)\), the stationary distribution becomes:

\[
\vec{x} = \begin{bmatrix}
s_1 s_2 \\
s_1 (1-s_2) \\
(1-s_1)s_2 \\
(1-s_1)(1-s_2)
\end{bmatrix}
\]

We may now define the relative propensity for unilateral escalation:

\[
\omega = s_1 (1-s_2) - (1-s_1)s_2 = q_1 (1-p_2) - q_2 (1-p_1)
\]

If \( \omega < 0 \) unilateral escalation is more common in equilibrium for the rebels than the government, if \( \omega > 0 \) unilateral restraint by the rebels is more common than unilateral restraint by the government, and if \( \omega = 0 \) there is no difference in the combatants’ likelihood to escalate.

In the more general case where we drop the independence restriction, we can obtain this statistic directly from the stationary distribution of transition matrix \( M \) as written in (1):

\[
\omega = \pi_2 - \pi_3
\]
References


Pomerantsev, Peter. 2014. “Can Ukraine Win Its Information War With Russia?” *The Atlantic*.


URL: [http://slon.ru/world/zachem_ukraine_ministerstvo_pravdy-1193446.xhtml](http://slon.ru/world/zachem_ukraine_ministerstvo_pravdy-1193446.xhtml)


