START

Quantifying Gray Zone Conflict: (De-)escalatory Trends in Gray Zone Conflicts in Colombia, Libya and Ukraine

Report to DHS S&T Office of University Programs and DoD Strategic Multilayer Assessment Branch

June 2017

National Consortium for the Study of Terrorism and Responses to Terrorism
A Department of Homeland Security Science and Technology Center of Excellence
Led by the University of Maryland

8400 Baltimore Ave., Suite 250 • College Park, MD 20742 • 301.405.6600
www.start.umd.edu

About This Report

The authors of this report are Barnett S. Koven, Senior Researcher at the National Consortium for the Study of Terrorism and Responses to Terrorism (START), Varun Piplani, Data Analyst at START, Steve Sin, Senior Researcher at START and Marcus Boyd, Senior Researcher at START. Questions about this report should be directed to Barnett S. Koven at bkoven@start.umd.edu.

This report is part of START project, "Shadows of Violence: Empirical Assessments of Threats, Coercion and Gray Zones," led by Amy Pate.

This research was supported by a Centers of Excellence Supplemental award from the Department of Homeland Security's Science and Technology Directorate's Office of University Programs, with funding provided by the Strategic Multilayer Assessment (SMA) Branch of the Department of Defense through grant award number 2012ST061CS0001-05 made to the National Consortium for the Study of Terrorism and Responses to Terrorism (START). The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security, the U.S. Department of Defense or START.

About START

START is supported in part by the Science and Technology Directorate of the U.S. Department of Homeland Security through a Center of Excellence program led by the University of Maryland. START uses state-of-the-art theories, methods and data from the social and behavioral sciences to improve understanding of the origins, dynamics and social and psychological impacts of terrorism. For more information, contact START at infostart@start.umd.edu or visit www.start.umd.edu.

Citations

To cite this report, please use this format:

Koven, Barnett S., Varun Piplani, Steve Sin, and Marcus Boyd. "Quantifying Gray Zone Conflict: (De-)escalatory Trends on Gray Zone Conflicts in Colombia, Libya and Ukraine," Report to DHS S&T Office of University Programs and DoD Strategic Multilayer Assessment Branch. College Park, MD: START, 2017.

Contents

Executive Summary

This report employs frequentist statistical analysis in order to model the effects of various factors, including the type of actors (state, violent non-state actor (VNSA) or civilian) involved and the prevalence of kinetic activity, on (de-)escalatory trends in Gray Zone conflicts. This is coupled with the development of a Bayesian Belief Network for predictive analysis of White, Gray and Black Zone behavior within Gray Zone conflicts.

Both sets of analyses utilize a version of the event-level data from the Worldwide Integrated Crisis Early Warning System (ICEWS). However, we heavily modified this data prior to running the analyses. Specifically, we recoded new variables of particular interest to the study of Gray Zone conflict, addressed erroneous and duplicate entries, and restructured the data in order to model temporal changes. This was accomplished using a hybrid process involving both automated recoding procedures and expert human coders.

Our procedure was applied to three diverse gray zone conflicts: Colombia (01 January 2002 to 19 September 2016), Libya[1] (01 January 2011 to 12 September 2016) and Ukraine (01 January 2014 to 12 September 2016). These conflicts all share two commonalities: they all entail a large amount of Gray Zone activity and myriad VNSAs. Nevertheless, the three cases vary in a number of important respects: the level of foreign involvement, the belligerents' motives, as well as their guiding ideologies, and their geographic location. Consequently, the results are highly likely to be generalizable to a diverse array of other Gray Zone conflicts.

Three principal findings hold across both methodological approaches and are apparent in multiple cases. First, contrary to popular belief, kinetic military operations are a key aspect of Gray Zone conflicts. While it is true that these events are relatively sparse (around 20% of all events depending on the case), they have substantial influence in shaping non-kinetic events. Second, while VNSAs are less proficient than states at identifying their adversaries (de-)escalation trends, the closer VNSAs are linked to states, the less that this is a problem. Finally, legitimacy matters. For this reason, both VNSA and state forces will moderate their behavior in order to avoid being perceived as the aggressor or engaging in more (easily visible) civilian victimization than their opponents.

[1] Data quality limitations precluded extending the Bayesian analysis to Libya.

Introduction

The National Consortium for the Study of Terrorism (START) has been tasked with providing support to the Special Operations Command (SOCOM) Gray Zone project undertaken as a Strategic Multilayer Assessment (SMA) initiative. Part of this support included the preparation of quantitative case studies covering three diverse Gray Zone conflicts: Colombia (01 January 2002 to 19 September 2016), Libya (01 January 2011 to 12 September 2016) and Ukraine (01 January 2014 to 12 September 2016). Each of the three quantitative case studies employ frequentist statistics aimed at elucidating (de)escalatory trends in Gray Zone conflicts involving violent non-state actors (VNSAs). The same data also undergirds a Bayesian network analysis with a geospatial component for Colombia and Ukraine. Data quality limitations precluded extending this approach to Libya. Each case study builds upon the following working definition of Gray Zones:

> *"The Gray Zone is a conceptual space between peace and war, occurring when actors purposefully use single or multiple instruments of power to achieve political-security objectives with activities that are typically ambiguous or cloud attribution and exceed the threshold of ordinary competition, yet intentionally fall below the level of large-scale direct military conflict, and threaten US and allied interests by challenging, undermining, or violating international customs, norms, or laws"* (Department of Defense Strategic Multi-Layer Assessment 2017).[2]

More specifically, this research utilizes frequentist statistical analysis to examine what drives activities by various types of belligerents (including state forces and violent non-state actors (VNSAs)) and civilians, across three White (peaceful competition below the level of Gray Zone conflict), Gray and Black (conflictual behavior exceeding the Gray Zone threshold) Zones. While the frequentist analysis is conducted across all three cases, data quality limitations limited the scope of the analysis in the Libyan case. Moreover, for the Colombian and Ukrainian cases where data quality was substantially better, we also employed a Bayesian analysis-based approach. Both sets of analyses utilize data that we recoded, cleaned and de-duplicated from the Worldwide Integrated Crisis Early Warning System (ICEWS).

The frequentist analysis leveraged a time-series, cross-sectional (TSCS) version of the data in order to examine how state, civilian and VNSA forces effect the Zonal preferences for activities by (other) VNSAs, as well as how state, civilian and VNSA forces effect the choice of White Zone activities by (other) states. The Bayesian analysis endeavored to build a predictive model that could estimate the probability of a Gray Zone even occurring conditional on the source and target of the event, as well as whether a given event was kinetic or not.

This report proceeds in seven sections. The first section justifies our case selection and provides background on each of the three cases. The subsequent section introduces the data. The third section

[2] While the analysis began before the January 2017 changes to this definition, the differences between the two definitions do not impact our approach, analysis or findings. The original definition read: *"The Gray Zone is a conceptual space between peace and war, occurring when actors purposefully use multiple instruments of power to achieve political-security objectives with activities that are ambiguous or cloud attribution and exceed the threshold of ordinary competition, yet fall below the level of large-scale direct military conflict, and threaten US and allied interests by challenging, undermining, or violating international customs, norms, or laws"* (Department of Defense Strategic Multi-Layer Assessment, "Gray Zone Effort Update," September 2016).

elaborates the research design for both the frequentist and Bayesian analysis. The next section presents the results from both sets of analyses. The fifth section provides a detailed discussion of key findings in each of the three cases. The penultimate section explores commonalities and differences across the cases. The final section concludes.

Case Selection[3]

As already indicated, the reports cover three distinct conflicts: Colombia, Libya and Ukraine. These cases share two crucial commonalities. First, they all entail an extensive amount of Gray activities. (Figures 1-3, below, depict the amount of Gray versus White and Black events for each of the three cases.) Second, all three conflicts involve substantial roles for multiple types of VNSAs. Nevertheless, the cases also diverge in numerous respects. First, while Colombia is largely an internal conflict, Libya has seen substantial foreign involvement and has become a proxy conflict for regional powers. Ukraine has seen the most extensive foreign involvement of the three. Whereas Libya began as a domestic uprising and only later became a hotbed for foreign belligerents, the Ukrainian crisis was precipitated by Russia. Second, Colombia and Libya involve armed competition for political power, whereas Ukraine is a secessionist conflict. Third, the belligerents' guiding ideologies vary across the conflicts. Colombia began as a conflict between Marxist insurgents and various conservative actors (such as traditional politicians and rightist paramilitary forces), but greed-driven motives have largely supplanted ideology. The Libyan crisis arose out of the Arab Spring uprisings, devolving into a civil war with international and domestic actors vying for control. Some actors are motivated by radical Islamist ideologies, whereas others are ideologically moderate but equally committed to obtaining political power. The Ukrainian case resulted from Russian designs on Ukrainian territory, in which ethnically Russian populations predominate. Ethnicity has played a large role in motivating this conflict, though other factors – such as support for European integration – are also salient. Finally, the cases represent three distinct regions: South America, North Africa and Eastern Europe. As a result of this variation, the selection of these three cases ensures that our findings are likely to be generalizable to other Gray Zone conflicts involving multiple types of VNSA forces.

Colombia[4]

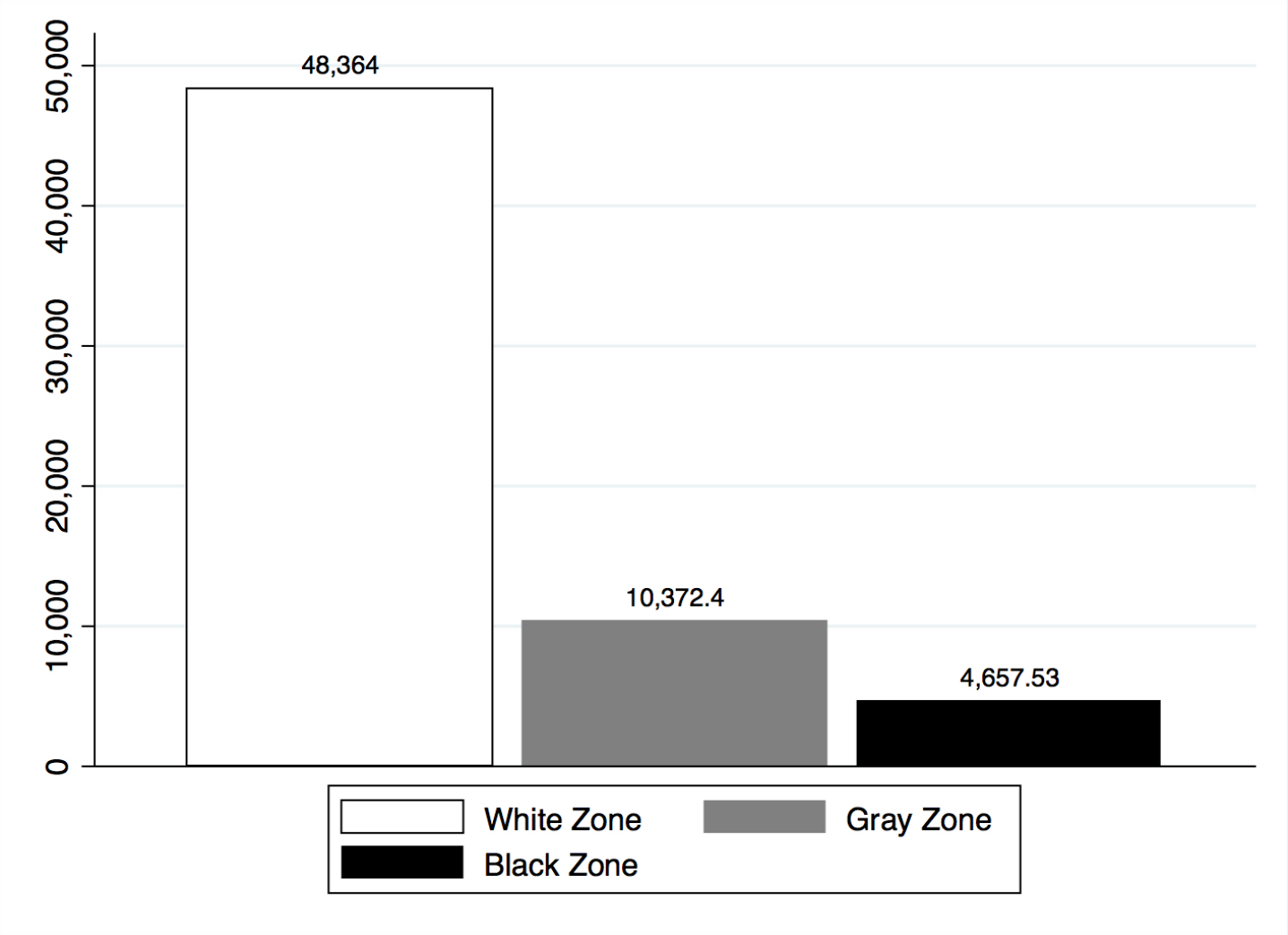
A power sharing agreement concluded a decade of civil war in 1958. The agreement established the National Front, and called for the two dominant political parties – the Liberals and the Conservatives – to alternate governing. While effective at ending the civil war, it also excluded the left from political power.

[3] This section is reproduced and adapted from Barnett S. Koven, "Demystifying Gray Zone Conflict: A Typology of Conflict Dyads and Instruments of Power in Colombia, Libya and Ukraine," report to DHS S&T Office of University Programs and DoD Strategic Multilayer Assessment Branch. College Park, MD: START, 2016.

[4] Barnett S. Koven, "Demystifying Gray Zone Conflict: A Typology of Conflict Dyads and Instruments of Power in Colombia, 2002-present," report to DHS S&T Office of University Programs and DoD Strategic Multilayer Assessment Branch (College Park, MD: START, 2016).

Consequently, six[5] distinct leftist insurgencies emerged.[6] An array of right wing paramilitary forces were stood-up to combat the insurgents. In 1997, the disparate paramilitary groups coalesced into a single entity, the United Self-Defense Forces of Colombia (*Autodefensas Unidas de Colombia*; AUC). In 2006 the AUC demobilized, though many former AUC combatants simply joined the ranks of Colombia's 16 organized criminal syndicates (*Bandas Criminales*; BACRIM). All of these actors, as well as state forces, have at various times cooperated and/or pursued peace, while fighting with each other at other times. The conflict, which has cost over 250,000 lives, while displacing millions more, continues to this day.

Figure 1: Total Weighted Activity by Zone in Colombia



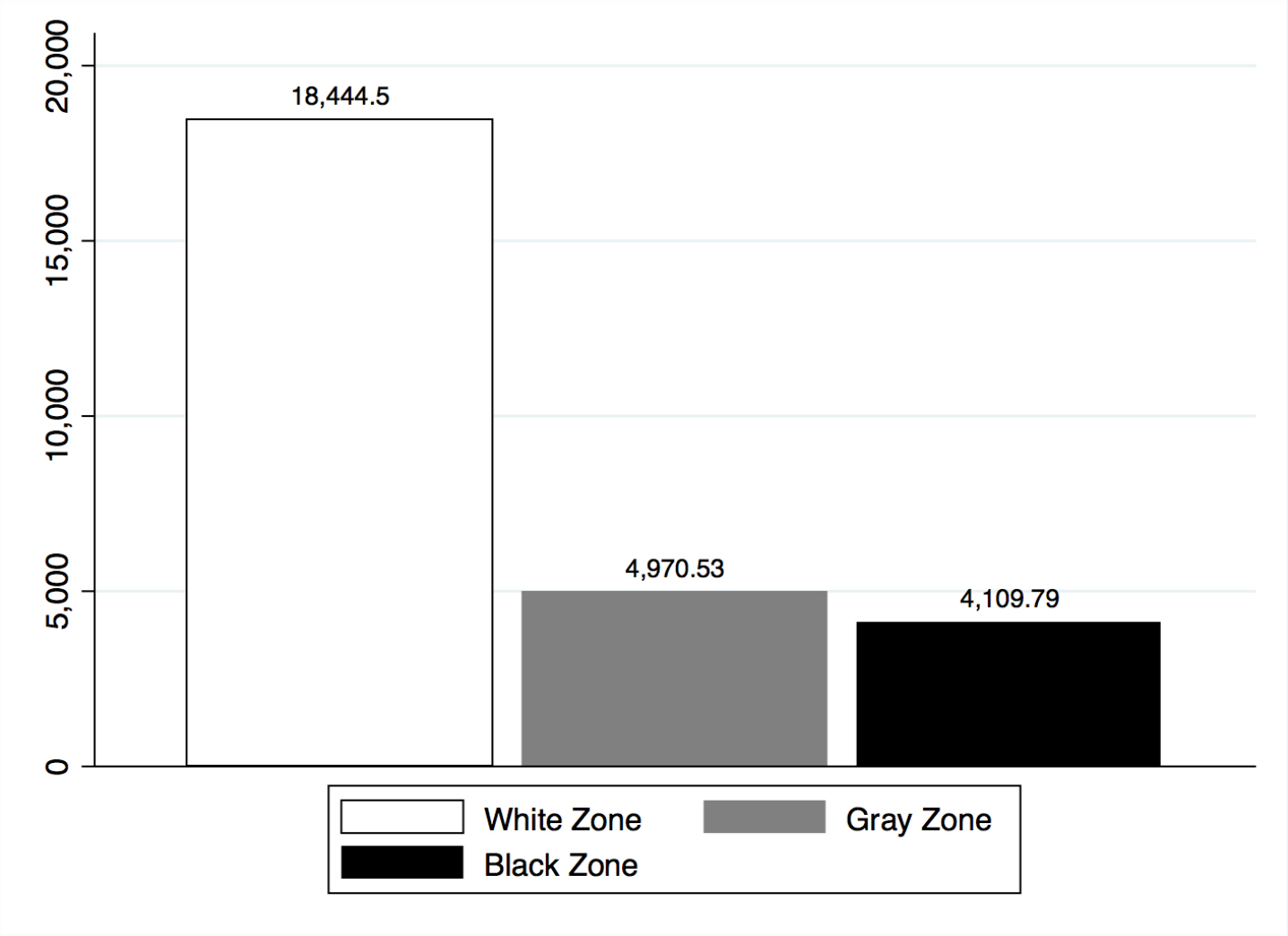
[5] The Revolutionary Armed Forces of Colombia (*Fuerzas Armadas Revolucionarias de Colombia*; FARC), the National Liberation Army (*Ejército de Liberación Nacional*; ELN), the 19th of April Movement (*Movimiento 19 de Abril*; M-19), the Popular Liberation Army (*Ejército Popular de Liberación*; EPL), the Quintín Lame Armed Movement (*Movimiento Armado Quintín Lame*; MAQL) and the Workers Revolutionary Party of Colombia (*Partido Revolucionario de los Trabajadores de Colombia*; PRT).
[6] The Colombia study focuses on the most recent phase (2002-present) of the Colombian conflict, during which just two (the FARC and the ELN) insurgent groups remain active.

Libya[7]

The 2011 Arab Spring uprisings quickly reached Libya and resulted in the demise of Colonel Muammar al-Qaddafi and his dictatorial regime. Unfortunately, this also left a power vacuum, and infighting between myriad VNSAs, which had cooperated to oust the Qaddafi regime, ensued. To quell the fighting, the National Transitional Council was established to facilitate democratic elections in July 2012. The General National Congress (GNC) proved victorious and was able to govern in relative peace. However new elections were held in June 2014, which saw the GNC's rival, the House of Representatives (HoR) take power despite the fact that only 16 percent of eligible voters turned out to the polls. The HoR deployed the Libyan National Army (LNA) in an effort to destroy its political opponents. This led to sustained armed conflict between the LNA and Libya Dawn, a coalition comprised of moderate and local Islamist forces loosely affiliated with the GNC. Libya Dawn and the GNC captured Tripoli and declared themselves as the new government. The HoR continued to claim legitimacy and re-established its government in Tobruk. In December 2015, a UN intervention led to the establishment of the Government of National Accord (GNA). However, neither the GNC nor the HoR have ratified the agreement establishing the GNA. This has led to a power struggle among the political organizations claiming legitimacy. The GNC, HoR and GNA all lack fully subordinated military forces; rather they are reliant on alliances of convenience with various armed factions. These groups facilitating the conflict between the aforementioned political actors are simultaneously engaged in their own private rivalries, which routinely result in violent clashes. In addition, quasi-governmental entities and both moderate and local Islamist VNSA forces are all engaged in conflict against the Islamic State of Iraq and the Levant (ISIL).

[7] Rachel A. Gabriel, and Mila A. Johns, "Demystifying Gray Zone Conflict: A Typology of Conflict Dyads and Instruments of Power in Libya, 2014-Present," report to DHS S&T Office of University Programs and DoD Strategic Multilayer Assessment Branch (College Park, MD: START, 2016).

Figure 2: Total Weighted Activity by Zone in Libya



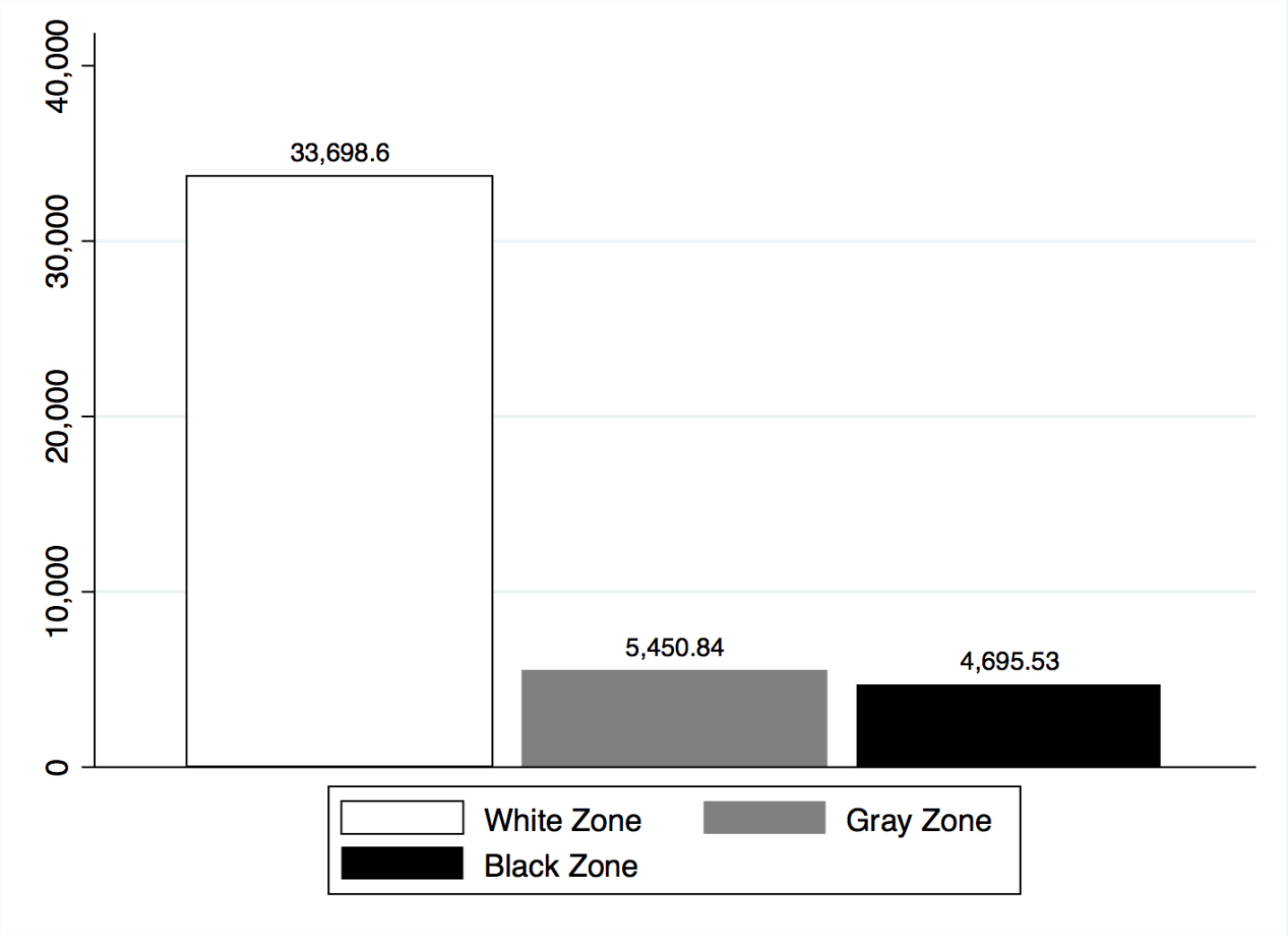
Ukraine[8]

The Ukrainian crisis began in late 2013, when President Viktor Yanukovych quashed a trade deal with the European Union, opting instead to pursue closer integration with Russia. Massive protests ensued and in February 2014, Yanukovych was impeached and fled into exile in Russia. Russia capitalized on the political crisis and orchestrated a referendum to annex the Crimean peninsula, home to a majority ethnic Russian population and the Black Sea Fleet. Simultaneously, anti-government protests emerged in the Donbas (comprised of the Donetsk and Luhansk regions of Eastern Ukraine). Violent confrontations ensued between pro-Russian and pro-Ukrainian mobs. Russian nationals, many of whom were connected to Russian security services, organized these clashes. In April 2014, pro-Russia militia forces, led by a Russian citizen, declared the establishment of the Donetsk People's Republic. The Luhansk People's Republic formed later that month. Regular Ukrainian military forces along with pro-Ukrainian volunteer

[8] Evgeny Finkel, "The Conflict in the Donbas between Gray and Black: The Importance of Perspective," report to DHS S&T Office of University Programs and DoD Strategic Multilayer Assessment Branch (College Park, MD: START, 2016).

formations attempted to reestablish state control. Local pro-Russian forces as well as Russian military personnel organized to counter the Ukrainian military and volunteer formations.

Figure 3: Total Weighted Activity by Zone in Ukraine



The Data

As noted, this research utilizes ICEWS data. This has required myriad substantial modifications to the original data. First, we developed a semi-automated approach to reclassifying each observation as White, Gray or Black, the type of actor initiating an event, and the intended target type (state, VNSA, civilian), along with other variables, such as if an event was kinetic or not. Second, we hand-coded stratified random samples of observations that contained missing data due to the fully automated collection methodology used to generate ICEWS and/or due to our semi-automated procedure for reclassifying variables. These hand-coded observations were then weighted to account for the remaining missing data, which was not hand-coded due to practical considerations. During this process, we also cleaned the dataset by removing any erroneous observations we discovered. Third, we developed a procedure to account for duplicate events, which were extremely common given that ICEWS was coded based on

media reports. Each of the aforementioned corrections were carefully validated using expertly trained human coders. We ran inter-coder reliable (ICR) tests to compare the automated aspects of our approach against a human coder and to ensure reliability between hand-coders. We also generated weighted variables for sensitivity analysis. Fifth, we generated a plethora of dyadic variables to model how events of various types (e.g., kinetic versus non-kinetic) across all three Zones undertaken by a given actor (e.g., state forces) affected actions of a particular type, Zone and actor. Finally, we transformed the event-level data into TSCS data. The TSCS data is disaggregated by month and administrative boundary. For Colombia, Ukraine and Libya we generate one version of the dataset disaggregated to the level of the first order subnational administrative boundary (the largest subnational administrative zone; for example, the departmental level in Colombia). For Colombia and Ukraine, we also generated a second version disaggregated to the level of the second order subnational administrative boundary (the second largest subnational administrative zone; for example, the municipal level in Colombia). Data quality limitations precluded following suit with the Libya data.

The resulting datasets for Colombia and Ukraine are therefore each of markedly different sizes. The first version of the data for Colombia includes 7,040 observations spanning 176 months and 40 departments. The second version includes 106,128 observations covering 176 months and 603 municipalities. Similarly, for Ukraine, the first version of the data includes 3,003 observations across 32 months and 91 oblasts (department-like areas). The second version contains 14,091 observations distributed across 32 months and 427 raions (municipality-like areas). As already indicated, for Libya, we were only able to generate the first version of the data. This version captures 1,518 observations spanning 68 months and 22 shabiyahs (department-like areas).

A detailed record and discussion of all modifications to the original ICEWS data can be found in Appendix A.

Dependent Variable

Insofar as we are interested in explaining how state, civilian and VNSA forces affect the Zonal preferences for activities by (other) VNSAs, as well, as how state, civilian and VNSA forces affect the choice of White Zone activities by (other) states, four different operationalizations of the dependent variable are required:

- $\text{White}_{\text{VNSA}}$ – A count of White Zone activity perpetrated by VNSAs.
- $\text{Gray}_{\text{VNSA}}$ – A count of Gray Zone activity perpetrated by VNSAs.
- $\text{Black}_{\text{VNSA}}$ – A count of Black Zone activity perpetrated by VNSAs.
- $\text{White}_{\text{state}}$ – A count of White Zone activity perpetrated by state forces.

Independent Variables

With respect to explaining VNSA Zonal preferences, the following independent variables are utilized:

- State-VNSA White – A count of White Zone activity perpetrated by states against VNSAs.
- State-VNSA Gray – A count of Gray Zone activity perpetrated by states against VNSAs.
- State-VNSA Black – A count of Black Zone activity perpetrated by states against VNSAs.
- Civilian-VNSA White – A count of White Zone activity perpetrated by civilians against VNSAs.

- Civilian-VNSA Gray – A count of Gray Zone activity perpetrated by civilians against VNSAs.
- VNSA-VNSA White – A count of White Zone activity perpetrated by VNSAs against other VNSAs.[9]
- VNSA-VNSA Gray – A count of Gray Zone activity perpetrated by VNSAs against other VNSAs.
- VNSA-VNSA Black – A count of Black Zone activity perpetrated by VNSAs against other VNSAs.
- State-Civilian White – A count of White Zone activity perpetrated by states against civilians.
- State-Civilian Gray – A count of Gray Zone activity perpetrated by states against civilians.
- State-Civilian Black – A count of Black Zone activity perpetrated by states against civilians.
- Kinetic – A count of kinetic activity.

As regards explaining the choice of White Zone activities by states, the following independent variables are utilized:

- VNSA-State White – A count of White Zone activity perpetrated by VNSAs against states.
- VNSA-State Gray – A count of Gray Zone activity perpetrated by VNSAs against states.
- VNSA-State Black – A count of Black Zone activity perpetrated by VNSAs against states.
- Civilian-State White – A count of White Zone activity perpetrated by civilians against states.
- Civilian-State Gray – A count of Gray Zone activity perpetrated by civilians against states.
- State-State White – A count of White Zone activity perpetrated by states against other states.
- State-State Gray – A count of Gray Zone activity perpetrated by states against other states.
- State-State Black – A count of Gray Zone activity perpetrated by states against other states.
- VNSA-Civilian White – A count of White Zone activity perpetrated by VNSAs against civilians.
- VNSA-Civilian Gray – A count of Gray Zone activity perpetrated by VNSAs against civilians.
- VNSA-Civilian Black – A count of Black Zone activity perpetrated by VNSAs against civilians.
- Kinetic – A count of kinetic activity.

Summary statistics are presented in appendix G.

Research Design

This section introduces the research design for both the frequentist and Bayesian analysis. Insofar as this analysis is exploratory in nature, no hypotheses are directly tested. As such, none are introduced in this section.

Frequentist Analysis

To estimate the impact of our covariates on the three Zones for VNSAs, and the White Zone for states, we use fixed and random effects models. Fixed effects models are useful because they control for unit-level heterogeneity. This is extremely helpful because this analysis does not account for a variety of factors at the unit-level that may impact (de-)escalation. These include economic, social or political conditions present in a given subnational administrative unit. The disadvantage of fixed effects models is that they only consider within-unit variation and ignore all between-unit variation. Put simply, they do not

[9] Variables, such as this one where the independent variable of interest is very similar to the dependent variable of interest are potentially problematic. Naturally, one would expect VNSA activity to predict VNSA activity. However, this issue is overcome by specifying temporal lags. Specifically, this variable, and others like it, examines the effect of VNSA activity in a given month on VNSA activity in the following month.

compare conditions across subnational administrative units. While useful, this ignores the important differences between administrative regions.

To account for this, we also test our findings using a random effects model, which estimates the impact of our covariates both within and between regions. While this does not deal entirely with the issue of unit-level heterogeneity, we gain the ability to compare the impact of our covariates on Zones across regions. By utilizing both fixed and random effects, we increase our confidence that the results are not an artifact of model choice. For our basic model, we use a random effects model at the most disaggregated subnational administrative level possible to maximize the amount of statistical power brought to bear on this question. However, the results from all other models are presented in Appendix H and explained below.

In addition to unit-level considerations, we also account for temporal dynamics in two ways. First, we include a lagged-dependent variable in all models. Second, we lag all of our covariates by one month so that the coefficient on each variable is the impact of that variable on White, Gray or Black Zone behavior in the next month. Finally, standard errors are clustered by subnational administrative unit in order to account for heteroskedasticity.

Our analysis therefore focuses on four models for the Colombian and Ukrainian cases. The first three models seek to explain Zonal preferences of VNSAs and the last model explores White Zone behavior by states:

1. Model 1: Explains White Zone activity by VNSAs. Therefore, the dependent variable is $White_{VNSA}$.
2. Model 2: Explains Gray Zone activity by VNSAs. Therefore, the dependent variable is $Gray_{VNSA}$.
3. Model 3: Explains Black Zone activity by VNSAs. Therefore, the dependent variable is $Black_{VNSA}$.
4. Model 4: Explains White Zone activity by states. Therefore, the dependent variable is $White_{state}$.

Each model is run four times to account for the two different subnational administrative levels and the choice of fixed versus random effects, and each version of the model is reported in its own table. In each country analysis, results that stood out as being particularly interesting are discussed in detail in the *Discussion* section, below. For these results, we also generated graphs using predicted probabilities in order to understand how increasing (or decreasing) values of a particular variable of interest affects the degree of White, Gray or Black Zone behavior by VNSAs, or the White Zone behavior of states. (In order to generate predicted probabilities, we dropped the lagged-dependent variable.[10])

For Libya, we were unable to parse state versus VNSA forces. This is the case for two reasons. First, as noted, there are multiple groups claiming to be the legitimate government of Libya. We considered coding the government as the claimant recognized by the international community – even though this has changed during the course of the conflict and our analysis. However, it was often impossible to determine

[10] Dropping the lagged DV does not substantively alter the basic findings for Colombia or Ukraine. However, in the case of Libya, dropping the lagged DV imparted substantive changes on some of the findings. We clearly identified these cases as model dependency undermines our confidence in these particular results.

which "government" was referenced in a given observation in the ICEWS data. Second, and even more problematic, we were unable to determine which militia forces were an extension of the internationally recognized government and which were in opposition to it. This was the case because alliances quickly and constantly shifted. Often insufficient detail also existed in the source material to determine which militia(s) were involved in a particular action. As such, we disaggregated source and target types by civilian or VNSA only. (Since no group exercised a monopoly on the legitimate use of force – which Max Weber noted as a defining condition of statehood – all political and associated militia groups are categorized as VNSAs (Weber 1918).) Consequently, our analysis is restricted to just the first three of the four aforementioned models utilized in Colombia and Ukraine.

However, given that Libya experienced a relatively high level (1,576 out of 43,039 events) of Gray Zone activity perpetrated by civilians, we chose to estimate a model 4'. Model 4' examines the use of Gray tactics by civilians:

Model 4': Explains Gray Zone activity by civilians. Therefore, the dependent variable is $\text{Gray}_{\text{civilian}}$.

Additionally, each model is run only two times insofar as the data is only disaggregated to the level of the first order subnational administrative boundary. The two model specifications account for the choice of fixed versus random effects. As with Colombia and Ukraine, each version of the model is reported in its own table and findings that were particularly interesting are discussed – along with accompanying figures – in the *Discussion* section, below.

Bayesian Analysis

For Colombia and Ukraine, we also explored the use of Bayesian belief network (BBN) analysis to determine if some level of predictive analysis is plausible from observing a handful of easily observable macro-level variables for each event. Unlike the aforementioned frequentist analysis, which aimed to estimate the impact of covariates on the Zones for VNSAs and the White Zone for states, the BBN analysis assesses the probability of an event occurring in a certain Zone by observing the presence (or absence) of specific variables, such as source and target type and the use of kinetic force.

BBNs are a form of probabilistic modeling that represents a system as a series of random variables and the dependencies among them. The advantage of BBN analysis is that it can handle complex dependencies or uncertainties that many other methods cannot. This is because it uses abductive reasoning rather than inductive or deductive reasoning (Jensen 1996).[11] In short, abductive reasoning can be explained as the "inference to the best explanation" (Sober 2013, 28). Furthermore, BBN analysis can be either qualitative or quantitative. Qualitative BBNs are used to graphically construct a Causal Network that represents the relationships between variables under examination. Quantitative approach to BBNs are used to determine the probabilities of variables occurring and their interconnections using Bayesian calculus.

[11] An example of this is that BBNs are excellent when attempting to analyze a situation when expert opinion is ambiguous, incomplete or uncertain.

In BBNs, variables are often known as "nodes" and the relationships that connect them are represented with arcs, signifying that the nodes are conditionally dependent. The absence of an arc between two nodes, on the other hand, denotes that the nodes are conditionally independent (Ayyub 2003). Given the data qualities, we selected three nodes to conduct the BBN analysis to test the plausibility of predicting the probability of a Gray Zone event. Our nodes were as follow:

- Independent Node A: Type of Source
- Independent Node B: Type of Target
- Independent Node C: Use of Kinetic Force
- Dependent Node: Gray Zone Event

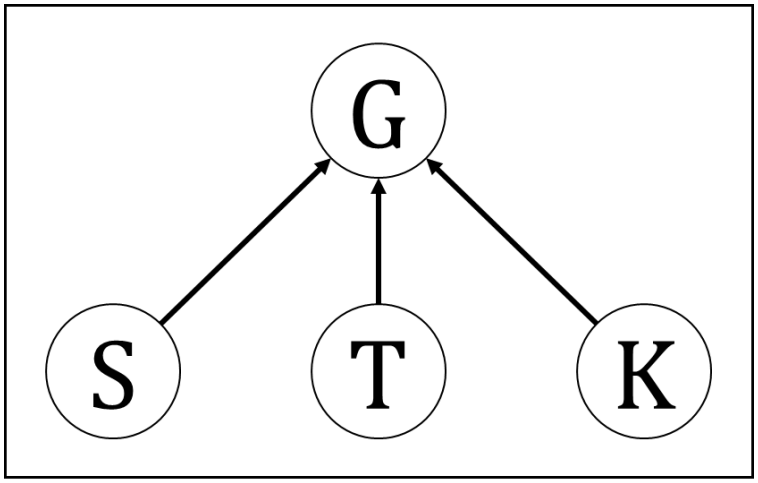
Given that the three independent nodes are conditionally independent – no one node effects the probability of occurrence of another node – the probability of Gray Zone event occurrence – P(G), given that Type of Source: P(S), Type of Target: P(T), and the Use of Kinetic Force: P(K) are known, can be written as shown in Equation 1, below.

Equation 1: Probability of Gray Zone BBN Equation

$$P(G|S,T,K) = P(G|S) + P(G|T) + (P(G|K)$$

Additionally, since the Source, Target, and Use of Kinetic Force all independently influence the probability of an event being Gray, we can graphically represent the relationships among these nodes. This is accomplished using a convergence structure diagram where the three conditionally independent nodes (Source, Target and Use of Kinetic Force) are parent nodes that each exert a certain amount of influence – independently of each other – on the probability of Gray Zone event occurrence, the child node. Figure 4, below, provides a graphic depiction of this BBN model.

Figure 4: Gray Zone BBN Structure



The BBN model shown above was utilized to conduct bivariate and multivariate plausibility tests for the Colombian and Ukrainian cases. Combining bivariate and multivariate tests helps to ensure consistency in the findings. Due to data quality limitations, Libya was not analyzed utilizing BBN.

Results

Colombia

VNSA Zonal Preferences

Table 1, below, presents the results for the models testing the impact of state, civilian and VNSA activities on the Zonal behavior of (other) VNSAs. Model 1 examines White Zone activity by VNSAs, Model 2 estimates Gray Zone activity by VNSAs and Model 3 looks at how different factors shape Black Zone activity by VNSAs.

Table 1: Colombian VNSA Zonal Preferences

Variable	Model 1: White		Model 2: Gray		Model 3: Black	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
State-VNSA White	0.02***	0.00	0.04***	0.00	-0.01***	0.00
State-VNSA Gray	0.03***	0.01	-0.04***	0.01	-0.06***	0.00
State-VNSA Black	-0.06***	0.00	-0.14***	0.00	0.03***	0.00
Civilian-VNSA White	0.01***	0.00	-0.02***	0.00	0.01***	0.00
Civilian-VNSA Gray	0.27***	0.02	-0.11***	0.03	-0.02	0.01
VNSA-VNSA White	-0.01	0.02	0.08***	0.03	-0.01	0.01
VNSA-VNSA Gray	-0.07***	0.01	0.22***	0.01	-0.05***	0.01
VNSA-VNSA Black	-0.11***	0.02	0.09***	0.03	0.00	0.01
State-Civilian White	0.04***	0.00	0.00	0.00	0.00	0.00
State-Civilian Gray	0.00	0.01	-0.01	0.01	0.02***	0.00
State-Civilian Black	0.07***	0.01	0.01	0.02	0.07***	0.01
Kinetic	0.54***	0.01	1.39***	0.01	0.23***	0.01
Lagged DV	0.10***	0.00	0.03***	0.00	-0.03***	0.00
Constant	0.01	0.01	0.05***	0.02	0.02***	0.01
Number of Obs.			106,128			
Number of Groups			603			

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Standard Errors clustered by country, using a Random Effects GLS Regression model. This model presents findings from the Colombia dataset at the second order subnational administrative level.

State-VNSA Variables

The use of White Zone tactics by states towards VNSAs is associated with higher levels of White Zone activity by VNSAs, and the effect is statistically significant at the 99 percent confidence level (Model 1). This finding is robust across all models in Table 1 as well as alternative specifications in Tables H.1-H.6. The use of Gray Zone tactics by states towards VNSAs is also correlated with higher White Zone activity by VNSAs, and the effect is statistically significant at the 99 percent confidence level. However, the effect is only statistically significant in three out of four models; it is not statistically significant in Table H.3. The use of Black Zone tactics by states is negatively correlated with White Zone activity by VNSAs, and the effect is statistically significant at the 99 percent confidence level and robust across all models in Table 1 as well as alternative specifications in Tables H.1-H.6.

The use of White Zone tactics by states towards VNSAs is also associated with higher levels of Gray Zone activity by VNSAs, and the effect is statistically significant at the 99 percent confidence level (Model 2). This finding is robust across all models in Table 1 as well as alternative specifications in Tables H.1-H.6. The use of Gray Zone tactics by states towards VNSAs does not have a statistically significant effect on Gray Zone tactics usage by VNSAs; a statistically significant finding is only observed in one of the four models, presented in Table 1. The use of Black Zone tactics by states is negatively correlated with Gray Zone activity by VNSAs, and the effect is statistically significant at the 99 percent confidence level and robust across all models in Table 1 as well as alternative specifications in Tables H.1-H.6.

The use of White Zone tactics by states towards VNSAs is associated with lower levels of Black Zone activity by VNSAs, and the effect is statistically significant at the 99 percent confidence level (Model 3). However, this finding is robust in three out of four models. In fact, the direction of the effect is reversed in the model presented in Table H.3. The use of Gray Zone tactics by states towards VNSAs reduces Black Zone activity by VNSAs, and the effect is statistically significant at the 99 percent confidence level and robust across all models. The use of Black Zone tactics by states increases Black Zone activity by VNSAs, and the effect is statistically significant at the 99 percent confidence level and robust across all models in Table 1 as well as alternative specifications in Tables H.1-H.6.

Civilian-VNSA Variables

The use of White and Gray tactics by civilians targeting VNSAs increases White Zone activity by VNSAs. These results are statistically significant at the 99 percent confidence level (Model 1). They are also robust across all models in Table 1 as well as alternative specifications in Tables H.1-H.6. The use of White Zone tactics by civilians towards VNSAs reduces Gray Zone activity by VNSAs, and the effects are statistically significant at the 99 percent confidence level and robust across all models in Table 1 as well as alternative specifications in Tables H.1-H.6 (Model 2). The use of Gray Zone tactics is weakly associated with lower Gray Zone activity by VNSAs – the results are statistically significant at the 99 percent confidence level in Tables 1 and H.1 but fail to achieve statistical significance in the models presented in Tables H.3 and H.5. The use of White Zone tactics by civilians increases the use of Black Zone activity by VNSAs, and the results are statistically significant at the 99 percent confidence level (Model 3). They are also robust across all models in Table 1 as well as alternative specifications in Tables H.1-H.6. Finally, the use of Gray Zone tactics by civilians has no discernable effect on Black Zone activity by VNSAs across all models.

VNSA-VNSA Variables

Zonal differences in the activities of VNSAs against other VNSAs also matter. The use of White Zone tactics by VNSAs towards other VNSAs has a weakly negative effect on the use on White Zone activity by VNSAs (Model 1). The effect is negative and statistically significant at the 99 percent confidence interval in two out of four models (Tables H.1 and H.5). Similarly, the use of Gray Zone tactics by VNSAs towards other VNSAs is weakly negatively correlated with White Zone activity by VNSAs. The effect is negative and statistically significant at the 99 percent confidence interval in two out of four models (Tables 1 and

H.1). The use of Black Zone activity between VNSAs reduces White Zone activity by other VNSAs. The effect is statistically significant at the 99 percent confidence level and robust in three out of four model specifications; the model presented in Table H.5 did not return statistically significant results.

The use of White Zone tactics by VNSAs towards other VNSAs is also associated with higher levels of Gray Zone activity by VNSAs, and the effect is statistically significant at the 99 percent confidence level. These findings are robust in three out of four models; the effect is reversed in Table H.1. The use of Gray Zone tactics by VNSAs against other VNSAs increases the use of Gray Zone tactics by VNSAs overall, and the effect is statistically significant at the 99 percent confidence level and robust across all models. The use of Black Zone tactics between VNSAs is also weakly associated with higher levels of overall Gray activity by VNSAs. This effect is statistically significant at the 99 percent confidence level in two out of four models; it does not achieve statistical significance in the models presented in Tables H.3 and H.5.

The use of White Zone tactics between VNSAs has no effect on Black Zone activity by VNSAs in any model. The use of Gray Zone tactics between VNSAs is strongly associated with lower levels of overall Black activity by VNSAs. These findings are statistically significant at the 99 percent confidence level and robust across all models. Interestingly, the use of Black Zone tactics between VNSAs does not have a statistically significant effect on overall usage of Black Zone tactics by VNSAs in any model.

State-Civilian Variables

The use of White, Gray and Black tactics by states against civilians is also sometimes relevant. We find that the use of White Zone tactics by states towards civilians is associated with higher levels of White Zone activity by VNSAs at the 99 percent confidence level, and the results are robust in three out of four models; statistically significant results are not returned in the model presented in Table H.5. Further, we find that the use of Gray Zone tactics by states towards civilians does not have a strong relationship with White Zone activity by VNSAs. Specifically, the effect is negative and statistically significant (at the 99 percent confidence level) in only one out of four models (Table H.1). The use of Black Zone tactics by states against civilians is correlated with higher levels of White Zone activity by VNSAs, and the effect is statistically significant at the 99 percent confidence level in two out of four models (Tables 1 and H.1).

The use of White and Gray Zone tactics by states towards civilians is weakly associated with lower levels of Gray Zone activity by VNSAs. This finding is statistically significant in only two out of four models; the results are not statistically significant in the models presented in Tables 1 and H.3. The use of Black Zone tactics by states towards civilians does not statistically significantly affect the use of Gray Zone tactics by VNSAs across any model.

The use of White Zone tactics by states towards civilians is also not associated with Black Zone activity by VNSAs. However, the use of Gray Zone tactics by states towards civilians is weakly associated with higher Black Zone activity by VNSAs. This finding is statistically significant at the 99 percent confidence level but only robust in two out of four models. (Tables 1 and H.1 present statistically significant results.) The use of Black Zone tactics by states against civilians is strongly associated with higher levels of Black Zone

activity by VNSAs overall – the results are statistically significant at the 99 percent confidence level and robust across all models.

Kinetic and Temporal Variables

The use of kinetic tactics is positively correlated with greater White, Gray and Black activity, and the results are statistically significant at the 99 percent confidence level for all models in Table 1 as well as alternative specifications in Tables H.1-H.6.

Finally, the lagged dependent variable is statistically significant at the 99 percent confidence level across all models, suggesting that time also shapes Zonal activity, which is unsurprising.

State White Zone Preferences

Table 2, below, illustrates the effect of state, civilian and VNSA activities on White Zone activity by (other) states.

Table 2: State White Zone Preferences in Colombia

Variable	Coefficient	SE
VNSA-State White	0.97***	0.03
VNSA-State Gray	-0.19***	0.03
VNSA-State Black	-1.45***	0.04
Civilian-State White	0.30***	0.02
Civilian-State Gray	0.09***	0.02
State-State White	0.35***	0.01
State-State Gray	2.25***	0.06
State-State Black	88.02***	3.64
VNSA-Civilian White	-0.15***	0.04
VNSA-Civilian Gray	0.13***	0.02
VNSA-Civilian Black	-0.98***	0.09
Kinetic	6.92***	0.06
Lagged DV	0.31***	0.00
Constant	0.47***	0.09
Number of Obs.	106128	
Number of Groups	603	

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed Tests). Reported Standard Errors clustered by country, using a Random Effects GLS Regression model. This model presents findings at the second order subnational administrative level.

VNSA-State Variables

As the table above shows, the use of White, Gray and Black tactics by VNSAs has a strong impact on the use of White Zone actions by states. The use of White Zone activities by VNSAs results in higher levels of

state White Zone activity. These results are statistically significant at the 99 percent confidence level and robust across all models in Table 1 and Tables H.1-H.6. The use of Gray and Black tactics by VNSAs results in lower levels of White Zone activity by states. Again, these results are statistically significant at the 99 percent confidence level and robust across all models in Table 1 and Tables H.1-H.6.

Civilian-State Variables

The use of White and Gray tactics by civilians against states are not a robust predictor of White Zone activity by states. The variables are positive and statistically significant in the basic model in Table 2, but the sign on these variables changes across different models. The use of White tactics by civilians towards states is positively and statistically significantly associated with state White Zone activity in Tables 2 and H.4, and negative and statistically significant in Tables H.2 and H.6. In short, the results are dependent on model choice between fixed and random effects. The use of Gray Zone tactics by civilians is positively associated with higher White Zone activity by states in only two out of four models. In other words, we do not find convincing evidence that civilians are statistically significant drivers of state White Zone activity.

State-State Variables

The Zonal choices of states against other states is also highly consequential. The use of White, Gray and Black Zone tactics by states against others states are positively correlated with greater White Zone activity by states. The results are statistically significant at the 99 percent confidence level and robust across all models.

VNSA-Civilian Variables

The use of White tactics by VNSAs towards civilians is not a robust predictor of White Zone activity by states – the results are negative and statistically significant for random effects models (Tables 2 and H.4), and positive and statistically significant in one fixed effects model (Table H.2). The use of Gray tactics by VNSAs towards civilians is weakly associated with higher levels of White Zone activity by states. Specifically, the results are positive and statistically significant at the 99 percent confidence level in two of four models (Tables 2 and H.2). The use of Black tactics by VNSAs towards civilians is strongly correlated with lower overall levels of White Zone activity by states. These findings are statistically significant at the 99 percent confidence level and robust across all models.

Kinetic and Temporal Variables

On the other hand, the use of Kinetic action is positively correlated with higher White Zone activity by states. These results are robust across all models and statistically significant at the 99 percent confidence level.

Finally, the lagged dependent variable is statistically significant at the 99 percent confidence level across all models, suggesting that time also shapes Zonal activity. Again, this is unsurprising.

Bayesian Analysis Results

Analyses of the Colombian dataset using the Bayesian relationship described in Equation 1, above, revealed that the model is able to correctly predict the occurrence of Gray Zone events approximately 79 percent of the time overall. The model is also able to correctly predict when events were not Gray approximately 63 percent of the time. More specifically, the model predicted the occurrence of Gray Zone events approximately 14 percent of the time when P(G|S,T,K) was 20 percent and below. When P(G|S,T,K) was between 21 and 40 percent, the model was able to accurately predict the outcome approximately 11 percent of the time. The model showed increased predictive power when P(G|S,T,K) was between 41 and 50 percent, accurately predicting the outcome approximately 23 percent of the time. However, the model's predictive power increased dramatically when P(G|S,T,K) was 51 percent and above, accurately predicting the outcome approximately 78 percent of the time when P(G|S,T,K) was between 51 and 60 percent, and approximately 87 percent of the time when P(G|S,T,K) was 61 percent and above. Table 3, below, provides a detailed breakdown of the frequencies for which the Bayesian model correctly predicted Gray Zone events for a given range of probabilities of a Gray Zone event (P(G|S,T,K)). We also calculated the probability of an event being Black or White using the same BBN model to examine the model's ability to accurately predict the occurrence of Black and White Zone events. The results show that the model was able to accurately predict the occurrence of a Black Zone event approximately 53 percent of the time overall and a White Zone event approximately 76 percent of the time overall. Table 3 also explores these results in more detail.

Table 3: Bayesian Belief Network Model Prediction Accuracies by Zone

Gray Zone Events		Black Zone Events		White Zone Events	
P(G|S,T,K)	Prediction Accuracy	P(B|S,T,K)	Prediction Accuracy	P(W|S,T,K)	Prediction Accuracy
Overall	79.37%	Overall	53.40%	Overall	76.45%
Non-Gray Zone	63.21%	Non-Black Zone	46.19%	Non-White Zone	16.62%
1% - 20%	14.53%	1% - 20%	3.33%	1% - 20%	84.89%
21% - 40%	10.99%	21% - 40%	35.27%	21% - 40%	80.85%
41% - 50%	23.28%	41% - 50%	n/a[12]	41% - 50%	98.37%
51% - 60%	78.05%	51% - 60%	n/a	51% - 60%	n/a
61% +	87.42%	61% +	n/a	61% +	n/a

In addition to calculating the probability of Gray Zone events when the Source, Target and Use of Kinetic Force are known: P(G|S,T,K), the level of influence that each conditionally independent parent node – P(G|S), P(G|T) and P(G|K) – has on the child node – P(G|S,T,K) – was also computed. This analysis revealed that the Use of Kinetic Force has the highest positive impact, increasing the probability that the event is a Gray Zone event by approximately 31 percent. The analysis also revealed that when the Source of the event is a VNSA, the probability that the event is a Gray Zone event increased by approximately 30 percent while the probability of an event being Gray decreased by approximately 32 percent if states is the Source. Table 4, below, provides a detailed accounting of each node's influence on the likelihood that an event is a Gray Zone event.

Table 4: Nodal Influence on the Probability of Events by Zone

Node	Gray Probability	Black Probability	White Probability
Source is a State	-32.52%[13]	3.60%	29.32%
Source is VNSA	30.24%	-0.52%	-33.06%
Source is Civilian	-1.86%	-15.16%	10.20%
Target is State	-5.67%	-9.09%	14.01%
Target is VNSA	-17.21%	8.80%	3.88%
Target is Civilian	18.40%	-0.10%	-19.42%
Use of Kinetic Force	31.03%	26.45%	-23.62%

[12] N/A for prediction accuracy denotes that no data at that level of probability was contained in the dataset.
[13] For each of the probabilities of the parent node – P(G|S), P(G|T) and P(G|K) – a negative probability can be interpreted as equivalent to 0. This is the case as a negative probability that the child node would occur given the parent node essentially means that parent node has no influence on the child node. Therefore, a negative probability for the parent node can be discarded in a two-node structure (one parent; one child). However, in a multi-node convergence structure where there are several conditionally independent parent nodes all influencing a single child node, the negative probability of a parent node becomes significant and should be included as negative value in the final calculation to ascertain the probability for the child node. This allows for increased accuracy by accounting for the relationships among the parent nodes and ensures that the calculated probability for the child node is not artificially inflated.

Ukraine

VNSA Zonal Preferences

Table 5, below, presents the results for the models testing the impact of state, civilian and VNSA activities on the Zonal behavior of (other) VNSAs. Model 1 examines White Zone activity by VNSAs, Model 2 estimates Gray Zone activity by VNSAs and Model 3 looks at how different factors shape Black Zone activity by VNSAs.

Table 5: Ukranian VNSA Zonal Preferences

Variable	Model 1: White		Model 2: Gray		Model 3: Black	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
State-VNSA White	0.05***	0.00	0.09***	0.01	0.22***	0.01
State-VNSA Gray	0.14***	0.01	0.38***	0.02	0.04**	0.02
State-VNSA Black	-0.01***	0.00	0.00	0.00	0.02***	0.00
Civilian-VNSA White	-0.19***	0.03	-0.34***	0.05	0.56***	0.06
Civilian-VNSA Gray	-0.22	0.44	14.92***	0.83	20.16***	1.03
VNSA-VNSA White	0.16***	0.05	0.25***	0.08	1.93***	0.10
VNSA-VNSA Gray	-0.01	0.02	-0.12***	0.04	-0.09*	0.05
VNSA-VNSA Black	0.07**	0.03	0.06	0.05	-0.34***	0.06
State-Civilian White	-0.02***	0.00	0.01*	0.01	-0.05***	0.01
State-Civilian Gray	-0.01**	0.00	-0.04***	0.01	0.02**	0.01
State-Civilian Black	-0.09***	0.01	-0.13***	0.03	-0.30***	0.03
Kinetic	0.24***	0.01	0.37***	0.02	0.02	0.02
Lagged DV	-0.02*	0.01	0.11***	0.01	0.27***	0.01
Constant	0.06**	0.03	0.14***	0.05	0.21***	0.06
Number of Obs.			13664			
Number of Groups			427			

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Standard Errors clustered by country, using a Random Effects GLS Regression model. This model presents findings from the Ukraine dataset at the second order subnational administrative level.

State-VNSA Variables

As the table above shows, the use of White Zone tactics by states towards VNSAs is associated with higher levels of White Zone activity by VNSAs (Model 1). The use of Gray Zone tactics by states towards VNSAs is also correlated with higher White Zone activity by VNSAs. The use of Black Zone tactics by states is negatively correlated with White Zone activity by VNSAs. These three effects are all statistically significant at the 99 percent confidence level and robust across all alternative model specifications presented in Tables 5, H.7, H.9 and H.10.

The use of White Zone tactics by states towards VNSAs does not appear to be associated with higher levels of Gray Zone tactics by VNSAs. While, two of the four models (see Tables 5 and H.7) returned positive and statistically significant coefficients, the results are not robust across the other two model

specifications. In one, statistically significant results were not observed (see Table H.9). In the other, the results are statistically significant; however, the coefficient has reversed sign and is now negative (see Table H.10). The use of Gray Zone tactics by states towards VSNAs increases the use of Gray Zone tactics by VNSAs, and the finding is statistically significant at the 99 percent confidence level. It is also robust across all alternative model specifications. The use of Black Zone tactics by states towards VNSAs appears to reduce the use of Gray Zone tactics by VNSAs. This effect is statistically significant at the 99 percent confidence level and is robust in three out of four models. In the fourth specification, the coefficient fails to achieve statistical significance (see Table 5).

The use of White Zone tactics by states towards VNSAs is also associated with higher levels of Black Zone activity by VNSAs, and the results are statistically significant at the 99 percent confidence level. They are also robust across all models. The relationship between state use of Gray and Black Zone and the VNSA responses is not clear. Two models show that the use of Gray Zone tactics by states against VNSAs reduces the use of Black Zone tactics by VNSAs. However, the results reverse sign in a third model (see Table 5) and fail to achieve statistical significance in a fourth model (see Table H.9). Consequently, this result is model dependent to the use of fixed versus random effects. Similarly, the impact of the use of Black Zone tactics by states against VNSAs on VNSA use of Black is unclear. In the two random effects models, it is positive; and in the two fixed effects models, it is negative.

Civilian-VNSA Variables

The use of White Zone tactics by civilians towards VNSAs results in lower levels of White Zone activity by VNSAs, and the result is statistically significant at the 99 percent confidence level. It is also robust across all model specifications. There is also weak evidence to suggest that the use of Gray Zone tactics by civilians towards VNSAs results in lower White Zone activity by VNSAs. However, this finding is only observed when the data is aggregated up to the first order subnational administrative boundary. In these two models (presented in Tables H.9 and H.10), the results are statistically significant at the 99 percent confidence level.

The use of White Zone tactics by civilians towards VNSAs results in lower levels of Gray Zone activity by VNSAs, and this result is statistically significant at the 99 percent confidence level. It is also robust across all model specifications. On the other hand, the use of Gray Zone tactics by civilians towards VNSAs results in higher levels of Gray Zone tactics by VNSAs, and this finding is statistically significant at the 99 percent confidence level. The finding is also robust across all models specifications. Importantly, the effect is substantively quite large (discussed in detail below).

The use of White Zone tactics by civilians towards VNSAs results in higher levels of Black Zone activity by VNSAs, and the results are statistically significant at the 99 percent confidence level across all model specifications. The use of Gray Zone tactics by civilians towards VNSAs are strongly associated with a sharp increase in the use of Black Zone tactics by VNSAs. These results are also statistically significant at the 99 percent confidence level and robust across all models.

VNSA-VNSA Variables

The use of White Zone tactics by VNSAs towards other VNSAs has a positive effect on overall use of White by VNSAs. This result is statistically significant at the 99 percent confidence level and robust across all models. However, we did not find a robust relationship between the use of Gray Zone tactics by VNSAs against other VNSAs and the use of White activities by VNSAs. Specifically, we failed to return statistically significant results in two of the model specifications. In a third specification, the coefficient is negative, but only statistically significant at the 90 percent confidence level (see Table H.7). In the final model, the results are positive and statistically significant at the 95 percent confidence level (see Table H.9). We also failed to find substantial evidence of a relationship between the use of Black tactics between VNSAs and higher overall usage of White activities by VNSAs. Specifically, we returned statistically significant results only when the data was aggregated to the second order subnational administrative level in the models presented in Tables 5 and H.7.

The use of White Zone tactics between VNSAs also does not have a robust relationship with the use of Gray Zone tactics by VNSAs. Again, two models fail to return statistically significant results. In the remaining two models, one returns a positive and statistically significant coefficient, and the other results in a negative and statistically significant coefficient (see Tables 5 and H.10, respectively). On the other hand, the use of Gray Zone tactics between VNSAs decreases overall Gray Zone activity by VNSAs, and the results are statistically significant at the 99 percent confidence level. This finding is robust across all models. The use of Black Zone tactics between VNSAs has a negative relationship with Gray Zone activity by VNSAs, but only when the data is aggregated up to the first order subnational administrative level (see Tables H.9 and H.10).

The use of White Zone tactics between VNSAs increases the use of Black Zone tactics by VNSAs, and the results are statistically significant at the 99 percent confidence level and robust across all models. Evidence also suggests that the use of Gray Zone tactics between VNSAs results in lower Black Zone activity overall. This latter finding is robust across three out of four models; the model presented in Table H.10 fails to return statistically significant results. However, in two of the three models in which it obtains statistical significance the findings are only statistically significant at the 90 percent confidence interval (see Tables 5 and H.9). In the fourth model, the results are statistically significant at the 95 percent confidence level (see Table H.7). The use of Black Zone tactics between VNSAs decreases the overall use of Black Zone tactics by VNSA. This finding is statistically significant at the 99 percent confidence level and robust across all models.

State-Civilian Variables

When the data is disaggregated to the second order subnational administrative level, our models suggest that the use of White Zone tactics by states towards civilians also lowers the use of White Zone tactics by VNSAs. However, these results are not robust to the analysis of data aggregated up to the first order subnational administrative level. The relationship between the use of Gray Zone tactics by states against civilians and White Zone activity by VNSAs is also unclear. The results are model dependent – they are negative and statistically significant in random effect models (Tables 5 and H.9) and insignificant in fixed

effects models (Tables H.7 and H.10). The use of Black Zone tactics by states towards civilians results in lower levels of White Zone activity by VNSAs. This finding is statistically significant at the 99 percent confidence level and robust across all models.

The use of White Zone tactics by states towards civilians does not have a robust effect on Gray Zone tactics by VNSAs. Specifically, the pertinent coefficient fails to achieve statistical significance in two of the models. In the model presented in Table 5, the results are positive and statistically significant. However, they are negative and statistically significant in the model presented in Table H.7. The use of Gray Zone tactics by states towards civilians results in lower levels of Gray Zone activity by VNSAs. This finding is statistically significant at the 99 percent confidence level and robust across all models. Evidence also suggests that the use of Black Zone tactics by states against civilians results in lower levels of Gray Zone activity by VNSAs. The results are statistically significant at the 99 percent confidence level across three of the four models (excluding the model presented in Table H.9).

The use of White Zone tactics by states towards civilians results in lower levels of Black Zone tactics by VNSAs. This result is statistically significant at the 99 percent confidence level and robust across all models. The relationship between the use of Gray Zone tactics by states against civilians and Black Zone activity by VNSAs is not clear. The results are model dependent. They are positive and statistically significant when the data is aggregated to the second order subnational administrative level, but negative and statistically significant when the data is aggregated up to the first order subnational administrative level. The use of Black Zone tactics by states against civilians results in lower levels of Black Zone activity by VNSAs. This finding is statistically significant at the 99 percent confidence level and robust across all models.

Kinetic and Temporal Variables

Kinetic actions have a positive and robust effect on the use of White and Gray Zone tactics by VNSAs. This finding is statistically significant at the 99 percent confidence level and robust across all models. The effect of kinetic activity on Black Zone tactics by VNSAs is unclear. Two of the models show a positive and statistically significant relationship. However, a third model fails to return statistically significant results (see Table 5), and the final model returns negative and statistically significant results (see Table H.7).

The lagged dependent variable is statistically significant (at the 99 percent confidence level in two models, the 90 percent confidence level in the model presented in Table 5, and at the 95 percent confidence level in the model presented in Table H.9). This suggests that time also shapes Zonal activity, which is to be expected.

State White Zone Preferences

Table 6, below, illustrates the effect of state, civilian and VNSA activities on White Zone activity by (other) states.

Table 6: State White Zone Preferences in Ukraine

Variable	Coefficient	SE
VNSA-State White	1.21***	0.12
VNSA-State Gray	-0.14**	0.07
VNSA-State Black	1.08***	0.04
Civilian-State White	0.48***	0.02
Civilian-State Gray	-0.25***	0.03
State-State White	-0.07**	0.03
State-State Gray	-0.76***	0.09
State-State Black	-1.04***	0.15
VNSA-Civilian White	-2.80***	0.29
VNSA-Civilian Gray	-0.52***	0.08
VNSA-Civilian Black	-3.08***	0.35
Kinetic	0.34***	0.12
Lagged DV	0.77***	0.03
Constant	0.08***	0.26
Number of Obs.	13664	
Number of Groups		

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Random Effects GLS Regression model. This model presents findings at the second order subnational administrative level.

VNSA-State Variables

The use of White Zone activity by VNSAs towards states results in higher state use of White Zone tactics, and the results are statistically significant at the 99 percent confidence level. This finding is robust across all models in Tables 6, H.8, H.11 and H.12. The use of Gray Zone tactics by VNSAs does not have a robust relationship with White Zone usage by states. The results are positive and statistically significant for the fixed effects models (Tables H.8 and H.12). However, the model presented in Table 6 returns negative and statistically significant results, while the model in Table H.11 fails to return statistically significant results. The use of Black Zone tactics by VNSAs results in higher use of White Zone approaches by states. This result is statistically significant at the 99 percent confidence level and robust across all models.

Civilian-State Variables

The use of White Zone tactics by civilians towards states is correlated with higher overall White Zone activity by states, and the result is statistically significant at the 99 percent confidence level and robust across all models. On the other hand, the use of Gray Zone tactics by civilians towards states is correlated with lower White Zone activity by states. Again, this result is statistically significant at the 99 percent confidence level and robust across all models.

State-State Variables

There is no robust relationship between the use of White Zone activity between states, and overall White Zone activity by states. The random effects models return negative and statistically significant results (Tables 6 and H.11), while the fixed effects models are positive and statistically significant (Table H.8) or fail to achieve statistical significance (Table H.12). However, the use of Gray or Black Zone activity by states against other states results in lower overall White Zone activity by states. The results are statistically significant at the 99 percent confidence level and robust across all models.

VNSA-Civilian Variables

The use of tactics by VNSAs across all three Zones (White, Gray or Black) towards civilians is correlated with lower White Zone activity by states. This finding is statistically significant at the 99 percent confidence level and robust across all models.

Kinetic and Temporal Variables

Kinetic events are positively correlated with higher White Zone usage by states. The results are statistically significantly at the 99 percent confidence level in three out of four models; statistically significant results are not returned in the model presented in Table H.11. Finally, the lagged dependent variable is statistically significant at the 99 percent confidence level, suggesting that time also shapes Zonal activity. Again, this is unsurprising.

Bayesian Analysis Results

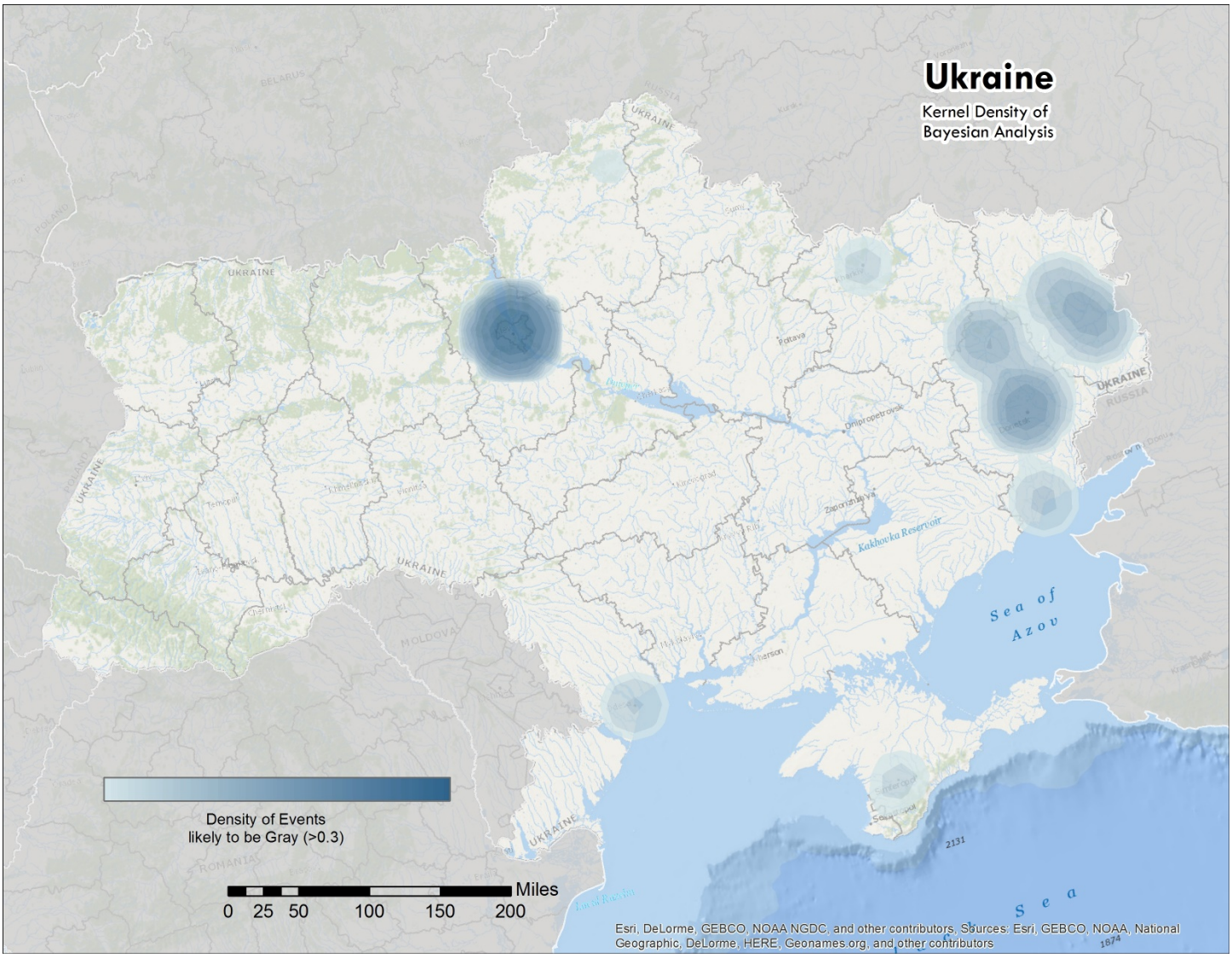
Analyses of the Ukraine dataset using the Bayesian relationship described in Equation 1, above, revealed that the model can accurately predict the occurrence of Gray Zone events approximately 75 percent of the time overall. For predicting non-Gray Zone events, the model was accurate approximately 67 percent of the time. Specifically, when P(G|S,T,K) was 20 percent or below, the prediction rate was approximately 10 percent. The prediction rate improved markedly to approximately 65 percent when P(G|S,T,K) was between 21 and 40 percent. Further improvement – to approximately 85 percent – occurred when P(G|S,T,K) was between 41 and 50 percent. The results also show that the model can accurately predict the occurrences of Black and White Zone events approximately 85 percent and 87 percent of the time, respectively, overall. Table 7, below, provides a detailed percentage breakdown for correctly predicted Gray Zone events for a given range of probability of Gray – P(G|S,T,K), Black – P(B|S,T,K) and White – P(W|S,T,K) Zone events.

Table 7: Bayesian Belief Network Model Prediction Accuracies by Zone in Ukraine

Gray Zone Events		Black Zone Events		White Zone Events	
P(G|S,T,K)	Prediction Accuracy	P(B|S,T,K)	Prediction Accuracy	P(W|S,T,K)	Prediction Accuracy
Overall	75.47%	Overall	85.18%	Overall	86.50%
Non-Gray Zone	67.46%	Non-Black Zone	75.22%	Non-White Zone	17.95%
1% - 20%	10.31%	1% - 20%	28.20%	1% - 20%	89.77%
21% - 40%	64.58%	21% - 40%	41.53%	21% - 40%	95.25%
41% - 50%	84.61%	41% - 50%	96.70%	41% - 50%	n/a
51% - 60%	n/a	51% - 60%	38.71%	51% - 60%	n/a
60% +	n/a	60% +	n/a	60% +	n/a

Figure 5, below, is a kernel density plot derived from the Bayesian analysis. It visualizes the density of predicted Gray Zone events, when P(G|S, T, K) is 30 percent or higher. Interestingly, the highest density of predicted Gray actions occur in the capital, Kiev (center left), with relatively fewer Gray events predicted in the Donbas (right) and even fewer in Crimea (bottom right). This makes sense, since coercive diplomatic activities, pernicious statements, etc., that tend to be Gray mostly originate from Kiev and target Moscow or vice versa. While the Donbas involves extensive Gray Zone activity, the conflict there has often escalated into the Black Zone. The paucity of Gray events in Crimea is likely a result of limited data due to the speed with which that conflict ended.

Figure 5, Kernel Density Plot of Predicted Gray Events



Analyses of the conditionally independent parent nodes' influence on the child node for Ukraine revealed that Use of Kinetic Force (or lack thereof) has the highest impact on the probability of an event being Gray, Black or White (21%). Table 8, below, provides a detailed account of each node's influence on the likelihood that an event occurs within a particular Zone.

Table 8: Each Node's Influence on Probability of Event by Zone in Ukraine

Node	Gray Probability	Black Probability	White Probability
Source is a State	-19.20%	0.59%	22.65%
Source is VNSA	20.29%	11.70%	-36.08%
Source is Civilian	14.34%	-11.51%	-9.67%
Target is State	-4.12%	-10.21%	14.90%
Target is VNSA	-0.30%	18.97%	-25.67%
Target is Civilian	5.39%	-1.52%	-4.30%
Use of Kinetic Force	21.33%	21.76%	-54.22%

Libya

VNSA Zonal Preferences

Table 9, below, presents the results for the models testing the impact of civilian and VNSA activities on the Zonal behavior of (other) VNSAs. Model 1 examines White Zone activity by VNSAs, Model 2 explores Gray Zone activity by VNSAs, and Model 3 looks at how different factors shape Black Zone activity by VNSAs.

Table 9: Libyan VNSA Zonal Preferences

Variable	White		Gray		Black	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
VNSA-VNSA White	-1.31***	0.24	0.10***	0.01	0.06***	0.02
VNSA-VNSA Gray	0.55***	0.18	0.38***	0.11	0.29***	0.10
VNSA-VNSA Black	0.59***	0.20	-0.20**	0.09	-1.35***	0.23
Civilian-VNSA White	0.31**	0.16	0.09	0.06	-0.04	0.07
Civilian-VNSA Gray	0.52*	0.30	-0.44***	0.13	-0.28*	0.14
Civilian-VNSA Black	-0.45	0.51	-0.54**	0.22	0.19	0.25
Civilian-Civilian White	4.14***	0.82	0.15	0.34	0.74*	0.42
Civilian-Civilian Gray	8.76***	1.75	1.28*	0.72	4.84***	0.86
Civilian-Civilian Black	-1.43	3.79	0.80	1.57	1.74	1.87
Kinetic	-0.72***	0.17	0.14*	0.08	-0.02	0.09
Lagged DV	1.91***	0.22	0.12	0.09	1.42***	0.23
Constant	1.22***	0.46	0.72***	0.19	0.66***	0.23
Number of Obs.			1496			
Number of Groups			22			

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Standard Errors clustered by country, using a Random Effects GLS Regression model. This model presents findings from the Libya dataset at the first order subnational administrative level.

VNSA-VNSA Variables

As the table above shows, the use of White Zone tactics by VNSAs towards other VNSAs results in lower overall White Zone activity by VNSAs. These results are statistically significant at the 99 percent

confidence level and robust across both models in Tables 9 and H.13. Table 9 also suggests that the use of Gray Zone tactics between VNSAs results in higher overall White Zone activity. While this result is positive and statistically significant at the 99 percent confidence interval in the random effects model (Table 9), the results from the fixed effects model (Table H.13) fail to achieve statistical significance. In short, this finding appears to be model dependent. The use of Black Zone tactics by VNSAs against other VNSAs results in higher overall White Zone activity by VNSAs. These findings are statistically significant at the 99 percent confidence level and robust across both models.

There is weak evidence to suggest that the use of White Zone tactics by VNSAs against other VNSAs results in higher overall Gray Zone activity. This result is positive and statistically significant at the 99 percent confidence level in the random effects model (Table 9) but fails to achieve statistical significance in the fixed effects model (Table H.13). The use of Gray Zone tactics between VNSAs results in higher overall Gray Zone activity by VNSAs. These findings are statistically significant at the 99 percent confidence level and robust across all models. The use of Black Zone tactics between VNSAs results in lower overall Gray Zone activity by VNSAs. These results are statistically significant at the 95 percent confidence level and robust across all models.

The use of White Zone tactics by VNSAs towards other VNSAs results in higher overall levels of Black Zone activity. Similarly, the use of Gray Zone tactics between VNSAs results in higher overall levels of Black Zone activity by VNSAs. Finally, the use of Black Zone tactics between VNSAs result in lower overall Black Zone activity by VNSAs. All three sets of results are statistically significant at the 99 percent confidence level and robust across all models.

Civilian-VNSA Variables

The use of White tactics by civilians towards VNSAs results in higher overall White activity by VNSAs. These results are statistically significant at the 95 percent confidence level and robust across all models. The use of Gray or Black tactics by civilians towards VNSAs do not seem to have a statistically significant impact on White Zone activity by VNSAs.

The use of White tactics by civilians towards VNSAs does not have a statistically significant impact on the use of Gray Zone tactics by VNSAs. The use of Gray Zone tactics by civilians towards VNSAs reduces overall Gray Zone behavior by VNSAs. These findings are statistically significant at the 99 percent confidence level and robust across all models. The use of Black Zone tactics by civilians towards VNSAs reduces overall Gray Zone behavior by VNSAs. These results are statistically significant at the 95 percent confidence level and robust across all models.

The use of White and Black Zone tactics by civilians towards VNSAs do not demonstrate a statistically significant impact on Black Zone activity by VNSAs. However, the use of Gray Zone approaches by civilians towards VNSAs reduces overall Black Zone behavior by VNSAs. These findings are only statistically significant at the 90 percent confidence level. Nevertheless, they are robust across both models.

Civilian-Civilian Variables

The use of White Zone tactics by civilians towards other civilians increases overall White Zone activity by VNSAs. These results are statistically significant at the 99 percent confidence level and robust across all models. The use of Gray Zone activities by civilians towards other civilians also increases overall White Zone activity by VNSAs. Again, these findings are statistically significant at the 99 percent confidence level and robust across all models. The use of Black Zone tactics by civilians against other civilians does not have a statistically significant impact on White Zone activity by VNSAs in our models.

The use of White and Black Zone tactics by civilians towards other civilians do not have a statistically significant effect on Gray Zone activity by VNSAs in our analyses. There is weak evidence to suggest that the use of Gray Zone tactics by civilians against other civilians results in higher overall Gray Zone activity by VNSAs. Specifically, the results are positive and statistically significant at the 90 percent confidence level in the random effects model (Table 9), but the fixed effects model (Table H.13) fails to return statistically significant results.

The use of White Zone tactics by civilians towards other civilians increases overall Black Zone activity by VNSAs. These results are statistically significant at the 90 percent confidence level and robust across all models. The use of Gray Zone tactics by civilians towards other civilians also increases overall Black Zone activity by VNSAs. These findings are statistically significant at the 99 percent confidence level and robust across all models. The use of Black Zone tactics by civilians against other civilians does not have a statistically significant impact on overall Black Zone activity by VNSAs in our models.

Kinetic and Temporal Variables

The use of kinetic tactics results in lower overall White Zone activity by VNSAs. These findings are statistically significant at the 99 percent confidence level and robust across all models. The use of kinetic approaches also results in higher overall Gray Zone activity by VNSAs. These results are statistically significant at the 90 percent confidence interval for the random effects model (Table 9) and at the 99 percent confidence level for the fixed effects model (Table H.13). The use of kinetic tactics does not appear to have a statistically significant effect on overall Black Zone activity by VNSAs in either model.

The lagged dependent variable is statistically significant at the 99 percent confidence level in both versions of the White and Black Zone models (Models 1 and 3). This suggests that temporal dynamics shape White and Black Zone behavior by VNSAs. However, the same variable fails to achieve statistical significance in the models for Gray Zone activity (Model 2). This appears to indicate that temporal dynamics do not affect Gray Zone behavior by VNSAs.

Civilian Gray Zone Preferences

Given our inability to distinguish between states and VNSAs, this subsection instead examines civilian Gray Zone preferences. Table 10, below, illustrates the effect of civilian and VNSA activities on the Gray Zone behavior of (other) civilian actors.

Table 10: Libyan Civilian Gray Zone Preferences

Variable	Coefficient	SE
VNSA-Civilian White	0.26***	0.05
VNSA-Civilian Gray	-0.04	0.05
VNSA-Civilian Black	-0.17*	0.10
VNSA-VNSA White	0.02**	0.01
VNSA-VNSA Gray	0.11**	0.05
VNSA-VNSA Black	-0.06	0.05
Civilian-Civilian White	-0.12	0.17
Civilian-Civilian Gray	1.07***	0.38
Civilian-Civilian Black	-0.15	0.82
Kinetic	0.00	0.04
Lagged DV	-0.03	0.06
Constant	0.33***	0.10
Number of Obs.	1496	
Number of Groups	22	

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Random Effects GLS Regression model. This model presents findings from the Libya dataset at the first order subnational administrative level.

VNSA-Civilian Variables

The use of White Zone tactics by VNSAs towards civilians results in higher overall Gray Zone behavior by civilians. These results are statistically significant at the 99 percent confidence level and robust across all models. However, the use of Gray Zone tactics by VNSAs on civilians does not have a statistically significant impact on Gray Zone behavior by civilians in our models. Finally, there is weak evidence to suggest that Black Zone tactics employed by VNSAs against civilians decrease Gray Zone activity by civilians. Specifically, the random effects model (Table 10) returns a negative coefficient that is statistically significant at the 90 percent confidence level. However, the results from the fixed effects model (Table H.14) fail to obtain statistical significance.

There is weak evidence to suggest that the use of White Zone tactics by VNSAs towards other VNSAs increases overall Gray Zone activity by civilians. The results from the random effects model (Table 10) are positive and statistically significant at the 95 percent confidence level. However, the findings from the fixed effects model (Table H.14) do not achieve statistical significance. There is also weak evidence to suggest that the use of Gray Zone tactics by VNSAs towards other VNSAs increases overall Gray Zone activity by civilians. The findings from the random effects model (Table 10) are positive and statistically significant at the 95 percent confidence level but fail to achieve statistical significance in the fixed effects model (Table H.14). Finally, there is weak evidence to suggest that the use of Black Zone tactics by VNSAs towards other VNSAs decreases overall Gray Zone activity by civilians. The fixed effects model (Table

H.14) returns a negative coefficient that is statistically significant at the 90 percent confidence level. However, the random effects model (Table 10) does not yield statistically significant results.

Civilian-Civilian Variables

The use of White and Black Zone tactics by civilians towards other civilians does not demonstrate a statistically significant impact on Gray Zone activity by civilians in our models. However, the use of Gray Zone tactics by civilians against other civilians increases overall Gray Zone activity by civilians. These results are statistically significant at the 99 percent confidence interval for the random effects model (Table 10) and at the 95 percent confidence interval in the fixed effects model (Table H.14).

Kinetic and Temporal Variables

Kinetic events do not appear to have a statistically significant impact on the Gray Zone behavior of civilians.

Finally, the lagged dependent variable fails to achieve statistical significance across both models. This suggests that temporal dynamics do not shape Gray Zone behavior by civilians.

Discussion

Colombia

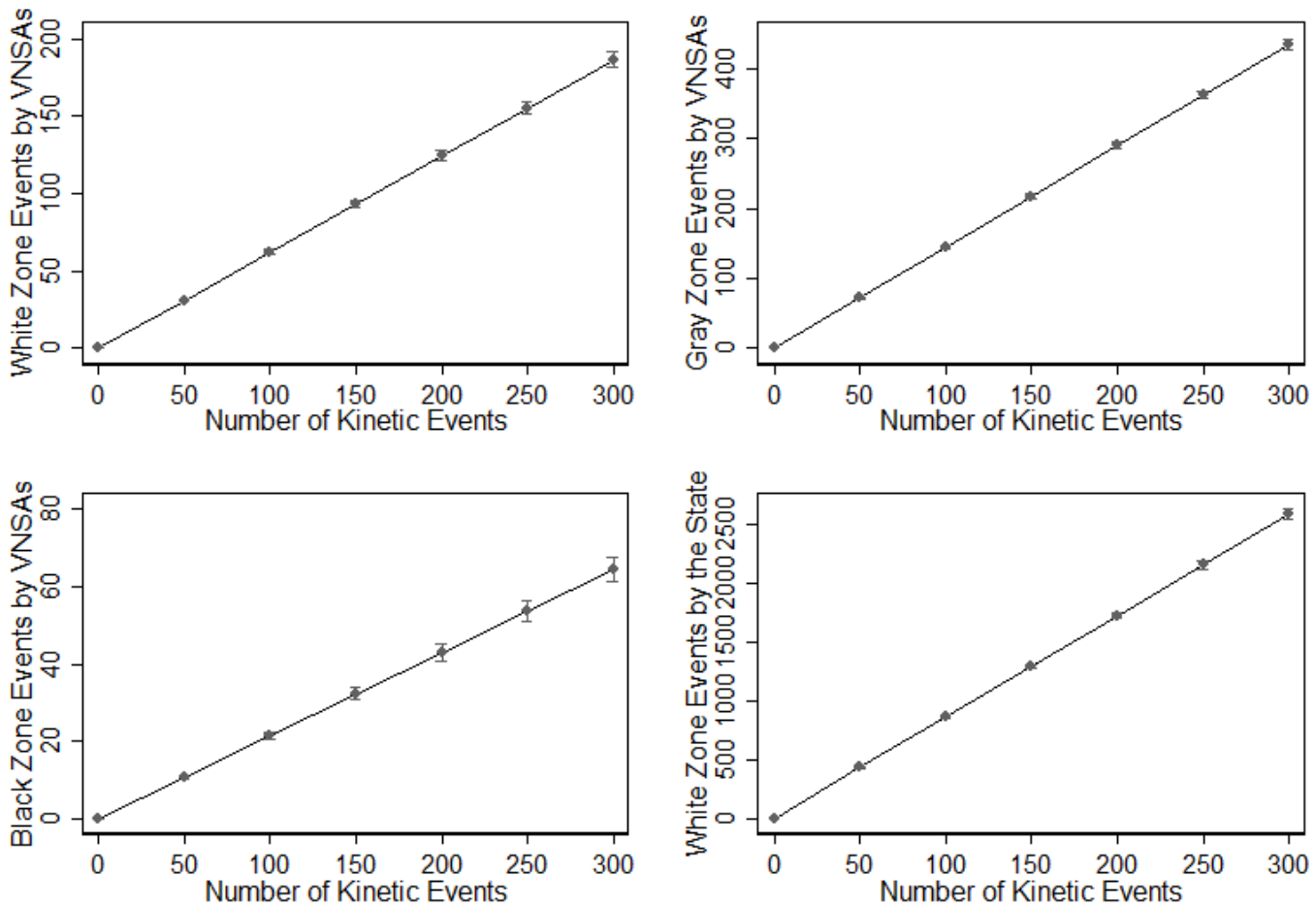
While the previous section discussed all of the findings, in this section, we focus in on key findings for detailed, substantive discussion. Specifically, we discuss four trends observed in the data. First, we found that kinetic activities are highly consequential. Unsurprisingly, they are associated with higher levels of Black activity, but they also seem to be driving greater levels of White and Gray Zone activity, including non-kinetic White and Gray Zone activity. Second, our analysis suggests that greater interaction among VNSAs drives higher levels of Gray Zone activity by VNSAs. In other words, the proliferation of VNSAs seems to results in higher overall levels of Gray Zone activity, relative to other Zones. Third, we found some evidence of signaling asymmetries between states and VNSAs. Specifically, we found that states are better at discerning between White and Gray Zone activities by VNSAs than VNSAs are at distinguishing between White and Gray Zone activity by states. Finally, we found that VNSAs respond strategically to state behavior. They de-escalate when states use Gray tactics, and escalate when states use Black. Taken together, the evidence suggests that even though VNSAs may struggle to distinguish between states' use of White and Gray, VNSAs nevertheless are able to distinguish between Gray and Black. Consequently, they mirror states response in order to improve their own legitimacy or rally civilian support.

The Importance of Kinetic Events

We find that kinetic activities are highly consequential. Specifically, kinetic events increase the likelihood of White, Gray and Black activities that are both kinetic and non-kinetic in nature by both states and VNSAs across all models. Substantively, our models predict that on average in Colombia, increasing kinetic events by 100 in one month in a given municipality is associated with an increase of 54 White Zone events, 139 Gray Zone events and 23 Black Zone events by VNSAs in that municipality during the

subsequent month. In terms of states' White Zone activity, increasing kinetic events by 100 in one month in a particular municipality is associated with an average increase of 692 White Zone events in that municipality during the subsequent month by the Colombian state. Figure 6, below, plots how different types of Zones change as the number of kinetic events rise.

Figure 6: Kinetic Events and Zonal Activity[14]



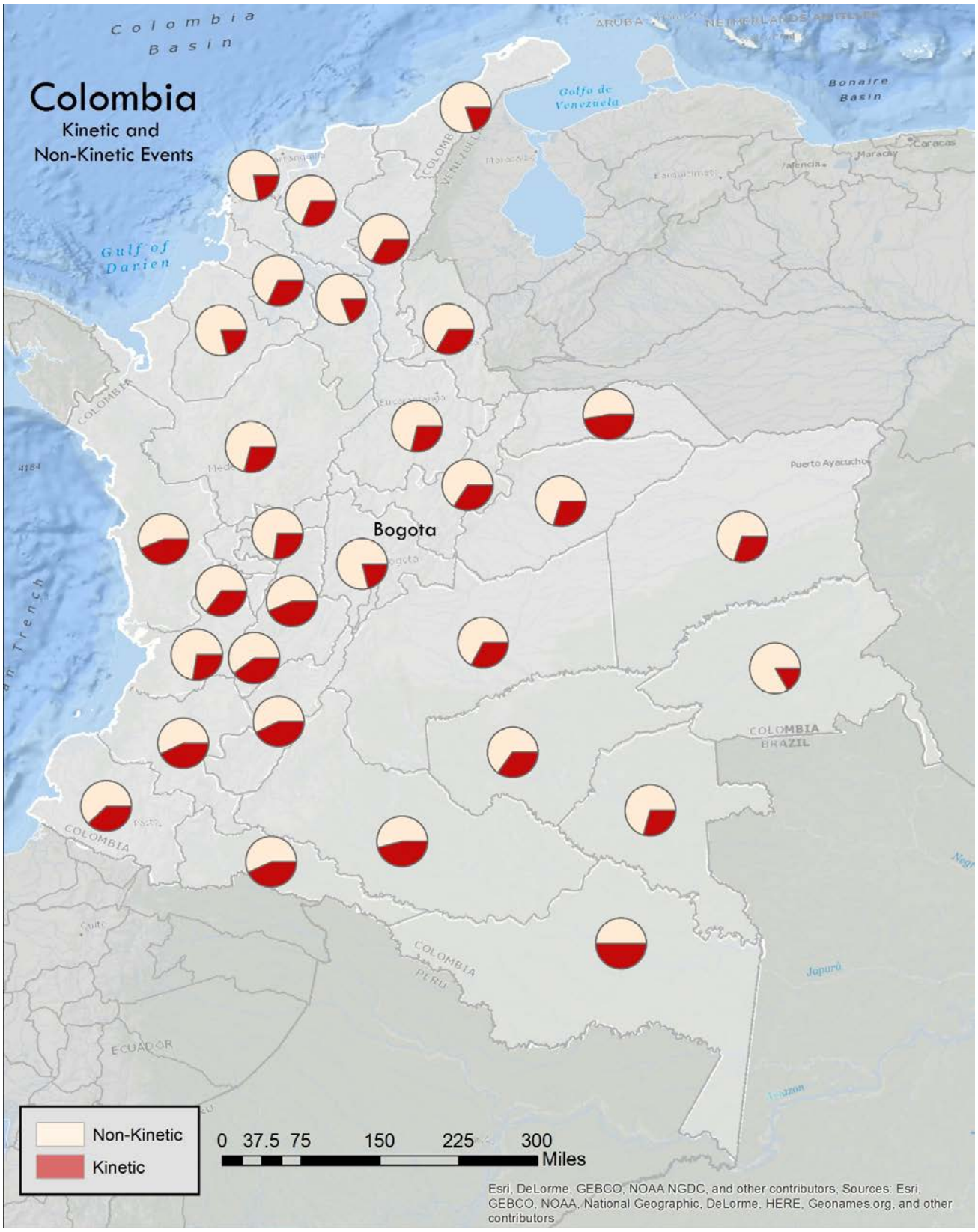
As the figure above shows, kinetic events have a sharp effect across all Zones. However, among VNSAs, their strongest effect is on Gray Zone activities. As kinetic events approach 300, Gray Zone tactics by VNSAs rise to above 400 monthly events while White Zone tactics approach 200. Black Zone activities rise to just above 60.

In other words, we find that kinetic activities are driving the use of a range of military and non-military instruments by VNSAs. Whereas the Gerasimov Doctrine suggests that the importance of the military

[14] Readers will note that the x-axis for various graphs are scaled differently. In some graphs, the values range from 0 to 100, while others span from 0 to as high as 2500. This is because the x-axes for each graph are based on the maximum weighted value observed in the data for that variable. We identified these maximum values from the summary statistics for each country.

instrument of power is severely diminished relative to other instruments, our results suggest that this is only superficially the case. It is true that kinetic events comprise only a small portion of the total observations. In Colombia, only approximately 23 percent of observations (or 17,458 of 76,681) were identified as kinetic events. The distribution of kinetic versus non-kinetic events across Colombia's 32 departments is visualized in Figure 7, below. The figure depicts the relative paucity of kinetic events. While kinetic activity is especially sparse in certain areas (e.g., the capital, Bogota), even where it is most prevalent (e.g., in the peripheries, where insurgent violence abounds) it never constitutes a majority of all events in any district. However, these relatively few kinetic events have an outsized effect in shaping the nature of non-kinetic events across non-military, and thus non-kinetic instruments.

Figure 7: Distribution of Kinetic versus Non-kinetic Events



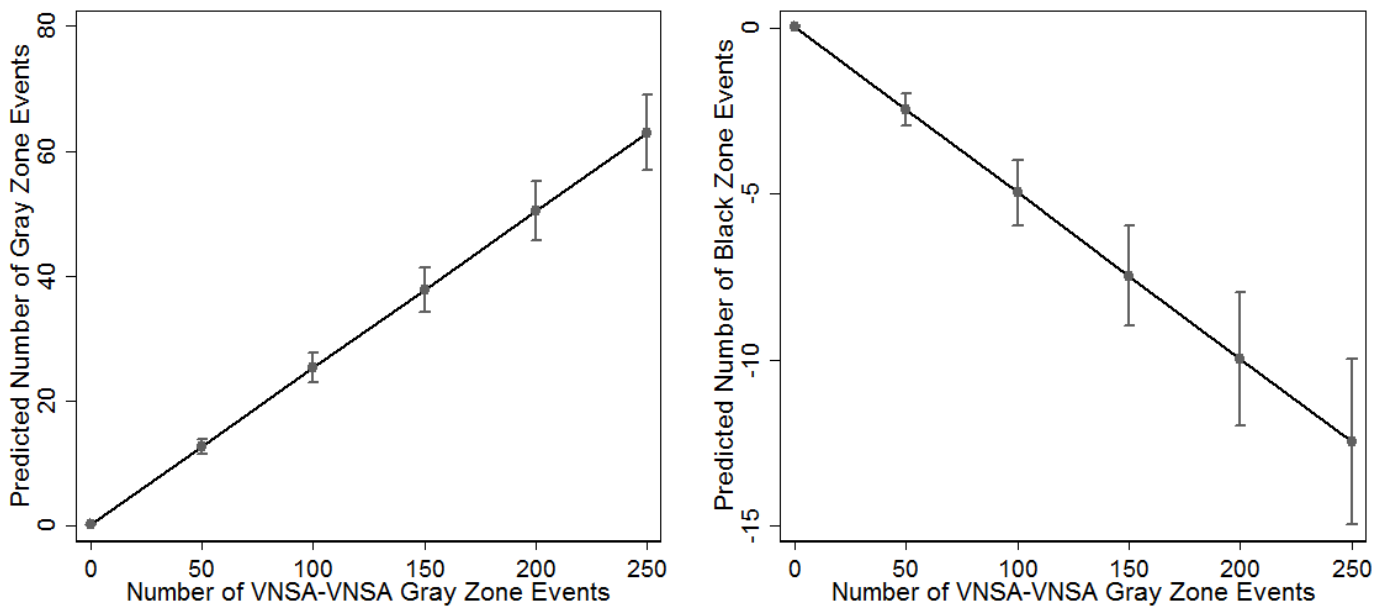
This finding, from the frequentist statistical analysis, is substantiated by the Bayesian analysis. In particular, the centrality of kinetic events in determining whether an event is Gray in the BBN may indicate that Gray Zone campaigns are accompanied by a credible kinetic threat. It may be the case that employing measured and limited kinetic force during the campaign increases said credibility and thus the chance of campaign success.

Moreover, the results may demonstrate that Gray Zone events are a foil for a much more kinetic-centric Black Zone strategy. Gray Zone activities may be designed to sway public opinion, thereby creating space for strategic maneuver. These approaches can mask the true intentions of the actor, including its willingness for escalation. As such, the model may be picking up on the hidden intentions of the source when it prominently features the Use of Kinetic Force as the most influential node acting on the probability of a Gray Zone event.

VNSAs Drive Higher Levels of Gray Zone Activity

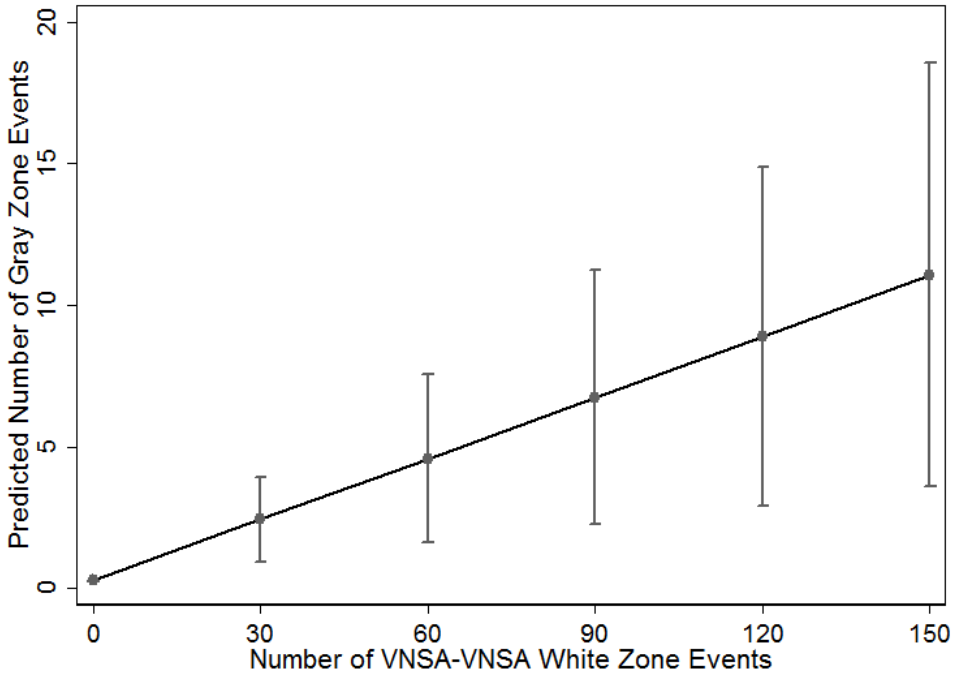
Our analysis suggests that greater interaction among VNSAs drives higher levels of Gray Zone activity by VNSAs. An increase of 100 Gray Zone events by VNSAs against other VNSAs in an average municipality and month is associated with a subsequent rise of Gray Zone behavior by VNSAs at a rate of 22 more Gray events in the subsequent municipality-month. However, the aforementioned increase in Gray events is associated with a drop in Black Zone behavior by VNSAs at a rate of five fewer Black events in the following municipality-month, as well as with a decline of seven White Zone events in the subsequent municipality-month. (However, the effect on White Zone activity is only robust in two out of four models.) Figure 8, below, plots the impact of greater VNSA-VNSA Gray Zone interaction on overall Gray and Black Zone activity by VNSAs.

Figure 8: Effects of VNSA-VNSA Gray Zone Interaction



As the figure above shows, greater VNSA-VNSA Gray Zone interaction is associated with rising Gray Zone behavior and falling Black Zone behavior. In addition, we also found that greater VNSA-VNSA White Zone interaction is associated with more overall Gray Zone usage by VNSAs. We observed that an increase of 100 White Zone events in an average municipality-month, where VNSAs were both the source and target, results in eight additional VNSA-driven Gray Zone events in the subsequent municipality-month. These results are consistent across three of the four models. Figure 9, below, plots the impact of White Zone tactics by VNSAs towards other VNSAs on overall Gray activity by VNSAs.

Figure 9: Effects of VNSA-VNSA White Zone Interaction



Together, these findings suggest that increasing VNSA-VNSA (White, Gray or Black) interaction, results in higher levels of Gray and lower levels of Black and White Zone behavior by VNSAs. The reality that more VNSA interaction results in more Gray activity is not surprising given that VNSAs are better able to exploit ambiguity than states. Indeed, this is why states often employ VNSA proxies in Gray Zone conflicts.

Again, this finding is validated by the Bayesian analysis. The nodal influence analysis reveals that VNSAs are much more likely to be the source of a Gray Zone event than states. Relatedly, it shows that states are less likely to be the source of Gray compared to Black or White Zone events. Since the Gray Zone encompasses a wide spectrum of activities, including activities like terrorism, this result may be reflective of the Colombian government's unwillingness to employ certain types of activities that fall within the range of Gray Zone conflict. On the other hand, the results may also be reflective of states' pursuit of Gray Zone activities through VNSA proxies, such as the United Self Defense Forces of Colombia (*Autodefensas Unidas de Colombia*; AUC).

States Are Better at Interpreting Signals than VNSAs

Our findings also suggest that states and VNSAs have asymmetric signaling capabilities. States are better at distinguishing between White and Gray VNSA actions (and responding in kind). First, we find state use of White Zone tactics targeted at VNSAs result in higher levels of VNSA White Zone activity. On average, an increase of 100 White Zone events by states in a given municipality-month is associated with an increase of two White Zone events by VNSAs. Given the fact that states use White Zone tactics statistically significantly more than VNSAs to begin with, this is not a trivial increase. However, we also find that on average, the aforementioned increase is also associated with two additional Gray Zone events by VNSAs. Figure 10, below, plots the effect of White Zone activity directed at VNSAs by states on White Zone and Gray Zone activity by VNSAs. As the graph shows, both White Zone and Gray Zone activity by VNSAs increase as a result of White Zone behavior by states. The effect is nearly twice as strong on Gray versus White Zone activity by VNSAs. This appears to indicate that VNSAs tend to misinterpret non-conflictual actions by states as being escalatory in nature; thus warranting an aggressive VNSA response. (However, this finding is only robust to three of the four model specifications.)

Figure 10: Effect of State White Zone Events on VNSA Actions

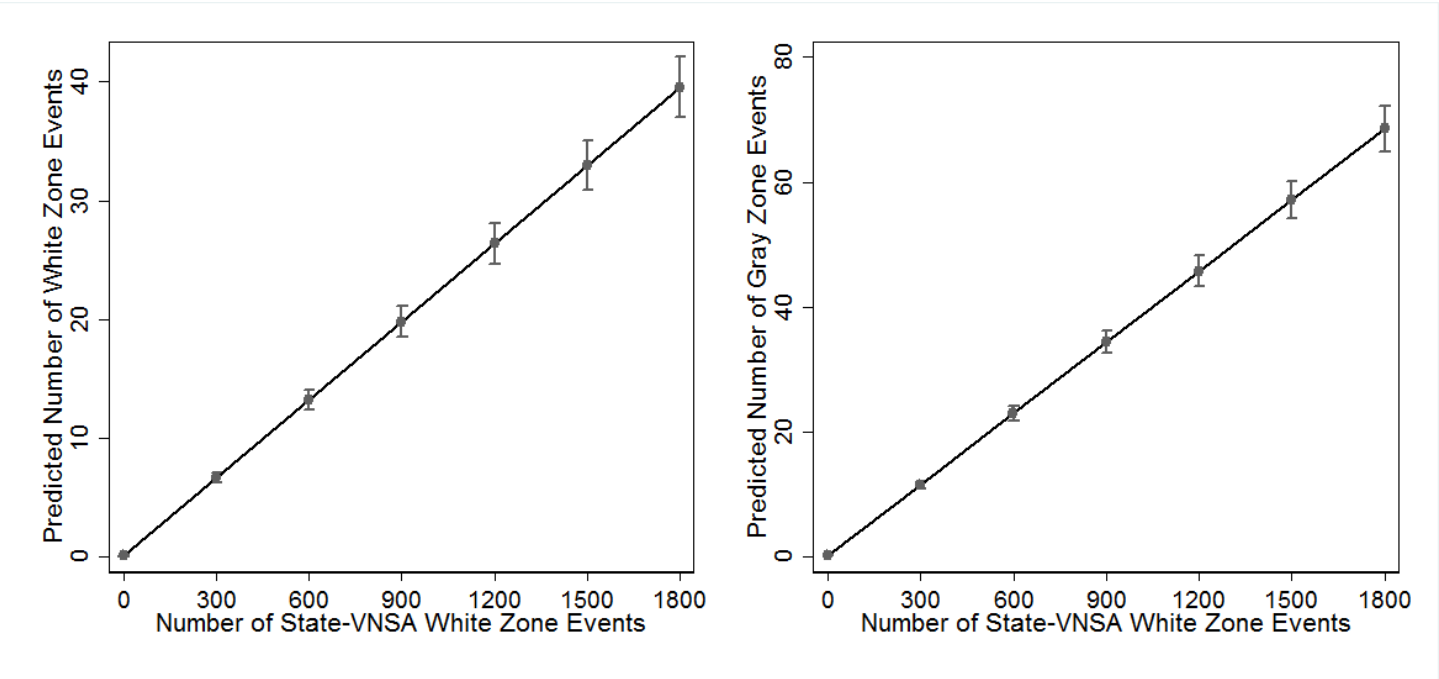
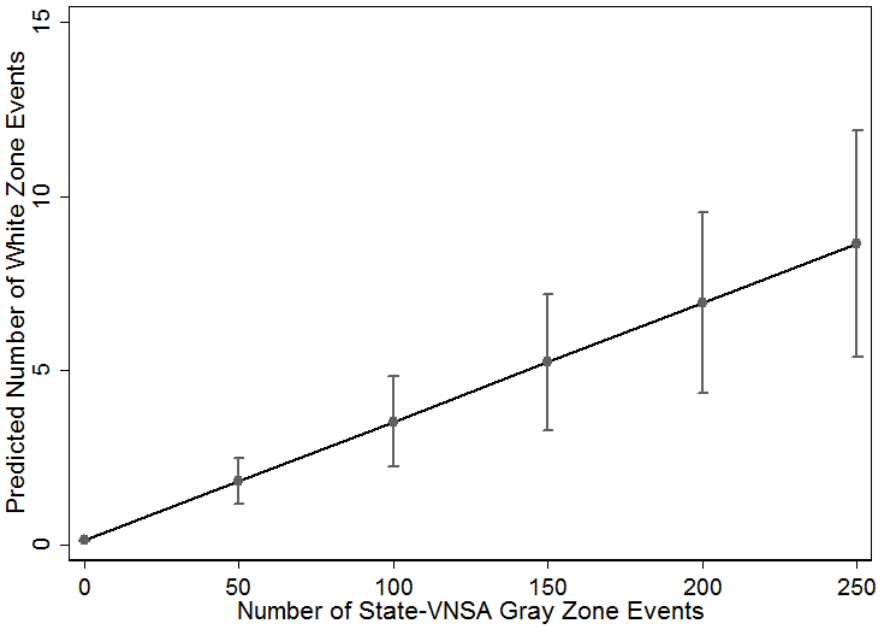


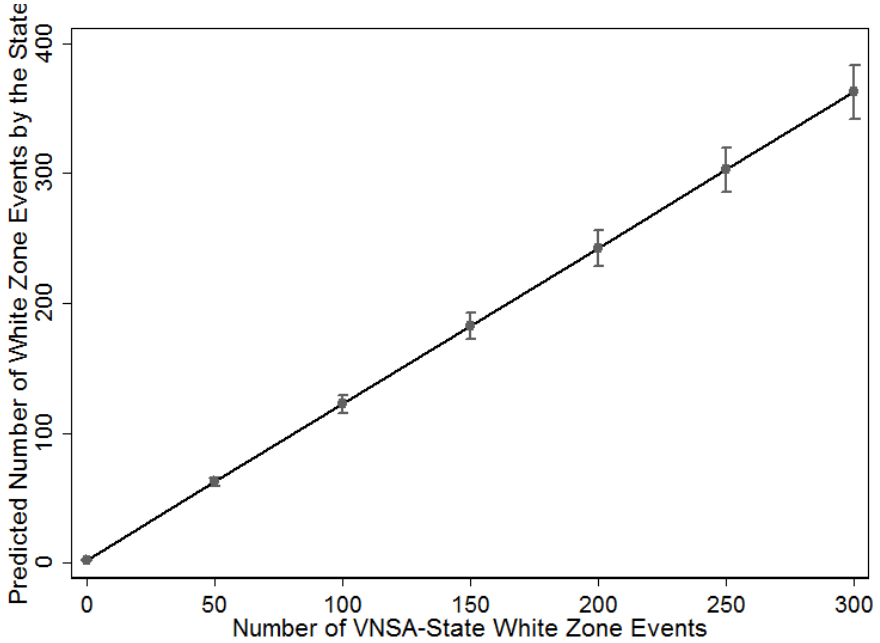
Figure 11, below, plots the effect of Gray Zone activity directed at VNSAs by states on White Zone activity by VNSAs. As the graph shows, VNSAs respond with higher levels of White Zone activity when states increasingly rely on Gray Zone tactics. Together, these findings suggest that VNSAs struggle to differentiate between White and Gray Zone tactics by states. In fact, it is possible that VNSAs may be interpreting White Zone activities by states as Gray Zone activity, and vice versa. Indeed, we found various instances of VNSAs accusing states of Gray Zone activity masked as White Zone activity. Our findings suggest that this misinterpretation may be driving VNSAs to respond with Gray Zone activities, even in response to conciliatory state actions.

Figure 11: Effect of State Gray Zone Events on VNSA Actions



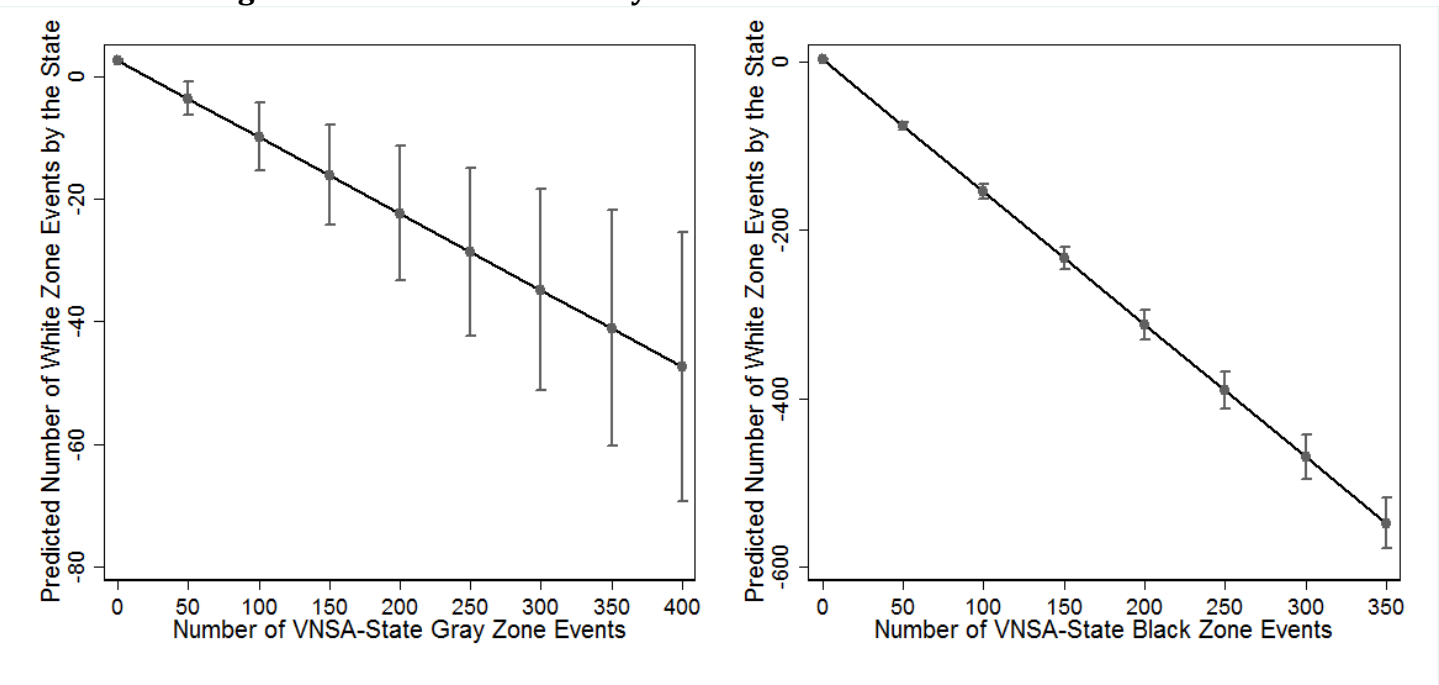
States, on the other hand, do not seem to have the same problem. In fact, we find that VNSA use of White Zone tactics towards states is positively associated with greater White activity by states. On average, an increase of 100 White Zone events by VNSAs towards states in a given municipality-month is associated with an increase of 97 White Zone activities by states in the subsequent municipality-month. Figure 12, below, depicts the effects of White Zone activity by VNSAs on White Zone activity by states. As the graph shows, VNSA use of White Zone tactics drives states to respond in kind.

Figure 12: Effect of VNSA White Zone Events on State Actions



Conversely, VNSA use of Gray and Black Zone activities reduces White Zone events by states. On average, an increase of 100 Gray Zone events by VNSAs towards states in a given municipality-month is associated with a decrease of 19 White Zone events by states, while an increase of 100 Black Zone events by VNSAs towards states in an average municipality-month is associated with a decline of 145 White Zone observations perpetrated by states. Figure 13, below, plots the impact of the use of Gray and Black Zone tactics by VNSAs on the use of White Zone tactics by states. As the figure shows, both Gray and Black behavior by VNSAs sharply drives down White Zone behavior by states. In other words, our analysis suggests that states are more likely to correctly classify VNSA behavior (and respond accordingly).

Figure 13: Effect of VNSA Gray and Black Zone Events on State Actions



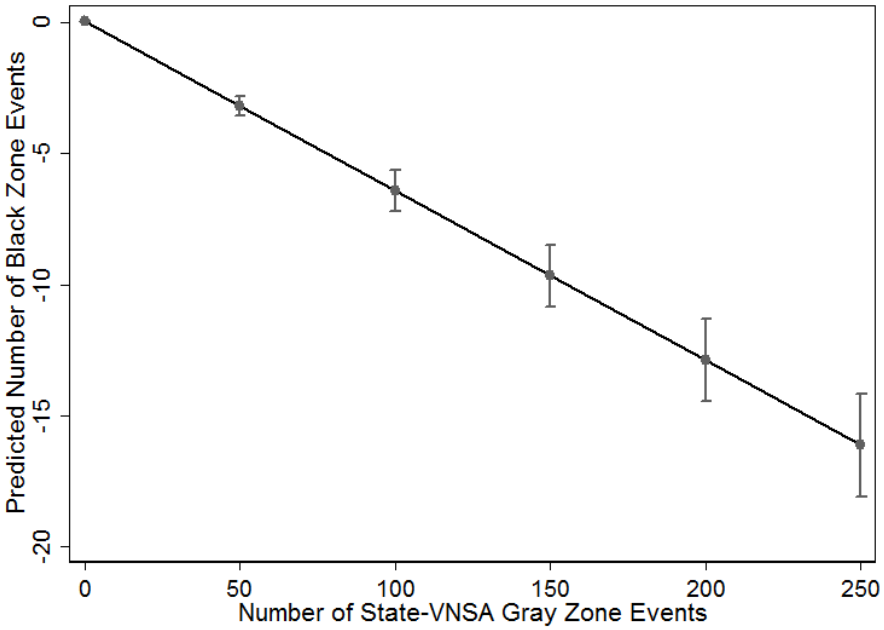
VNSA Mirroring of State Escalatory (Gray to Black) and De-escalatory (Black to Gray) Trends

We also observed that VNSAs respond strategically to the use of Gray and Black Zone tactics by states. Specifically, VNSAs de-escalate when states use Gray, and escalate when states use Black. In other words, the VNSA response often mirrors states.

We found that on average, an increase of 100 Gray Zone events where states is the source and VNSAs are the target in a given municipality-month resulted in six fewer VNSA-driven Black Zone events per month. Figure 14, below, plots the impact of the use of Gray tactics by states against VNSAs on their decision to engage in Black Zone activities. As the figure shows, VNSAs tend towards de-escalation in response to Gray Zone behavior by states. Recall also that Figure 11 shows that states use of Gray activity is associated with higher levels of White Zone behavior by VNSAs. Together, these findings suggest that state use of Gray Zone tactics result in de-escalation by VNSAs. One possibility as to why this happens might be that VNSAs have incentives to reduce Black Zone activity when states also shifts to Gray Zone tactics so as to not lose legitimacy. Specifically, in the Colombian case, numerous VNSAs engaged in diplomatic and informational activities clearly designed to improve both their domestic and international

legitimacy. If states is prosecuting the conflict in the Gray Zone, its conflictual actions towards VNSAs are likely to be ambiguous. If VNSAs respond in an overt, Black Zone manner, they may be (incorrectly) perceived as the aggressors, which could adversely affect how they are viewed by both international and domestic audiences.

Figure 14: Effect of State Gray Zone Events on VNSA Actions



In contrast to Gray Zone behavior, the use of Black Zone tactics by states is associated with escalation by VNSAs. On average, an increase of 100 Black Zone tactics by states aimed at VNSAs in a given municipality-month results in an increase of three Black Zone events by VNSAs in the subsequent municipality-month. Figure 15, below, plots the impact of Black Zone tactics by states on the number of Black Zone events by VNSAs. As the figure shows, VNSAs escalate towards open conflict as a response to the use of Black Zone tactics by states. In fact, this finding, that state use of Black results in escalation, is also consistent with the observed decrease in the use of White and Gray activities by VNSAs in municipality-months following escalation by states.

Figure 15: Effect of State Black Zone Events on VNSA Black Zone Actions

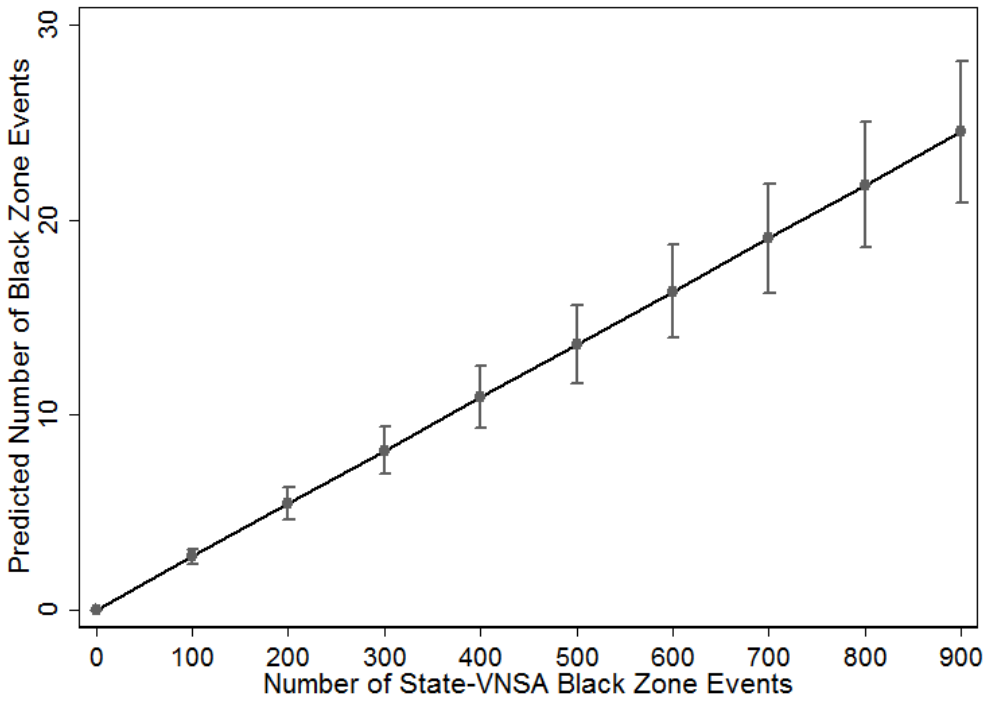
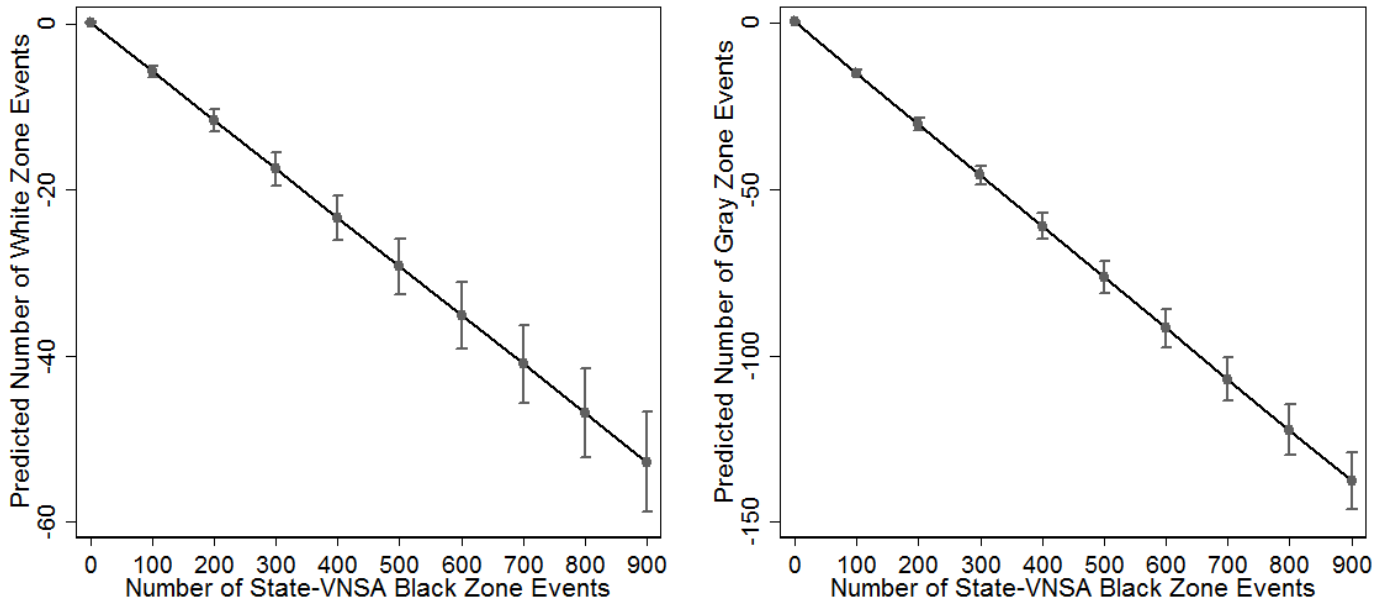


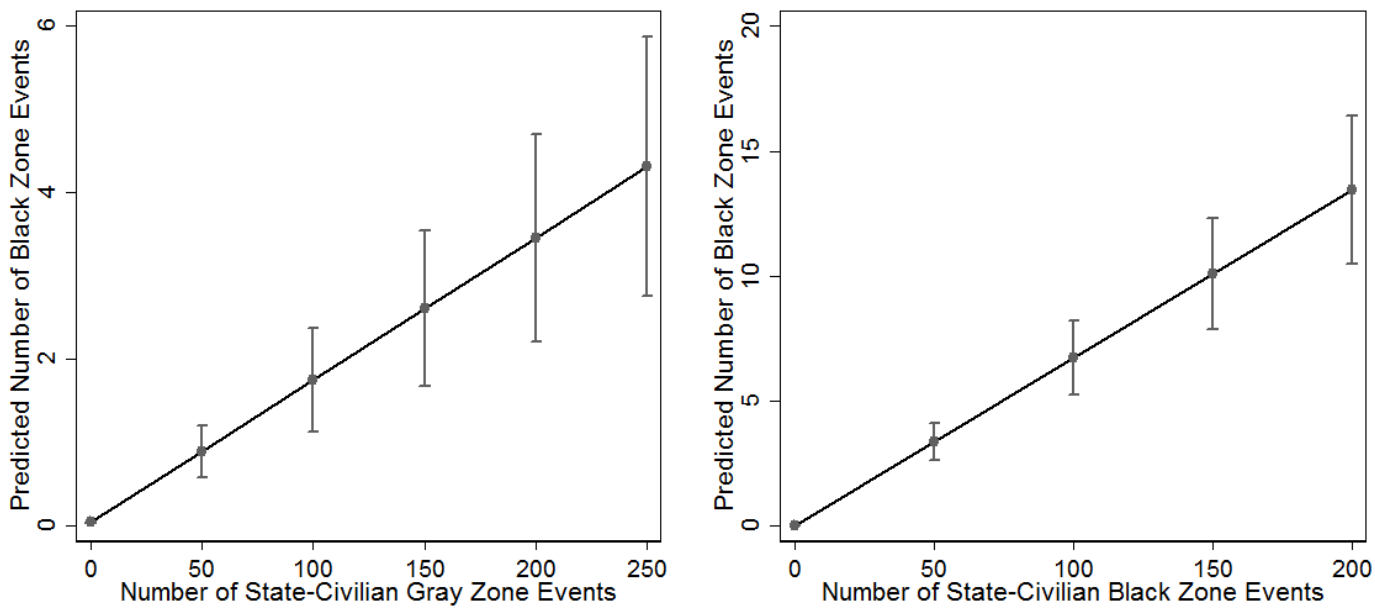
Figure 16, below graphs the impact of Black Zone activities by states on the number of White and Gray Zone events by VNSAs. As the graph shows, increasing reliance on Black Zone tactics by states results in a decrease in VNSA utilization of both White and Gray Zone tactics. The effect is higher for Gray versus White Zone activities. Together, these findings suggest that when states increase their reliance on Black Zone tactics against VNSAs, VNSAs respond in kind, and shift from White and Gray Zone tactics towards open, Black Zone, conflict. In these cases, VNSAs were not likely to be as concerned about legitimacy, given that states' conflictual activities were easily apparent.

Figure 16: Effect of State Black Zone Events on VNSA White and Gray Zone Actions



This aforementioned reality is consistent with our finding that state escalation towards civilians also resulted in VNSA escalation. One potential explanation for this dynamic is that state abuses against civilians provide a clear justification for conflictual activities undertaken by VNSA targeting states. Specifically, we find that on average an increase of 100 Gray Zone events by states targeting civilians results in two additional Black actions by VNSAs targeting states in the subsequent municipality-month. (However, this finding is only robust across two of the four model specifications.) Second, we find that on average, an increase of 100 state-initiated Black Zone events targeting civilians in a given municipality-month results in an increase of seven Black Zone events by VNSAs in the subsequent municipality-month. Figure 17, below, plots the impact of state use of Gray and Black Zone actions targeted towards civilians on the Black Zone behavior of VNSAs.

Figure 17: Effect of State Civilian Targeting on VNSA Actions



Finally, the nodal influence analysis shows that a civilian target is a strong predictor of Gray Zone events, but not for White or Black Zone actions. This result suggests that the intention of Gray Zone events in Colombia – regardless of the source – is to sway the perceptions and opinions of the civilian population. This finding is consistent with the canon of literature on intrastate violence, which finds that gaining the support of civilian populations is critical to victory. Furthermore, it is congruent with the aforementioned discussion regarding the importance of legitimacy and its effects on (de-)escalatory trends towards civilians.

Ukraine

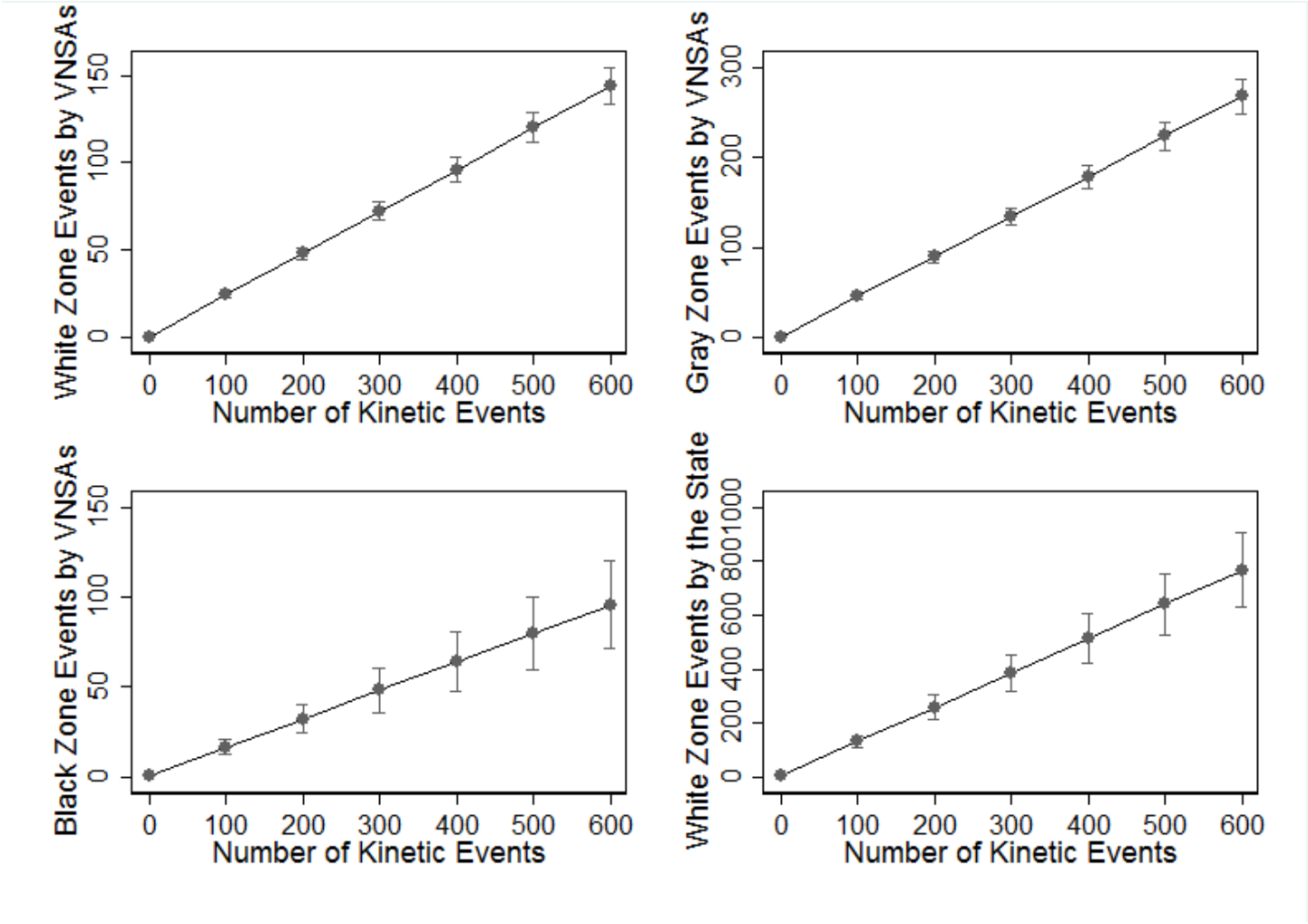
While all of the results for Ukraine have already been described, this section details key findings. First, we observed that kinetic events are highly consequential. Kinetic activities seem to be driving higher levels of White and Gray Zone activity by VNSAs and the Ukrainian and Russian states, including non-kinetic White and Gray Zone activity. Second, we observed that while Russian forces have relied heavily on proxy forces in the form of militias, etc., whenever Ukrainian state and VNSA forces succeeded in sufficiently attriting the capabilities of Russian proxies, Russian state forces directly intervened. Third, VNSAs in the Ukrainian conflict do not suffer substantially from asymmetries in interpreting and responding to state-initiated White Zone activities. Finally, our findings suggest that civilians are a key part of the story in the Ukrainian conflict. We found that VNSAs respond to state (de-)escalation towards civilians by following suit. However, when civilians agitate or take actions against VNSAs, we found that VNSAs respond very aggressively. VNSA actions directed against civilians also affects state behavior.

The Importance of Kinetic Events

We find that kinetic activities are highly consequential. Specifically, kinetic events increase the likelihood of White and Gray activities that are both kinetic and non-kinetic in nature by both states and VNSAs

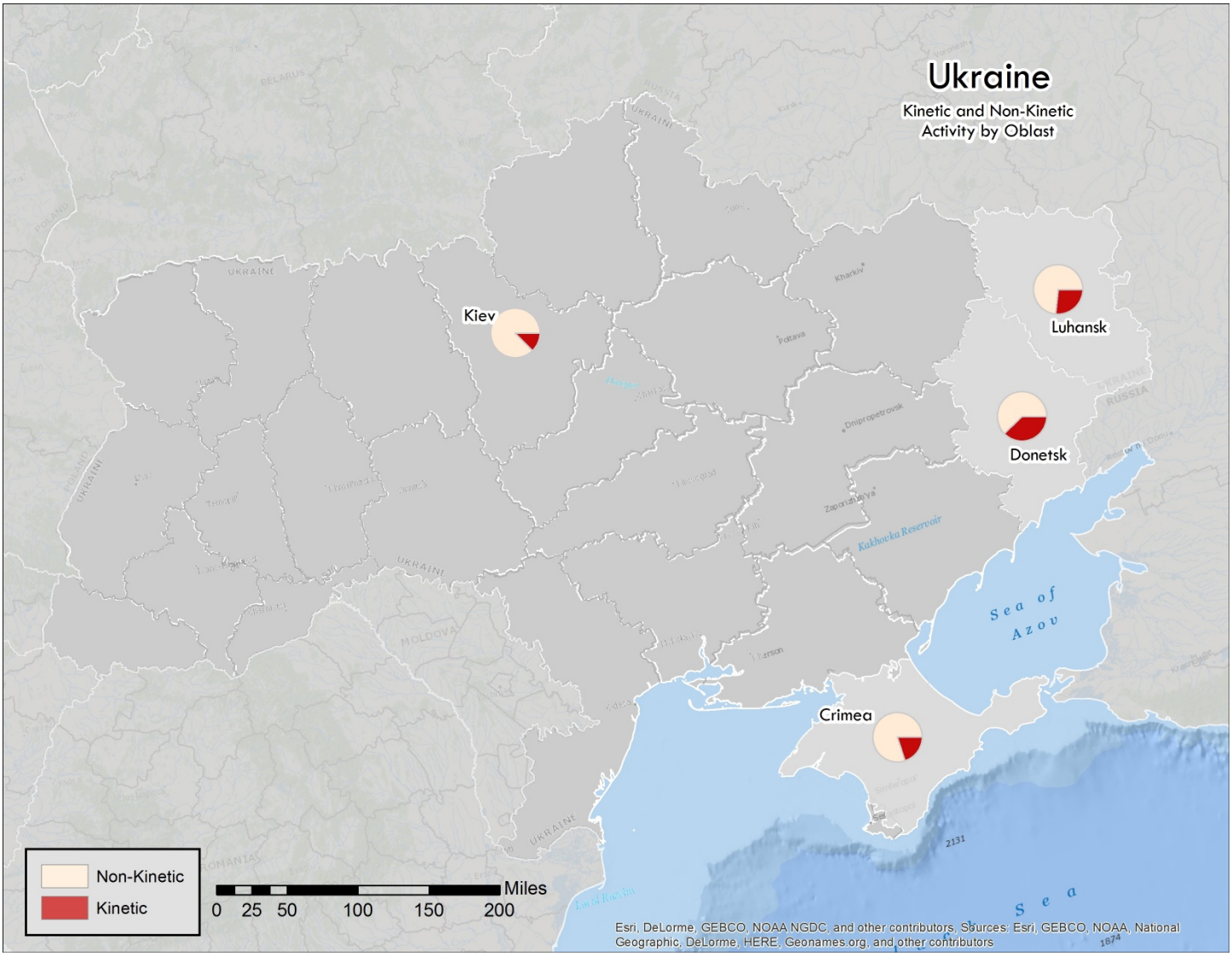
across all models. Substantively, our models predict that on average in Ukraine, an increase of 100 kinetic events in a given district-month is associated with an increase of 24 White Zone events and 37 Gray Zone events by VNSAs in that district during the subsequent month. In terms of states' White Zone activity, an increase of 100 kinetic events in a given district-month is associated with an average increase of 34 White Zone events during the subsequent district-month. Figure 18, below, plots how different types of Zones change as the number of kinetic events rise.

Figure 18: Kinetic Events and Zonal Activity



As the figure above shows, kinetic events have a sharp effect on all Zones. However, among VNSAs, their strongest effect is on Gray and Black Zone activities. As kinetic events approach 300, Gray and Black Zone tactics by VNSAs rise to around 200 monthly events while White Zone tactics approach 100 monthly events. In other words, kinetic events are driving the use of a range of military and non-military instruments by VNSAs. This is the case despite the fact that kinetic events in the Ukrainian conflict comprise only a small share of all events. Specifically, just 17.2 percent (or 11,705) of all 67,899 events are kinetic. Figure 19, below, depicts the distribution of kinetic and non-kinetic events in Crimea, the Donbas (comprised of the Luhansk and Donetsk districts or *oblasts*) and in the capital, Kiev. Even, in the Donetsk, which experienced the most fighting, kinetic activity comprises only a minority of all events.

Figure 19: Distribution of Kinetic versus Non-kinetic Events



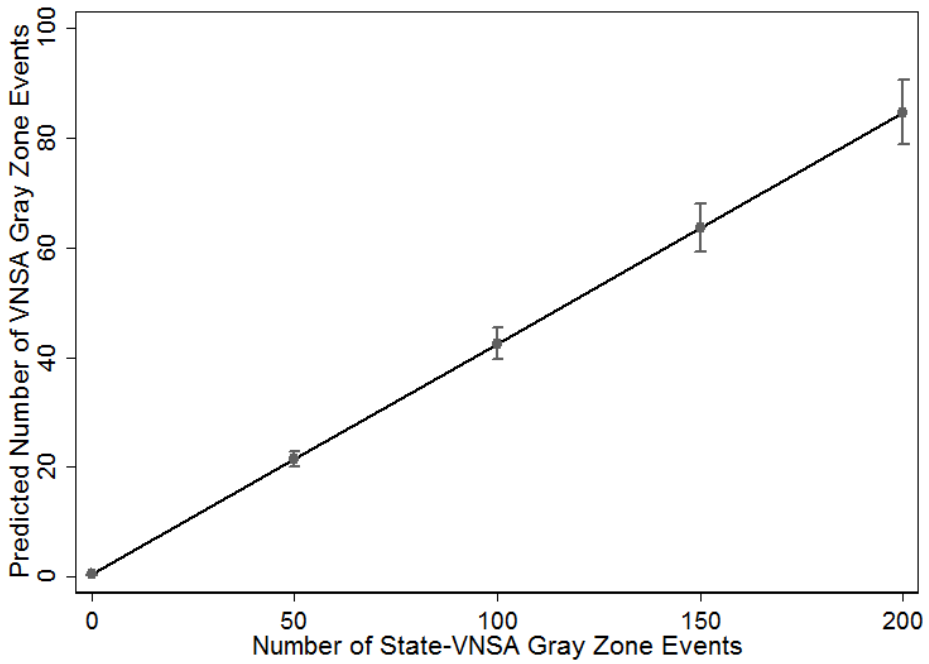
Similarly, the Bayesian analysis demonstrates the centrality of the Use of Kinetic Force node in predicting Gray Zone events. The Bayesian results seem to indicate that belligerents are employing measured and limited kinetic force during their Gray Zone operations as a credible threat to increase the chances of campaign success. It is also possible that Gray Zone activities are being employed to mask, or at least reshape, how Black Zone activities are interpreted.

VNSAs Adequately Interpret State Signals

VNSAs in the Ukrainian conflict seemed to be at least somewhat adept at correctly interpreting state actions and responding in kind. First, we observed that VNSAs are good at identifying White Zone activity by states and responding with proportionally more White Zone actions. On average, an increase of 100 White Zone events by states targeted at VNSAs in a given district-month is associated with an increase of five White Zone events by VNSAs during the following district-month.

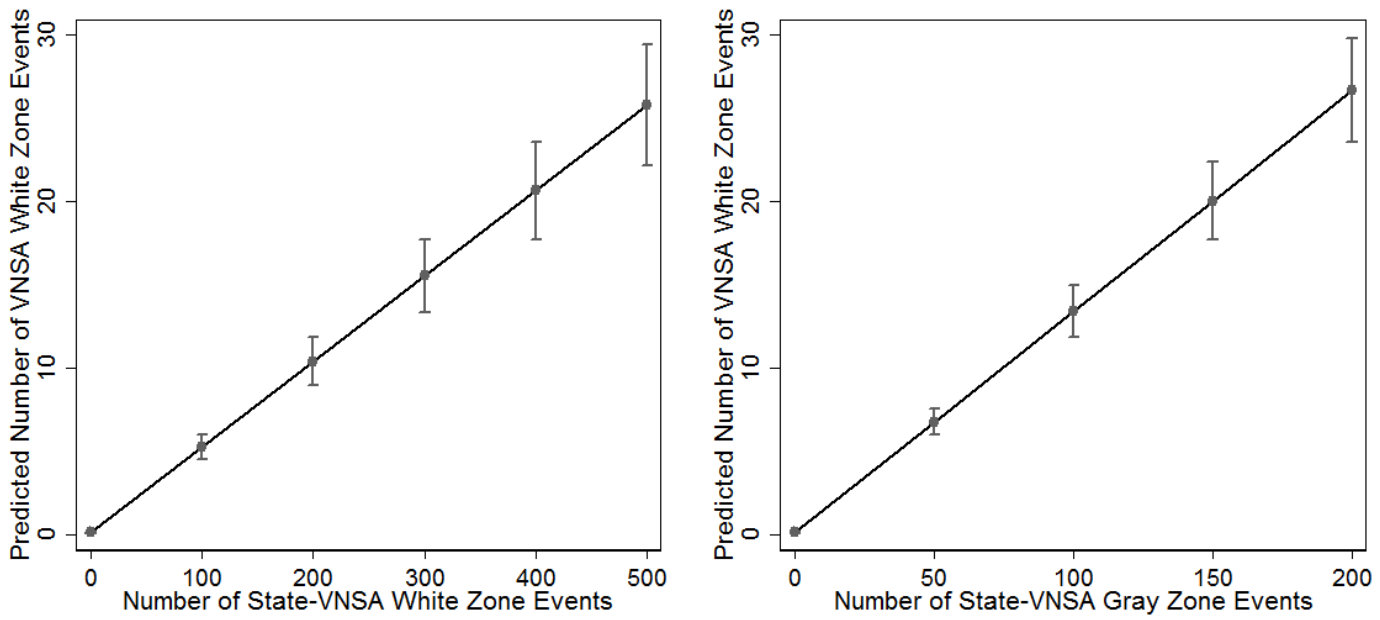
We similarly found that the use of Gray Zone tactics by statess against VNSAs results in higher use of Gray Zone activity by VNSAs. On average, an increase of 100 monthly Gray Zone actions by states in a given district-month precipitated an increase of 38 Gray Zone events conducted by VNSAs in the subsequent district-month. Figure 20, below, graphs the impact of Gray Zone events by states on Gray Zone behavior by VNSAs.

Figure 20: Effect of State Gray Zone Events on VNSA Gray Actions



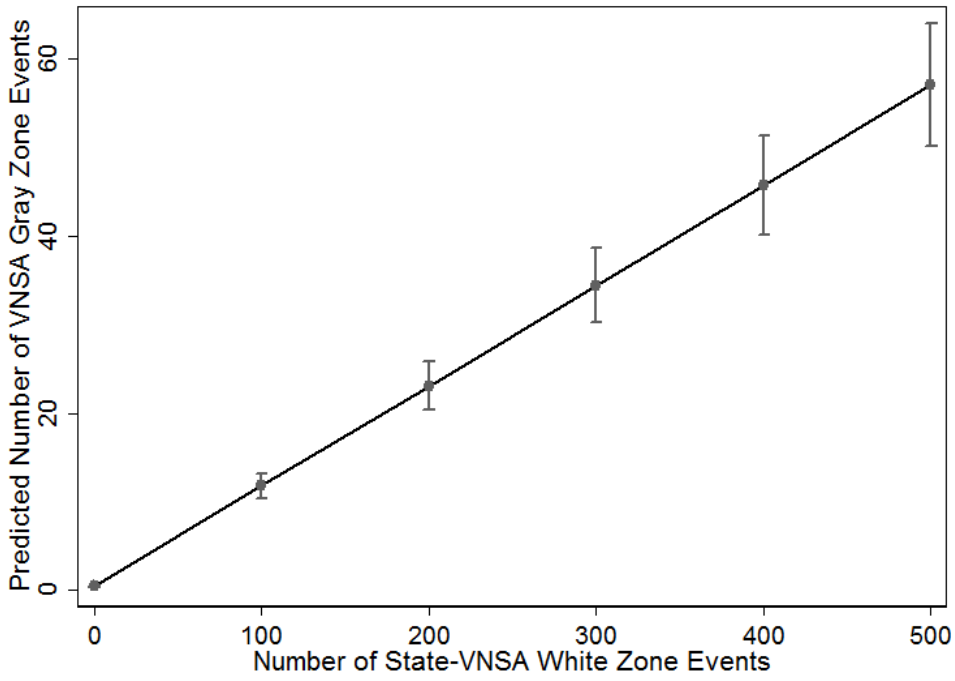
Nevertheless, VNSAs at least sometimes experience difficulty discerning between White and Gray Zone tactics employed by states. For example, an increase of 100 Gray Zone events by states targeted at VNSAs in a given district-month results in an increase of 14 White Zone events by VNSAs in the subsequent district-month. Figure 21, below, plots the effect of White and Gray Zone tactics by states on the White Zone behavior of VNSAs.

Figure 21: Effect of State White and Gray Zone Events on VNSA White Actions



Along similar lines, the use of White Zone tactics by states is also associated with an increase in Gray Zone tactics by VNSAs. On average, an increase of 100 White Zone actions by states in a given district-month results in a rise of nine Gray Zone events perpetrated by VNSAs in the subsequent district-month. Figure 22, below, plots the impact of White Zone tactics by states on Gray Zone activity by VNSAs.

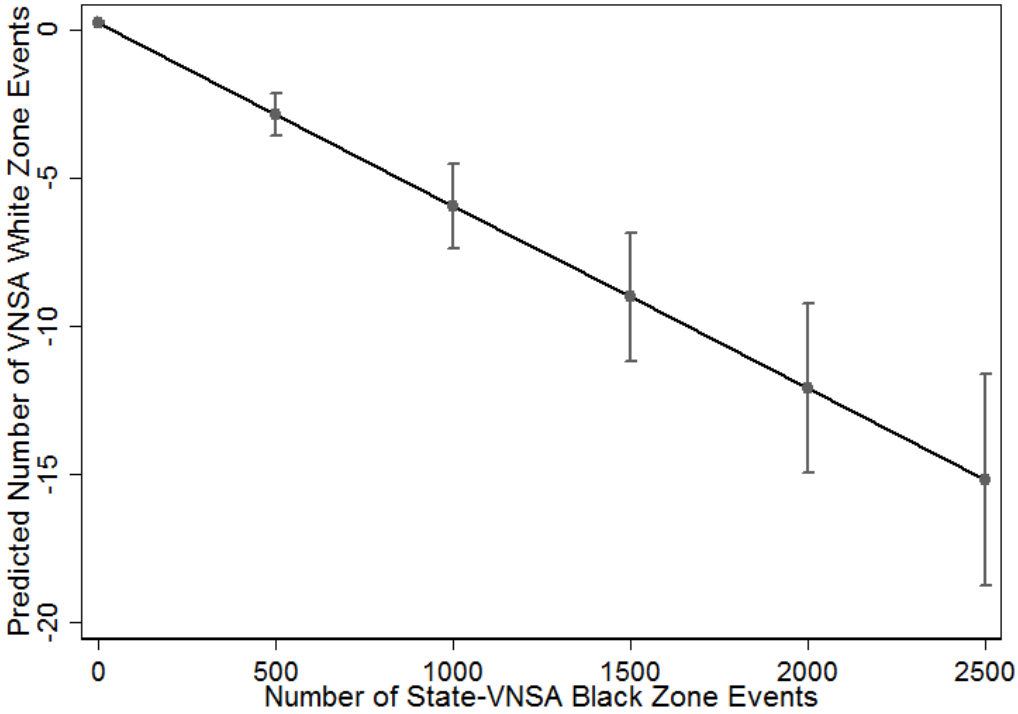
Figure 22: Effect of State White Zone Events on VNSA Gray Actions



The two figures above show that while VNSAs respond with some White Zone activity to White Zone usage by states, they also seem to respond with Gray Zone tactics. Conversely, they seem to respond with White Zone tactics to Gray Zone usage by states. Taken together, these findings show that VNSAs are sometimes able to correctly distinguish between White and Gray state actions but occasionally also experience difficulty.

This finding does not extend to Black Zone activity. Our findings suggest that the use of Black Zone tactics against VNSAs by states precipitates reduced VNSA reliance on White Zone activity. On average, an increase of 100 Black Zone events in a given district-month results in one fewer monthly White Zone action by VNSAs in the following district-month. Figure 23, below, plots the impact of Black Zone tactics by states on White Zone behavior by VNSAs. As the graph shows, the use of Black approaches by states has a dampening effect on White Zone activity by VNSAs. Even though this effect is the weakest, the direction and statistical significance clearly indicate that VNSAs do not struggle to identify and respond to Black Zone activities on the part of states. Given the overt nature of Black actions, this is unsurprising.

Figure 23: Effect of State Black Zone Events on VNSA White Actions



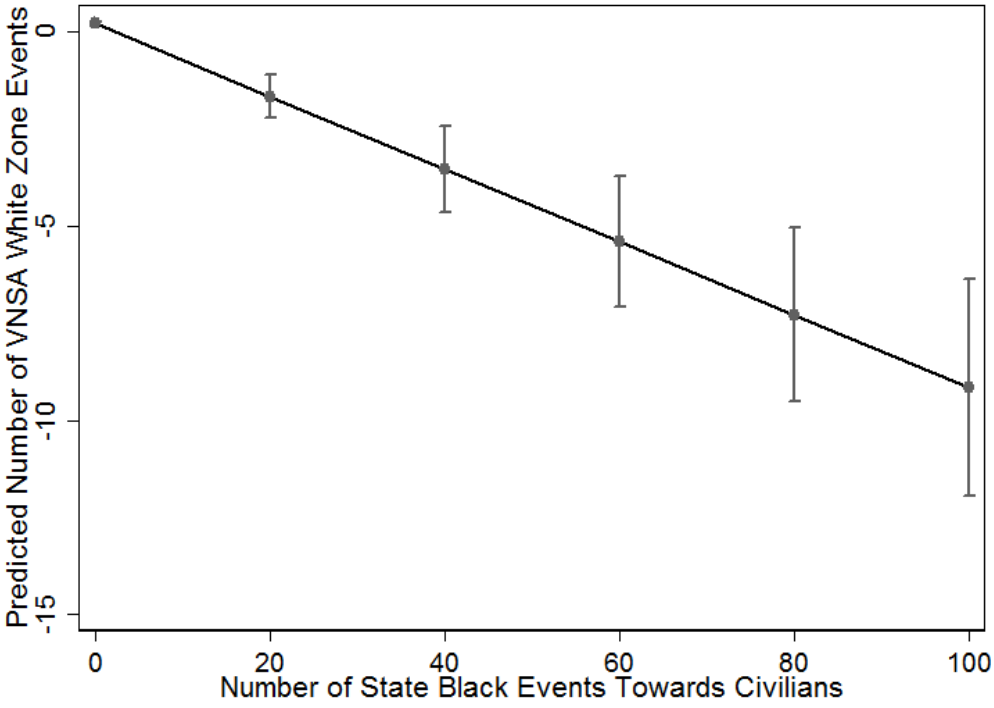
The aforementioned findings speak to the relative ease that VNSAs have in interpreting state behavior. This is likely the case insofar as VNSAs in the Ukrainian conflict closely collaborate with their state partners. The Bayesian results, similarly demonstrate that VNSAs are much more likely than states to be the source of both Gray and Black activities. This finding is consistent with the reality that Russian state forces have only directly intervened militarily when their VNSA proxies risked defeat. Similarly, Ukraine

was often forced to rely primarily on militia forces due to the poor state of combat readiness of its armed forces (Finkel 2016).

Civilian Victimization and Escalatory Trends

Our findings suggest that civilians are a key part of the story in the Ukrainian conflict. Specifically, we noted that VNSAs respond to state escalation and de-escalation towards civilians with their own escalation or de-escalation. On average, we found that an increase of 100 Black Zone events by states against civilians in a given district-month is associated with a decline of nine White Zone events by VNSAs in the subsequent district-month. Figure 24, below, plots the impact of Black Zone tactics by states against civilians on White Zone usage by VNSAs. As the figure shows, VNSAs escalate in response to the use of Black Zone tactics by states against civilians. It is likely that civilian victimization by states provides VNSAs with a justification that legitimates violence against states, thereby ameliorating the need to obscure their activities and thus prioritize Gray Zone tactics. Consequently, escalation is unsurprising.

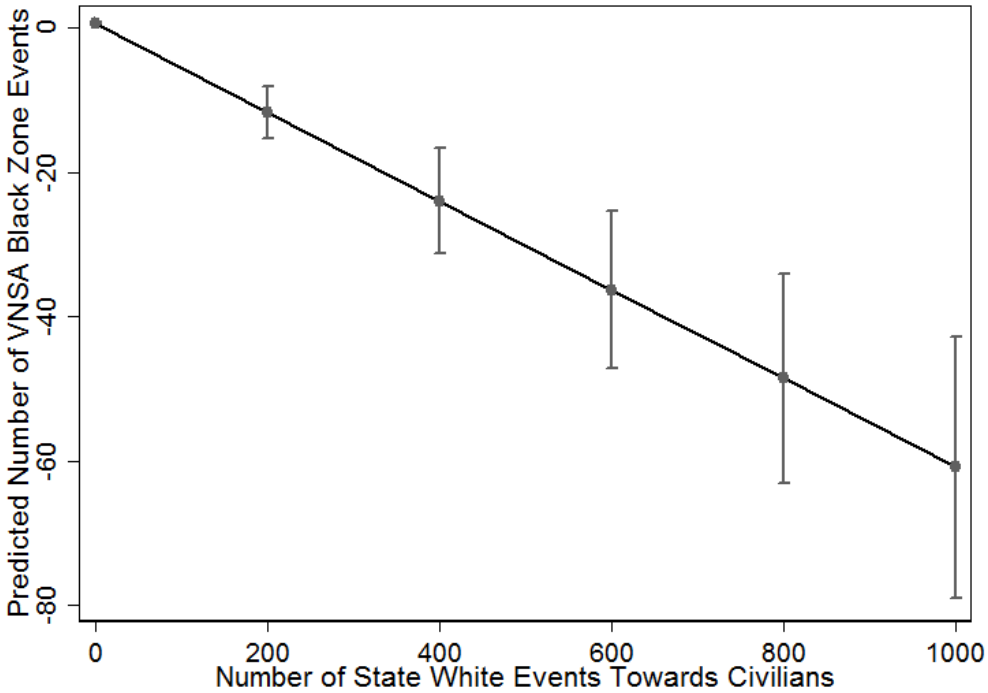
Figure 24: Effect of State Black Zone Events on VNSA White Actions



Similarly, we find that VNSAs de-escalate towards states in response to greater White Zone usage by states towards civilians. On average, we found that for each additional 100 White Zone events by states towards civilians in a given district-month, there is an associated drop of five Black Zone events by VNSAs per month at the district level. Figure 25, below, plots the impact of state White Zone tactics towards civilians on Black Zone activity by VNSAs. As the figure shows, we find some evidence that VNSAs de-escalate in response to state de-escalation towards civilians. It is possible this de-escalation is also a tactic meant to maintain or improve legitimacy. Specifically, as states victimize civilians less, the

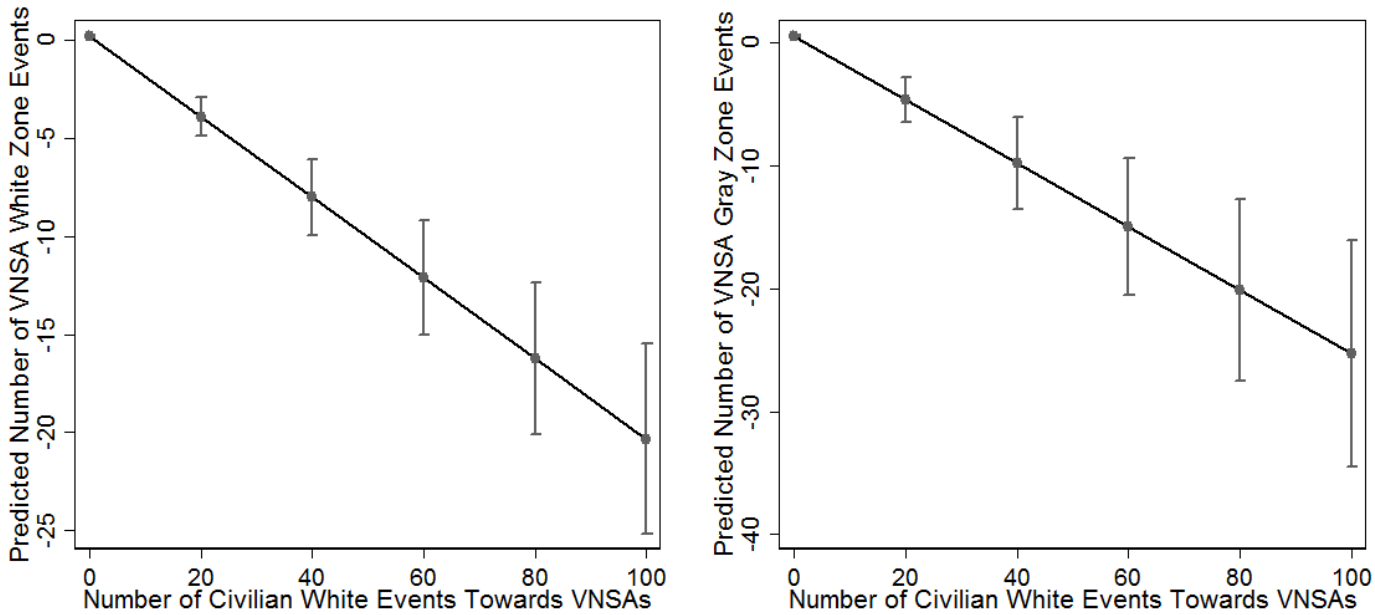
justification for the use of force by VNSAs against states is weakened, forcing VNSAs to decrease overt (i.e., Black Zone) violent attacks on states.

Figure 25: Effect of State White Zone Events on VNSA Black Actions



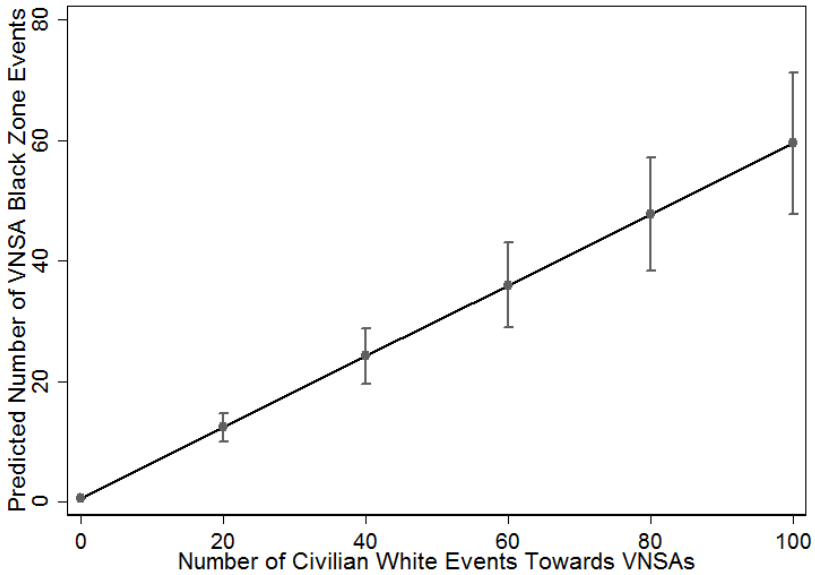
Unfortunately, we also observed substantial willingness on the part of VNSAs to target civilians. Specifically, we found that an increase of 100 White Zone events by civilians towards VNSAs in a given district-month, correlated with 19 less White Zone events by VNSAs in the subsequent district-month. We also found that for an average district-month increase of 100 White Zone events by civilians towards VNSAs, there was also an associated drop of 34 Gray Zone events by VNSAs in the subsequent district-month. Figure 26, below, plots the impact of White Zone tactics by civilians towards VNSAs on White and Gray Zone activity by VNSAs.

Figure 26: Effect of Civilian White Zone Events on VNSA White and Gray Actions



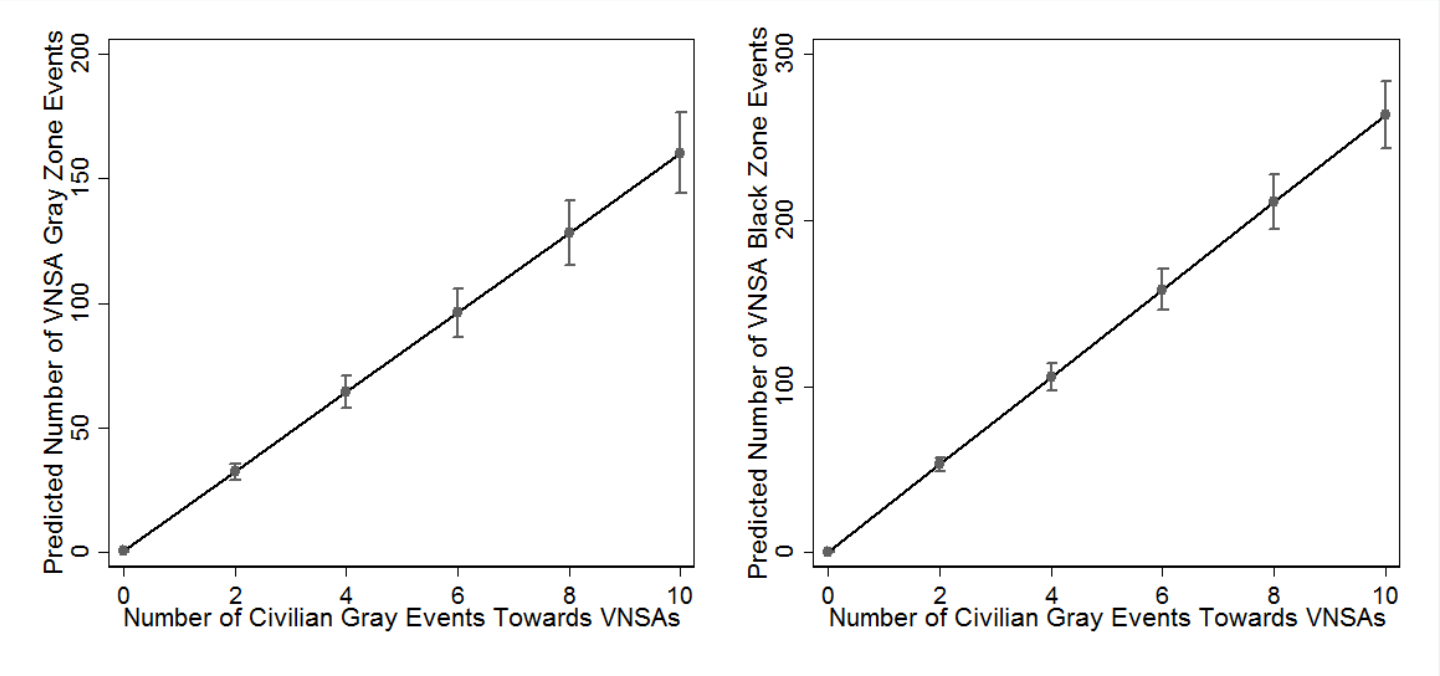
The two plots can be interpreted in multiple ways. It could be that VNSAs are simply disengaging as a result of civilian actions. It could also be the case that any action by civilians toward VNSAs is seen as provocation, which results in escalation. Unfortunately, our findings suggest the latter explanation is most probable. Specifically, we find that the aforementioned decline in White and Gray activities are more than compensated for in an increase in Black Zone actions. For example, an increase of 100 White Zone events by civilians towards VNSAs in a given district-month resulted in an increase of 56 Black Zone tactics by VNSAs in the subsequent district-month. Figure 27, below, shows the impact of White Zone agitation by civilians towards VNSAs on Black Zone usage by VNSAs.

Figure 27: Effect of Civilian White Zone Events on VNSA Black Actions



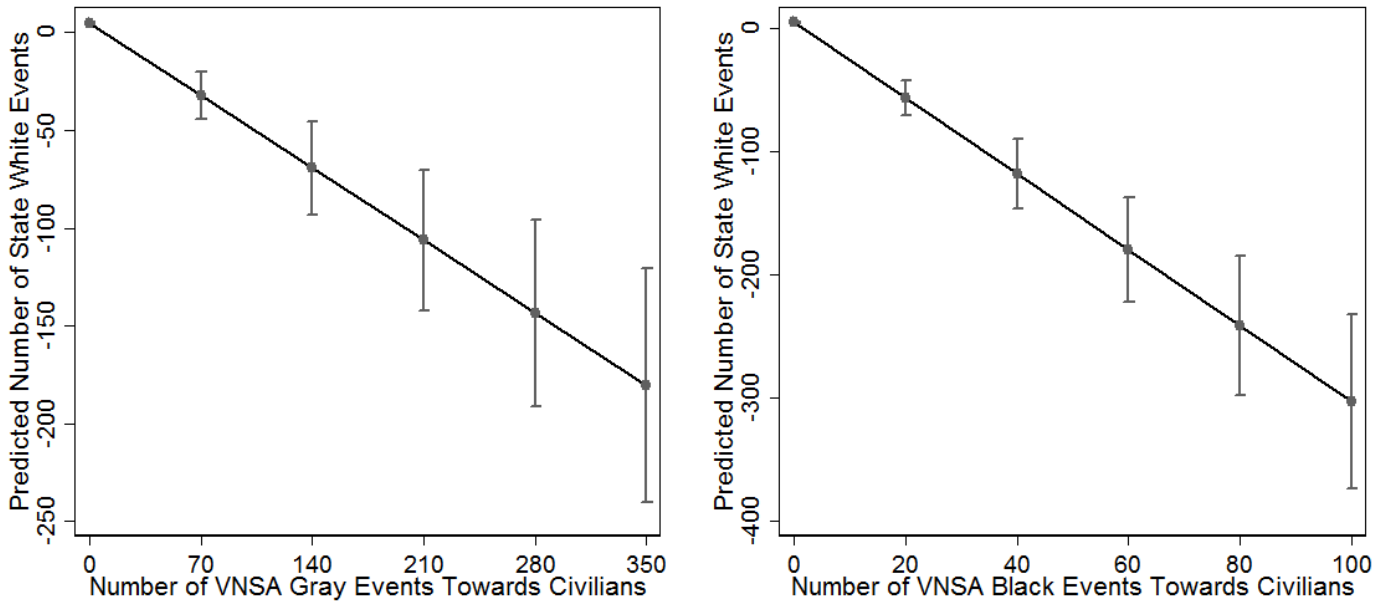
We similarly observed that for an increase of 100 Gray Zone events by civilians against VNSAs in a given district-month, VNSAs perpetrated 1,492 more Gray Zone events in the following district-month. Moreover, the same increase of 100 Gray Zone events by civilians against VNSAs resulted in 2,016 additional Black Zone events undertaken by VNSAs in the subsequent district-month. Before continuing, it is worth noting that the exceedingly large effect sizes are likely a byproduct of the difficulties distinguishing civilians and VNSA perpetrators of Gray Zone events given the often biased nature of media coverage of this conflict. It is also worth noting that these results are consistent with the Bayesian analysis, which shows that civilian-initiated events increase the probability of an event being Gray. Figure 28, below, plots the impact of Gray Zone use by civilians against VNSAs on Gray and Black actions by VNSAs.

Figure 28: Effect of Civilian Gray Zone Events on VNSA Gray and Black Actions



On the other hand, our results also suggest that states respond to civilian victimization by VNSAs by escalating. We noted that both Gray and Black Zone VNSA activity directed at civilians results in lower White Zone activity by states. Specifically, an increase of 100 Gray Zone events by VNSAs directed at civilians in a given district-month is correlated with a drop of 52 White Zone events by states in the subsequent district-month. Similarly, an increase of 100 Black Zone events by VNSAs against civilians in a given district month is associated with a drop of 308 White Zone events by states in the following district-month. Figure 29, below, plots the impact of Gray and Black Zone activity by VNSAs towards civilians on the White Zone behavior of states. As the figure shows, states respond to VNSA-civilian interaction with escalation of their own. Again, civilian victimization may provide state forces with a pretext for escalation against VNSAs. Taken together with the evidence of VNSA escalation against civilians, these findings paint a rather stark picture of civilians being caught in the middle of this conflict.

Figure 29: Effect of VNSA-Civilian Events on State Actions



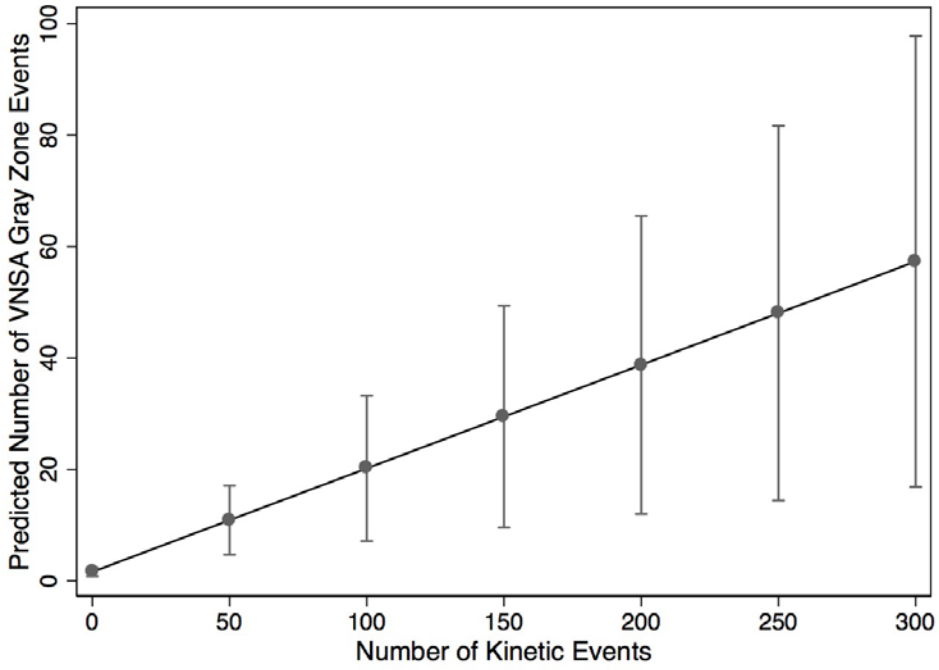
Libya

While the previous section presented all of the Libya findings, this section is devoted to further exploration of a key finding. Specifically, we observed that kinetic events also drive increased non-kinetic activity. However, this effect appears to be limited to White and Gray Zone events.

The Importance of Kinetic Events

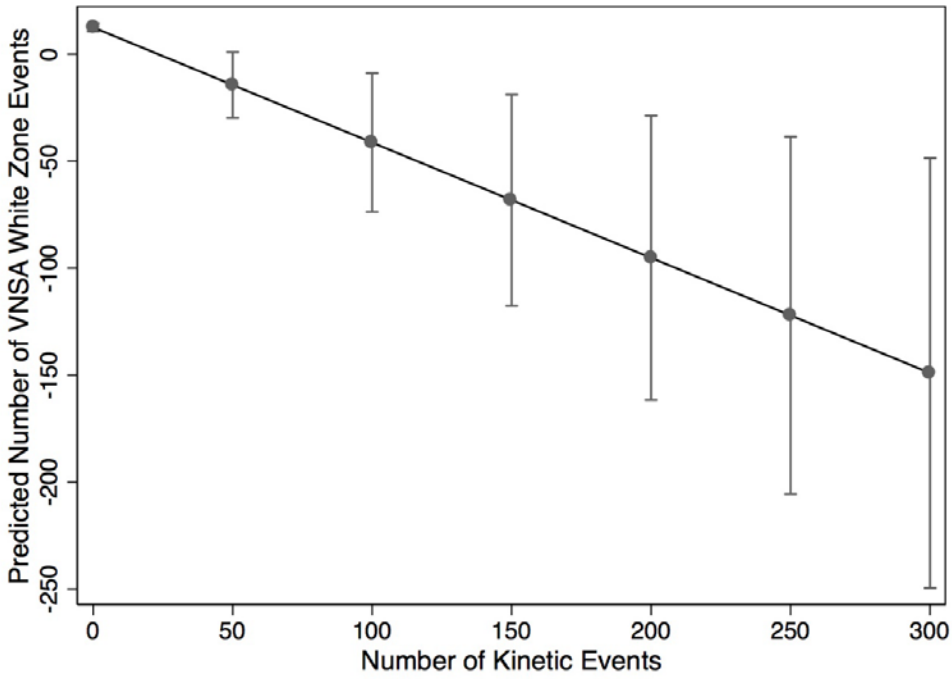
Our results demonstrate that kinetic events effect both kinetic and non-kinetic events. Specifically, kinetic activity increases the likelihood of Gray Zone actions that are both kinetic and non-kinetic in nature by VNSAs across all models. Substantively, our models predict that an increase of 100 kinetic events in a given departmental-month is associated with 36 additional Gray Zone events perpetrated by VNSAs in that department during the subsequent month. Figure 30, below, charts the impact of kinetic events on the Gray Zone behavior of VNSAs. As the figure shows, kinetic events drive higher levels of Gray Zone behavior by VNSAs.

Figure 30: Kinetic Events and Gray Zone Activity



In addition, the same increase of 100 kinetic events, in an average departmental-month, results in a reduction of 38 White Zone events by VNSAs in the following departmental-month. Figure 31, below, graphically depicts this result.

Figure 31: Kinetic Events and White Zone Activity



However, we did not observe a change in the amount of VNSA-initiated Black Zone activity as a result of increased kinetic actions.

Conclusion

This research yielded three primary conclusions that are consistent across both frequentist and Bayesian statistical approaches and that appear in multiple, distinct cases of Gray Zone conflict. First, despite the commonly held belief that Gray Zone conflicts rely more heavily on the diplomatic, information, economic, financial, intelligence and legal versus the military instruments of power, it appears that kinetic military operations have an outsized effect. While kinetic operations represent only about a fifth of all observations, they heavily influence non-kinetic activity. Specifically, in Colombia, kinetic events correlated with subsequently increased activity in the White, Gray and Black Zones. Similarly, in Ukraine kinetic activity appeared to be driving both more White and more Gray Zone events. Finally, in Libya, kinetic observations corresponded to increased Gray Zone actions and decreased White Zone events in the following month. While the way in which different types of kinetic actions shape both kinetic and non-kinetic responses is beyond the scope of this investigation, it is clear that kinetic events are highly consequential in all three cases.

Second, in both Colombia and Ukraine, we found VNSAs to be less adept than states at interpreting their adversaries' moves towards (de-)escalation. The primary difficulty occurred in shifts from the White to the Gray Zone (escalation) and vice versa (de-escalation). Shifts to and from the Black Zone tended to be far less ambiguous, and thus easy to correctly interpret. Nevertheless, in Ukraine these asymmetries were sharply reduced relative to Colombia. This is not surprising given that, on average, VNSA forces tended to be much more closely linked to the (Ukrainian or Russian) government than in Colombia.

Third, in both Colombia and Ukraine both VNSA and state forces appear to be concerned about domestic and/or international legitimacy. For example, in Colombia VNSAs de-escalate when states uses predominately Gray approaches to targeting them, perhaps in order to avoid being seen as the aggressor if they are relying on predominately Black (easily observable) activities and states is not. However, as states inflicts violence on VNSAs in an overt manner, they respond in kind. Similarly, in Ukraine, numerous belligerents have proven willing to victimize civilians. However, no side wants to appear to be disproportionately targeting civilians. As such, if state forces begin curbing or at least obscuring their victimization of civilians, VNSAs respond in kind.

Appendix A: Methodology

Introducing the Data

In order to understand how states and VNSAs behave in Gray Zone conflicts, we leverage Lockheed Martin Corporation's Integrated Crisis Early Warning System (ICEWS) data. The data is structured at the event-actor level, with each observation providing detailed information on each actor involved in a given event. The ICEWS dataset includes the following variables:

- StoryID – a unique identifier for each news story coded in the dataset. Note that multiple events are often coded from a single story. These events all share the same StoryID, but each of them will have a unique EventID (see below).
- EventID – a unique identifier for each event-actor combination. Specifically, each observation in the dataset involves a distinct combination of actors involved in a single action. For example, a single altercation wherein insurgents targeted police and civilians would result in two distinct observations being coded, each with its own EventID. The first observation would code insurgents targeting police and the second would code insurgents targeting civilians. Both would share the same StoryID. As such, EventID – unlikely StoryID – is a unique identifier for each, distinct observation in the dataset.
- EventDate –the date of the event.
- PublicationDate – the date the story was reported. This does not always match the EventDate because, in some cases, news stories were published in the days following an event.
- Headline – the title of the news report being coded.
- Sentence – the relevant text of the news report from which the coding was based.
- SentenceNumber – the order of the sentence being coded. For example, if the action being coded is described in the third sentence of a report, the SentenceNumber would be recorded as 3.
- Publisher – the name of the publisher of the news report from which the observations were coded.
- Latitude – latitudinal coordinates for the event location.
- Longitude – longitudinal coordinates for the event location.
- EventType – each event was assigned an EventType category by ICEWS, based on the CAMEO (Conflict and Mediation Event Observations) coding scheme developed by Schrodt and Yilmaz (2007).[15] Thus, each event in the ICEWS data is coded as one of 256 CAMEO categories that range from cooperative to hostile actions. All CAMEO categories are listed in Appendix C.[16] For example, EventTypes included actions like "Use conventional military force," "make statement" and "arrest, detain or charge with legal action".
- Score – the CAMEO score for the EventType, ranging from -10 (most hostile) to 10 (most cooperative).
- SourceCountry – the country from which the event/action was initiated.
- SourceSectors – the category of the actor initiating the event. For instance, categories include "dissident," "rebel," "government," "civilian" and "national party."

[15] The codebook for the CAMEO coding scheme can be found at: http://eventdata.parusanalytics.com/cameo.dir/CAMEO.CDB.09b5.pdf.

[16] The ICEWS datasets also included EventTypes that are not in the basic ICEWS codebook. This issue is discussed in greater detail in the *Automated Coding* section below, and *all* EventType categories (those in the original ICEWS codebook and those that are not) are listed in Appendix C.

- Source – the name of the actor initiating the event.
- TargetCountry – the country targeted by the event/action.
- TargetSectors – the category that the target actor belongs. The categories are identical to those used for the SourceSectors variable.
- Target – the name of the actor targeted by the event/action.

Given our interest in Colombia, Libya and Ukraine, we obtained the ICEWS data for these countries for the aforementioned time periods. However, we also procured data for Cuba from 28 January 2005 through 27 September 2016, in order to incorporate events relating to peace negotiations between two distinct VNSAs and the Colombian government, which were hosted in Cuba during this time. Second, given that Russia precipitated the Ukrainian crisis and remains extensively involved, we also obtained Russian data from 02 January 2014 to 30 June 2016. Observations from these two datasets that did not involve Colombia and Ukraine, respectively, were dropped. While other actors are at least tangentially involved in all three conflicts, severe data quality limitations coupled with practical considerations precluded including all but these two highly consequential third parties.

Implementation Plan: Definitions, and Process Preview

In order to conduct our analysis, we took raw ICEWS data and then cleaned and transformed it into key variables of interest. First, we created a trifold typology of actor-type. That is, we defined actors in our dataset as belonging to one of three categories: government, VNSA and civilian. Each actor in each observation was assigned to one of these three categories.

Second, following the definition from the project background, we classified Gray Zone activities as those which meet any of three criteria:

1.) Exceeds "ordinary competition" [e.g. peaceful economic competition such as increasing oil prices, as opposed to threatening to cut off crucial natural gas supplies to specific countries and irrespective of price] yet falls below the threshold of large-scale direct military conflict.
2.) Suffers from problems with attribution, and where the activity is ambiguous.
3.) Undermines/violates international norms/laws.

By contrast, White and Black Zone activity do not meet any of these three criterion. In determining whether an action was White versus Black, the first component of the definition is critical. White Zone activities do not exceed "ordinary competition," whereas Black Zone actions exceed this threshold and the threshold of large-scale direct military conflict. We reviewed the CAMEO codebook and assigned each type of activity as being White, Gray or Black accordingly. However, we also found that some activities were ambiguous, and depending on the context, could be either White or Gray, or they could be Gray or Black. These activities were automatically coded as White/Gray, or Gray/Black and marked for hand-coding (see below). Further details on definitions and coding of the variables, summary statistics as well as our assumptions for each variable are outlined in the *Automated Coding* section below.

Because this data was originally collected in an entirely automated manner, one of the risks was that the events were coded incorrectly. To the extent that this was the case, automating the transformation of the

raw ICEWS data into usable variables for this study might have systematically biased our results. In order to test whether our automatically generated variables were correctly coded, we selected random samples of the data from the five countries. We then had a research assistant hand-code each variable for all observations in the random sample. We did this so that we could test how accurate the automated procedure we developed was. In other words, how much did the automated coding match coding done by a human coder? After the manual hand-coding was complete, we conducted a series of inter-coder reliability (ICR) tests, in order to estimate the degree of agreement between our automated procedure and a trained researcher. Our first test involved using percentage agreement, asking simply how often the hand-coder agreed with the machine's coding. Second, we used more sophisticated measures for inter-coder reliability, using Krippendorff's Alpha (K-alpha) coefficient to measure levels of agreement between the automated process and hand-coded sample. Details on the methodology behind the ICR process, along with the results, are detailed in the *Round 1 ICR Test Results* section below.

Once we obtained an acceptable level of agreement between the machine and human coders, we moved into the next phase of the analysis. Specifically, a key issue with the datasets generated from the raw ICEWS data was that a substantial number of observations had missing source and target data, and/or ambiguous Zones (e.g., White/Gray or Gray/Black). Details on this segment of the data are provided in Table A.1 below. Instead of ignoring this missing data, we randomly selected a sub-sample of it and employed hand-coders to review these observations and fill in source, target and Zone information manually. Further details on the sample selection strategy, the training of coders and the hand-coding process are outlined in the *Hand-Coding Missing Data* section below.

In addition, to ensure consistency between the different coders and cases, we assigned each coder a main sample to hand-code and a smaller overlapping sample that we used to generate a second round of ICR scores. These results are outlined in the *ICR Round 2 Results* section below.

Once the coders had manually corrected the sample of missing data, we merged this sample back into the dataset for the country to which the sample belonged and dropped the remaining missing data. Subsequently, we tackled an additional problem that affects event-level datasets generated through news reports, duplicate news stories. Duplication of events results in a large number of redundant observations in these datasets. As a result, we used a variety of sampling and hand-coding strategies to estimate the degree of duplication in our datasets. We explain the issues with duplicate stories and our strategy for correcting such issues in the *Accounting for Duplicate Events* section below. We also came across a variety of methodological issues with ICEWS data. While we could deal with some of them, we could necessarily account for all of them. We have listed these issues, as well as our approach to dealing with some of them, in the *Methodological Issues with ICEWS Data* section.

Finally, once all hand-coded datasets (for missing source, target and Zone information, as well as de-duplication datasets) were correctly merged back in, we generated new variables for analysis that included weights based on the random samples. Specifically, the hand-coded observations were weighted

in order to compensate for the loss of the dropped data. The *Weighting Strategy for Key Variables and Duplicates* section below provides a detailed discussion of the weighting approach we employed.

Automated Coding: Key Variables and Assumptions

For the purposes of our project, we generated 13 variables from the ICEWS data:

1. Sourcetype – Nominal. Denotes the identity of the perpetrator of the event or action. It can be one of three categories: "Government," "Violent Non-State Actors" or "Civilian."
2. Targettype – Nominal. Denotes the identity of the target of the event or action. It can also be one of three categories: "Government," "Violent Non-State Actors" or "Civilian."
3. White – Binary. Coded as 1 whenever the event or action constitutes White Zone activity, 0 otherwise.
4. Gray – Binary. Coded as 1 whenever the event or action constitutes Gray Zone activity, 0 otherwise.
5. Black – Binary. Coded as 1 whenever the event or action constitutes Black Zone activity, 0 otherwise.
6. White/Gray – Binary. Coded as 1 whenever the event is ambiguous between White and Gray, 0 otherwise. That is, it is 1 whenever the event cannot automatically be coded as White or Gray based on our automated procedure. These cases were flagged for review and hand-coding by a trained researcher. Consequently, the final dataset utilized in the analysis does not include White/Gray observations.
7. Gray/Black – Binary. Coded as 1 whenever the event is ambiguous between Gray and Black, 0 otherwise. That is, it is 1 whenever the event cannot automatically be coded as Gray or Black based on our automated procedure. These cases were flagged for review and hand-coding by a trained researcher. Consequently, the final dataset utilized in the analysis does not include Gray/Black observations.
8. Statesource – Binary. Coded as 1 whenever the source of the event or action is a government, 0 otherwise. This includes any individual government (and not just the government of the country being studied), a coalition of multiple nations or international governmental organizations (IGOs). While IGOs are not technically governments, they are comprised of and responsible to governments. As such, we felt this coding was justifiable.
9. Statetarget – Binary. Coded as 1 whenever the target of the event or action is a government, 0 otherwise. We defined government in the same manner as for Statesource (above).
10. HNSource – Binary. Coded as 1 whenever the source of the event or action is the country of the case study in question, 0 otherwise. For example, for the Colombia case study, this means 1 whenever the source is Colombian, 0 for foreign sources. However, since Cuba and Russia were only included in the study given their relevance to Colombia and Ukraine, respectively, and because the relevant data from these two countries was merged with the data for Colombia and Ukraine, respectively, HNSource was coded as 1 in these cases if the country that was the source of the event or action was Colombia (and not Cuba) or Ukraine (and not Russia), 0 otherwise.
11. HNTarget – Binary. Coded as 1 whenever the target of the event or action is the country of the case study in question, 0 otherwise. For example, for the Colombia case study, this means 1 whenever the target is Colombian, 0 for foreign targets. However, since Cuba and Russia were only included in the study given their relevance to Colombia and Ukraine, respectively, and because the relevant data from these two countries was merged with the data for Colombia and Ukraine, respectively, HNSource was

coded as 1 in these cases if the country that was the target of the event or action was Colombia (and not Cuba) or Ukraine (and not Russia), 0 otherwise.

12. Kinetic – Binary. Coded as 1 whenever the event or action involves the use of force, 0 otherwise.
13. Civtarget – Binary. Coded as 1 whenever the target of the intended event or action is civilian, 0 otherwise. Civilian collateral damage is coded as a 0. For example, if the source text indicates that two civilians were caught in the crossfire between the military and insurgents, this variable is coded as 0. While civilians were victimized, they were clearly not the intended victim.

First, we tried to code the actors involved in the event, that is, the source and target of the action or event. In order to code the sourcetype and targettype variables, we leveraged the same keywords used by ICEWS for classifying source, sourcesectors, target and targetsectors variables. Each of these variables used a variety of keywords to identify the perpetrator and target of the action or event. Using these keywords, we made a determination as to whether the source and target were government, VNSA or civilian. Media and media personnel were coded as civilian unless explicitly identified as state-owned media or officials thereof. Overall, the majority of the keywords used were universal across the country cases. However, some keywords were unique to each country. For example, "Rebel" was a keyword for VNSAs across all countries, but "Bloque" was a unique keyword for VNSAs in Colombia (given that the FARC refers to different detachments of fighters as "Bloques" or Blocks). This sometimes included names of specific actors. For example, "Dmytro Yarosh" was a keyword for VNSAs in the Ukraine case. Yarosh is the leader of the far-right militant group *Right Sector* in Ukraine (Shuster 2014). Similarly, "Castro" was a keyword for government in Cuba, and "Ansar al-Sharia" was a keyword for VNSAs in Libya. All keywords used to identify the three sourcetypes and targettypes are listed in Appendix B.

Second, we tried to assign a Zone coding to each action or event in the data. In order to code White, Gray and Black Zones, we used the EventType category from the ICEWS data. As explained in the *Implementation Plan* section above, this category has been populated using the CAMEO coding scheme. Each event in ICEWS is coded as one of 256 CAMEO categories that range from cooperative to hostile actions. We assigned a Zone to each CAMEO category – the Zone coding for all CAMEO categories is available in Appendix C. However, 19 of those categories are ambiguous – 11 categories could be either White or Gray, and eight categories could be either Gray or Black. All of these categories were instead given an *ambiguous Zone* coding; either White/Gray for events that might be White or Gray or Gray/Black for events that might be Gray or Black. We selected a stratified random sub-sample of these cases, and then employed researchers to hand-code these ambiguous Zones manually. Table A.1 below summarizes the extent of ambiguous Zones per country, and the *Hand-Coding Missing Data* section below details the training and hand-coding of our sub-sample of ambiguous Zones.

Third, we identified several CAMEO-like categories in the ICEWS data that were being used in the country datasets that did not exist in the main codebook. All of these categories were manually added and assigned a Zone classification. These manually added categories are also listed in Appendix C, along with the basic CAMEO categories. Finally, we assigned a *kinetic* coding to all CAMEO categories that involved the active use of force. All categories coded as kinetic are also listed in Appendix C.

While our automation allowed us to code a majority of the data, we have missing data where the source and target were not clear from the keywords. We also have missing Zone codings where the Zone is ambiguous. Table A.1 below provides summary statistics for the various variables we have automatically coded, and previews the extent of missing and ambiguous data that remained after running our automated re-coding procedure.

Table A.1: Summary Statistics and Missing Data

	Colombia	Ukraine	Libya*	Russia	Cuba
Source Type					
Government	38,787	30,899		15,704	314
Violent Non-State Actor	20,551	3,726		795	47
Civilian	8,070	8,762	8,377	1,059	82
Unknown Source	10,467	9,635		356	525
Total	*77,875*	*53,022*		*17,914*	*968*
Target Type					
Government	30,899	30,261		15,475	424
Violent Non-State Actor	16,064	3,051		611	388
Civilian	17,443	10,222	8,034	1,349	31
Unknown Target	13,469	9,488		479	125
Total	*77,875*	*53,022*		*17,914*	*968*
Zone					
White	52,356	37,150	25,806	14,170	932
Gray	11,860	4,751	3,254	1,032	24
Black	5,630	3,598	4,334	1,478	2
White/Gray	6,587	4,419	3,766	1,044	10
Gray/Black	1,442	3,104	6,101	190	0
Total	*77,875*	*53,022*	*43,261*	*16,680*	*968*

*For Libya we only have summary statistics for civilian source and target data. The below section, *Additional Caveats and Assumptions*, explains why we did not code government and violent non-state actor sources and targets.

Additional Caveats and Assumptions

In the Libya case, we did not code government and violent non-state actor source and targets. Because of the fluid nature of the conflict on the ground, and the fact that there are multiple local entities claiming to be the legitimate government of Libya, there is a great deal of ambiguity over which groups should be considered government and which should be considered violent non-state actors. Additionally, there have been groups claiming to be governments that existed earlier but no longer meaningfully exist. Instead, they have morphed into new entities, and it is hard to determine whether these can be considered the "government" or are better coded as violent non-state actors now. Finally, it was often

difficult to precisely identify the actors involved from the text (sentences and headlines) provided in the ICEWS data because the descriptions were too vague. For example, many incidents would be described as "gunmen shooting dead four men [in attack]" (sic).[17] It is also not usually possible to determine whether any given news source that refers to "the government" is actually referring to the internationally recognized government at the time or not. While we could not parse government versus VNSAs, we were able to identify civilians based on the same keywords as the other country datasets. The keywords used for identifying civilians are discussed in Appendix B.

There is a second exception to the Libya case. For all other countries, the use of "unconventional violence" was coded as Gray Zone conflict, because the tactics that ICEWS coded as such tended to violate international laws and norms. However, in the Libya case we coded the use of "unconventional violence," as assigned by ICEWS, as ambiguous between Gray and Black. Unlike the other four cases, many events that were coded as "unconventional violence" in Libya were actually conventional in nature. For example, many of these observations actually involved large-scale direct conflict, such as protracted gun battles between organized troops, massive artillery fires or sieges. These types of tactics align much more closely with our conceptions of Black Zone actions. In other instances, small-scale and unorganized events were coded as "unconventional violence." Samples of these ambiguous cases were then selected for hand-coding.

We also made a few other assumptions in the automated coding. All observations that involved governments providing material support to VNSAs were coded as occurring in the Gray Zone. Additionally, in the cases of Russia and Ukraine, we observed that a substantial share of the source and target variables were listed as simply "Russia" or "Ukraine" for source and target, and "/N" for SourceSectors and TargetSectors, thus providing insufficient context to automatically code the identity of the actors. In the Ukrainian data, 17.9 percent (9,477 of 53,022 observations) simply listed "Ukraine" as the source, and 7.2 percent (3,796 of 53,022 observations) listed "Russia" as the target. In the Russian data, 45.1 percent (8,082 of 17,914 observations) listed "Russia" as the source and 53.6 percent (9,599 of 17,914 observations) listed "Ukraine" as the target.

In these cases, we suspected that the use of the country names was meant to identify government actors (that is, the Russian and Ukrainian governments, respectively). But we could not make this assumption outright without the risk of introducing bias if we were wrong. In order to test this assumption, we generated random samples of 100 observations each from the Ukrainian and Russian datasets, to see how many observations of this format were indeed referring to states. We found that in the Ukrainian data 100 percent of the observations that listed the source as "Ukraine" and 100 percent of the observations that noted "Russia" as the target, referred to the Ukrainian and Russian governments, respectively. In the Russian data, 96 percent of observations that listed "Russia" for the source, and 99 percent of observations that listed "Ukraine" for the target referred to the Russian and Ukrainian governments, respectively. Based on this, we felt justified in adopting this assumption and coding this data accordingly.

[17] EventID 21026867 in the Libya dataset.

Round 1 ICR Test Results

Our next step was then to run ICR tests between the coding done utilizing our automated approach and a trained human coder, to confirm that our automation procedure was correctly coding events. After that, we generated a random sample of the missing data, to be coded manually.

Recall that we used keywords to identify the source and target, and used general EventType categories to denote White, Gray and Black Zone activity. There are two potential risks associated with this strategy. First, the keywords being used by the ICEWS data to code the source and target may be inaccurate; that is, the sentence and headline may suggest a source or target different from the one coded in the source and target variables. Second, the EventType coded by ICEWS may be inaccurate, with the sentence and headline implying a different event type. Both of these types of errors would affect our analysis.

Ideally, we would read through every observation manually and ensure that the coding was accurate. However, this option is not practical. Thus, the alternative we pursued was to estimate how likely it was that we were incorrectly coding the source, target or Zone variables. This would then undergird sensitivity analysis. In order to do this, we ran ICR tests to estimate the level of agreement between hand-coded and automatically generated observations for all key variables in a sub-sample of the data.

First, we generated random samples of 100 observations each for all countries. Second, we had a coder hand-code every observation for all variables. In other words, they read through all ICEWS generated variables, in addition to the sentence and headline, and then coded the *sourcetype, targettype*, *Zone*, whether or not a state was the source and target, etc. Third, we then generated automated codings, which took ICEWS' initial codings at face value, for the same variables. Fourth, we estimated the degree of agreement between the hand-coded and automated observations. However, we dropped two types of observations from the ICR analysis:

a) Observations that were erroneous in the raw ICEWS data. We marked observations as erroneous if not even the event-type was coded correctly based on the sentence and headline
b) Observations that implied ambiguous Zone activity. That is, where it was not clear whether the activity was White or Gray, or Gray or Black simply from the EventType, since these observations were going to require hand-coding regardless of other errors' existence.

There are various measures of ICR used in the social sciences. For our purposes, we calculated ICR in two ways. First, we used a simple percentage of agreement measure. However, some scholars argue that percentage agreement alone is not enough and that researchers should use more sophisticated methods for estimating ICR (see for example, Neuendorf 2002). Consequently, we used K-alpha reliability coefficients, which have emerged as a gold standard (Joyce 2013). K-alpha ranges from 0 to 1, with 0 indicating perfect disagreement and 1 indicating perfect agreement. It is important to note that the coefficient penalizes lack of variation and other factors beyond simple agreement and disagreement. As such, a coefficient of 0.90 should not be interpreted as indicating that 90 percent of observations were coded the same in both samples. Because other factors were likely penalized, a coefficient of 0.90 could be indicative of more than 90 percent agreement in coding across samples. Similarly, a coefficient of 0

may indicate perfect disagreement, but it could also be returned even in cases of perfect agreement if there is no variation in a variable across all coders. According to Krippendorff, "[w]hen all coders use only one category, there is no variation and hence no evidence of reliability" (2004, 425). In other words, if there is no variation in the variable, the algorithm assumes that the variable is unreliable and scores it 0[18] (Krippendorff 2011). But this may not always be the case. As such, a bit of care is warranted when interpreting K-alpha coefficients.

So how should coefficients for K-alpha be interpreted? Krippendorff (2004) indicates that 0.66 is the minimum acceptable standard for ICR, and coefficients above 0.80 are considered extremely reliable (Neuendorf 2002; Krippendorff 2004; Joyce 2013). The results of the first round of ICR tests for all countries are reported in Tables A.2 and A.3 below.

Table A.2: ICR Round 1: Percentage Agreement between Our Automated Approach and Hand-coders

Variable	Colombia	Ukraine	Libya	Russia	Cuba
White Zone	89.4	86.3	82.9	88.5	94.7
Gray Zone	95.7	87.4	90.0	79.5	97.3
Black Zone	96.8	92.6	91.4	98.7	100
State Source	90.4	85.3	*n/a*	96.2	36.0
State Target	89.4	83.2	*n/a*	92.3	84.0
Host Nation Target	95.7	94.7	95.7	98.7	100
Kinetic	95.7	94.7	98.6	98.7	100
Civilian Target	97.9	92.6	81.4	96.2	98.7
Erroneous	*8*	*14*	*30*	*22*	*25*

Table A.3: ICR Round 1 K-alpha Degree of Agreement between Our Automated Approach and Hand-coders

Variable	Colombia	Ukraine*	Libya	Russia	Cuba
White Zone	0.75	0.92	0.81	0.68	0.00
Gray Zone	0.83	0.80	0.73	0.50	0.00
Black Zone	0.81	0.93	0.94	0.92	1.00
State Source	0.80	0.71	*n/a*	0.75	0.00
State Target	0.85	0.72	*n/a*	0.66	0.92
Host Nation Target	0.96	0.95	0.92	0.00	1.00
Kinetic	0.89	0.91	0.97	0.93	1.00
Civilian Target	0.91	0.82	0.78	0.65	0.00
Erroneous	*8*	*14*	*30*	*22*	*25*

*For Ukraine, we conducted a second round of ICR tests for the Zonal coding. This was necessary because the coder struggled with the Gray Zone concept, which forced us to sharpen our definitions and adjust coder-training. The second round of results are reported for Ukraine. Ukraine was hand-coded first, by

[18] As discussed below, a 1 may occasionally be coded in the event that there is no variation.

the first hand-coder we hired. As such all other coders and codings benefited from the sharpened definition and revised training from the outset.

As the tables above indicate, there is a high degree of agreement between the machine and the manual coders on our variables in the three key country cases – Colombia, Ukraine and Libya. The results of the percentage agreements are generally high, never falling below 80 percent for the main country cases. Black Zone activity is the most reliable, followed by White Zone, and then Gray Zone activity. As expected, the Krippendorff's Alpha scores provide more nuanced results. The highest degree of agreement seems to be on whether the host nation was a target or not, whether the event was a kinetic event or not and whether the event constitutes Black Zone activity or not. This intuitively makes sense, since those factors are usually highly visible. There is also a relatively high degree of agreement on whether an activity was White or Gray Zone activity or not. Given the nature of Gray Zone conflict, we believe this makes sense as well. In the Libya case, there are no ICR scores reported for state source and target, because we did not code source and targets for Libya. Finally, the number of erroneous observations, which were dropped, ranged from 8 to 30, with the Colombia data being the most reliable and the Libya data being the least reliable. On average, 18.2 percent of the randomly sampled data was erroneous.

The results for the ancillary country studies – Russia and Cuba – are different from the main cases. First, in these countries, there was no variation in many of the variables. In the case of Cuban Zones, almost all activity was diplomatic (again, Cuba was facilitating peace negotiations), and therefore, there was no variation on the White and Gray Zone variables. Since all activity was between governments and VNSAs, there were also no civilian targets. Host Nation Target and Kinetic also had no variation; they were both 0 for all observations. Because the data itself has no variation in all of these cases, the question of agreement or disagreement between the machine and coder is not important. The K-Alpha score of 0 or 1 here, therefore, is not a result of actual ICR between machine and coder, but an artificial result of the fact that this variable is always 0. The same is true for the Host Nation Target variable for Russia.

Additionally, most of the results for Russia are poor. This is because the Russia data was of relatively lower quality. Further, while almost all of the results are at or above the minimum acceptable threshold of 0.66, the Gray Zone variable has a K-alpha score of just 0.50. (Civilian Target also missed the 0.66 cutoff, but only by one hundredth of a point.) This is due to numerous discrepancies that emerged between the human coders and our automation procedure for Russian Gray Zone activities.

In our random sample, in 10 out of 72 observations (or 13.9%) the automated process coded Gray events (as identified by a human coder) as White.[19] In other words, we found that the automation procedure systematically underestimates the degree of Gray Zone activity in Russia. We did not find similar issues in any of the other cases, which leads us to believe that this was a product of Russian expertise at masking Gray Zone operations as White activities. Specifically, we found that the automated procedure tended to miscode Russian statements couched in diplomatic niceties as White, when in fact Russia was engaged in Gray Zone information operations or other Gray activities. For example, we found cases where Russian

[19] 28 observations were dropped because they were either erroneous, or suggested ambiguous Zone activity.

government officials were accusing Kiev of violating ceasefire agreements or conducting military operations against civilians but were coded as normal diplomatic criticism or the making of statements.[20] Similarly, we had instances of Russian military mobilization along the borders with Ukraine, which were coded as normal police activity in the dataset.[21] In all these cases, the machine coded these activities as White, when they should really be coded as Gray. Given Russia's mastery of Gray Zone information operations, this result is not all that surprising.

Improving our automated coding approach to adjust for this is not possible given that these errors are largely a result of ICEWS' automated coding scheme, which sometimes but not always miscoded the EventType for these types of observations. Unfortunately, we could not identify patterns as to when and why ICEWS coded these types of events incorrectly. Manually reviewing every possibly affected observation was also not feasible. Our first-cut analysis of random samples from Russia suggests that approximately 14 percent of White Zone activity was in fact Gray Zone activity. Fortunately, because we know the direction and magnitude of the bias, we can easily conduct sensitivity analysis. Specifically, we re-ran our analyses after making the assumption that five percent of White Zone activity was in fact Gray (and vice versa) to determine if changing the distribution of White and Gray Zone activity altered the findings.

Hand-Coding Missing Data

Table A.1 above shows the level of missing source/target data and ambiguous Zones affecting each case. Ideally, we would manually read through each observation and code the missing source, target and Zone information. Since this is not practical, we were forced to come up with alternatives. One alternative strategy would be to drop the missing data entirely. However, because certain event types are systematically more likely to be missing data, simply dropping these observations would systematically bias our analysis. Consequently, we chose to hand-code a subset of the missing data. In our analysis, we then weighted this hand-coded data, such that it can substitute for the rest of the missing data. We used a stratified random sampling strategy to select a sample of missing data. For each country, we generated two random samples:

1. We generated a sample of 5,000 observations with missing source, target, ambiguous Zones and/or any combination thereof. We selected an n of 5,000 because we wanted to maximize the number of observations with missing and ambiguous data that we could correct manually, given practical constraints. In Colombia, which had the largest dataset at over 77k observations, a random sample of 5,000 observations with missing information represents over 15 percent of all observations with missing data. Table A.4 below provides a snapshot of the size of the random sample, relative to the total population of missing data for all countries. Given how relatively large our hand-coded random samples were, we are confident that our random samples are representative of the overall population of missing data in the datasets.

[20] EventIDs 21238204 and 21259722, respectively, in the Russia dataset.
[21] EventID 21168232 in the Russia dataset.

2. Out of this sample, we created a second random sub-sample of 500 observations (or 10 percent of the sample). This was created for the purpose of conducting ICR tests, to ensure consistency across human coders.

Each coder was assigned the full sample for one country, and a second sub-sample for a different country, so that we could estimate and ensure ICR. The random sampling strategy for generating our samples is explained in detail below.

Sampling Strategy for Hand-Coding

For each country, we generated a sample of data to code based on the missing categories. First, we identified all observations with missing source and/or target data and/or ambiguous Zone codings. We oversampled observations with multiple missing or ambiguous variables, in order to maximize the impact of human intervention on the quality of the data. Though we were sure to also include numerous observations where only one variable was missing or ambiguous. Second, we selected a random sub-sample of these observations, stratified by source, target, White/Gray and Gray/Black. For each strata, we picked a sample that was proportional to the total population of missing data. That is, if 15 percent of the missing data entailed a missing source, then 15 percent of the sub-sample to be hand-coded would focus on fixing missing sources.

To better understand our approach, take Colombia as an example. In Colombia's case, 2,902 observations contained missing or ambiguous information for multiple variables. We hand-coded all of these, since fixing these observations enabled between two to three missing or ambiguously coded variables to be corrected in roughly the same amount of time as observations with only one missing variable. Since we decided to hand-code 5,000 observations for each country, this left us with the task of creating a second sub-sample of 2,098 observations with missing or ambiguous data for a single variable. We selected observations stratified by source, target, White/Gray and Gray/Black, such that the proportions of observations missing one variable versus another in the sub-sample were identical to the distribution of missing or ambiguous variables in the country dataset from which the sub-sample belonged.

In the Colombia dataset of 77,875 total observations, 31,665 observations had missing data – and within that 75.6 percent (23,936 observations) had missing source or target data, and 24.4 percent (7,729) entailed ambiguously coded Zones. Thus, 75.6 percent of the 2,098 sub-sample, or 1,573 observations needed to be a random sample of observations with missing source/target data, and 24.4 percent of 2,098, or 525 observations needed to be a random sample of observations with Zones marked White/Gray or Gray/Black.

Second, in the Colombia dataset, of all 23,936 observations with missing source or target data, 43.7 percent (10,467) were missing source data and 56.3 percent (13,469) were missing target data. Therefore, we ensured that 43.7 percent of the 1,573 observations in the sub-sample or 687 observations were a random sample of events with the source missing, and 56.3 percent of 1,573 or 886 observations were a random sample of events with the target missing. Similarly, in the Colombia dataset, of all 7,729 observations with ambiguously coded Zones, 82 percent were White/Gray (6,587 observations), and 18

percent were Gray/Black (1,442 observations). Therefore, 82 percent of 525 observations in the sub-sample, or 430 observations, was a random sample of events that were ambiguous between White and Gray, and 18 percent of 525, or 95 observations, was a random sample of events that were ambiguous between Gray and Black. This created a Colombia sample for hand-coding with the following characteristics:

1. *Sub-sample 1*: 2,902 observations with multiple missing or ambiguously coded variables. This represents the entire population of data with more than one missing or ambiguous variable.
2. *Sub-sample 2*: 2,098 randomly selected observations with at least one missing variable of interest:
    a. 886 observations with missing target information.
    b. 687 observations with missing source information.
    c. 430 observations with a White/Gray ambiguous Zone.
    d. 95 observations with a Gray/Black ambiguous Zone.

We followed a similar strategy for Ukraine and generated two sub-samples for hand-coding. In Libya's case, recall that we did not code sources and targets. As a result, the Libya sample was only stratified on the two types of ambiguous Zones. In the full Libya dataset, of 43,261 observations, 9,867 had ambiguous Zones. Of these, 61.8 percent (6,101) were White/Gray, and 38.2 percent (3,766) were Gray/Black. Thus, 61.8 percent of the 5,000 sample for hand-coding, or 3,090 observations were a random selection of events that were ambiguous between White and Gray Zone. And 38.2 percent of 5,000 observations, or 1,910 were a random selection of events that were ambiguous between Gray and Black Zone.

In the case of Russia, of the 17,914 observations in the Russia dataset, only 1,967 observations had missing data or ambiguous Zones. Therefore, we decided to hand-code all of them. Similarly, the Cuba dataset was small enough that we could hand-code all missing data. Of the 968 observations in the Cuba dataset, 580 were missing data or had ambiguously coded Zones.[22] Table A.4 below provide a breakdown of the hand-coded samples for each country.

Table A.4: Hand-Coded Samples by Country

Hand-Coded Samples	**Colombia**	**Ukraine**	**Libya**	**Russia**	**Cuba**
Sub-Sample 1					
Hand-coded Multiple Missing or Ambiguous Variables	2,902	3,221	*n/a*	1,967	580
Population of Multiple Missing or Ambiguous Variables	2,902	3,221	*n/a*	1,967	580
Sub-Sample 2					
Hand-coded Missing Target Only Variables	886	638	*n/a*	*n/a*	*n/a*
Hand-coded Missing Source Only Variables	687	639	*n/a*	*n/a*	*n/a*
Hand-coded Ambiguous White/Gray Only Variables	430	295	1,910	*n/a*	*n/a*
Hand-coded Ambiguous Gray/Black Only Variables	95	207	3,090	*n/a*	*n/a*
Total Sample Size	*5,000*	*5,000*	*5,000*	*1,967*	*580*
Population with Missing or Ambiguous Variables	*31,965*	*26,608*	*9,867*	*1,967*	*580*

[22] As explained previously, the Russia and Cuba datasets are comparatively small because only the observations where the target was Ukraine and Colombia, respectively, were pulled out of the larger ICEWS data for those countries and time periods.

The Hand-Coding Process

We began training four human coders on the project by sharing our codebook and background documents on Gray Zones and the larger project. Each coder was assigned to one (or in one case, two) primary case(s) and to a separate (or in one case, two separate) ICR sample(s). The primary samples included 5,000 observations for the three primary cases and 1,967 observations for Russia and 580 for Cuba. We then drew random samples from within our primary samples that included at least[23] 10 percent of the observations, for the purpose of testing ICR. These ICR samples were then assigned to a coder that was not engaged coding the main sample for that country.

During the first week of hand-coding, we had the coders hand-code approximately 200 observations and then convened at the end of the week to discuss the coding, and any questions they had. That first group meeting resulted in the creation of a living document of rules and exceptions to the rules for coding, which was constantly updated as coding progressed. We also created a second living document specifically dedicated to rules for differentiating between Zones. Both of these documents are included in Appendix D and Appendix E. Subsequently, we would meet at the end of each week to ensure consistent progress and to discuss the particular cases where the coders were unsure of the correct coding. The number of questions raised during and the length of the meetings decreased precipitously as time progressed.

The coders were told to finish the first 500 observations from their main dataset and then the ICR sample before completing the remainder of their main datasets. Even though the ICR samples were randomly drawn from the main samples, we reordered the main samples to ensure that the 500 randomly selected observations, which were duplicated in the ICR samples, appeared first. This was intended to enable us to check for issues and correct any inconsistencies in the coding (fortunately none arose) through retraining, adjusting rules and re-coding erroneously coded observations, etc., before too much effort had been wasted. Of course, this means that the ICR scores do not account for inconsistencies due to coder fatigue or other issues (though we took other measures to ensure coding remained consistent such as limiting coders to 20 hours a week and consistently discussing the coding rules at our weekly meetings). However, we felt that the most issues would arise at the beginning (a trend which was subsequently confirmed in our weekly coder meetings) and as such this would constitute a harder test for ICR and preserve sufficient resources such that we would have been able to make adjustments had they been necessary.

As the ICR round two results presented in the *ICR Round 2 Results* section below shows, our hand-coding process yielded excellent results. The numbers suggest that the coders developed a high degree of shared

[23] The ICR samples for the three primary cases were 500 observations, or 10 percent. Because the Russia and Cuba samples were much smaller, we drew sub-samples for ICR that exceeded 10 percent of the main sample size. For Cuba, the ICR sample was 100 observations, which was roughly 17 percent of the primary sample (580), and for Russia, the ICR sample of 500 observations was roughly 25 percent of the primary sample for hand-coding.

understanding of what constitutes Gray (versus White or Black) Zone activity and how to identify the sources and targets for individual events.

Additional Caveats and Assumptions for Hand-Coding

The instructions given to coders for the purposes of coding Zones, and identifying missing source and targets are outlined in Appendix D and Appendix E. Observations that were erroneous, where the EventType is unsupported by the headline or the sentence, were marked as such and dropped before the analysis was performed. Observations where the source or target was incorrectly coded based on the headline or sentence were manually corrected. Events that were clearly unrelated to the conflict, such as soccer matches (which were commonly coded as kinetic events in Colombia given the martial language used to describe the games), protests unrelated to the conflict and crime that was not clearly related[24] were dropped as well. Protests that related to the conflict were included. If coverage of these protests did not mention a specific target, we assumed that they were targeted at the government. We also dropped events that posed entirely hypothetical scenarios that did not actually take place. For example, there were events in the Colombia dataset which were coded as "sign formal agreement" when the sentence described a potential peace agreement between the FARC and the Colombian government.[25] Because this agreement had not actually been signed at the time of the event, ICEWS miscoded the EventType. On the other hand, there were events in our dataset where the EventType was coded as an action, but there was ambiguity in the sentence and headline over whether the action had already taken place or not. In such cases, if we had evidence (from news sources or other events at the same time) that the action described by the EventType had indeed taken place, then we coded the observation, even though the description of the case in the sentence/headline was technically hypothetical. For example, the EventType in some cases in the Ukraine dataset was coded as "Use conventional military force" between Russia and Ukraine when the description in the sentence and headline suggested that violence was imminent because Russian troops were mobilizing or the opposing forces were setting up artillery.[26] However, nowhere in the sentence or headline do the reports claim onset of violence. Nonetheless, in these cases, we know from the news as well as other events coded at the same time that violence between Russia and Ukraine was indeed ongoing. Therefore, these are relevant cases, which we included.

We also dropped events that were based on RSS feeds and that lacked sufficient context. In other words, we dropped events that were based on a sentence or headline that was actually a group of headlines with no underlying details on any one particular event. Finally, we identified various rules for what constituted White, Gray and Black Zone activity, and how to differentiate between the Zones (please refer to Appendix E for further details).

[24] That said, organized criminal activity (like narcotrafficking in Colombia) that relates to the conflict was included. We particularly erred on the side of caution when coding narcotrafficking in Colombia given the importance of narcotics for fueling the conflict there.
[25] See, for example, EventID 24948754 in the Colombia dataset.
[26] See, for example, EventID 21638635 in the Russia dataset.

ICR Round 2 Results

The purpose of the conducting a second round of ICR tests was to estimate the degree of agreement between the different coders to whom we assigned a sub-sample to hand-code. As with round one, we conducted two rounds of ICR tests, estimating the degree of agreement between hand-coders using two methods – percentage agreement and using K-Alpha scores. As before, the minimum acceptable K-alpha score was 0.66. Coefficients above 0.80 are considered especially reliable. Tables A.5 and A.6 below summarize the results of the second round of ICR tests:

Table A.5: ICR Round 2 Percentage Agreement between Hand-coders

Variable	Colombia	Ukraine	Libya	Russia	Cuba
White Zone	97.9	91.6	96.7	91.8	100
Gray Zone	97.6	90.2	87.9	89.0	100
Black Zone	99.5	98.3	90.8	97.8	100
Source Type	99.7	97.3	*n/a*	99.2	98.9
Target Type	88.5	89.7	*n/a*	99.8	100
Erroneous	*118*	*22*	*21*	*36*	*9*

Table A.6: ICR Round 2 K-alpha Degree of Agreement between Hand-coders

Variable	Colombia	Ukraine	Libya	Russia	Cuba
White Zone	0.93	0.83	0.92	0.80	1.00
Gray Zone	0.92	0.72	0.76	0.78	1.00
Black Zone	0.91	0.96	0.71	0.91	1.00
Source Type	0.99	0.95	*n/a*	0.98	1.00
Target Type	0.82	0.84	*n/a*	0.99	1.00
Erroneous	*118*	*22*	*21*	*36*	*9*

As the above tables show, the results from the second round of ICR tests were extremely promising. Colombia was near perfect across the board, with a high degree of agreement on the Zonal classifications, as well as the identity of the actors involved. In both Ukraine and Russia's case, there was slightly more ambiguity around what constituted Gray Zone activity. Given the nature of the Gray Zone in general, as well as Russian expertise in operating within the Gray Zone in particular, this result is logical. Finally, the results for Cuba show almost universal agreement among coders, which also makes sense since there were only 100 observations in that sample, and most of the data deals with diplomatic negotiation (White Zone activity). In other words, just as a lack of variation drove the results for Cuba in Round 1 of ICR, we found the same lack of variation driving the results in Round 2. Following the previous round of ICR tests, erroneous observations (where the event was irrelevant to the conflict, or the EventType itself was incorrect) were dropped.

The one concerning result from the round two ICR results was the large number of dropped/erroneous observations in the Colombia sample. While the others hovered between four and nine percent of erroneous observations, the Colombia sample had approximately 24 percent (118) observations marked

as erroneous. The high error rate gives us reasons for concern in the Colombia data because it implies that potentially over 20 percent of the observations in the full sample may be erroneous (irrelevant or wrong EventType). In some ways, this is unsurprising because a lot of news about Colombia is unrelated to the conflict (the same is far less true for Ukraine and especially Libya).

After probing the erroneous observations further, we found that 99 of 118 (or 84%) erroneous cases involved White Zone activity. In order to estimate the degree to which White Zone activity in Colombia is being coded incorrectly, we drew a random sample of 1500 (or 2.9% of the 52,356) White Zone observations, which had not previously been hand-coded. Using this sample, a human coder identified the erroneous observations and marked the reason why they were erroneous. There are three reasons why an event may be erroneous:

A. Wrong EventType: where the sentence/headline implied a Zone entirely different from the one ICEWS coded.
B. Gray Zone activity: where the EventType implied White Zone Activity (making statements, for example) but the content actually indicated Gray Zone activity.
C. Irrelevant events: where the event is correctly coded as White Zone, but it refers to activity that is irrelevant to the conflict (such as economic summits with countries uninvolved in the conflict).

Table A.7 below provides a breakdown of the types of errors we found in the analysis.

Table A.7: White Zone Sample Analysis Error Breakdown

Error-type	N
Wrong EventType	66
Gray coded as White	27
Irrelevant Event	193
Total Errors	*286*
Sample Size	*1500*

The results suggest that in Colombia, the data is slightly overestimating the degree of White Zone activity for two reasons. First, the results show that 286 observations, or 19.1 percent of the White Zone sample, was erroneous observations. Of these, the vast majority (67.5%of erroneous observations and 12.9% of the entire sample) are not incorrectly coded, but irrelevant to the conflict. We believe that irrelevant observations are likely to be distributed non-randomly in the Colombia data, with respect to Zone only. In other words, a lot more irrelevant activity is likely to be White Zone than Gray or Black Zone – because often, the irrelevant activities involved diplomacy, international trade or other features of a functioning country and economy – which are all White Zone activities.

Second, we also found that 27 observations (9.4% of erroneous observations or 1.8% of the White Zone sample) are actually Gray Zone activity masked as White Zone activity. Put together, this suggests that the data is might be overestimating White Zone activity and slightly underestimating the degree of Gray Zone

activity. To account for this possibility, we conducted sensitivity analysis across all of our country studies using the methods described in the *Sensitivity Analysis* section below.

Finally, there were also 66 observations (23.1% of erroneous observations and 4.4% of the White sample) that were erroneous because the EventType was miscoded. This issue is widely known to affect all global event-coded datasets (Wang et al. 2016). We did not find any reason to believe this affects White Zone activity any more than any other Zone. While this issue is certainly a concern, we do not believe it affects the findings substantially.

Methodological Issues with ICEWS Data

ICEWS is a widely respected database of global events among scholars and policymakers and has been used in a wide range of studies utilizing event data (Schrodt 2012b; Schrodt and Brackle 2012; Ward et al. 2013; Schein et al. 2015; Hegre et al. 2017). Yet, emerging research on the methodological issues particular to automated coding algorithms has highlighted a variety of issues with the ICEWS model (Wang et al. 2016). Briefly, they found that global event coding datasets suffer from high levels of duplication, resulting from repeated reporting of historical events; that the changing number of news sources used by such datasets may introduces bias in the distribution of events being reported; and that the text processing algorithms used for event coding often misclassify the events being reported (Wang et al. 2016, p 1502). We also found these issues during our cleaning and analysis of the data, and we discuss these issues in greater detail below. However, we also document other issues that have not yet been discussed in sufficient detail by the academic and policy communities.

First, we found that translation issues were a serious problem for data from non-English speaking countries. There were serious translation errors in the Spanish-to-English translations (Colombia and Cuba), Arabic-to-English translations (Libya) and in Russian or Ukrainian-to-English translations (Russia and Ukraine). These issues impacted the EventType, source and target coding, which in turn affects the accuracy of any variable generated using the raw data.[27] We corrected for this to a limited degree through hand-coding a sub-sample of missing data and then overweighting it in the analysis.

Wang et al. (2016) already address the fact that most automated datasets suffer from duplicate entries as a result of re-reporting of events. In addition, we also found the converse to be an issue, often a single source sentence referred to multiple events. ICEWS did not always correctly code each individual event from a given source. Take, for example, instances where the sentence or headline referred to violence among armed groups that resulted in police arrests.[28] In these cases, the sentence would sometimes only code the violence between VNSAs and not the subsequent arrests or vice versa. As a result, the dataset is actually missing some relevant events that were reported in a sentence recounting multiple events. Nevertheless, ICEWS appears to have caught both or all events in a given source sentence most of the time.

[27] For instance, see EventID 22156726 in the Libya dataset, where the sentence and headline translations were essentially unintelligible.
[28] See for example, EventIDs 19514976 and 18670194 in the Libya dataset as cases where the multiple events are implied in the sentence, but ICEWS does not code them all.

A related issue was the fact that many events were simply being coded from running RSS feeds from major news sources like Reuters or BBC. Specifically, the headline or sentence would actually be a large batch of headlines with no more context from which to accurately code the details of the event. The machine simply picked one of the many headlines and coded the whole event from that. But further probing of such events found that headlines often did not provide enough details to code source, target and EventType with confidence. Such events had to be dropped from the analysis as well.

Another major issue in the data was the fact that source and targets were frequently wrong, missing or reversed. Even within our limited samples, we found a number of cases where the description (sentence and headlines) did not justify the actual source or target already coded by ICEWS. This was not the case for the sub-sample of observations with missing source, target of Zone information because of course, in those cases, ICEWS had no values coded. However, we had a number of observations per country-sample where the already coded information was incorrect. For that sub-sample, we corrected the miscoded observations as much as possible, resulting in several hundred corrections per sample. Table A.8 below provides a breakdown of the number of observations we could fix, and the number of observations we could not fix (or were irrelevant) and thus dropped.

Table A.8: Breakdown of Fixed and Dropped Errors in Hand-Coded Data by Country

	Colombia	**Cuba**	**Ukraine**	**Russia**	**Libya**
Total Sample	5000	580	5000	1967	5000
Total Dropped Erroneous	1472	54	424	136	163
Source	96	3	242	55	n/a
Target	608	8	510	108	n/a
Zone	75	21	46	234	n/a
Total Fixed	*779*	*32*	*798*	*397*	*n/a*

As the above table shows, we fixed a number of observations for Colombia, Ukraine and Russia. The Cuba sample was quite small – and so were the number of corrections. We were unable to do this for Libya because in Libya, we only sampled observations with missing Zone information for hand-coding, and we did not code source and target information there. As a result, , every Zone was fixed in Libya – since White, Gray and Black Zone codings were missing in the entire sub-sample of 5000 observations.

Additionally, source and target country were also frequently wrong. Specifically, source and targets would often be identified as officials or VNSAs from one country when in reality they belonged to a different country entirely. This is less of an issue for us because we have limited our analysis between state and non-state actors (with less emphasis on *which* state they belong to). However, if states are being misidentified often, the data cannot then be used to accurately measure the frequency with which particular states utilize different tools of cooperation or conflict. Similarly, missing, reversed or incorrect actor-types would affect the degree to which ICEWS could accurate measure the frequency with which different actors utilize tools of cooperation and conflict.

There were also a variety of issues related to source and target keywords. The first issue was that source and target keywords are not universal. The countries in our dataset had country-specific keywords that were not found in other countries – as a result, ICEWS struggled to identify the source and targets in these cases. For example, Ansar al-Sharia is an Islamist militia in Libya and should have been identified as a VNSA in the data, but was often listed with keywords used to identify civilians, suggesting that the algorithm for ICEWS may be struggling to pick up country-specific nuances. We also found multiple instances of VNSAs using aliases for individual or groups within cartels and paramilitary groups, without any mention of affiliation. This often led to miscoding of many VNSAs as civilians.[29] In short, ICEWS experiences difficulty identify source and target actors correctly. Of course, as a result, this affected our coding based on the raw ICEWS data. As with the issue of translation errors, we tried to account for this through our hand-coding strategy. We also built prevalent, but country-specific, terms, like Ansar al-Sharia into our automated recoding procedure even though they were not included by ICEWS. See Appendix B for a complete list of these terms.

This was also evident in the fact that acronyms were often incorrectly applied. For example, in the Russia dataset, an event was always sourced or targeted at the "Democratic Party of Russia" whenever the sentence or headline mentioned the "DPR." However, in Russia's case, *DPR* always refers to the "Donetsk People's Republic," the separatist entity in Ukraine that is in conflict with the Ukrainian government. So while the automated coding identified DPR as a political party, it was in fact, a VNSA. Similarly, "international community" was always coded as "unidentified state actor" when sometimes, the phrase referred to international civil society and sometimes to foreign (often Western) governments.

ICEWS also had issues with the CAMEO score coding. The ICEWS codebook suggests that their algorithm assigns a CAMEO score to each event based on the coding scheme developed by Schrodt and Yilmaz (2007). However, we found that in every single country dataset we examined, there existed CAMEO categories that did not exist in the CAMEO codebook. Colombia, Libya, Ukraine all had new categories that were exclusive to that country alone. Russia and Cuba had categories that existed in the main country datasets (that is, in Colombia, Ukraine or Libya) but not in the original CAMEO codebook. It is possible that the algorithm combined or concatenated existing CAMEO categories and created new ones. Either way, this is problematic for researchers interested in cross-national comparisons that use the EventType or CAMEO variables, if some categories are being used per country that do not exist in others. This is especially problematic since these additions do not appear to be documented in a supplemental codebook or appendix.[30] We incorporated all of these unique events into our coding. The list of new EventTypes is provided in Appendix C below.

Additionally, ICEWS struggled with geolocation. One of the major selling point of the ICEWS data was the fact that each observation includes longitude and latitude coordinates, thus allowing scholars and

[29] See, for instance, EventID 18999874 in the Colombia dataset.
[30] A subsequent version of the CAMEO codebook is now available, and it seems to have incorporated the new categories. This version of the codebook is available at: http://eventdata.parusanalytics.com/cameo.dir/CAMEO.Manual.1.1b3.pdf.

policymakers to map major crises geographically. However, we found that the location coding for events was incorrect in a variety of instances. In many cases, the geolocation coordinates were referring to the city from which the news organization was reporting (such as major cities or state capitals like Kiev, Tripoli, Havana or Bogota), even though the event occurred elsewhere. In other cases, the geolocation coordinates referred to the first major city mentioned in the sentence, even though the event occurred elsewhere.

Another major issue in ICEWS was the fact that hypothetical statements or news reports were coded as empirical reality. For example, in the Russia dataset, one event refers to Putin's warning Israel against selling weapons to Ukraine – this event was coded as Russia "cooperating militarily" with Ukraine. Even beyond getting the actual actors incorrect, this is not an instance of military cooperation between anyone, but rather potential cooperation between Israel and Ukraine.[31] In Libya, there are several cases where the government authorizes arrests of those involved in kidnappings and unconventional violence, and these potential future arrests have been coded as having occurred.[32] Finally, the ICEWS data also sometimes codes news reports of rumors as fact. For instance, there were instances in the Russia data where Russian Foreign Minister Sergei Lavrov had "heard" rumors that Crimean authorities were providing material and organizational assistance to those citizens that wanted to relocate from Crimea. This was coded as Crimea "providing aid" to Ukrainian citizens, even though there is evidence to suggest that Ukrainian authorities as well as self-proclaimed authorities have failed to provide adequate aid to internally displaced people in Crimea (United Nations 2015).[33]

Finally, ICEWS' algorithm struggled to identify and account for political bias. We found two different types of biases in the data. First, in many cases, ICEWS was unable to differentiate between information warfare (i.e., the strategic use of language, propaganda or information) and factual coverage. In other words, there were a variety of events where the subjects of the story (such as Russian Foreign Minister Lavrov above) used language for strategic ends, but ICEWS treated the statement as objective reporting. In the above case, this was a problem because Lavrov was quoting a rumor that got coded as fact.

But more insidious versions of this bias problem actually affected the very EventType or actor-types coded. In Russia's case, material aid provided by Russia to pro-Russian rebels in Crimea and Eastern Ukraine was often couched in the language of providing aid or humanitarian relief to civilians or "refugees" by Russian state officials or Crimean VNSAs.[34] Further, both governments and VNSAs frequently flung accusations of violations, human rights abuses, etc., in order to shame or delegitimize their adversaries. Similarly, these instances were all treated as objective reporting of events by ICEWS. In one instance, pro-Russian rebels in Donbass accused Ukraine of violating a ceasefire, and ICEWS coded this even as Ukraine "violat[ing a] ceasefire" against VNSAs.[35]

---

[31] EventID 22846970 in the Russia dataset.
[32] See, for instance, Event ID 18834861 in the Libya dataset.
[33] EventID 21051571 in the Russia dataset.
[34] See, for instance, EventID 21678912 or 20979005 in the Russia dataset.
[35] See EventID 24280950 in the Russia dataset.

This problem extends beyond the Russia-Ukraine conflict. In Libya, for instance, General Khalifa Haftar routinely refers to all of his adversaries as "terrorists" or "Islamist extremists" in his statements to the press.[36] Consequently, ICEWS labeled the targets of these statements as "extremists" or "terrorists." However, Haftar is not only fighting Islamists but also various other political opponents. He strategically refers to all of his enemies as "terrorists" and "extremists," regardless of their tactics, motivations, ideology, etc., in order to strengthen domestic and especially international support for his faction (Gomati 2014).

Second, the ICEWS data does not account for political bias in the way the sentence and headlines are written. Specifically, the algorithm does not account for the fact that many news sources themselves use politically charged language to describe events and actors, such that the resulting coding of the actor or event reflects the bias of the news source, rather than empirical fact. This is distinct from the above discussion because while the bias in those cases was injected by a subject being quoted or summarized in a news story; in this case, it is the news outlet or journalist that is introducing bias. Libya provides a good illustration of biased reporting.

In Libya, various armed groups were described by some news sources as "Islamist forces" when they were not all Islamist ideologically. For example, some news outlets in Libya frequently described Libya Dawn as "Islamist forces."[37] However, while some elements of Libya Dawn were indeed Islamist militias, Libya Dawn also frequently included non-Islamist forces (Fitzgerald and Toaldo 2017). In fact, Libya Dawn extensively fought against Islamic State (IS).

In Sirte, former Dawn elements allied with the internationally recognized government to militarily oppose IS. In Sirte in particular, we noticed substantial bias in the way the forces and events are described. Western news agencies, like the BBC, called Libya Dawn "pro-government forces." At the same time, other news outlets, including numerous Arabic dailies branded Libya Dawn as "Islamist forces." For the former cases, ICEWS coded Libya Dawn as government forces. In the latter ones, ICEWS coded them as militias or armed bands.[38] In the Ukraine-Russia conflict, Russian news agencies often referred to Russian military aid to VNSAs in Crimea and eastern Ukraine as "humanitarian aid," and ICEWS coded it as such.[39]

Accounting for Duplicate Events

In the previous section, we briefly discussed an issue also identified by others regarding large-n automated datasets. Specifically, these datasets have a large number of duplicate entries for events, usually as a result of multiple news sources reporting the same event, or follow-up stories on the same event by the same news source. On the one hand, it is possible that one event appearing multiple times in

[36] See, for example, EventID 22740256.
[37] EventID 21952007, for instance.
[38] EventID 23056245 refers to the BBC version of the event, while EventID 22933356 is the same event, described by Al-Wasat Online, a Bahranian Arabic daily. Note the different characterizations of Libya Dawn.
[39] See EventID 21413327, for instance. The publisher of the news report, GTRK Ingushetiya, is a regional state-owned news channel in Russia (Diatchkova 2008, 218).

the dataset may have some advantages. One could argue that not all events are created equal, and if events are reoccurring as a result of greater media coverage, perhaps that is acceptable by giving the event greater weight than other events in the data (Wang et al. 2016).

However, there are three issues with assuming that more media coverage of an event is a result of greater importance of an event in a conflict. First, certain events garner greater attention in Western media, but that does not mean that those events are consequential to a conflict. For example, the kidnapping of 276 schoolgirls in 2014 by Boko Haram in Northeastern Nigeria resulted in hundreds of news stories in Western media (Abubakar 2017). In another major automated dataset using news stories, GDELT, this resulted in nearly 300 distinct observations in the data. However, this event neither resulted in any substantial action by Western (or other foreign governments) except the expression of moral outrage, nor did it change how either the Nigerian government or the subversives prosecuted the conflict.

Second, even if the events do have out-sized effects on the conflict, there is no way to determine if the size of the effect is equivalent to the factor by which the event is repeated in the data. If the effect is even larger, repetition is still resulting in underweighting the event. If the effect is smaller, then repetition is biasing the findings. The larger concern is that there is no way for the authors or the readers to know which scenario might be accurate. Relatedly, media coverage itself is not unbiased, but a result of various business factors that shape media reporting at any given time. For example, while the events in Benghazi were arguably quite important to the conflict in Libya, it is hard to argue that they were more important than the seizure of oil ports by other violent non-state actors – but the former was covered far more extensively than the other in Western media, for domestic political reasons within the United States (Friedersdorf 2013). Finally, news coverage of any event wanes over time as media fatigue steps in. This will create time-trends in the events, not because the conflict itself is evolving, but rather because reporting of the events is waning. Riccihiardi (2008) shows, for example, how drastically media fatigue shaped the reporting on the Second Gulf War.

We believe that the conservative approach to using news reporting to understand conflict is to make no assumptions about the importance of events using the size of media coverage. Consequently, we used a variety of strategies to reduce the impact of duplicate stories in the data.

Our Approach to De-duplication

The ideal approach to *de-duplication* would be to read through all events in the dataset and remove all duplicate entries. Of course, practical constraints make such an approach unfeasible. Our alternative was to generate random samples of potentially duplicate events (see below for details on how potentially duplicate events were identified) for each country, such that they would be representative of trends in duplication within the full datasets. We then manually hand-coded each event in the sample for the actual level of duplication. The level of duplication ranged on a scale of 0 through 4, from most to least duplicate:

0. Almost certainly a duplicate
1. Somewhat certain it is a duplicate

2. Not sure whether it is a duplicate
3. Somewhat certain it is *not* a duplicate
4. Almost certainly *not* a duplicate

We used a scale instead of a binary (duplicate or not) approach because there was uncertainty in the data over whether an event was a duplicate or not, for various reasons. First, news reporting can often be quite vague, such that from the limited text available from ICEWS, it is unclear whether two different stories refer to the same event or not.

Second, there were cases where there was ambiguity over whether the news reports referred to the same event repeatedly, or to unique events within a larger, developing story. For example, there was a case in the Colombia data where the news reported free trade agreement negotiations between the United States and Colombia in May 2006, where the government reported that the potential deal would be signed in August 2006. However, a potential duplicate event was identified in October 2006, which reported the Colombian government as saying that a free trade agreement with the United States would be finalized in November 2006.[40] In this case, it is clearly a different point in time in an ongoing story, but it is not clear if there is enough progress to make it a truly distinct point in the timeline. As a result, we coded the second case as "likely" original (a score of 3) rather than definitely original (score of 4).

Sampling Strategy for De-Duplication

We generated two random samples of potentially duplicate events per country, resulting in a total of 10 random samples of over 300 observations each, to be hand-coded. We selected this sample size because it was large enough to allow the samples to be representatives, and yet small enough to allow coders to manually read (and keep in mind) all of the events and decide which ones were repeated stories and which ones were not.

In order to identify potential duplicates, we sorted the data first on location, assuming that repeated events are likely to be clustered around the same locations. We then sorted the data a second time by source-type, assuming that duplicates are likely to identify the same sources. We sorted the data a third time by target-type, again, assuming that duplicates are likely to identify the same targets. We then sorted the data a fourth and final time by event-type, assuming that duplicates are likely to have the same event-type. We then generated a variable, possible duplicate, which grouped together all observations that had the same values for location, source, target and event-type. A brief first-cut glance through all of our datasets showed that indeed, this strategy did group together duplicate entries in the data. We then randomized these groups of duplicate stories and selected two random samples per country.

First, we found that across our countries, roughly 50 percent of the data had approximately 15-20 or fewer potential duplicates in the data. That is, for 50 percent of observations, a story was reported as one event and then 15-20 (or fewer) observations had the exact same values for location, source, target and event-type. Therefore, the first random sample selected events with few to moderate (15-20) number of

[40] See for example, EventID: 9453833 and EventID 10160025. In fact, this deal was first recorded in an observation captured by ICEWS in 2003 (eventID 5289300).

duplicates, approximately representing the bottom 50 percent of the distribution of incidents with duplicate entries in the data. The advantage of this approach was that within an n of approximately 300, this allowed us to investigate 15-25 unique cases of duplicates, and estimate the level of duplication within each case, and this represented the majority of the data across our countries.

However, this strategy did not identify the most egregious of the duplicate cases. For example, the Benghazi story in Libya, or the arrest of Ukrainian Army pilot Nadiya Savchenko in the Ukraine-Russia conflict where the same event was reported hundreds or even thousands of times. To account for these cases, we conducted a second round of de-duplication, focusing on the worst offenders, or the other end of the potential duplicates distribution. While we chose the bottom 50 percent of the distribution earlier, we focused on the top 15 percent of the distribution of potential duplicates in the second round. This was explicitly designed to account for those key stories like Benghazi or Savchenko that were subjected to intense media focus and were at the core of the duplication critiques of machine learning datasets by various scholars.

These events had anywhere between 400 to over 3,100 potential duplicate observations in their respective country dataset. Since we could not review a grouping of 3,100 observations and hope to keep them all in mind at once, in order to identify duplicates we drew random samples of 15-20 stories per grouping. Thus, our round two de-duplication samples were comprised of 15-20 groups of potential duplicates. Table A.9 below depicts our results from the first round of duplicate analysis, and Table A.10 below summarizes our results from the second round of duplicate analysis.

Table A.9: De-duplication Round 1 Country Results by Type of Duplicates?

Type of Duplicate	Colombia	Ukraine	Libya	Russia	Cuba
Definitely Duplicate	85 (26.2%)	134 (36.4%)	129 (40.4%)	161 (44.4%)	121 (32%)
Likely Duplicate	6 (1.9%)	21 (5.7%)	11 (3.5%)	20 (5.5%)	15 (4%)
Likely Original	17 (5.3%)	23 (6.3%)	7 (2.2%)	15 (4.1%)	12 (3.2%)
Definitely Original	212 (65.4%)	159 (43.2%)	149 (46.7%)	148 (40.8%)	225 (59.5%)
Unclear	4 (1.2%)	31 (8.4%)	23 (7.2%)	19 (5.2%)	5 (1.3%)
N	*324*	*368*	*319*	*363*	*378*

Table A.10: De-duplication Round 2 Country Results by Type of Duplicates for Top Offenders?

Type of Duplicate	Colombia	Ukraine	Libya	Russia	Cuba
Definitely Duplicate	8 (2.7%)	25 (8.3%)	29 (9.7%)	42 (14%)	9 (20%)
Likely Duplicate	3 (1%)	14 (4.7%)	27 (9%)	7 (2.3%)	4 (8.9%)
Likely Original	18 (6%)	15 (5%)	38 (12.7%)	30 (10%)	0 (0%)
Definitely Original	258 (86%)	223 (74.3%)	131 (43.7%)	158 (52.7%)	30 (66.7%)
Unclear	13 (4.3%)	23 (7.7%)	75 (25%)	63 (21%)	2 (4.4%)
N	*300*	*300*	*300*	*300*	*45*

The next section explains our weighting strategy, where we weighted all of our hand-coded samples – on ambiguous Zones, missing source and target data, and the duplicate samples.

Weighting Strategy for Key Variables and Duplicates

Once the hand-coding was complete, our next step was to merge all the data back into the country datasets. This included all hand-coded samples, as well as samples testing for duplicates. The hand-coded observations, which were random samples drawn from the larger population of missing source, target and Zone information, now had complete and verified information. Consequently, they were weighted, such that they would analytically represent the population of missing source, target and Zone information in the country dataset. We accomplished this using the following three-step process.

First, we dropped those hand-coded observations identified by the coders as erroneous or irrelevant from the dataset, as well as those observations from the entire dataset without geocoded information (because we could not identify a location for them). Additionally, we also dropped the all observations with missing source, target and Zone information from the main dataset, since the weighted data will account for this sub-sample.

Second, we added weights to the source, target and Zone information in the hand-coded data such that those observations would analytically represent the population of missing data in the country dataset. The weight for each variable was calculated such that each hand-coded observation accounted for all missing observations for that variable. For instance, in the case of Ukraine, after dropping all erroneous observations, our hand-coded random sample included 3,425 valid observations out of a total of 7,521 observations with ambiguous Zone data in the entire dataset. Dividing the total population by the sample gives us an index of 2.196, which represents the weight each hand-coded observation must have to account for the full population of 7,521.

In the cases of Cuba and Russia, however, we did not assign add any additional weight to the hand-coded observations, because in these cases, we hand-coded the entire population of observations with missing information. As such, they will all carry the default weight of 1 in the analysis. In the case of Libya, we did not code any source and target information; thus, no weights needed to be calculated for source and target variables. We did, however, employ the same procedure for ambiguous Zones in Libya. Zone, source and target indices for each country are summarized in Table A.11 below.

In addition to the variables for Zone, source and target, we also generated a variety of composite variables (described in detail in Appendix F). Specifically, the former sets of variables are monadic – that is, they provide aggregate information about the Zone, source and target individually. As a result, these variables have *only* one weight each (e.g., the Zone variables only have the weights to account for all ambiguous Zones, the source variables only have weights to account for missing sources). The composite variables, on the other hand, take *all* information – Zone, source and target into account together. Consequently, the composite variables were assigned the combined weight for ambiguous Zones, missing source and missing target information, using the following formula:

Composite weight index = Ambiguous Zone Weight x *Missing Source Weight* x *Missing Target Weight*

For example, in the Ukraine data, all composite variables for hand-coded observations were weighted by an index of 51.666 (which is Zone weight (2.196) x source weight (6.726) x target weight (3.498) = 51.666). While the aforementioned weight is large, this is a function of the poor initial data quality. Composite weight indices for each country are also provided in Table A.11 below.

The third step in our strategy involved weighting our data to account for duplicates. The issues with duplication have been described in the *Accounting for Duplicate Events* section above. As explained, we generated two random samples of possible duplicates per country. The first random sample selected events with a small to moderate number of duplicates. The second random sample incorporate events with many potential duplicates. Our human coders examined each set of potential duplicate entries for every story in the random samples, and coded every observation on a scale of 0 to 4, with 0 representing definite duplicates, and 4 representing definite originals. We then rescaled this coding scheme between 0 and 1, with 0 representing definite duplicates and 1 representing definite originals. Once rescaled, we calculated the average level of duplication in each of the random samples for every country. Next, all potential duplicates (for which a particular random sample was representative) were multiplied by the average level of duplication for that random sample. The index of duplication for the bottom 50 percent of duplicates was defined as Base 1. The index of duplication for the top 15 percent was defined as Base 2. Both are summarized for each country in Table A.11 below.

For instance, in the Russia data, approximately 55 percent of all observations in the dataset had 20 or fewer duplicates. In our random sample of this data, we found that the average level of duplication was 0.479. Thus, we assume that in our full sample, for all observations with 20 or fewer duplicates, the average level of duplication is 0.479. Thus, we multiplied all observations with 20 or fewer duplicated by 0.479. We followed the same procedure for Base 1 and Base 2 calculations for all countries.

We did not sample roughly the middle-to-upper 35 percent of the distribution of potential duplicates (since Base 1 covered the bottom 50% and Base 2 looked at the top 15%). However, in every single case, we found that the level of duplication dropped as the number of identified duplicates for an incident rose. While this may seem counter-intuitive, it actually makes sense. Protracted events often result in repeated actions by parties that may be part of one ongoing conflict-process, but are discrete events. As such, in stories with hundreds of potential duplicates, coders often found that events that were identified as duplicates actually referred to discrete points within a larger story (often over a long period of time), rather than really being a duplicate report of just one incident.

Therefore, for those observations in the middle-to-upper 35 percent of the distribution of potential duplicates, we generated an average level of duplication by taking a weighted average of Base 1 and Base 2. We thus assume that the level of duplication in the middle is a weighted average of the level of duplication in approximately the bottom 50 percent and the top 15 percent. We defined this middle range as Base 3, also summarized for each country in Table A.11 below.

For example, in the Russia data, approximately 30 percent of all observations in the data had between 21 and 143 potential duplicate reports for one event. To calculate the weight for the 30 percent of observations that made up Base 3, we took a weighted average (0.521) of Base 1 (0.479) and Base 2 (0.713), and multiplied each observation in Base 3 (those with more than 20 and less than 144 duplicates) by it. Table A.11 below gives a summary of the weights used for each sample – ambiguous Zone, missing source and target, and duplicates, for each country.

Table A.11: Weights for Zone, Source, Target and Levels of Duplication for All Countries

Weights	**Colombia**		**Cuba**		**Ukraine**		**Russia**		**Libya**	
Ambiguous Zones										
Sample Size	6994		8		7521		1114		9683	
(Population)	(2416)		(8)		(3425)		(1114)		(4828)	
Index	2.895		1.000		2.196		1.000		2.005	
Missing Source										
Sample Size	9638		479		9612		287			
(Population)	(1890)		(479)		(1429)		(287)		N\A	
Index	5.099		1.000		6.726		1.000			
Missing Target										
Sample Size	12458		79		9475		383			
(Population)	(1383)		(79)		(2709)		(383)		N\A	
Index	9.008		1.000		3.498		1.000			
Composite Index[41]	132.973		1.000		51.666		1.000		2.005	
Weights for Duplicates	Duplicate Range	Index	Duplicate Range	Index	Duplicate Range	Index	Duplicate Range	Index	Duplicate Range	Index
Base 1 – Min Duplicates	1 – 24	0.698	1 – 9	0.636	1 – 25	0.535	1 – 20	0.479	1 – 15	0.528
Base 2 – Max Duplicates	Above 190	0.929	Above 19	0.711	Above 334	0.831	Above 143	0.713	Above 93	0.679
Base 3 – Mid-level Duplicates	25 – 190	0.775	10 – 19	0.669	26 – 334	0.631	21 – 143	0.521	16 – 93	0.534
Erroneous Dropped	1731		45		422		91		163	

Importantly, all observations with no potential duplicates in the data were given a weight of 1. Put simply, observations without potential duplicates were, in effect, not weighted, while observations with duplicates were underweighted. The more duplicates, the heavier the underweighting.

Sensitivity Analysis

We also tested how sensitive our findings are to our assumptions. One of the key concerns in examining Gray Zone conflict is the difficulty in identifying and differentiating between White and Gray Zone tactics.

[41] The composite index was calculated using the following formula: Composite index = *Ambiguous Zone Weight* x *Missing Source Weight* x *Missing Target Weight.*

We found that Gray Zone activity by state and VNSAs masquerades as White Zone activity (such as information warfare being coded as routine diplomacy). Thus, our key sensitivity analysis will revolve around testing whether our findings change if a greater share of events are coded as Gray instead of White.

In order to do this, we generated alternate Zone variables, in which we over-weighted observations coded as Gray Zone events and proportionately underweight observations coded as White Zone events. Our first cut analysis will examine how findings change if five percent more observations are actually Gray and five percent less are White. We accomplished this by generating weighted variables. Specifically, we multiplied Gray Zone codings by 1.05, and White Zone codings by 0.95.

Further sensitivity analyses could also include changing the ratio of Gray and Black Zone activity, or the proportion of source and target information coded as government versus VNSA. For the latter, however, we found a low level of error in the hand-codings, thus making this additional analysis less important.

Similarly, we found a relatively low level of erroneous or irrelevant observations in the data. More importantly, while we have estimates of what the level of error is in the missing data, we do not have estimates of the level of error in the country datasets. Thus, we cannot run sensitivity analyses by dropping or underweighting observations in the full sample as erroneous using estimates from the hand-coding. Fortunately, as noted the levels of erroneous data in the hand-coded samples was relatively low. Moreover, because these samples were comprised of problematic observations, it is likely that the proportion of erroneous observations in the country datasets are systematically lower.

Appendix B: Source and Target Codings

A complete list of keywords can be found in Schrodt (2012a). Below we list only the keywords that we utilized.

Government Keywords

- Government
- Military
- Police
- Parties/Major Party/Minor Party
- Centrist/Center Right/Center Left
- Ministry
- IGO/IGOs/International Governmental Organization
- Presidential
- Main Opposition
- State Media

Violent Non-State Actor Keywords

- Rebel
- Armed Gang
- Insurgent
- Militia
- Guerrilla
- Armed Band
- Revolutionary
- Paramilitary
- Combatant
- Violent Group
- Militant
- Extremist
- Criminal
- Fundamentalists
- Gangs
- Kidnapper
- Separatists

Civilian Keywords

- Citizen
- Children
- Civilian
- Media/News/Newspaper/News Editor
- Social

- Employee
- Minority
- Student
- Nongovernmental Organizations (country or sector)/NGO/NGOs
- Lawyer
- Deserter
- Business
- Protestor
- Labor Union
- Activist
- Professor/Economist
- Refugee
- Farm Worker

Country-Specific Keywords by Actor Type

- Colombia:
  - Government:
    - David Andrews
  - VNSA:
    - Bloque
    - Death Squad
  - Civilian:
    - N/A
- Ukraine:
  - Government:
    - Tymoshenko
  - VNSA:
    - N/A
  - Civilian:
    - N/A
- Libya:
  - Government:
    - Media Personnel/Spokesman/Spokesperson
  - VNSA:
    - N/A
  - Civilian:
    - Al Jazeera
    - Saif al-Islam al-Qaddafi
    - Bishop
    - Christian
    - Hindu

        - Media Personnel (International)
        - BBC
        - Scientist
        - Fishermen
- Russia:
    - Government:
        - Liberal Democratic Party of Russia
        - Leonid Kuchma
    - VNSA:
        - Democratic Party of Russia (actually referring to the separatist, Donetsk People's Republic, due to the fact that both share the initials DPR, but only the former is in ICEWS dictionary)
        - Dmytro Yarosh
        - Sergey Aksyonov
    - Civilian:
        - N/A
- Cuba:
    - Government:
        - Castro
    - VNSA:
        - Luis Hernando Gómez
    - Civilian:
        - Exiles

Appendix C: Zone and Kinetic Codings

White Zone Activities

- Exhibit Force Posture
- Demonstrate Military Or Police Power
- Increase Police Alert Status
- Increase Military Alert Status
- Mobilize Or Increase Police Power
- Mobilize Or Increase Armed Forces
- Threaten To Expel Or Withdraw Peacekeepers
- Threaten With Violent Repression
- Threaten To Use Military Force
- Threaten Blockade
- Threaten Occupation
- Threaten Conventional Attack
- Give Ultimatum
- Threaten
- Threaten Non-Force
- Threaten To Reduce Or Stop Aid
- Threaten To Boycott, Embargo, Or Sanction
- Threaten To Reduce Or Break Relations
- Threaten With Administrative Sanctions
- Threaten To Impose Restrictions On Freedoms Of Speech And Expression
- Threaten To Ban Political Parties Or Politicians
- Threaten To Impose Curfew
- Threaten To Impose State Of Emergency Or Martial Law
- Threaten Collective Dissent
- Threaten To Halt Negotiations
- Threaten To Halt Mediation
- Demand
- Demand Information, Investigation
- Demand Policy Support
- Demand Aid, Protection, Or Peacekeeping
- Make Empathetic Comment
- Express Accord
- Appeal For Cooperation
- Appeal For Diplomatic Cooperation
- Appeal For Material Cooperation
- Appeal For Policy Support
- Appeal For Aid
- Appeal For Economic Aid
- Appeal For Military Aid
- Appeal For Humanitarian Aid
- Appeal For Military Protection Or Peacekeeping
- Praise Or Endorse
- Engage In Diplomatic Cooperation
- Defend Verbally
- Rally Support On Behalf Of
- Appeal To Others To Meet Or Negotiate
- Appeal To Others To Settle Dispute
- Appeal To Others To Engage In Mediation
- Express Intent To Cooperate
- Express Intent To Meet Or Negotiate
- Express Intent To Provide Policy Support
- Express Intent To Settle Dispute
- Express Intent To Mediate
- Mediate
- Yield
- Ease Administrative Sanctions
- Ease Restrictions On Freedoms Of Speech And Expression
- Ease Ban On Political Parties Or Politicians
- Ease Curfew
- Ease State Of Emergency Or Martial Law

- Demand Political Reform
- Demand Change In Leadership
- Demand Policy Change
- Demand Rights
- Demand Change In Institutions, Regime
- Demand Mediation
- Demand Withdrawal
- Demand Ceasefire
- Demand Meeting, Negotiation
- Reject Proposal To Meet, Discuss, Or Negotiate
- Reject Mediation
- Threaten
- Reject
- Reject Proposal
- Reject Ceasefire, Withdrawal
- Reject Peacekeeping
- Reject Settlement
- Reject Request For Material Aid
- Reject Demands For Political Reform
- Reject Demands For Change In Leadership
- Reject Demands For Policy Change
- Reject Demand For Rights
- Reject Demand Change In Institutions, Regime
- Reduce Relations
- Investigate
- Investigate Crime, Corruption
- Investigate Human Rights Abuses
- Investigate Military Action
- Ease Popular Protest
- Accede To Demands For Political Reform
- Accede To Demands For Change In Leadership
- Accede To Demands For Change In Policy
- Accede To Demands For Rights
- Accede To Demands For Change In Institutions
- Express Intent To Engage In Material Cooperation
- Express Intent To Cooperate Economically
- Express Intent To Cooperate Militarily
- Express Intent To Provide Aid
- Express Intent To Provide Economic Aid
- Express Intent To Provide Military Aid
- Express Intent To Provide Humanitarian Aid
- Express Intent To Provide Military Protection Or Peacekeeping
- Grant Diplomatic Recognition
- Engage In Material Cooperation
- Cooperate Economically
- Express Intent To Bring Political Reform
- Express Intent To Change Leadership
- Express Intent To Change Policy
- Express Intent To Provide Rights
- Express Intent To Change Institutions, Regime
- Express Intent To Yield, Not Specified Below
- Express Intent To Ease Administrative Sanctions
- Express Intent To Stop Protests
- Express Intent To Accede To Political Demands
- Express Intent To Release Persons Or Property
- Express Intent To Ease Economic Sanctions, Boycott, Or Embargo

- Investigate War Crimes
- Disapprove
- Criticize Or Denounce
- Accuse
- Accuse Of Crime, Corruption
- Accuse Of Espionage, Treason
- Complain Officially
- Bring Lawsuit Against
- Make Pessimistic Comment
- Appeal For Political Reform
- Appeal For Change In Leadership
- Appeal For Policy Change
- Appeal For Rights
- Appeal For Change In Institutions, Regime
- Appeal To Yield
- Decline Comment
- Make Public Statement
- Make Statement
- Consider Policy Option
- Acknowledge Or Claim Responsibility
- Engage In Symbolic Act
- Make Optimistic Comment
- Consult
- Discuss By Telephone
- Make A Visit
- Meet At A 'Third' Location
- Host A Visit
- Appeal

Colombia-specific ICEWS Codings

- Accede To Requests Or Demands For Political Reform
- Appeal For De-Escalation Of Military Engagement
- Appeal For Diplomatic Cooperation (Such As Policy Support)
- Appeal For Easing Of Administrative Sanctions
- Express Intent Allow International Involvement (Not Mediation)
- Express Intent To De-Escalate Military Engagement
- Express Intent To Accept Mediation
- Engage In Negotiation
- Apologize
- Forgive
- Share Intelligence Or Information
- Provide Aid
- Grant Asylum
- Return, Release
- Return, Release Person(S)
- Return, Release Property
- Ease Economic Sanctions, Boycott, Embargo
- Engage In Judicial Cooperation
- Provide Economic Aid
- Provide Humanitarian Aid
- Sign Formal Agreement
- Provide Military Aid
- Provide Military Protection Or Peacekeeping
- Allow International Involvement
- Receive Deployment Of Peacekeepers
- Receive Inspectors
- Allow Delivery Of Humanitarian Aid
- De-Escalate Military Engagement
- Declare Truce, Ceasefire
- Ease Military Blockade
- Demobilize Armed Forces
- Retreat Or Surrender Militarily
- Demand That Target Yields
- Demonstrate For Leadership Change
- Ease Restrictions On Political Freedoms
- Express Intent To Cooperate On Intelligence

- Appeal For Easing Of Political Dissent
- Appeal For Economic Cooperation
- Appeal For Intelligence
- Appeal For Judicial Cooperation
- Appeal For Military Cooperation
- Appeal For Release Of Persons Or Property
- Appeal To Engage In Or Accept Mediation
- Demand De-Escalation Of Military Engagement
- Demand Diplomatic Cooperation (Such As Policy Support)
- Demand Economic Aid
- Demand Intelligence Cooperation
- Demand Judicial Cooperation
- Demand Material Cooperation
- Demand Military Aid
- Demand Release Of Persons Or Property
- Demand Settling Of Dispute
- Express Intent To Cooperate On Judicial Matters
- Express Intent To Engage In Diplomatic Cooperation (Such As Policy Support)
- Express Intent To Institute Political Reform
- Express Intent To Provide Material Aid
- Make An Appeal Or Request
- Reject Economic Cooperation
- Reject Judicial Cooperation
- Reject Material Cooperation
- Reject Plan, Agreement To Settle Dispute
- Reject Request For Economic Aid
- Reject Request For Rights
- Threaten With Military Force
- Threaten With Political Dissent, Protest
- Threaten With Repression
- Threaten With Sanctions, Boycott, Embargo

Ukraine-specific ICEWS Codings

- Allow Humanitarian Access
- Appeal for Easing of Economic Sanctions, Boycott, or Embargo
- Demand Humanitarian Aid
- Demand Easing of Administrative Sanctions
- Threaten With Restrictions On Political Freedoms

Libya-specific ICEWS Codings

- Demand Material Aid
- Reject Request for Change in Institution

Gray Zone Activities

- Kill By Physical Assault
- Conduct Suicide, Car, Or Other Non-Military Bombing
- Carry Out Suicide Bombing
- Carry Out Car Bombing
- Carry Out Roadside Bombing
- Assassinate
- Expel Or Withdraw Aid Agencies
- Halt Mediation
- Coerce
- Protest
- Conduct Hunger Strike
- Conduct Hunger Strike For Change In Leadership

- Use Violent Repression
- Assault
- Use Unconventional Violence
- Abduct, Hijack, Or Take Hostage
- Physically Assault
- Sexually Assault
- Torture
- Impose Embargo, Boycott, Or Sanctions
- Use As Human Shield
- Attempt To Assassinate
- Obstruct Passage, Block
- Obstruct Passage To Demand Change In Leadership
- Obstruct Passage To Demand Policy Change
- Obstruct Passage To Demand Rights
- Obstruct Passage To Demand Change In Institutions, Regime
- Protest Violently, Riot
- Obstruct Passage To Demand Change In Leadership
- Obstruct Passage To Demand Policy Change
- Obstruct Passage To Demand Rights
- Obstruct Passage To Demand Change In Institutions, Regime
- Threaten Unconventional Violence
- Threaten Attack With WMDs
- Expel Or Withdraw
- Expel Or Withdraw Peacekeepers
- Expel Or Withdraw Inspectors, Observers
- Conduct Hunger Strike For Policy Change
- Conduct Hunger Strike For Rights
- Conduct Hunger Strike For Change In Institutions, Regime
- Conduct Strike Or Boycott
- Conduct Strike Or Boycott For Change In Leadership
- Conduct Strike Or Boycott For Policy Change
- Conduct Strike Or Boycott For Rights
- Conduct Strike Or Boycott For Change In Institutions, Regime
- Halt Negotiations
- Reduce Or Stop Aid
- Reduce Or Stop Economic Assistance
- Reduce Or Stop Military Assistance
- Reduce Or Stop Humanitarian Assistance
- Defy Norms, Law
- Reject Accusation, Deny Responsibility
- Veto
- Impose Administrative Sanctions
- Impose Restrictions On Freedoms Of Speech And Expression
- Ban Political Parties Or Politicians
- Impose Curfew
- Impose State Of Emergency Or Martial Law
- Expel Or Deport Individuals
- Reduce Or Break Diplomatic Relations
- Rally Opposition Against

Colombia-specific ICEWS Keywords

- Deny Responsibility
- Engage In Violent Protest For Policy Change
- Reduce Or Stop Material Aid
- Refuse To Ease Administrative Sanctions

- Impose Restrictions On Political Freedom

Ukraine-specific ICEWS Keywords

- Obstruct Passage To Demand Leadership Change

Libya-specific ICEWS Keywords

- Conduct Strike or Boycott for Leadership Change

Black Zone Activities

- Fight
- Use Conventional Military Force
- Fight With Artillery And Tanks
- Employ Aerial Weapons
- Engage In Ethnic Cleansing
- Use Weapons of Mass Destruction
- Use Chemical, Biological or Radiological Weapons
- Detonate Nuclear Weapons
- Impose Blockade, Restrict Movement
- Occupy Territory
- Engage In Mass Expulsion

White/Gray Ambiguous Activities

- Engage In Popular Protest
- Demonstrate Or Rally
- Demonstrate For Change In Leadership
- Demonstrate For Policy Change
- Demonstrate For Rights
- Demonstrate For Change In Institutions, Regime
- Arrest, Detain, Or Charge With Legal Action
- Accuse Of Human Rights Abuses
- Accuse Of Aggression
- Accuse Of War Crimes
- Cooperate Militarily

Colombia-specific ICEWS Keywords

- Refuse To Yield

Gray/Black Ambiguous Activities

- Fight With Small Arms And Light Weapons
- Engage In Unconventional Mass Violence
- Engage In Mass Killings
- Violate Ceasefire
- Seize Or Damage Property

- Confiscate Property
- Destroy Property

Colombia-specific ICEWS Keywords

- Refuse To De-Escalate Military Engagement
- Refuse To Release Persons Or Property
- Use Tactics Of Violent Repression

Kinetic Event-Types

- Kill By Physical Assault
- Conduct Suicide, Car, Or Other Non-Military Bombing
- Carry Out Suicide Bombing
- Carry Out Car Bombing
- Carry Out Roadside Bombing
- Assassinate
- Fight
- Use Conventional Military Force
- Fight With Small Arms And Light Weapons
- Fight With Artillery And Tanks
- Employ Aerial Weapons
- Engage In Unconventional Mass Violence
- Engage In Mass Killings
- Engage In Ethnic Cleansing
- Use Weapons Of Mass Destruction
- Use Chemical, Biological, Or Radiological Weapons
- Detonate Nuclear Weapons
- Impose Blockade, Restrict Movement
- Occupy Territory
- Violate Ceasefire
- Engage In Mass Expulsion
- Seize Or Damage Property
- Confiscate Property
- Destroy Property
- Use Violent Repression
- Assault
- Use Unconventional Violence
- Abduct, Hijack, Or Take Hostage
- Physically Assault
- Sexually Assault
- Torture
- Use As Human Shield
- Attempt To Assassinate
- Protest Violently, Riot
- Use Tactics Of Violent Repression
- Fight With Artillery And Tanks

Appendix D: Manual Recoding Instructions

Background

- The objective of the larger Gray Zone project is to explore conceptually how the presence/involvement of non-state actors impacts the development and evolution of Gray Zone conflicts in three countries – Colombia, Ukraine and Libya. You have been given some conceptual readings on what exactly 'Gray Zones' are. Before you can go any further, it is important for you to have done this reading – to understand what it is that the project is trying to accomplish.
    - Please refer back to these three Gray Zone criteria:
        1. Exceeds "ordinary competition" yet falls below the threshold of large-scale direct military conflict.
        2. Problems with attribution/activity is ambiguous.
        3. Undermines/violates international norms/laws.
    - Please keep in mind that Gray Zone activity can be violent OR non-violent.
    - When trying to determine if an activity is Gray vs. White or Gray vs. Black, it is helpful to consider whether the action is escalatory or de-escalatory in nature.
        - If you are coding Black or Gray, then de-escalatory events are likely Gray and escalatory events are likely Black.
        - If you are coding White or Gray, then de-escalatory events are likely White and escalatory events are likely Gray.
- In order to understand how states and non-state actors behave in Gray Zone conflicts, we have a dataset called ICEWS, which is essentially a dataset of events that have occurred in these three countries over a period of several years (each country has a different time period). Each event is listed in your dataset as its own row (i.e. the event is the unit of analysis).
- For each of these events, we have the following basic information:
    - EventDate (Column B) – the date the event was recorded.
    - Source (Column D) – The actor initiating the event.
    - Source-Sectors (Column E) – More details about the identity of the Source.
    - Event-type (Column F) – the type of action being initiated/committed.
    - Target Country (Column H) – The country on the receiving end of the action.
    - Target (Column I) – the actor on the receiving end of the action.
    - Target-Sector (Column J) – More details about the identity of the Target.
    - Sentence (Column R) – The original sentence from which the entire entry was machine coded. This should be the first point of reference in order to make the coding decisions required of you.
        - PLEASE NOTE: Many sentences include multiple events, try to determine which event the sentence has been machine coded on where applicable before recoding variables. In the event that both the source and target are missing and there is not enough information to determine which event is being coded in a sentence recording multiple events and where there is sufficient information to completely code both events mentioned in the sentence, alternate. For the first occurrence, code the first event listed,

for the second occurrence, the second event, etc. If however, only one event contains sufficient information for you to completely code all missing variables as identified by the instruction variable, code that event.

- Using this information, you will edit Columns S through AB.
- **Note**: Ignore the columns for: EventID, SourceCountry, Score, StoryID, Publisher, SentenceNumber, PublicationDate, Latitude and Longitude. Also ignore the columns for CoderSample, ICRSample and Rand. In fact, we highly recommend that you simply "hide" these columns in Excel while you are working on your datasets (this will make your task easier!). Be sure to "unhide" them when you are done.

Introducing Your Task

- The basic issue that we have so far is that for some of the events, we do not have complete, accurate or clear information.
- In some cases, the source is not clear, such as when the Source is listed as "Colombia" or "/N". In other cases, the target is not clear, such as when the Target is listed as "Unspecified Actor" or "Unspecified Forces." Keep in mind that many unclear categories repeat across source and target – you will see "unspecified forces" in both target and source categories.
- In yet other cases, we cannot tell simply from the Event-Type category whether the event should be considered a "White Zone" interaction, a "Gray Zone" interaction or a "Black Zone" interaction. Some actions for example, could be White or Gray depending on the specifics of the event (e.g. protests are an example of normal, peaceful interactions and as such are White. Violent demonstrations, on the other hand, exceed the limits of normal interactions. As such, they are Gray.) So activities below this threshold should be coded as White, and above this threshold should be coded as Black or Gray.
- The difference between Gray and Black activity is the upper limit – does the activity count more as a 'large-scale conflict' or not? If the activity is no longer "in-between" or suffering from ambiguity but rather a large-scale violent activity, it should be coded as Black.
- All of the activities listed as "whitegray" or "grayblack" in your datasets are ambiguous – that is, you cannot code a White, Gray or Black Zone purely by reading the event-type. In each of the cases, you will have to read the actual sentence (from the Sentence variable, Column R) and make a determination as to the Zone. Be sure that you can defend your decision if asked about it.
- Once you have decided which column the event belongs in, change the White, Gray, and Black Zone (Columns U, V and W) coding in that cell accordingly. Each event can only belong in one of the three columns – so make sure only one of them is coded as 1 while the others are coded 0.
- Once you have made the change, go ahead and recode the "grayblack" or "whitegray" columns (AA and AB) as 0. When you are done with your dataset, these columns should be 0 for every single event (unless there is insufficient information to make a determination for one or more events), just as every single event should have either a White, Gray or Black coding.
- Detailed instructions on coding are provided in the supplemental sheet titled "SMA SOCOM Gray Zone Coding Guidelines." Follow these coding rules consistently.

Fixing Source and Target Type

- Along with ambiguous Zones, many events in your dataset also do not have clear source and target information.
- For events where the source is not clear, the variable SourceType (Column S) has an empty cell. For each event where the target is not clear, the variable TargetType (Column T) has an empty cell.
- Sources and targets can be one of three categories: government, violent non-state actor or civilian. We need you to read the source and source-sector (or target and target-sector in the latter case) and the sentence *carefully*, to make a determination about what kind of category belongs in each empty cell.
    - Government is defined broadly to include the various branches of government: police, armed forces and security services, individual government and political officials, state owned enterprises, etc.
    - For statements and denunciations, the SourceType is the actor making the statement, the TargetType is the subject of the statement. For example, the Colombian National Police denounced the FARC for their use of child soldiers – SourceType: Government (Colombian National Police); TargetType: Violent Non-State Actors (FARC).
- For each empty cell in these columns, you must make a determination about what kind of source or target we are dealing with. Fill in the cell with the appropriate category.
- Be careful to use the correct spelling and case (type in "Government", not "government") – otherwise, we will have trouble reading them into Stata efficiently.
- Once you have correctly filled in the columns, double-check the coding for StateSource (Column X), StateTarget (Column Y) and HNTarget (Column Z).
- StateSource should be 1 if the source is a "Government" source – otherwise, it should be 0. StateTarget should be 1 if the target is a "Government" target – otherwise, it should be 0. HNTarget should be 1 if the target is your country (e.g. Colombia when working on the Colombia dataset; the Libya coder will not deal with Source and Target), otherwise it should be 0.

Follow the "Instructions" Variable

- We have sorted the data in the datasets according to what needs to be fixed in each event. The instructions column (Column AF) at the end of the dataset lists what exactly you should fix in that row. We recommend that you read this column for each event first.

Insufficient Information to Code and Questions

- If you are confident that there is insufficient information to make a determination for a given event for one or more variables, simply code what variables you can (if any) and do not change the rest. Simply move on.
- There will be times that you are uncertain about how to code some or all events for a given variable. That is fine. Please make note of this by coding a 1 in the variable review column and type in any notes to remind yourself what the concern was. We will go over these as a group during our weekly meetings, as it will help increase confidence on the most difficult/ambiguous entries and also will be instructive in ensuring that everyone is coding uniformly.

Assignments

- You will each be assigned a stratified random sample of approximately 5,000 observations from one country. You will receive this sample in the form of two Excel spreadsheets. The first contains 500 observations that are to be completed first. Please notify Varun [Piplani] and I [Barnett S. Koven] when you have completed the first 100 and then when you have completed the first 500. We will spot check at those intervals. You do not need to stop and wait for us to do so before continuing. Once the first 500 observations are coded, move on to the main dataset containing approximately 4,500 observations. Again, let us know when you are done.
- In addition, you will each receive 500 observations from a second country. These will enable us to calculate inter-coder reliability scores and ensure that the approach is uniform across coders. Like with the above, notify Varun and I [Barnett] when you complete the first 100 observations and again when you complete all 500.

Saving and Securing your Data:

- This data is Sensitive but Unclassified (SBU); specifically it is marked For Official Use Only (FOUO). This designation requires the following:
  1. Please work from and store the data on START's network drive (see below bullet) when you are in the office. When you are not in the office, please remotely access the network drive. You will need to download the data to your machine. Any files you download need to be uploaded to the network drive when you are done and then deleted from your personal machine. While you are working and have the data on your machine, it must be in a password protected file and/or your computer must be locked whenever you step away. Do not use your personal cloud storage (e.g. DropBox) for this data.
  2. Do not use this data for your own personal purposes, if you are interested in doing so talk to me and I can tell you where to find a version of the data that is available for public use (unfortunately that version is embargoed for two years and does not include all of the variables of this version).
  3. Do not share this data with anyone outside of our team:
     - Barnett S. Koven
     - Varun Piplani
     - Michelle Haragos
     - James Siebens
     - Evan Romanko
     - Rachel Gabriel (possibly; Rachel is involved in qualitative aspects of this project and may become more involved on the quantitative side as well as we progress; I will notify you all if she requires access to this data).
- I have set up a folder for each of you (with your name on it) under the network drive: Research Data (U:)/START SMA/ICEWS Hourlies. This drive will automatically appear when you log in on any START computer.
- To access the network drive remotely, please use GlobalScape at remote.bsos.umd.edu. Your UMD Directory ID and Password are your login credentials.

Appendix E: Coding Guidelines

- Three Primary Gray Zone Criteria:
  1. Exceeds "ordinary competition" yet falls below the threshold of large-scale direct military conflict.
  2. Problems with attribution/activity is ambiguous.
  3. Undermines/violates international norms/laws.
- Gray Zone activity can be violent OR non-violent.
- When trying to determine if an activity is Gray vs. White or Gray vs. Black, it is helpful to consider whether the action is escalatory or de-escalatory in nature.
  - If you are coding Black or Gray, then de-escalatory events are likely Gray and escalatory events are likely Black.
  - If you are coding White or Gray, then de-escalatory events are likely White and escalatory events are likely Gray.
- Tactic-Centric approach:
  - Prioritize the tactic(s) used over the actor employing the tactic.
  - Governments and Non-State Actors can employ White, Gray and Black tactics.
- There are many cases of large scale violence in the datasets that are ambiguous between Gray and Black – for example, "fighting with small arms and light weapons," "engage in unconventional mass violence," "engage in mass killings" etc. All of these cases can be either Gray or Black.
- **The Black vs. Gray distinction should be based on the tactics being used**. If the attack or tactic is what conventional military forces regularly do (or are supposed to do), then it is likely Black. If, on the other hand, the tactic being used is not something regular militaries do (or are supposed to do), then it is likely Gray. For example, using artillery, or fighting with small arms and weapons by large military formations is more likely Black than Gray. On the other hand, IEDs and suicide attacks are likely Gray.
- Consider the following distinctions: Direct combat or less than combat (IED, assassination, etc.)? Known combatants or unknown? War crimes or normal military operations?
  - **Military operations conducted by groups, or direct combat between known/identified enemy combatants, are generally Black.**
    - FARC attacked a police checkpoint with small arms=Black
    - Government troops exchanged fire with separatist forces=Black
    - The town was attacked/shelled/captured by the separatist militia=Black
  - **War crimes, terrorism, combat where the identity of one side is unknown, or attacks by individuals are generally Gray.**
    - Two police were killed in a drive-by shooting by a motorcyclist=Gray
    - Government troops executed 9 suspected militants=Gray
    - The assassin shot the mayor twice before he was killed by police=Gray
- Be particularly careful about using the number of casualties or injuries as an indicator of Gray vs. Black. For instance, we can imagine a mechanized conflict (involving conventional military forces) that moved against each other and fired upon each other, in which very few casualties occurred. On the other hand, we can imagine a suicide bombing killing a lot of people. The first scenario is mass-scale event without mass casualties, and the second scenario is not 'mass-scale' in the same way, and

still involves mass casualties. The first scenario is Black, and the second is Gray. It is fine to keep the number of fatalities in mind to think about the scope of the damage being done, but the primary focus should be on what *kind* of attack we are talking about.

- If the sentence gives no description of tactics, and the sentence just says that two forces clashed, then code it as Black. If there is not enough information to imply some level of organization by the two sides, or if you have evidence of lack of organization, then code it as Gray. Also, if there is a lot of ambiguity about who is doing the attacking, lean towards coding it as Gray.
- If there are events that are nonviolent, and they are protests or events related to bread and butter issues that are unrelated to the conflict, then they should be dropped from the analysis. In these cases, the events should be marked N/A.
    - If they are violent, they cannot be dropped because we cannot make the assumption that VNSAs had no role in encouraging the events/violence (a common tactic).
- If the event is a protest or rally, and the target is not clear, the default should be to consider the target the government. Generally, protests and rallies are meant to influence the government.
- If the event-type is obviously wrong, code it N/A.
- Drug-related issues in Colombia are generally to be included because drugs are a part of the overall conflict. However, if you can make a determination that the event was an isolated event (single individual, etc.), those can be dropped. However, if the drug event implies some involvement of organized groups, include it.
- When two violent non-state actors agree to coordinate military activity, this should be coded as Gray.
- If the sentence is posing a hypothetical based on something that has already happened, then you can code it. If it the sentence is entirely hypothetical, then code it as N/A.
- If the sentence is composed of a whole bunch of headlines (e.g. an RSS feed) and it is not clear which headline is being used for the coding, then mark it as N/A.

For Libya Coders Only

- For Libya, we are not actually worrying about source and target, because it is extremely hard to distinguish between state and non-state actors in Libya at the moment. As a result, all of our focus in Libya will be on the ambiguous Zonal codings.
- The other key difference in Libya is that "unconventional violence" is being included in the Gray/Black ambiguous category. In other countries, instances of unconventional violence were treated as Gray because that is by definition a Gray activity. However, in Libya's case, a lot of large-scale conflict is being coded as "unconventional" for some reason. At the same time, small-scale fighting is also being coded as unconventional.
- So "Use unconventional violence" is not itself an indicator of Gray activity in Libya. As such, we will code this activity as Gray if the sentence specifically mentions the use of unconventional tactics such as:
    - Suicide bombs
    - Car bombs
    - Violence against civilians
    - Hostage taking

    - Assassinations
    - Other activities that are atypical in the case of Black Zone armed confrontation in Libya specifically
- In cases where "use unconventional violence" refers to general fighting between groups or 'clashes':
    - When there is an acknowledged conflict between groups, even if the groups are not specifically identified (e.g., "clashes" between armed groups), these events should be coded as Black.
        - For example, "Armed groups in Benghazi clashed with forces loyal to General Haftar" would be coded as Black.
    - Where there is significant ambiguity about the actors involved or their affiliations with a specific organization, these observations should be coded as Gray.

Event Type: Arrest, Detain or Charge

- It is difficult to identify whether arrests in Libya serve as attempts to crush political opposition, and thus if accusations are legitimate. This is partially due to the fact that it is unlikely that political opponents will receive a fair trial.
    - Arrests that are related to political events or conflicts should be coded as Gray.
        - For example, the arrest of Saif el-Qaddafi would be coded as Gray.
- In situations where opposing factions are engaged in ongoing conflict, mass arrests of militants or individuals on the opposing side could constitute war by other means.
    - If the individual is identified as a member of a group and has been implicated in a specific illegal activity, the event should be coded as White.
    - Arrests of 'suspected' members of certain groups should be coded as Gray.
- In situations where individuals have been arrested and charged for specific crimes that actually occurred, these observations should be coded as White.
    - For example, the sentence "the police arrested gang members for possession of narcotics" should be coded as White.

Event Type: Cooperate Militarily

- In situations where factions (political and militant) are cooperating in the context of politicized conflicts, these observations should be coded as Gray.
    - For example, Haftar's cooperation with militias against "terrorists" in Benghazi should be coded as Gray because a) he is a renegade general and b) his campaign paints all political opposition as "terrorists".
- Where Libya's various factions are cooperating against IS, this should be coded as White.
    - The internationally recognized government has sanctioned the fight against IS.

Examples by Event Type for All Coders

- "Arrest, detain, or charge with legal action" by VNSAs are generally Gray. (Libya is a partial exception to this, as discussed above.)
- Government (sourcetype) arrests are to be coded as White UNLESS clearly extrajudicial or arbitrary:

  - Police arrested suspected rebels=White.
  - Judge sentenced convicted cartel leader=White.
  - Judge sentenced rebel leader *in absentia*=Gray.
- In the case of arrests and detentions, code arresting foreign nationals as Gray unless there is clear evidence to suggest that the arrest was legal or that the arrest is not relevant to this project. (Many of the arrests of foreign nationals that were included in the data are clearly related to the conflict and appear to be politically motivated.)
- In the case of arrests and detentions, if both the perpetrators of the crime and/or the people arrested for those crimes are unclear it is likely Gray.
- If the arrest appears to be politically motivated, then code it as Gray.
- "Make statement":
  - Statements of objective fact are White.
  - If the purpose is clearly to attack or delegitimizing another other actor, then it is Gray.
  - If the purpose is to accuse another actor of Gray activity, the accusation is also likely Gray.
  - If the purpose is to deny responsibility for Gray activity, the denial is also likely Gray.

Review/Comment Codes

In the notes column, the following format should be used to indicate problems or corrections:

- Mark a "1" in the review column in order to flag an observation for further review at our weekly meeting or by Barnett or Varun.
- Mark N/A in the comments column to note that an observation is completely irrelevant; unrelated to research topic.
  - E.g. Petty crime involving individuals only and where sophistication of crimes does not indicate organized actor involvement.
- "N/A?" " in Comments column: unclear if notes are relevant to research topic.
  - Add any instructions or details as required (e.g. "N/A? No context").
- VariableName in Comments column: specifies which variable needs to be reviewed.
  - Add any instructions or details as required after standard error format (e.g. "whitegray" unclear if targettype=civilian or VNSA").
- (VariableName)=(corrected value) in Comments column: Format to indicate suggested corrections for other data included in the analysis outside the instructions that appear incorrectly coded. Flag for review.
  - sourcetype=civ
  - targettype=gov
  - sourcetype=VNSA
  - For dummy variables, write the name of the correct dummy category (e.g. "White")
- If both coders independently (did not discuss with each other beforehand) coded N/A for an event, then the event should be automatically coded as N/A. If both coders independently (did not discuss with each other beforehand) coded an observation as White or Gray or Black, then the coding should automatically be retained.

- Once an error has been confirmed, finalize the correction by indicating them in the columns to the far-right of the dataset:
  - **In the Csourcetype and Ctargettype columns:**
    - "Gov" indicates Government.
    - "Civ" indicates Civilian.
    - "VNSA" indicates Violent Non-State Actors.
  - **In the Cwhite_gray_black column:**
    - "W" indicates White.
    - "G" indicates Gray.
    - "B" indicates Black.
  - **In the Error_NA column:**
    - "1" indicates the observation contains an error that should exclude it from the analysis.

Appendix F: Codebook

Event-level Datasets

Original ICEWS Variables

EventID – A unique identifier for each event-actor combination. Specifically, each observation in the dataset involves a distinct combination of actors involved in a single action. For example, a single altercation wherein insurgents targeted police and civilians would result in two distinct observations being coded, each with its own EventID. The first observation would code insurgents targeting police and the second would code insurgents targeting civilians during the same StoryID. As such, EventID, unlikely StoryID, is a unique identifier for each, distinct observation in the dataset.

EventDate – The date of the event identified by EventID.

SourceCountry – The country from which the event/action was initiated.

Source – The name of the actor initiating the event.

SourceSectors – The category of the actor initiating the event. For instance, categories included phrases like "dissident," "rebel," "government," "civilian," "national party," etc.

EventType – Each event was assigned an eventtype category by ICEWS, based on the CAMEO (Conflict and Mediation Event Observations) coding scheme developed by Schrodt (2007). Thus, each event in the ICEWS data is coded as one of 256 CAMEO categories that range from cooperative to hostile actions. All CAMEO categories are listed in Appendix C. For example, eventtypes included actions like "use conventional military force," "make statement," "arrest, detain or charge with legal action," etc.

Score – The CAMEO score for the eventtype, ranging from -10 (most hostile) to 10 (most cooperative), based on the coding scheme developed by Schrodt (2007).

TargetCountry – The country targeted by the event/action.

Target – The name of the actor targeted by the event/action.

TargetSectors – The category that the target actor belongs to. Categories included the same set of phrases used to populate the sourcesectors variable.

StoryID – A unique identifier for each news story coded in the dataset. Note that multiple events are often coded from a single story. These events all share the same StoryID, but each of them will have a unique EventID.

Publisher – The name of the publisher of the news report from which the observations were coded.

SentenceNumber – The order of the sentence being coded. For example, if the action being coded is described in the third sentence of a report, the SentenceNumber would be recorded as 3.

Headline – The title of the news report being coded.

PublicationDate – The date the story was reported. This does not always match the EventDate because, in some cases, news stories were published in the days following an event.

Latitude – Latitudinal coordinates for geolocation.

Longitude – Longitudinal coordinates for geolocation.

Sentence – The relevant text of the news report that was coded.

SMA Gray Zone Aggregate Variables

Sourcetypefin – Denotes the identity of the perpetrator of the event or action. It can be one of three categories: "Government," "Violent Non-State Actors (VNSA)" or "Civilian." The category "Government", refers to not only the central government of all state parties to the conflict, but also to state security forces (e.g. the national army) and civilian governmental agencies. Politicians, who are currently in government office are also coded as such. "VNSAs" are armed entities that are distinct from states (e.g. subversive groups, militia forces, etc.), even if they often collaborate with "Government" forces. Individual people not affiliated with either of the former categories and are coded as "Civilians". As are, non-violent civil society groups, press (excluding state-owned media), etc.

Targettypefin – Denotes the identity of the target of the event or action. It can also be one of three categories: "Government," "VNSA" or "Civilian".

Grayfin – Coded as 1 whenever the event or action constitutes Gray Zone activity, 0 otherwise. Gray Zone activity meets at least one of three criteria:

1.) Exceeds "ordinary cooperation" [e.g. peaceful economic competition such as increasing oil prices, as opposed to threatening to cut off crucial natural gas supplies to specific countries and irrespective of price] yet falls below the threshold of large-scale direct military conflict.
2.) Suffers from problems with attribution, and where the activity is ambiguous.
3.) Undermines/violates international norms/laws.

For example, hit-and-run style attacks by unknown perpetrators, were common in the Colombia data. This type of event is Gray because it both exceeds "ordinary competition", and because it suffers from problems with attribution.

Whitefin – Coded as 1 whenever the event or action constitutes White Zone activity, 0 otherwise. White Zone events do not meet any of the aforementioned categories. Competition remains "ordinary",

attribution is clear and norms/law violations are absent. Diplomatic engagements and the signing of accords between states is an example. Non-violent protest is another.

Blackfin – Coded as 1 whenever the event or action constitutes Black Zone activity, 0 otherwise. Black Zone events also do not meet any of the above categories. However, in this case, events do not fall below the threshold of large-scale direct military conflict, suffer from attribution problems or violate international norms/laws. For example, large-scale fighting that might use artillery or motorized and mechanized forces between two clearly identified belligerents would be coded as Black.

Statesourcefin – Coded as 1 whenever the source of the event or action is a government, 0 otherwise. This includes any government (and not just the government of the country being studied), multiple governments as a source (e.g., a coalition of multiple nations) or international governmental organizations (e.g., the United Nations).

Statetargetfin –Coded as 1 whenever the target of the event or action is a government, 0 otherwise. This includes any government (and not just the government of the country being studied) or multiple governments as a target (e.g., a coalition of multiple nations) or international governmental organizations (e.g., the United Nations).

VNSASourcefin – Coded as 1 whenever the source of the event is a VNSA, 0 otherwise.

VNSATargetfin – Coded as 1 whenever the target of the event is a VNSA, 0 otherwise.

Civsourcefin – Coded as 1 whenever the source of the event is civilian, 0 otherwise. This includes individual civilians as well as civil society groups, non-governmental organizations, and journalists from non-state owned media outlets.

Civtargetfin – Coded as 1 whenever the target of the intended event or action is civilian, 0 otherwise. Civilian collateral damage is coded as a 0. For example, if the source text indicates that two civilians were caught in the crossfire between the military and insurgents, this variable is coded as 0. While civilians were victimized, they were clearly not the intended victim.

StateW – Coded as 1 whenever the source of the event is a government and the event or action constitutes White Zone activity, 0 otherwise.

StateG – Coded as 1 whenever the source of the event is a government and the event or action constitutes Gray Zone activity, 0 otherwise.

StateB – Coded as 1 whenever the source of the event is a government and the event or action constitutes Black Zone activity, 0 otherwise.

VNSAW – Coded as 1 whenever the source of the event is a VNSA and the event or action constitutes White Zone activity, 0 otherwise.

VNSAG – Coded as 1 whenever the source of the event is a VNSA and the event or action constitutes Gray Zone activity, 0 otherwise.

VNSAB – Coded as 1 whenever the source of the event is a VNSA and the event or action constitutes Black Zone activity, 0 otherwise.

CivW – Coded as 1 whenever the source of the event is civilian and the event or action constitutes White Zone activity, 0 otherwise.

CivG – Coded as 1 whenever the source of the event is a government and the event or action constitutes Gray Zone activity, 0 otherwise.

CivB - Coded as 1 whenever the source of the event is a government and the event or action constitutes Black Zone activity, 0 otherwise.

HNSource – Coded as 1 whenever the source of the event or action is the country of the case study in question, 0 otherwise. For example, for the Colombia case study, this means 1 whenever the source is Colombian, 0 for foreign sources. However, since Cuba and Russia were only included in the study given their relevance to Colombia and Ukraine, respectively, and because the relevant data from these two countries was merged with the data for Colombia and Ukraine, respectively, HNSource was coded as 1 in these cases if the country that was the source of the event or action was Colombia (and not Cuba) or Ukraine (and not Russia), 0 otherwise.

HNTarget – Coded as 1 whenever the target of the event or action is the country of the case study in question, 0 otherwise. The coding of this variable follows the same convention as for HNSource, above.

Kinetic – Coded as 1 whenever the event or action involves the active use of force, 0 otherwise.

Eventyear – The year of the reported event.

Monthyear – The month and year of the reported event.

Eventdate2 – The exact date of the reported event.

Tothandcode – Coded as 1 if the event was part of one of the random samples hand-coded by researchers, 0 otherwise.

Location Variables

NAME_0 – The name of the country in which the event took place.

ISO – Country abbreviation for the country in which the event took place.

ID_1 – A unique identifier for the location within the country at the largest available subnational administrative level. For example, the departmental level in Colombia.

NAME_1 – The universally accepted geoname of the location in ID_1.

VARNAME_1 – The variant name of the location in ID_1. Unlike the geoname, which strips accents, the variant name includes them.

NL_NAME_1 – The name of the location in ID_1, written in the native language of the country.

TYPE_1 – The type of administrative division in ID_1, in the native language.

ENGTYPE_1 – The type of administrative division in ID_1, in English.

ID_2 – A unique identifier for the location within the country at the second largest available subnational administrative level. For example, municipality in Colombia.

NAME_2 – The universally accepted geoname of the location in ID_2.

VARNAME_2 – The variant name of the location in ID_2. Unlike the geoname, which strips accents, the variant name includes them.

NL_NAME_2 – The name of the location in ID_2, written in the native language of the country.

TYPE_2 – The type of administrative division in ID_2, in the native language.

ENGTYPE_2 – The type of administrative division in ID_1, in English.

Shape_Leng – The length of the perimeter in kilometers of each subnational administrative unit.

Aggregate Variables with Weights

A detailed discussion of our weighting procedure for the following variables can be found in the appendices to Koven, Barnett S., Varun Piplani, Steve Sin, and Marcus A. Boyd. "Quantifying Gray Zone Conflict: (De-)escalatory Trends in Gray Zone Conflicts in Colombia, Libya and Ukraine," Report to the U.S. Department of Homeland Security Science and Technology Office of University Programs and the U.S. Department of Defense Strategic Multi-layer Assessment Branch (College Park, MD: START 2017).

whitewt – This is a weighted version of whitefin, based on corrections derived from hand-coded random samples. Coded as 1 whenever the event or action constitutes White Zone activity and the event is not in the hand-coded sample. If the event is within the hand-coded sample, this variable carries a value between 0 and 1 based on the weights. The variable is coded as 0 if the event is not White Zone.

Graywt – This is a weighted version of grayfin. This variable follows the same coding convention as whitewt, described above.

blackwt – This is a weighted version of blackfin. This variable follows the same coding convention as whitewt, described above.

statesourcewt – This is a weighted version of the statesource variable based on corrections derived from hand-coded random samples. Coded as 1 whenever the source is a state organization and the event is not in the hand-coded sample. If the event is within the hand-coded sample, this variable carries a value between 0 and 1 based on the weights. The variable is coded as 0 if the source is not state.

statetargetwt – This is a weighted version of the statetarget variable. This variable follows the same coding convention as statesourcewt, described above.

vnsasourcewt – This is a weighted version of the vnsasource variable based on corrections derived from hand-coded random samples. Coded as 1 whenever the source is a VNSA and the event is not in the hand-coded sample. If the event is within the hand-coded sample, this variable carries a value between 0 and 1 based on the weights. The variable is 0 if the source is not VNSA.

vnsatargetwt – This is a weighted version of the vnsatarget variable. This variable follows the same coding convention as vnsasourcewt, described above.

civsourcewt – This is a weighted version of the civsource variable based on corrections derived from hand-coded random samples. Coded as 1 whenever the source is civilian and the event is not in the hand-coded sample. If the event is from the hand-coded sample, this variable carries a value between 0 and 1 based on the weights. The variable is 0 if the source is not civilian.

civtargetwt – This is a weighted version of the civtarget variable. This variable follows the same coding convention as civsourcewt, described above.

StateWwt - This is a weighted version of the StateW variable based on corrections derived from hand-coded random samples. Coded as 1 whenever the source of the event is a government, the event or action constitutes White Zone activity and the event is not in the hand-coded sample. If the event is from the hand-coded sample, this variable carries a value between 0 and 1 based on the weights. The variable is 0 if the source is not government and the event is not White Zone.

StateGwt – This is a weighted version of the StateG variable. This variable follows the same coding convention as StateWwt, described above.

StateBwt – This is a weighted version of the StateB variable. This variable follows the same coding convention as StateWwt, described above.

VNSAWwt – This is a weighted version of the VNSAW variable. This variable follows the same coding convention as StateWwt, described above.

VNSAGwt – This is a weighted version of the VNSAG variable. This variable follows the same coding convention as StateWwt, described above.

VNSABwt – This is a weighted version of the VNSAB variable. This variable follows the same coding convention as StateWwt, described above.

CivWwt – This is a weighted version of the CivW variable. This variable follows the same coding convention as StateWwt, described above.

CivGwt – This is a weighted version of the CivG variable. This variable follows the same coding convention as StateWwt, described above.

CivBwt – This is a weighted version of the CivB variable. This variable follows the same coding convention as StateWwt, described above.

hnsourcewt – This is a weighted version of the hnsource variable based on the hand-coded random samples. This is a weighted version of the civsource variable based on corrections derived from hand-coded random samples. Coded as 1 whenever the sourcecountry was "Colombia" in the Colombia dataset, "Ukraine" in the Ukraine dataset and "Libya" in the Libya dataset, and the event is not in the hand-coded sample. If the event is within the hand-coded sample, this variable carries a value between 0 and 1 based on the weights. The variable is 0 if the sourcecountry is not the host-nation, as explained above.

hntargetwt – This is a weighted version of the hntarget variable based on corrections derived from hand-coded random samples. This variable follows the same coding convention as hnsourcewt, described above, with the exception that this variable focuses on targetcountry instead of sourcecountry.

kineticwt – This is a weighted version of the kinetic variable based on the hand-coded random samples. Coded as 1 whenever the event or action involves the active use of force, and the event is not in hand-coded sample. If the event is within the hand-coded sample, this variable carries a value between 0 and 1 based on the weights. The variable is 0 if the event does not involve the active use of force.

graywt_higray – This is a weighted version of the Gray Zone variable for sensitivity analysis. This variable artificially inflates Gray Zone codings by 5 percent and is to be used in combination with whitewt_higray, described below. A detailed discussion of our sensitivity analysis can be found in the above appendices.

whitewt_higray – This is a weighted version of the White Zone variable for sensitivity analysis. This variable artificially deflates White Zone codings by 5 percent and is to be used in combination with graywt_higray, described above.

graywt_hiwhite – This is a weighted version of the Gray Zone variable for sensitivity analysis. This variable artificially deflates Gray Zone codings by 5 percent and is to be used in combination with whitewt_hiwhite, described below.

whitewt_hiwhite – This is a weighted version of the White Zone variable for sensitivity analysis. This variable artificially inflates White Zone codings by 5 percent and is to be used in combination with graywt_hiwhite, described above.

Composite Variables

In order to undergird dyadic analysis aimed at determining how actions of a certain type (e.g., a unique source, target, Zone and kinetic combination) affect the nature of actions by other actors, we generated a series of composite variables. These variables are named so that source, target, Zone and kinetic information would be apparent in the name itself. Table F.1, below, summarizes the naming convention.

Table F.1: Composite Variables Naming Convention

Source	Abbreviation	Target	Abbreviation	Zone	Abbreviation	Kinetic
State	st	State	st	White	w	k
VNSA	vnsa	VNSA	vnsa	Gray	g	K
Civilian	civ	Civilian	civ	Black	b	k

For example, ststw denotes state source, state target, White Zone, non-kinetic, whereas vnsavnsabk indicates VNSA source, VNSA target, Black Zone, kinetic.

The entire list of composite variables is broken down by source below.

State Source: ststw ststwk ststg ststgk ststb ststbk stvnsaw stvnsawk stvnsag stvnsagk stvnsab stvnsabk stcivw stcivwk stcivg stcivgk stcivb stcivbk

VNSA Source: vnsastw vnsastwk vnsastg vnsastgk vnsastb vnsastbk vnsavnsaw vnsavnsawk vnsavnsag vnsavnsagk vnsavnsab vnsavnsabk vnsacivw vnsacivwk vnsacivg vnsacivgk vnsacivb vnsacivbk

Civilian Source: civstw civstwk civstg civstgk civstb civstb civstbk civvnsaw civvnsawk civvnsag civvnsagk civvnsab civvnsabk civcivw civcivwk civcivg civcivgk civcivb civcivbk

Composite Variables with Weights

These variables all have the suffix *wt*, which denotes that they have been weighted. These variables are weighted versions of the composite variables from the above section. They employ weighted corrections for source, target and Zone information, which was derived from hand-coded random samples. As already noted, a detailed discussion of the weighting strategy is included in the above appendices.

The entire list of weighted composite variables is broken down by source below.

State Source: ststwwt ststwkwt ststgwt ststgkwt ststbwt ststbkwt stvnsawwt stvnsawkwt stvnsagwt stvnsagkwt stvnsabwt stvnsabkwt stcivwwt stcivwkwt stcivgwt stcivgkwt stcivbwt stcivbkwt

VNSA Source: vnsastwwt vnsastwkwt vnsastgwt vnsastgkwt vnsastbwt vnsastbkwt vnsavnsawwt vnsavnsawkwt vnsavnsagwt vnsavnsagkwt vnsavnsabwt vnsavnsabkwt vnsacivwwt vnsacivwkwt vnsacivgwt vnsacivgkwt vnsacivbwt vnsacivbkwt

Civilian Source: civstwwt civstwkwt civstgwt civstgkwt civstbwt civstbkwt civvnsawwt civvnsawkwt civvnsagwt civvnsagkwt civvnsabwt civvnsabkwt civcivwwt civcivwkwt civcivgwt civcivgkwt civcivbwt civcivbkwt

Time-series, Cross-sectional Datasets

We also collapsed the data by month and subnational administrative boundaries to create time-series, cross-sectional versions of the data.

SMA Gray Zone Aggregate Variables

monthyear – A unique identifier for each month/year time period.

NAME_0 – The name of the country.

ID_0 – A unique numerical identifier for each country.

ID_1 – A unique identifier for the location within each country at the largest available subnational administrative level. For example, the departmental level in Colombia.

Newid - This variable exists in datasets that combined data from multiple countries (Colombia and Cuba, as well as Ukraine and Russia). In such datasets, newid assigns a unique ID number to each administrative unit in the dataset. For example, for the Colombia and Cuba data, newid assigns a unique ID number to each *Department* in Colombia and Cuba at the ID_1 level, and a unique ID number to each *Municipality* in Colombia and Cuba at the ID_2 level. For the Ukraine and Russia data, newid assigns a unique ID number to each region or city at the ID_1 level, and a unique ID number to each *Raion* (or district), at the ID_2 level.

NAME_1 – The universally accepted geoname of the location in ID_1.

VARNAME_1 – The variant name of the location in ID_1. Unlike the geoname, which strips accents, the variant name includes them.

NL_NAME_1 – The name of the location in ID_1, written in the native language of the country.

TYPE_1 – The type of administrative division in ID_1, in the native language.

ENGTYPE_1 – The type of administrative division in ID_1, in English.

ID_2 – A unique identifier for the location within the country at the second largest subnational administrative level. For example, municipality level in Colombia.

NAME_2 – The universally accepted geoname of the location in ID_2.

VARNAME_2 – The variant name of the location in ID_2. Unlike the geoname, which strips accents, the variant name includes them.

NL_NAME_2 – The name of the location in ID_2, written in the native language of the country.

TYPE_2 – The type of administrative division in ID_2, in the native language.

ENGTYPE_2 – The type of administrative division in ID_1, in English.

whitefin – A monthly sum of all White Zone events for each location.

grayfin – A monthly sum of all Gray Zone events for each location.

blackfin – A monthly sum of all Black Zone events for each location.

statesourcefin – A monthly sum of all state-initiated events for each location.

statetargetfin – A monthly sum of all events that targeted a state for each location.

vnsasource – A monthly sum of all VNSA-initiated events for each location.

vnsatarget – A monthly sum of all events that targeted VNSAs for each location.

civsourcefin – A monthly sum of all civilian-initiated events for each location.

civtargetfin – A monthly sum of all events that targeted civilians (civilian collateral damage is excluded) for each location.

StateW – A monthly sum of all events where the source was state and the event or action constituted White Zone activity for each location.

StateG – A monthly sum of all events where the source was state and the event or action constituted Gray Zone activity for each location.

StateB – A monthly sum of all events where the source was state and the event or action constituted Black Zone activity for each location.

VNSAW – A monthly sum of all events where the source was VNSA and the event or action constituted White Zone activity for each location.

VNSAG – A monthly sum of all events where the source was VNSA and the event or action constituted Gray Zone activity for each location.

VNSAB – A monthly sum of all events where the source was VNSA and the event or action constituted Black Zone activity for each location.

CivW – A monthly sum of all events where the source was civilian and the event or action constituted White Zone activity for each location.

CivG – A monthly sum of all events where the source was civilian and the event or action constituted Gray Zone activity for each location.

CivB - A monthly sum of all events where the source was civilian and the event or action constituted Black Zone activity for each location.

hnsource – A monthly sum of all events where the host nation was the source of the event, for each location.

hntarget – A monthly sum of all events where the host nation was the target of the event, for each location.

Kinetic – A monthly sum of all kinetic events (e.g. raids but not making statements) per location.

Aggregate Variables with Weights

whitewt – A monthly sum of all weighted White Zone events for each location.

graywt – A monthly sum of all weighted Gray Zone events for each location.

blackwt – A monthly sum of all weighted Black Zone events for each location.

statesourcewt – A monthly sum of all weighted state-initiated events for each location.

statetargetwt – A monthly sum of all weighted events that targeted states.

vnsasourcewt – A monthly sum of all weighted VNSA-initiated events for each location.

vnsatargetwt – A monthly sum of all weighted events that targeted VNSAs for each location.

civsourcewt – A monthly sum of all weighted civilian-initiated events for each location.

civtargetwt – A monthly sum of all weighted events that targeted civilians for each location (civilian collateral damage is excluded).

StateWwt – A monthly sum of all weighted events where the source was state and the event or action constituted White Zone activity for each location.

StateGwt – A monthly sum of all weighted events where the source was state and the event or action constituted Gray Zone activity for each location.

StateBwt – A monthly sum of all weighted events where the source was state and the event or action constituted Black Zone activity for each location.

VNSAWwt – A monthly sum of all weighted events where the source was VNSA and the event or action constituted White Zone activity for each location.

VNSAGwt – A monthly sum of all weighted events where the source was VNSA and the event or action constituted Gray Zone activity for each location.

VNSABwt – A monthly sum of all weighted events where the source was VNSA and the event or action constituted Black Zone activity for each location.

CivWwt – A monthly sum of all weighted events where the source was civilian and the event or action constituted White Zone activity for each location.

CivGwt – A monthly sum of all weighted events where the source was civilian and the event or action constituted Gray Zone activity for each location.

CivBwt - A monthly sum of all weighted events where the source was civilian and the event or action constituted Black Zone activity for each location.

hnsourcewt – A monthly sum of all weighted events where the host nation was the source of the event, for each location.

hntargetwt – A monthly sum of all weighted events where the host nation was the target of the event, for each location.

kineticwt – A monthly sum of all weighted kinetic events (e.g. raids but not making statements) for each location.

graywt_higray – A monthly sum of all weighted Gray Zone events per location. This variable is artificially inflated by 5 percent and is to be used in combination with whitewt_higray, below, for sensitivity analysis.

whitewt_higray – A monthly sum of all weighted White Zone events per location. This variable is artificially deflated by 5 percent and is to be used in combination with graywt_higray, above, for sensitivity analysis.

graywt_hiwhite – A monthly sum of all weighted Gray Zone events per location. This variable is artificially deflated by 5 percent and is to be used in combination with whitewt_hiwhite, below, for sensitivity analysis.

whitewt_hiwhite – A monthly sum of all weighted White Zone events per location per. This variable is artificially inflated by 5 percent and is to be used in combination with graywt_hiwhite, above, for sensitivity analysis.

Composite Variables

In order to undergird dyadic analysis aimed at determining how actions of a certain type (e.g. a unique source, target, Zone and kinetic combination) affect the nature of actions by other actors, we generated a series of composite variables. These variables are named so that source, target, Zone and kinetic information would be apparent in the name itself. Importantly, each composite variable captures the monthly sum of activities involving a specific source, target, Zone and kinetic or non-kinetic activity. Table F.2, below, summarizes the naming convention.

Table F.2: Composite Variables Naming Convention

Source	Abbreviation	Target	Abbreviation	Zone	Abbreviation	Kinetic
State	st	State	St	White	w	k
VNSA	vnsa	VNSA	Vnsa	Gray	g	K

Civilian	civ	Civilian	Civ	Black	b		k

For example, ststw captures the monthly sum of all activity with state source, state target, White Zone, non-kinetic coding, whereas vnsavnsabk does the same for all events with VNSA source, VNSA target, Black Zone, kinetic coding.

The entire list of composite variables is broken down by source below.

State Source: ststw ststwk ststg ststgk ststb ststbk stvnsaw stvnsawk stvnsag stvnsagk stvnsab stvnsabk stcivw stcivwk stcivg stcivgk stcivb stcivbk

VNSA Source: vnsastw vnsastwk vnsastg vnsastgk vnsastb vnsastbk vnsavnsaw vnsavnsawk vnsavnsag vnsavnsagk vnsavnsab vnsavnsabk vnsacivw vnsacivwk vnsacivg vnsacivgk vnsacivb vnsacivbk

Civilian Source: civstw civstwk civstg civstgk civstb civstb civstbk civvnsaw civvnsawk civvnsag civvnsagk civvnsab civvnsabk civcivw civcivwk civcivg civcivgk civcivb civcivbk

Composite Variables with Weights

These variables all have the suffix *wt*, which denotes that they have been weighted. These variables are weighted versions of the composite variables from the above section, which captures the monthly sum of activities involving a specific source, target, Zone and kinetic or non-kinetic activity. They employ weighted corrections for source, target and Zone information, which was derived from hand-coded random samples.

The entire list of weighted composite variables is broken down by source below.

State Source: ststwwt ststwkwt ststgwt ststgkwt ststbwt ststbkwt stvnsawwt stvnsawkwt stvnsagwt stvnsagkwt stvnsabwt stvnsabkwt stcivwwt stcivwkwt stcivgwt stcivgkwt stcivbwt stcivbkwt

VNSA Source: vnsastwwt vnsastwkwt vnsastgwt vnsastgkwt vnsastbwt vnsastbkwt vnsavnsawwt vnsavnsawkwt vnsavnsagwt vnsavnsagkwt vnsavnsabwt vnsavnsabkwt vnsacivwwt vnsacivwkwt vnsacivgwt vnsacivgkwt vnsacivbwt vnsacivbkwt

Civilian Source: civstwwt civstwkwt civstgwt civstgkwt civstbwt civstbkwt civvnsawwt civvnsawkwt civvnsagwt civvnsagkwt civvnsabwt civvnsabkwt civcivwwt civcivwkwt civcivgwt civcivgkwt civcivbwt civcivbkwt

Appendix G: Summary Statistics

Table G.1: Colombian Municipal-level Summary Statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Total VNSA White	106,731	0.14	4.4	0.0	596.8
Total VNSA Gray	106,731	0.28	6.6	0.0	536.2
Total VNSA Black	106,731	0.05	2.4	0.0	322.1
Total State White	106,731	2.69	46.6	0.0	2598.5
State-VNSA White	106,731	1.33	23.7	0.0	1871.0
State-VNSA Gray	106,731	0.03	1.9	0.0	206.1
State-VNSA Black	106,731	0.10	4.2	0.0	880.8
State-State White	106,731	0.90	21.5	0.0	1621.7
State-State Gray	106,731	0.02	1.6	0.0	229.5
State-State Black	106,731	0.00	0.0	0.0	3.0
State-Civilian White	106,731	0.46	10.0	0.0	747.2
State-Civilian Gray	106,731	0.04	2.3	0.0	248.5
State-Civilian Black	106,731	0.01	1.0	0.0	190.0
Civilian-VNSA White	106,731	0.26	15.3	0.0	2993.7
Civilian-VNSA Gray	106,731	0.01	0.6	0.0	134.4
Civilian-State White	106,731	0.19	6.1	0.0	510.2
Civilian-State Gray	106,731	0.07	4.9	0.0	1343.5
VNSA-VNSA White	106,731	0.01	0.7	0.0	133.0
VNSA-VNSA Gray	106,731	0.02	1.5	0.0	237.5
VNSA-VNSA Black	106,731	0.01	0.7	0.0	133.0
VNSA-State White	106,731	0.09	3.0	0.0	285.3
VNSA-State Gray	106,731	0.11	3.6	0.0	385.4
VNSA-State Black	106,731	0.03	2.0	0.0	320.7
VNSA-Civilian White	106,731	0.04	2.5	0.0	554.0
VNSA-Civilian Gray	106,731	0.16	4.5	0.0	431.3
VNSA-Civilian Black	106,731	0.01	1.0	0.0	208.4
Kinetic	106,731	0.13	2.3	0.0	266.9

Table G.2: Colombian Department-level Summary Statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Total VNSA White	7,080	2.11	17.1	0.0	596.8
Total VNSA Gray	7,080	4.23	25.9	0.0	536.2
Total VNSA Black	7,080	0.70	9.4	0.0	322.1
Total State White	7,080	40.61	180.9	0.0	2675.5
State-VNSA White	7,080	20.08	93.1	0.0	1871.0
State-VNSA Gray	7,080	0.48	7.5	0.0	206.1
State-VNSA Black	7,080	1.44	16.1	0.0	880.8
State-State White	7,080	13.59	83.4	0.0	1714.5
State-State Gray	7,080	0.30	6.2	0.0	229.5
State-State Black	7,080	0.01	0.1	0.0	3.0
State-Civilian White	7,080	6.94	38.9	0.0	747.2
State-Civilian Gray	7,080	0.58	9.0	0.0	248.5
State-Civilian Black	7,080	0.19	3.7	0.0	190.0
Civilian-VNSA White	7,080	3.91	59.2	0.0	2993.7
Civilian-VNSA Gray	7,080	0.08	2.2	0.0	134.4
Civilian-State White	7,080	2.87	24.1	0.0	605.8
Civilian-State Gray	7,080	1.01	19.2	0.0	1343.5
VNSA-VNSA White	7,080	0.17	2.7	0.0	133.0
VNSA-VNSA Gray	7,080	0.27	5.7	0.0	237.5
VNSA-VNSA Black	7,080	0.08	2.6	0.0	133.0
VNSA-State White	7,080	1.32	11.6	0.0	285.3
VNSA-State Gray	7,080	1.62	13.8	0.0	385.4
VNSA-State Black	7,080	0.46	8.0	0.0	320.7
VNSA-Civilian White	7,080	0.63	9.6	0.0	554.0
VNSA-Civilian Gray	7,080	2.34	17.6	0.0	432.3
VNSA-Civilian Black	7,080	0.16	4.0	0.0	208.4
Kinetic	7,080	1.99	8.9	0.0	267.6

Table G.3: Ukrainian District-level Summary Statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Total VNSA White	14,091	0.19	3.42	0	166.69
Total VNSA Gray	14,091	0.53	7.07	0	234.00
Total VNSA Black	14,091	0.63	8.21	0	290.24
Total State White	14,091	4.74	73.47	0	2940.74
State-VNSA White	14,091	0.66	8.30	0	445.31
State-VNSA Gray	14,091	0.23	3.65	0	122.09
State-VNSA Black	14,091	2.40	39.83	0	1902.36
State-State White	14,091	3.30	58.50	0	2743.64
State-State Gray	14,091	0.23	4.35	0	209.30
State-State Black	14,091	0.09	2.05	0	138.49
State-Civilian White	14,091	0.78	13.65	0	920.77
State-Civilian Gray	14,091	0.82	20.65	0	1841.24
State-Civilian Black	14,091	0.10	2.01	0	87.21
Civilian-VNSA White	14,091	0.04	1.29	0	89.10
Civilian-VNSA Gray	14,091	0.00	0.07	0	3.41
Civilian-State White	14,091	1.85	49.46	0	4335.79
Civilian-State Gray	14,091	0.84	30.47	0	2629.36
VNSA-VNSA White	14,091	0.01	0.61	0	53.74
VNSA-VNSA Gray	14,091	0.03	1.13	0	55.28
VNSA-VNSA Black	14,091	0.01	0.93	0	110.57
VNSA-State White	14,091	0.15	2.68	0	104.55
VNSA-State Gray	14,091	0.34	5.04	0	201.68
VNSA-State Black	14,091	0.60	7.86	0	262.70
VNSA-Civilian White	14,091	0.03	0.99	0	55.82
VNSA-Civilian Gray	14,091	0.16	3.40	0	166.81
VNSA-Civilian Black	14,091	0.02	0.76	0	55.28
Kinetic	14,091	0.50	8.11	0	583.31

Table G.4: Ukrainian Region-level Summary Statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Total VNSA White	3,003	0.91	7.61	0	169.76
Total VNSA Gray	3,003	2.49	18.25	0	481.68
Total VNSA Black	3,003	2.95	21.97	0	458.90
Total State White	3,003	22.23	159.83	0	2941.74
State-VNSA White	3,003	3.10	19.32	0	445.31
State-VNSA Gray	3,003	1.10	8.58	0	164.42
State-VNSA Black	3,003	11.27	107.42	0	2357.76
State-State White	3,003	15.50	126.58	0	2743.64
State-State Gray	3,003	1.08	9.38	0	209.30
State-State Black	3,003	0.40	4.56	0	141.52
State-Civilian White	3,003	3.64	29.66	0	923.45
State-Civilian Gray	3,003	3.86	45.23	0	1841.24
State-Civilian Black	3,003	0.47	4.66	0	90.60
Civilian-VNSA White	3,003	0.18	2.80	0	89.10
Civilian-VNSA Gray	3,003	0.01	0.15	0	3.94
Civilian-State White	3,003	8.68	107.33	0	4337.86
Civilian-State Gray	3,003	3.96	66.40	0	2630.43
VNSA-VNSA White	3,003	0.06	1.32	0	53.74
VNSA-VNSA Gray	3,003	0.13	2.44	0	55.28
VNSA-VNSA Black	3,003	0.04	2.02	0	110.57
VNSA-State White	3,003	0.70	5.91	0	105.62
VNSA-State Gray	3,003	1.59	12.72	0	288.91
VNSA-State Black	3,003	2.84	21.17	0	457.90
VNSA-Civilian White	3,003	0.15	2.17	0	57.43
VNSA-Civilian Gray	3,003	0.77	8.64	0	312.31
VNSA-Civilian Black	3,003	0.08	1.92	0	79.31
Kinetic	3,003	2.36	18.97	0	583.31

Table G.5: Libyan District-level Summary Statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Total VNSA White	1,518	9.73	35.89	0	501.33
Total VNSA Gray	1,518	2.49	9.47	0	134.73
Total VNSA Black	1,518	2.42	10.66	0	216.08
Total Civilian Gray	1,518	0.79	4.06	0	81.38
VNSA-VNSA White	1,518	8.24	30.84	0	426.64
VNSA-VNSA Gray	1,518	1.38	5.77	0	105.80
VNSA-VNSA Black	1,518	1.95	9.02	0	193.25
Civilian-VNSA White	1,518	2.17	8.75	0	163.24
Civilian-VNSA Gray	1,518	0.74	3.85	0	80.33
Civilian-VNSA Black	1,518	0.28	1.39	0	20.20
Civilian-Civilian White	1,518	0.25	1.12	0	23.86
Civilian-Civilian Gray	1,518	0.05	0.36	0	7.24
Civilian-Civilian Black	1,518	0.01	0.12	0	2.12
VNSA-Civilian White	1,518	1.49	5.64	0	87.15
VNSA-Civilian Gray	1,518	1.10	4.60	0	67.32
VNSA-Civilian Black	1,518	0.47	2.30	0	44.55
Kinetic	1,518	4.88	18.36	0	267.28

Appendix H: Alternative Model Specifications

Table H.1: Colombian VNSA Zonal Preferences at the Municipal Level (Fixed Effects)

Variable	White		Gray		Black	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
State-VNSA White	0.00***	0.00	0.02***	0.00	-0.01***	0.00
State-VNSA Gray	0.06***	0.01	-0.01	0.01	-0.06***	0.00
State-VNSA Black	-0.02***	0.00	-0.11***	0.00	0.03***	0.00
Civilian-VNSA White	0.01***	0.00	-0.02***	0.00	0.01***	0.00
Civilian-VNSA Gray	0.23***	0.02	-0.15***	0.03	-0.02	0.01
VNSA-VNSA White	-0.12***	0.02	-0.05**	0.03	-0.01	0.01
VNSA-VNSA Gray	-0.06***	0.01	0.24***	0.01	-0.05***	0.01
VNSA-VNSA Black	-0.08***	0.02	0.12***	0.03	0.00	0.01
State-Civilian White	0.01***	0.00	-0.04***	0.00	0.00	0.00
State-Civilian Gray	-0.03***	0.01	-0.04***	0.01	0.02***	0.00
State-Civilian Black	0.05***	0.01	-0.01	0.02	0.07***	0.01
Kinetic	0.17***	0.01	1.00***	0.02	0.22***	0.01
Lagged DV	0.08***	0.00	0.02***	0.00	-0.04***	0.00
Constant	0.10***	0.01	0.15***	0.02	0.02***	0.01
Number of Obs.			106128			
Number of Groups			603			

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Fixed Effects Regression model. This model presents findings at the second order subnational administrative level.

Table H.2: State White Zone Preferences at the Municipal Level in Colombia (Fixed Effects)

Variable	Coefficient	SE
VNSA-State White	0.45***	0.03
VNSA-State Gray	-0.48***	0.02
VNSA-State Black	-0.93***	0.04
Civilian-State White	-0.16***	0.02
Civilian-State Gray	0.08***	0.02
State-State White	0.23***	0.01
State-State Gray	1.81***	0.05
State-State Black	42.70***	3.28
VNSA-Civilian White	0.23***	0.03
VNSA-Civilian Gray	0.14***	0.02
VNSA-Civilian Black	-0.66***	0.08
Kinetic	4.13***	0.05
Lagged DV	0.08***	0.00
Constant	1.73***	0.08
Number of Obs.	106128	
Number of Groups	603	

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Fixed Effects Regression model. This model presents findings at the second order subnational administrative level.

Table H.3: Colombian VNSA Zonal Preferences at the Departmental Level (Random Effects)

Variable	White		Gray		Black	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
State-VNSA White	0.02***	0.00	0.04***	0.00	0.00***	0.00
State-VNSA Gray	0.03	0.03	-0.04	0.04	-0.07***	0.02
State-VNSA Black	-0.06***	0.01	-0.17***	0.02	0.03***	0.01
Civilian-VNSA White	0.01***	0.00	-0.02***	0.00	0.01***	0.00
Civilian-VNSA Gray	0.26***	0.08	-0.11	0.11	-0.01	0.05
VNSA-VNSA White	-0.03	0.07	0.41***	0.10	-0.01	0.04
VNSA-VNSA Gray	0.05	0.03	0.46***	0.05	-0.05***	0.02
VNSA-VNSA Black	-0.12*	0.07	0.03	0.10	0.00	0.04
State-Civilian White	0.04***	0.01	-0.01	0.01	0.00	0.00
State-Civilian Gray	0.02	0.02	-0.03	0.03	0.01	0.01
State-Civilian Black	0.08	0.05	0.00	0.07	0.07**	0.03
Kinetic	0.50***	0.03	1.37***	0.05	0.22***	0.02
Lagged DV	0.10***	0.01	0.03**	0.01	-0.04***	0.01
Constant	0.14	0.18	0.78***	0.26	0.30***	0.11
Number of Obs.			7040			
Number of Groups			40			

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Random Effects GLS Regression model. This model presents findings at the first order subnational administrative level.

Table H.4: State White Zone Preferences at the Departmental Level in Colombia (Random Effects)

Variable	Coefficient	SE
VNSA-State White	1.05***	0.14
VNSA-State Gray	-0.28***	0.11
VNSA-State Black	-1.14***	0.17
Civilian-State White	0.29***	0.06
Civilian-State Gray	-0.07	0.07
State-State White	0.32***	0.03
State-State Gray	2.38***	0.22
State-State Black	100.21***	14.59
VNSA-Civilian White	-0.27*	0.14
VNSA-Civilian Gray	0.04	0.09
VNSA-Civilian Black	-1.18***	0.34
Kinetic	6.66***	0.22
Lagged DV	0.34***	0.02
Constant	7.30***	1.35
Number of Obs.	7040	
Number of Groups	40	

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Random Effects GLS Regression model. This model presents findings at the first order subnational administrative level.

Table H.5: Colombian VNSA Zonal Preferences at the Departmental Level (Fixed Effects)

Variable	White		Gray		Black	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
State-VNSA White	0.01**	0.00	0.02***	0.00	-0.01***	0.00
State-VNSA Gray	0.06**	0.02	-0.01	0.04	-0.07***	0.02
State-VNSA Black	-0.03**	0.01	-0.14***	0.02	0.03***	0.01
Civilian-VNSA White	0.01***	0.00	-0.02***	0.00	0.01***	0.00
Civilian-VNSA Gray	0.22***	0.08	-0.16	0.11	-0.01	0.05
VNSA-VNSA White	-0.12*	0.07	0.30***	0.10	-0.01	0.04
VNSA-VNSA Gray	0.05	0.03	0.46***	0.05	-0.06***	0.02
VNSA-VNSA Black	-0.09	0.07	0.06	0.10	0.01	0.04
State-Civilian White	0.01	0.01	-0.04***	0.01	0.00	0.00
State-Civilian Gray	-0.02	0.02	-0.06**	0.03	0.01	0.01
State-Civilian Black	0.05	0.05	-0.03	0.07	0.06**	0.03
Kinetic	0.16***	0.04	1.03***	0.06	0.22***	0.02
Lagged DV	0.08***	0.01	0.02*	0.01	-0.05***	0.01
Constant	1.41***	0.19	2.04***	0.27	0.32***	0.12
Number of Obs.			7040			
Number of Groups			40			

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Fixed Effects Regression model. This model presents findings at the first order subnational administrative level.

Table H.6: State White Zone Preferences at the Departmental Level in Colombia (Fixed Effects)

Variable	Coefficient	SE
VNSA-State White	0.48***	0.12
VNSA-State Gray	-0.52***	0.10
VNSA-State Black	-0.64***	0.16
Civilian-State White	-0.17***	0.06
Civilian-State Gray	0.00	0.07
State-State White	0.21***	0.03
State-State Gray	1.92***	0.20
State-State Black	62.69***	13.14
VNSA-Civilian White	0.12	0.13
VNSA-Civilian Gray	0.06	0.08
VNSA-Civilian Black	-0.93***	0.30
Kinetic	3.93***	0.21
Lagged DV	0.10***	0.02
Constant	26.00***	1.29
Number of Obs.	7040	
Number of Groups	40	

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Fixed Effects Regression model. This model presents findings at the first order subnational administrative level.

Table H.7: Ukrainian VNSA Zonal Preferences at the District Level (Fixed Effects)

Variable	White		Gray		Black	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
State-VNSA White	0.02***	0.00	0.02**	0.01	0.12***	0.01
State-VNSA Gray	0.07***	0.01	0.29***	0.02	-0.09***	0.02
State-VNSA Black	-0.02***	0.00	-0.03***	0.00	0.00**	0.00
Civilian-VNSA White	-0.11***	0.03	-0.15***	0.05	0.77***	0.06
Civilian-VNSA Gray	0.29	0.43	15.60***	0.82	22.08***	1.02
VNSA-VNSA White	0.18***	0.05	0.08	0.08	1.73***	0.10
VNSA-VNSA Gray	-0.04*	0.02	-0.10**	0.04	-0.11**	0.05
VNSA-VNSA Black	0.07**	0.03	0.03	0.05	-0.31***	0.06
State-Civilian White	-0.02***	0.00	-0.01*	0.01	-0.10***	0.01
State-Civilian Gray	0.00	0.00	-0.03***	0.01	0.03***	0.01
State-Civilian Black	-0.12***	0.01	-0.19***	0.03	-0.31***	0.03
Kinetic	0.19***	0.01	0.32***	0.02	-0.04*	0.02
Lagged DV	-0.10***	0.01	0.03***	0.01	0.17***	0.01
Constant	0.17***	0.02	0.36***	0.05	0.48***	0.06
Number of Obs.	13664					
Number of Groups	427					

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Fixed Effects Regression model. This model presents findings at the second order subnational administrative level.

Table H.8: State White Zone Preferences at the District Level in Ukraine (Fixed Effects)

Variable	Coefficient	SE
VNSA-State White	1.12***	0.11
VNSA-State Gray	0.12*	0.06
VNSA-State Black	0.85***	0.04
Civilian-State White	0.44***	0.02
Civilian-State Gray	-0.51***	0.03
State-State White	0.12***	0.03
State-State Gray	-0.92***	0.08
State-State Black	-0.32**	0.13
VNSA-Civilian White	-1.93***	0.26
VNSA-Civilian Gray	-0.68***	0.08
VNSA-Civilian Black	-3.18***	0.32
Kinetic	2.21***	0.13
Lagged DV	0.19***	0.03
Constant	1.66***	0.23
Number of Obs.	13664	
Number of Groups	427	

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Fixed Effects Regression model. This model presents findings at the second order subnational administrative level.

Table H.9: Ukrainian VNSA Zonal Preferences at the Regional Level (Random Effects)

Variable	White		Gray		Black	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
State-VNSA White	0.07***	0.01	0.03	0.02	0.40***	0.02
State-VNSA Gray	0.13***	0.02	0.58***	0.04	0.04	0.04
State-VNSA Black	-0.01***	0.00	-0.02***	0.00	0.02***	0.00
Civilian-VNSA White	-0.24***	0.06	-0.47***	0.11	0.46***	0.13
Civilian-VNSA Gray	-2.58***	0.97	13.68***	2.05	19.28***	2.36
VNSA-VNSA White	0.18*	0.10	-0.16	0.19	1.46***	0.22
VNSA-VNSA Gray	0.10**	0.05	-0.51***	0.11	-0.22*	0.13
VNSA-VNSA Black	-0.01	0.06	-0.52***	0.12	-0.68***	0.15
State-Civilian White	0.00	0.01	0.01	0.02	-0.20***	0.02
State-Civilian Gray	-0.02***	0.01	-0.10***	0.01	-0.06***	0.02
State-Civilian Black	-0.11***	0.03	-0.09	0.07	-0.41***	0.08
Kinetic	0.26***	0.02	0.53***	0.04	0.45***	0.04
Lagged DV	-0.06**	0.02	0.20***	0.02	0.23***	0.02
Constant	0.27**	0.12	0.57**	0.25	0.52*	0.30
Number of Obs.			2912			
Number of Groups			91			

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Random Effects GLS Regression model. This model presents findings at the first order subnational administrative level.

Table H.10: Ukrainian VNSA Zonal Preferences at the Regional Level (Fixed Effects)

Variable	White		Gray		Black	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
State-VNSA White	0.04***	0.01	-0.09***	0.02	0.28***	0.02
State-VNSA Gray	0.07***	0.02	0.40***	0.04	-0.15***	0.04
State-VNSA Black	-0.02***	0.00	-0.05***	0.00	-0.02***	0.00
Civilian-VNSA White	-0.18***	0.06	-0.22**	0.11	0.71***	0.13
Civilian-VNSA Gray	-1.91**	0.95	15.93***	1.97	21.44***	2.30
VNSA-VNSA White	0.19*	0.10	-0.54***	0.19	0.90***	0.22
VNSA-VNSA Gray	0.07	0.05	-0.45***	0.10	-0.12	0.12
VNSA-VNSA Black	-0.02	0.06	-0.50***	0.12	-0.64***	0.14
State-Civilian White	-0.01	0.01	-0.01	0.02	-0.22***	0.02
State-Civilian Gray	-0.01	0.01	-0.08***	0.01	-0.06***	0.02
State-Civilian Black	-0.16***	0.03	-0.26***	0.07	-0.51***	0.08
Kinetic	0.23***	0.02	0.46***	0.04	0.38***	0.05
Lagged DV	-0.13***	0.02	0.12***	0.02	0.13***	0.02
Constant	0.68***	0.12	1.96***	0.26	2.04***	0.30
Number of Obs.			2912			
Number of Groups			91			

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Random Effects GLS Regression model. This model presents findings at the first order subnational administrative level.

Table H.11: State White Zone Preferences at the Regional Level in Ukraine (Random Effects)

Variable	Coefficient	SE
VNSA-State White	1.11***	0.27
VNSA-State Gray	0.12	0.14
VNSA-State Black	0.86***	0.08
Civilian-State White	0.49***	0.04
Civilian-State Gray	-0.16***	0.06
State-State White	-0.12*	0.07
State-State Gray	-0.54***	0.20
State-State Black	-1.11***	0.33
VNSA-Civilian White	-2.28***	0.64
VNSA-Civilian Gray	-0.76***	0.17
VNSA-Civilian Black	-1.72***	0.66
Kinetic	-0.27	0.26
Lagged DV	0.82***	0.06
Constant	1.07	1.23
Number of Obs.	2912	
Number of Groups	91	

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Fixed Effects Regression model. This model presents findings at the first order subnational administrative level.

Table H.12: State White Zone Preferences at the Regional Level in Ukraine (Fixed Effects)

Variable	Coefficient	SE
VNSA-State White	0.96***	0.25
VNSA-State Gray	0.33***	0.13
VNSA-State Black	0.73***	0.08
Civilian-State White	0.44***	0.03
Civilian-State Gray	-0.42***	0.06
State-State White	0.05	0.07
State-State Gray	-0.70***	0.18
State-State Black	-0.25	0.30
VNSA-Civilian White	-1.49**	0.58
VNSA-Civilian Gray	-1.00***	0.15
VNSA-Civilian Black	-2.44***	0.61
Kinetic	1.61***	0.27
Lagged DV	0.26***	0.06
Constant	8.26***	1.14
Number of Obs.	2912	
Number of Groups	91	

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Fixed Effects Regression model. This model presents findings at the first order subnational administrative level.

Table H.13: Libyan VNSA Zonal Preferences at the Departmental Level (Fixed Effects)

Variable	White		Gray		Black	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
VNSA-VNSA White	-0.86***	0.23	0.01	0.01	0.04**	0.02
VNSA-VNSA Gray	-0.10	0.17	0.31***	0.11	0.22**	0.11
VNSA-VNSA Black	0.43**	0.19	-0.37***	0.09	-1.38***	0.23
Civilian-VNSA White	0.33**	0.15	0.06	0.06	-0.05	0.07
Civilian-VNSA Gray	0.45	0.28	-0.56***	0.12	-0.28*	0.14
Civilian-VNSA Black	-0.29	0.49	-0.65***	0.20	0.17	0.25
Civilian-Civilian White	3.61***	0.78	-0.08	0.32	0.68*	0.42
Civilian-Civilian Gray	6.62***	1.67	0.42	0.67	4.59***	0.86
Civilian-Civilian Black	-5.25	3.61	-1.40	1.47	1.22	1.86
Kinetic	-0.38**	0.17	0.36***	0.08	-0.02	0.09
Lagged DV	1.28***	0.22	-0.15	0.09	1.44***	0.23
Constant	3.47***	0.46	1.78***	0.19	0.98***	0.24
Number of Obs.	1496					
Number of Groups	22					

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Standard Errors clustered by country, using a Fixed Effects Regression model. This model presents findings from the Libya dataset at the first order subnational administrative level.

Table H.14: Libyan Civilian Gray Zone Preferences at the Departmental Level (Fixed Effects)

Variable	Coefficient	SE
VNSA-Civilian White	0.24***	0.05
VNSA-Civilian Gray	-0.06	0.05
VNSA-Civilian Black	-0.14	0.10
VNSA-VNSA White	0.00	0.01
VNSA-VNSA Gray	0.03	0.05
VNSA-VNSA Black	-0.08*	0.05
Civilian-Civilian White	-0.21	0.17
Civilian-Civilian Gray	0.75**	0.37
Civilian-Civilian Black	-0.76	0.80
Kinetic	0.03	0.04
Lagged DV	-0.06	0.06
Constant	0.65***	0.10
Number of Obs.	1496	
Number of Groups	22	

* p <0.1; ** p<0.05; *** p<0.01 (two-tailed tests). Reported Standard Errors clustered by country, using a Fixed Effects Regression model. This model presents findings from the Libya dataset at the first order subnational administrative level.

References

Abubakar, Ali Abare. 2017. "Three Years Later, Nearly 200 Nigerian Schoolgirls Still Missing." *USA Today*. (April 13). https://www.usatoday.com/story/news/world/2017/04/13/chibok-nigeria-girls-still-missing-boko-haram/100319108/.

Ayyub, Bilal M. 2003. *Risk Analysis in Engineering and Economics* 579. Boca Raton, FL: Chapman & Hall/CRC.

Department of Defense Strategic Multi-Layer Assessment. 2017. "Updated Gray Zone Definition Based off SME Feedback v2." (January).

–. 2016. "Gray Zone Effort Update." (September).

Diatchkova, Galina. 2008. "Indigenous Media as an Important Resource for Russia's Indigenous Peoples." Global Indigenous Media: Cultures, Poetics, and Politics. Pamela Wilson and Michelle Stewart, editors. Durham, NC: Duke University Press. 214-31.

Finkel, Evgeny. 2016. "The Conflict in the Donbas between Gray and Black: The Importance of Perspective." Report to DHS S&T Office of University Programs and DoD Strategic Multilayer Assessment Branch (College Park, MD: START).

Fitzgerald, Mary, and Mattia Toaldo. 2017. "Mapping Libya's Factions: A Quick Guide to Libya's Main Players." ECFR MENA Programme. *European Council on Foreign Relations*. http://www.ecfr.eu/mena/mappinglibya.

Friedersdorf, Conor. 2013. "Whopper of the Year: 'The Mainstream Media' Ignored Benghazi." *The Atlantic*. (August 5). https://www.theatlantic.com/politics/archive/2013/08/whopper-of-the-year-the-mainstream-media-ignored-benghazi/278351/.

Håvard Hegre, Nils W Metternich, Håvard Mokleiv Nygård, Julian Wucherpfennig, Daina Chiba, and Kristian Skrede Gleditsch. 2017. "The Shape of Things to Come? Expanding the Inequality and Grievance Model for Civil War Forecasts with Event Data." *Journal of Peace Research* 54(2): 275-97.

Gabriel, Rachel A., and Mila A. Johns. 2016. "Demystifying Gray Zone Conflict: A Typology of Conflict Dyads and Instruments of Power in Libya, 2014-Present." Report to DHS S&T Office of University Programs and DoD Strategic Multilayer Assessment Branch (College Park, MD: START).

Gomati, Anas El. 2014. "Khalifa Haftar: Fighting Terrorism or Pursuing Political Power," *Al Jazeera English.* (June 10). http://www.aljazeera.com/indepth/opinion/2014/06/khalifa-hifter-operation-dignity-20146108259233889.html.

Jensen, Finn V. 1996. *An Introduction to Bayesian Networks* 210. London: UCL Press.

Joyce, Mary. 2013. "Picking the Best Intercoder Reliability Statistic for Your Digital Activism Content Analysis." *Digital Activism Research Project*. (May 11). http://digital-activism.org/2013/05/picking-the-best-intercoder-reliability-statistic-for-your-digital-activism-content-analysis/.

Koven, Barnett S. 2016. "Demystifying Gray Zone Conflict: A Typology of Conflict Dyads and Instruments of Power in Colombia, 2002-present." Report to DHS S&T Office of University Programs and DoD Strategic Multilayer Assessment Branch (College Park, MD: START).

–. 2016. "Demystifying Gray Zone Conflict: A Typology of Conflict Dyads and Instruments of Power in Colombia, Libya and Ukraine." Report to DHS S&T Office of University Programs and DoD Strategic Multilayer Assessment Branch (College Park, MD: START).

Krippendorff, Klaus. 2011. Computing Krippendorff's Alpha-Reliability. Retrieved from http://repository.upenn.edu/asc_papers/43.

–. 2004. Reliability in Content Analysis: Some Common Misconceptions and Recommendations. *Human Communication Research 30*: 411-33.

Michael D. Ward, Andreas Beger, Josh Cutler, Matthew Dickenson, Cassy Dorff, and Ben Radford. 2013. "Comparing GDELT and ICEWS Event Data." *Analysis 21*: 267-97.

Neuendorf, Kimberly A. 2002. *The Content Analysis Guidebook.* Thousand Oaks, CA: Sage.

Office of the United Nations High Commissioner for Human Rights. 2015. *Report On The Human Rights Situation In Ukraine: 1 December 2014 to 15 February 2015*. New York: United Nations. http://www.ohchr.org/Documents/Countries/UA/9thOHCHRreportUkraine.pdf.

Ricchiardi, Sherry. 2008. "Whatever Happened To Iraq? How The Media Lost Interest In A Long-Running War With No End In Sight." *American Journalism Review*. 30(3): 20.

Schein, Aaron, John Paisley, David M. Blei, and Hanna Wallach. 2015. "Bayesian Poisson Tensor Factorization for Inferring Multilateral Relations from Sparse Dyadic Event Counts." KDD '15 Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Sydney, Australia. (August 10 – 13). http://delivery.acm.org/10.1145/2790000/2783414/p1045-schein.pdf.

Schrodt, Philip A. 2012a. "CAMEO: Conflict and Mediation Event Observations Event and Actor Codebook." Event Data Project, Pennsylvania State University. http://eventdata.parusanalytics.com/cameo.dir/CAMEO.Manual.1.1b3.pdf.

–. 2012b. "Precedents, Progress, and Prospects in Political Event Data." *International Interactions* 38(4): 546-69.

Schrodt, Philip A., and David Van Brackle. 2012. "Automated Coding of Political Event Data." *Handbook of Computational Approaches to Counterterrorism*. V.S. Subrahmanian, editor. New York: Springer: 23-49.

Schrodt, Philip A., and Omur Yilmaz. 2007. "CAMEO: Conflict and Mediation Event Observations Codebook." Center for International Political Analysis, University of Kansas. http://eventdata.parusanalytics.com/cameo.dir/CAMEO.CDB.09b5.pdf.

Shuster, Simon. 2014. "Exclusive: Leader of Far-Right Ukrainian Militant Group Talks Revolution with TIME." Time. February 4. http://time.com/4493/ukraine-dmitri-yarosh-kiev/.

Sober, Elliott. 2013. *Core Questions in Philosophy: A Text with Readings* 6th ed. Boston: Pearson Education.

Weber, Max. 1918. "Politics as a Vocation." Lecture, Free Students Union, Bavaria, Germany, January 28.

Wang, Wei, Ryan Kennedy, David Lazer, and Naren Ramakrishnan. 2016. "Growing Pains for Global Monitoring of Societal Events." *Science* 353(6307): 1502-3.