DOI:
10.1093/jogss/ogy015

Document Version
Peer reviewed version

Citation for published version (APA):
Jäger, K. (2018). The Limits of Studying Networks Via Event Data: Evidence from the ICEWS Dataset. *The Journal of Global Security Studies* , *3*(4), 498-511. https://doi.org/10.1093/jogss/ogy015

The Limits of Studying Networks with Event Data:

Evidence from the ICEWS Dataset

ABSTRACT

Machine-coded event datasets have become popular in conflict research. I argue that systematic media biases render news-based event data unsuitable for studying anti-government networks of insurgents and political parties. Insurgent networks are too secretive to be captured by media reports, whereas alliances among regular political parties are too constant to be considered newsworthy. I analyze the data accuracy of Metternich et al.'s (2013) network study of insurgents and political parties in Thailand, which is based on the most comprehensive event dataset currently available – the ICEWS project. Based on simple evaluation criteria, I show that most of the network data entries are incorrect, leading to a depiction of the networks that is unrelated to real-world cleavages in Thailand. While my hand-coded event dataset captures relatively more network-relevant information than ICEWS, the comparison confirms that journalists specifically underreport cooperative events among insurgents and parties. In addition, the ICEWS project provides unreliable counts of conflictual events in Thailand. Using alternative conflict measurements from the Deep South Watch dataset, and a dummy variable based on established periods of unrest, I show that violent activities in Thailand's Deep South declined during periods of conflict between pro- and anti-Thaksin groups. Conflicts were unrelated to network fragmentation, contradicting Metternich et al.'s primary finding.

KEYWORDS

Event Data, Media Bias, Machine Coding, Network Analysis, ICEWS, Intra-State Conflict, Thailand.

1. Introduction

Recent years have witnessed a rapid increase in the proliferation of large-scale event datasets for quantitative analysis. The "Big Data" revolution poses a great opportunity to advance conflict research, but has been accompanied by discussions of systematic media and machine-coding biases that plague existing event datasets (Earl et al. 2004; Althaus et al. 2011; Schrodt 2012; Hammond and Weidmann 2014; Baum and Zhukov 2015; Weidmann 2015, 2016). In this article, I focus on media bias, which has negative consequences for social network analysis. I show that insurgent networks are too secretive to be accurately depicted by media reports, and activities of regular political parties are often too well-known and constant over time to be considered newsworthy. As a consequence, media-based event data fails to draw an accurate picture of insurgent networks and regular party competition.

I illustrate the shortcomings of media-based event data by analyzing the accuracy of Metternich et al.'s (2013) study of the anti-government network in Thailand, which comprises insurgents, social movements and political parties. The authors find that a fragmented anti-government network was associated with a greater number of conflictual events, based on daily event data from Lockheed Martin's International Crisis Early Warning System (ICEWS). By using simple criteria to evaluate the ICEWS event observations, I find that about 80 percent of the entries used in their analysis are inaccurate, and that the network placement of groups is unrelated to real-world political cleavages in Thailand. Furthermore, Metternich et al.'s (2013) results cannot be replicated with alternative conflict measurements such as the Deep South Watch dataset. My analysis shows that network fragmentation, measured by the lowest eigenvalue of the anti-government network, does not correspond with more violence, and that conflicts in Thailand's

Deep South actually declined during periods of political unrest between pro- and anti-Thaksin groups.

The implications of this evaluation go beyond Thailand. What distinguishes ICEWS from other event-data projects is its scale. It utilized 6.5 million media reports from over 75 regional and international news sources over the period 1998-2006 to construct a raw dataset of daily events for 29 countries. In a subsequent version, ICEWS coverage was extended to create a global dataset with about 200 news sources (O'Brien 2010; Schrodt 2012, 547-8). Metternich et al. (2013) was the first study that utilized the expanded dataset for a country study, and ICEWS has recently been made available to the public (Boschee et al. 2015). Given its status as "the current gold standard for event data" (Metternich et al. 2013, 901), my findings imply that less comprehensive event datasets may suffer from the same problems.

Following a brief discussion on media bias, the article proceeds as follows: I conduct several analyses that draw into question both the reliability of ICEWS and the main finding of Metternich et al. I analyze Metternich et al.'s network analysis and the problems of false negatives and positives in ICEWS that result from media bias. I draw particular attention to the media's sparse documentation of cooperative events among insurgents and political parties: Such interactions are critical to obtaining a complete picture of the anti-government network. Metternich et al. use cooperative events relative to conflictual ones to construct their network measurement. Subsequently, to illustrate the unreliability of ICEWS more concretely, I discuss qualitative scholarship that shows real-world Thailand to be quite different than ICEWS' version.

After looking in-depth at Metternich et al.'s findings, I turn to my own hand-coded event dataset based on the news database LexisNexis in order to evaluate whether the detected problems were caused by the automated coding scheme or by systematic media underreporting. My analysis confirms the pervasiveness of false negatives - media reports systematically fail to adequately capture the interactions of anti-government groups, including both insurgents and political groups. Finally, I directly challenge Metternich et al.'s main finding on the relationship between anti-government network fragmentation and conflict by replicating their analysis with two alternative conflict measurements. The article concludes with the suggestion that researchers should be cautious when relying exclusively on machine-coded event data for social network analysis.

2. Problems with Using Media-based Event Data for Network Studies

Media reports suffer from both description and selection biases. A description bias exists if media reports lack sufficient information to represent an event accurately. A selection bias exists if not all relevant events are reported (Weidmann 2015, 1130). Previous studies discuss how such systematic biases could lead to an incorrect event count, which regularly functions as the dependent variable in conflict studies. Althaus et al. (2011) show that the reporting of U.S. war deaths was unrelated to how many U.S. soldiers actually died in five major conflicts. Relatedly, Hammond and Weidmann (2014) and Weidmann (2015) show that machine-coded datasets produce an incorrect geo-location of events because journalists often report from urban centers and mention a country's capital more often than the event's precise location in their stories.

As conflict research has moved towards disaggregation and studying micro-level complexity, the relations of local agents and their network structures have increasingly been used as independent variables to explain popular mobilization, rebel recruitment, or the occurrence of violent events (Cederman and Gleditsch 2009; Kalyvas 2012; Balcells and Justino 2014).

Gleditsch, Metternich, and Ruggeri (2014, 305) suggest that "many researchers have become interested in less aggregated data on conflict processes and turned to event data that may allow studying interaction by arranging specific actions as temporal sequences." I argue, however, that news-based event data is unsuitable for constructing network variables of insurgent groups and political parties for conflict studies, which might include centrality measures like betweenness, closeness, and eigenvectors. Journalists systematically underreport facts that are too secretive (description bias) or too mundane (selection bias). In the analysis below, I expect the former bias to affect reports on insurgent networks, whereas the latter affects reporting on regular party competition. Both types of bias generate false negatives in ICEWS.

Events only make it into the news if they are observed, and they are not equally visible. Weidmann (2015, 1134-5) finds that events in remote or dangerous locations are underreported because they are less accessible for observers. Thompson (2007, 4) describes the reasons for the limited media coverage of the Rwandan civil war in the following way: "They [U.S. correspondents] had scarcely arrived when their editors gave them all orders to come home. […] I picked up snatches of their conversations: 'Too dangerous, not enough interest […] deep Africa, you know […] middle of nowhere.'" In the case of insurgents' networks (i.e., groups that engage in illegal anti-government action), reporting is not only dangerous for the journalist, but it is also essential for the survival of insurgent networks that their organization structure remains unknown to security authorities and the public. Media reports cannot be an exogenous source of objectivity due to their public nature. Even if investigative media reports were indeed able to reveal insurgents' networks, insurgents would abandon these ties immediately as security authorities would try to use the information to destroy the insurgent groups.[1] Thus, accurate reporting would affect the behavior of insurgent organizations.

The illegal nature of Muslim insurgent groups in Thailand's Deep South forces them to operate in decentralized local cells. This problem likely applies to insurgent networks in other countries, all of whom must operate in an opaque environment beyond the reach of the media. Consequently, journalists rely on less direct sources for information on insurgent networks, particularly the security authorities. But security authorities have incentives to withhold information that if publicly known would thwart their anti-insurgent campaigns. Thus, they tend to disseminate stories to the public which serve their objectives. For instance, in 2007 Thai security authorities distributed selected information to newspapers to establish the questionable belief that Barisan Revolusi Nasional (BRN) was the leading organization in the insurgency. By contrast, academic research suggests that the insurgent groups were decentralized units without demonstrated links to older hierarchical insurgent organizations. Their hierarchies and interactions remain largely unknown (Askew 2010; Croissant 2007, 4-6; McCargo 2008, 168-81; Srisompob and McCargo 2010, 169-70). Media reports are thus unable to reveal the true network structure of insurgent groups.[2]

While insurgents' networks are generally too secretive to be objectively captured in news-based data, the opposite is the case for alliances among regular political parties or social movements, which in themselves lack the newsworthiness to be regularly reported. When journalists decide which material to include in reports, they tend to prefer certain characteristics

[1] Additionally, McCargo (2000) shows that the Thai national media tended to lack investigative reporting and provincial coverage. Reports often tended to favor the political inclination of media owners.
[2] In addition, the political alignment of insurgent groups influences reporting. Baum and Zhukov (2015) find that the media in democracies (non-democracies) had a pro-challenger (incumbent) bias in their coverage of the Libyan civil war.

of information. As already implied by the term "news," the empirical evidence suggests that magazines, newspapers, or news agencies want to report information that is novel, unexpected, or surprising (Earl et al. 2004, 73-4; Groeling and Baum 2008, 1068; Baum and Zhukov 2015, 385).

Party networks or alliances of groups within a party family appear to be stable over time. From the perspective of media reports, then, the stability of party system alliances should lead to a systematic underreporting of these ties, because journalists do not consider them newsworthy. It is unlikely that they will regularly report on constant relationships, which comprise trivial and obvious information. For example, Baum and Groeling (2010) show that media reports are more likely to cover criticism of the U.S. president if it comes from his own party rather than from the opposition. The underreporting of expected political ties is a major problem for news-based network analysis, which requires that cooperation between allied groups be regularly mentioned in media reports in order to place these groups relatively closer to each other in the network. But reports tend to discuss party alliances more often if they are changing, or if one of the groups generates newsworthy events. As a consequence for network analysis, any association between network activity and conflicts may be exaggerated, reflecting journalists' inclination to report more often on political groups involved in upheaval of some sort.

Besides the media description and selection biases, Schrodt (2012, 552-7) highlights that machine coding, albeit technologically advancing, is susceptible to including duplicates from different news sources as well as irrelevant stories. Machine-coding protocols that include widely used language might capture events that are unrelated to phenomena of interest, which has been "the bane of the automated processing of event data source texts from the beginning" (Schrodt 2012, 553). The automated coding scheme also could mistake a generic name for a network group name or confuse unrelated verbs with interaction between groups. The problems created by duplicate or irrelevant stories are false positives – data entries that wrongly become part of a dataset.

<<< TABLE 1 >>>

Table 1 summarizes the discussion of the consequences of media biases for machine-coded event data accuracy. False positives generated in the machine-coded data tend to be relatively easy to rectify, as they require only the removal of incorrect entries. By contrast, the false negatives in a dataset – generated by description and selection biases – are more difficult to rectify because researchers would have to add correct entries to the dataset. Such a task requires qualitative case knowledge about political groups, which is often beyond the scope of large-scale data projects. Without correcting for false negatives, it is likely that the media biases cause systematic underreporting of cooperative interactions among insurgents and political parties.

3. Evaluating the Network Data Accuracy of ICEWS for Thailand

Metternich et al. (2013) utilizes ICEWS to study how the characteristics of the anti-government network affect conflict in Thailand. The authors claim that more conflict occurs when the opposition is fragmented and programmatically polarized. Metternich et al. (2013, 901) use the following selection criteria to construct the anti-government network: A domestic group becomes a member of the anti-government network if it is involved in more conflictual than cooperative interactions with the government in a given month, and drops out of the network if it is inactive for 24 months or if the number of cooperative interactions with the government exceeds the number of conflictual ones.

Through this automated process, 24 groups emerge as potential members of the anti-government network, which includes one party twice under both its English and Thai name (Thai Nation Party/Chart Thai), and groups that do not exist in Thailand, such as the "Outruled Democratic Party," based in the Vietnamese city of *Thái* Bình. Metternich and colleagues use the recorded cooperative events between groups of the anti-government network to conduct a latent space analysis (Hoff, Raftery, and Handcock 2002) estimating the position of groups in a two-dimensional space. The authors claim that the group placements "reflect their closeness to one another, preferences for similar policies, and propensity to interact cooperatively" (Metternich et al. 2013, 902). The authors use the latent positions to calculate their main explanatory variable, the lowest eigenvalues of the network, which measures the structure of the anti-government network. Metternich et al. (2013) find that small eigenvalues (i.e., a fragmented network structure) are associated with more conflictual events between the government and the anti-government network.

3.1 Detecting False Negatives in ICEWS

Metternich et al. (2013, 894) correctly state that "there are two ongoing internal political conflicts in Thailand," but they do not show if these cleavages are adequately represented by their network placement. The first, color-coded conflict centers upon support or opposition to Thaksin Shinawatra, Thailand's former and beleaguered Prime Minister. The main contention exists between the anti-Thaksin forces, the Democrat Party and the yellow shirts (People's Alliance for Democracy - PAD) on one side, and the Thaksin supporters, the Pheu Thai Party (PTP) and their social movement, the red shirts (United Front for Democracy against Dictatorship - UDD), on the other (Metternich et al. 2013, 894-5).[3] The conflict erupted in 2006 after Thaksin tried to break the power of "network monarchy," which inadvertently caused the emergence of a broad royalist movement against him (McCargo 2005, 2008). Survey research suggests that the color-coded conflict was driven by different concepts of democratic legitimacy. Red shirts supported the majoritarian elements of democracy and put salience on reducing economic inequality, while yellow shirts emphasized checks and balances and anti-corruption efforts (Jäger 2012; Jäger 2017, 337-9; Naruemon and McCargo 2011). Aside from the PTP and the Democrat Party, Thailand's other political parties do not have substantial numbers of supporters. These parties "come and go like soap bubbles" as they are "classic election machines and often preoccupied with achieving material gains for their leaders" (Croissant and Völkel 2012, 246-9). Such parties can engage in verbal conflict, but they do not have the potential to confront the government on the street, which is ignored by Metternich et al.[4]

The second conflict comprises a violent Muslim insurgency against the Thai state. The insurgents seek a higher level of autonomy or ultimate separation for Thailand's Muslim-majority Deep South (Metternich et al. 2013, 895-6). The authors implicitly separate the Muslim groups into known insurgent groups (BRN, Council of Muslim People, PULO I, PULO II, Gerakan Mujahidin Islam Patani - GMIP) and a general category (Southern Rebel Movement). Insurgent groups are clearly outside of the Thai political system and mainstream political parties do not

[3] The pro-Thaksin side also includes the predecessors of PTP, the People's Power Party (PPP) and Thai Rak Thai (TRT). The Democrats and yellow shirts ended their alliance under the Abhisit administration.

[4] Moreover, Metternich et al. do not include the Assembly of the Poor as a member of the anti-government network in the dataset. The Assembly of the Poor gradually lost momentum after the first election of Thaksin, but it was the main civil-society group to organize mass protests against the government before the color-coded conflict emerged (Narut 2013).

cooperate with them (Croissant 2007; McCargo 2008). This fact puts into immediate doubt Metternich et al.'s model which allows for cooperation among all groups.

Based on the basic characteristics of the two conflicts in Thailand, I propose the following two minimum necessary benchmarks to gauge ICEWS' ability to correctly capture the relevant cleavages:

1. ICEWS should capture more cooperative events for groups within the red pro-Thaksin movement, within the yellow anti-Thaksin movement, and within the Muslim insurgency relative to all other groups.
2. As a consequence, the groups within the pro-Thaksin movement, within the anti-Thaksin movement, and within the Muslim insurgency should be clustered closer to each other in a two-dimensional spatial analysis relative to all other groups, while the distance between these three should be comparatively large.

The media description and selection biases suggest that cooperation between political allies within the pro- and anti-Thaksin movements is not newsworthy enough, while cooperation among Muslim insurgents is too secretive, to be captured by media reports. Indeed, I find that ICEWS captures very few cooperative events within the pro- and anti-Thaksin groups, and the Muslim insurgents: among the 1,241 cooperative events of the dataset (yielding a group average of 103.4), the red shirts are twice coded in cooperation with the pro-Thaksin PPP when it was the governing party in November 2008, and 48 times thereafter when PPP was already dissolved. PTP, the successor of PPP as the pro-Thaksin party, is never coded in cooperation with its own social movement group. The anti-Thaksin Democrat Party and the yellow shirts have 11 cooperative events between March 2006 and October 2008. No cooperation is detected among Muslim insurgent groups. The rebel group PULO I is coded 102 times as in cooperation with the generic "Southern Rebel Movement." Thus, ICEWS appears to suffer from false negatives as it was unable to capture a substantial amount of cooperative events between groups that are closely aligned to each other in reality.

<<< FIGURE 1 >>>

Figure 1 shows the average placement of these groups in the two-dimensional political space. The pro-Thaksin groups, the anti-Thaksin groups, and the Muslim insurgents should be placed closer to each other, relative to the distance between these three movements.[5] However, Figure 1 shows that the ICEWS placement does not reflect actual alignment patterns of groups involved in the two Thai conflicts examined here. The Muslim insurgent groups have an average Euclidean distance of 0.87 from each other compared to a lower average distance of 0.73 from all Thai political groups. The average distance between a Thaksin-affiliated party and the red shirts is 0.76 – but it is 0.70 between red and yellow groups. Only the distance between the Democrats and yellow shirts has a reasonably low average value of 0.45, as ICEWS was able to record 11 cooperative events between them. But the lowest average distance in the dataset, 0.37, occurs between the pro-Thaksin PPP and the anti-Thaksin yellow shirts, while the largest average distance of 1.54 stretches between the pro-Thaksin PTP and its own movement, the red shirts.

[5] As described by Metternich et al. (2013, 903), "actors that are close in space have more cooperative interactions with each other and have similar cooperative behavior patterns with other actors in the network."

3.2 Detecting False Positives in ICEWS

To illustrate the problem of false positives presented by machine-coded data, I subject Metternich et al.’s (2013) dataset to the following simple, common-sense rules from real-world Thailand to detect false positives in the form of 1) actors falsely identified as members in the anti-government network, and 2) events falsely coded as “cooperative” within the anti-government network:

1. A group cannot be part of the anti-government network if it is actually part of the government coalition.
2. A group cannot be part of the anti-government network if it has not been established yet or if it has ceased to exist.
3. A group cannot be part of the anti-government network if it is unidentifiable or a foreign group.
4. Cooperative events between groups are not possible if at least one group falls under criteria 1 to 3.
5. General or unspecified group categories such as “Southern Rebel Movement” or “Unspecified Groups” are excluded from the analysis, as one cannot assume that they refer to a unitary actor over time.

<<< FIGURE 2 >>>

Figure 2 shows that large portions of the dataset are based on false positives: Between January 2001 and December 2010, there are 1,416 entries that identify anti-government network members. According to Criteria 1 to 3, 763 entries (53.9 percent) are incorrectly placed in the network. Furthermore, 109 entries (7.7 percent) and 108 entries (7.6 percent) fall into the general categories of "Southern Rebel Movement" and "Unspecified Groups."

For cooperative events – which are crucial to calculate the network's lowest eigenvalue – the error rate increases further. The dataset includes 1,241 cooperative events between two groups, of which 978 events (78.8 percent) are incorrect according to Rule 4. There are also 167 cooperative events (13.5 percent) involving the general category of "Southern Rebel Movement." This indicates that ICEWS suffers substantially from false positives.[6]

3.3 Qualitative evidence

The problems detected in the quantitative analysis can be illustrated by looking at the historical record of periods of political unrest in Thailand: In February 2006, mass demonstrations by the yellow shirts against the TRT government forced Thaksin to announce snap elections that were boycotted by all opposition parties. The largest opposition party, the Democrats, played an important role in supporting the protests of the yellow shirts and in organizing the boycott (Sinpeng 2014, 159-60). However, this is exactly the month in which the Democrats drop out of the anti-government network in the dataset. At the same time, Thaksin's governing TRT erroneously becomes part of the anti-government network and is coded as engaging in cooperative events with the boycotting Mahachon Party, the Muslim insurgents, and even the yellow shirts.

[6] There were 14 cooperative events (1.1 percent) between groups on opposing sides of the color-coded conflict (cooperation between pro- and anti-Thaksin groups) and of the Southern conflict (cooperation between insurgent groups and the political parties of the dataset). Such cooperation could be caused by verb confusion, in which interactive words, such as negotiate or urge, were incorrectly coded as cooperative events. The Online Supplemental Material includes a detailed discussion of the evaluation.

Between March and May 2010, "the largest mass protest ever in contemporary Thai history" by the red shirts took place (Sinpeng 2014, 164). In April 2010, amidst the red-shirt siege of Bangkok to overthrow the Democrat-led government, ICEWS drops the red shirts from the anti-government network. The pro-Thaksin political party is correctly included in the anti-government network, albeit three times instead of once (as the defunct TRT and PPP, and the current PT). Additionally, it is coded as in-cooperation with its own predecessor, and with the anti-Thaksin yellow shirts, which mobilized against the red-shirt protest in the real world (Hookway 2010).

The ICEWS conflict count reveals similar problems with accuracy. Metternich et al. (2013, 907-8) divide the year 2010 evenly into three separate periods (January-April, May-August, September-December) for their out-of-sample prediction exercise, thereby splitting up the period of massive continuous red shirt anti-government protests in Bangkok that happened between March 14 and May 19, 2010. Metternich and colleagues ignore the event in their presentation, but the red-shirt protest resulted in the worst episode of violence in Thailand in over three decades with 85 deaths and about 2,000 injured (Dalpino 2011). Yet ICEWS does not adequately reflect this violent period: As shown in Figure 5, the conflict count during this time is relatively low and declines even further during the red-shirt siege, leading to the incorrect conclusion that their network model "does very well in the first prediction period" (Metternich et al. 2013, 907).

3.4 Hand-coding Cooperative Events from Media Reports to Detect False Negatives

While ICEWS was unable to adequately depict Thailand's cleavages, it is possible that the outcome was caused by the automated coding scheme rather than by systematic underreporting by journalists. Therefore, I compiled a hand-coded dataset of cooperative events between known insurgent groups, between yellow shirts and the Democrat Party and between red shirts and the pro-Thaksin parties. The search was conducted for periods when these groups were in opposition in order to measure anti-government activity using LexisNexis, which provides access to 31,480 sources. About 97.9 percent of the sources used to obtain the 139,784 ICEWS entries for Thailand over the period 2001-2010 are covered by LexisNexis. This includes all major international and regional sources that are covered by ICEWS, such as Agence France-Presse, Associated Press, BBC, Xinhua News Agency, Thai News Service, Bangkok Post, and The Nation.

My search yielded a total of 13,064 news stories on insurgents in southern Thailand over the period 2001-2010. As in Metternich et al. (2013, 901), most of the reports use a generic term for insurgents and rebels. The known insurgent groups BRN, PULO, or GMIP are mentioned in 934 reports; 25 (0.2 percent) of the reports discuss cooperation between insurgent groups. Nearly all alleged cooperation is based on third-hand sources that contradict each other in some cases. For example, one report conjectures that the three known groups coordinated in an instance of violent activity (Xinhua General News Service 2004), while another report claims that GMIP was responsible for the violence and does not cooperate with the other two groups (Mills 2004). Nine reports either claim that Rumpun Kampurua Kencil (RKK) is a subgroup, the leading organization, or a part of BRN-Coordinate, and six reports contain information on meetings and statements of insurgent groups – including a secret meeting that was held four months earlier (Bangkok Post 2006). The analysis confirms that media reports do not provide reliable insight into insurgent networks and that cooperative events are systematically underreported.

Media-based data generates problematic network depictions of the political groups of the color-coded conflict, as well. The red shirt alliance had a closer and more cooperative relationship with corresponding pro-Thaksin parties compared to the alliance between yellow shirts and

Democrats (Sinpeng 2014): The red shirts shared leadership and organizational structure with the PPP/PTP: nearly all red-shirt leaders were members of parliament for PPP/PTP and the red shirts were actively involved in the PPP/PTP election campaigns. The anti-Thaksin alliance also saw leadership overlap, but the organizations remained independent, enabling a fast break-up of this alliance once the Democrats were in power. Thus, the different group proximities between the color movements and their corresponding political parties allow for a straightforward test of the proposed media bias: quantitative network analysis should reveal relatively more cooperative events among the red alliance and place the red shirts and PPP/PTP closer together compared to the yellow shirts and Democrats. However, I predict that media reports will mention the alliance between yellow shirt and Democrats relatively more often because the stronger organizational overlap amongst the red alliance members would result in fewer changes and thus be less newsworthy.

In the analysis below, I distinguish between narrow and broad cooperation within the yellow and red alliances. The distinction allows for a better evaluation of whether underreporting occurs for machine-coded as well as for hand-coded datasets. Narrow cooperation exists if a report directly states an allegiance between two groups with their names in a sentence. Well-developed automated coding schemes should be able to capture narrow cooperative events. By contrast, broad cooperation is not easily detectable by automated coding schemes. Here, broad cooperation refers to cooperative events that are described in vague terms, expressed in goal or policy congruence between groups, or not directly stated in one sentence.[7]

[7] The verbs used for identifying narrow cooperation are abet, agrees, aid, ally, align, assist, associate, back, comprise of, endorse, facilitate, help, join, lead, link, mobilize, share, support, sympathize, ties with, work with, welcome. Examples are: "In the case of our contemporary politics, the conflict on the surface seems to revolve around two main

<<< FIGURES 3 AND 4 >>>

Figures 3 and 4 show the total number of reports and cooperative events for both alliances over time. For the yellow shirts, there are 9,627 reports that mention affiliated groups of which 174 (1.8 percent) depict events of broad cooperation and 35 (0.4 percent) depict events of narrow cooperation with the Democrats. The average per month figures are 534.8 total reports with 9.67 reports indicating an instance of broad cooperation, and 1.94 showing narrow cooperation. For the red shirts, the total reports mentioning the alliance sums to 8,813, of which 113 (1.3 percent) signal broad cooperation and 25 (0.3 percent) show narrow cooperation with PPP/PTP. This equates to a monthly average of 251.8 total reports, 3.2 broad cooperation events, and 0.7 narrow cooperation events.[8] The outcome suggests that only a small fraction of media reports touch on the alliances of the yellow or red shirts. Yet, these numbers of narrow and broad cooperation were higher than detected in Metternich et al. As discussed in section 3.1, ICEWS captures only two cooperative events between the red shirts and PPP when it was in government, and eleven cooperative events between the yellow shirts and the Democrats. Thus, the failure of ICEWS to detect cooperative

factions - the United Front of Democracy Against Dictatorship (UDD) and the People Power party-led government, versus the People's Alliance for Democracy (PAD) and the opposition Democrat party" (Thai News Service 2008); "Long-running antigovernment protests by anti-Thaksin activists of the People's Alliance for Democracy, backed by the Democrat Party, led to the closure of Bangkok's two main airports for nearly 10 days late last month" (Japan Economic Newswire 2008); "But the opposition Pheu Thai Party, which is closely associated with the reds, has decided to opt out of the constitution amendment process" (The Nation 2009).
Examples for broad cooperation are: "The former Bangkok governor reaffirmed that the PAD remained committed to securing Mr. Thaksin's resignation, and to petition for a royally-appointed interim prime minister as a replacement. He also condemned the disruption of an opposition Democrat Party rally by supporters of the prime minister in Mr Thaksin's northern hometown, Chiang Mai" (Thai News Service 2006); "The protesters and the Democratic Alliance Against Dictatorship plan a major rally on Thursday to oust the Abhisit government after the Pheu Thai Party missed toppling it during the censure debate" (Thai News Service 2009).

[8] Duplicates were excluded from the analysis. One cooperative event for Southern insurgents, six for the yellow shirts, and three for the red shirts are based on the following LexisNexis sources that were not covered by ICEWS: The Advertiser, IHS Global Insight, Malaysia General News, and Morning Star.

events appears to be at least partly the result of the automated coding scheme used by Metternich et al. (2013).[9]

While the hand-coded analysis captures more instances of cooperation than the automatic coding scheme, it records less cooperation overall among the red alliance compared to the yellow alliance. This would wrongly place yellow shirts and Democrats closer to each other relative to red shirts and PPP/PTP in a network analysis. Two aspects of the media bias appear to explain this outcome: First, as the yellow shirts and Democrat Party were two separate organizations, journalists reported events including both groups relatively more often in the same report. 1,245 of the 9,627 reports (12.9 percent) on the yellow shirts also included information on the Democrat Party, while only 365 of the 8,813 reports (4.1 percent) on the red shirts included PPP/PTP. As a consequence, there was a higher probability that reports discussed cooperative events among yellow shirts and Democrats compared to the red shirts and PPP/PTP.

Second, the data also suffers from the fact that the yellow shirts were relatively more often in the political spotlight for the period of observation. This increased the likelihood that reports extensively discussed their movement. The cooperative events between the red shirts and PPP/PTP, on the other hand, were unequally distributed over the period of observation. The frequency of reported cooperation skyrocketed when mass protests by red shirts challenged the government, even though the close organizational overlap within the red alliance did not change over time (Sinpeng 2014). The hand-coded dataset further illustrates the fundamental problem of news-based network analysis that any apparent association between network activity and political

[9] Metternich et al. (2013) used the coding programs JABARI and TABARI. The current version of ICEWS is based on the BBN ACCENT event coder (Boschee et al. 2015).

events could be caused by the tendency of journalists to discuss groups more often when they are at the center of political drama.

3.5 Using Alternative Conflict Measurements

This section analyzes whether Metternich et al.'s (2013) empirical results can be replicated with two alternative measurements of the dependent variable; the counts of conflictual events. I use alternative data sources for each of the two discussed conflicts in Thailand: The Deep South Watch dataset (DSW) for the Southern conflict (Srisompob 2011), and a dummy variable for the four periods of political unrest that comprise the color-coded conflict (CCC). The DSW is based on local Thai-language newspapers and police radio scanners, comprising a hand-coded monthly dataset of conflict incidents going back to the inception of the insurgency in January 2004. The DSW is the most comprehensive dataset for violence in Thailand's Deep Southern provinces.

As described by McCargo (2010, 8), the color-coded conflict "is a much more conventional political stand-off where electoral contestation and relatively peaceful protests have generally predominated." Thus, the CCC dummy measures the four periods of extraordinary color-coded conflict, which are: 1) February to September 2006, when anti-Thaksin protests by the yellow shirts skyrocketed in February and culminated in the military coup of September 19, 2006 (Ockey 2007); 2) August to December 2008, when the yellow shirts started their assault on the pro-Thaksin government on August 26 by seizing government buildings and airports. This period ended when the government was overthrown on December 15, 2008 (Kitti 2009); 3) April 2009, when red-shirt protesters stormed an ASEAN Summit in Pattaya, followed by violent demonstrations in Bangkok and an assassination attempt of PAD leader Sondhi (Kitti 2010); 4) March and May 2010, which comprised the second red-shirt uprising in Bangkok (Dalpino 2011).

Besides replacing the dependent variable with the alternative conflict measurements, the empirical analysis also evaluates whether there was an association between both conflicts by regressing the color-coded conflict on the conflict count for insurgent violence and vice versa. This procedure allows for a straightforward test of Metternich et al.'s (2013) hypothesis that the structure of the anti-government network matters for conflict levels in Thailand, as these were the only groups capable of substantial anti-government action in real-world Thailand. Metternich et al. (2013) claim that network fragmentation should be associated with more incidents because actors do not have a tendency to free ride on the anti-government efforts of groups that are far apart in social space. The argument is based on the assumption that "efforts of one actor to challenge the government should decrease the state's ability to fight other antigovernment groups" (Metternich et al. 2013, 896). The required depletion of state resources suggests that only extraordinary anti-government actions can affect the propensity of ideologically dissimilar anti-government groups to strategically engage in conflicts. In the Thai context, the argument implies an interaction between the conflicts in the Deep South and the color-coded conflict; i.e., a strategic response by red or yellow shirts to violence in the Deep South and vice versa. In fact, Srisompob and McCargo (2010, 159) suggest that there might have been a strategic response by Southern insurgents, at least in the aftermath of the military coup in 2006.

<<< FIGURE 5 >>>

Figure 5 compares Metternich et al.'s ICEWS dataset with the DSW dataset and with the four periods of extraordinary yellow and red mobilization in the color-coded conflict. ICEWS captures only a small fraction of conflict incidents uncovered by the DSW. The correlation coefficient between ICEWS and DSW is 0.51, suggesting that ICEWS tends to track the DSW, but the coverage remains modest. Figure 5 confirms Srisompob and McCargo's (2010) hypothesis that Southern conflicts declined substantially after the army introduced new security measures in July 2007. Table 2 replicates Metternich et al.'s (2013) main model. The ICEWS conflict variable is replaced as the dependent variable by the DSW conflict measurement and, subsequently, by the CCC dummy variable. In additional model specifications, the CCC dummy also functions as additional explanatory variables ("Other Political Conflict") to explain DSW conflict counts, and vice versa, in order to evaluate whether there was an association between both conflicts.

<<< TABLE 2 >>>

The first model in Table 2 is the original model by Metternich et al. (2013). They show that small eigenvalues of the anti-government network are significantly associated with more conflicts between anti-government groups and the government. When replacing the dependent variable with the DSW measurement in Model 2, Metternich et al.'s network measurement becomes insignificant. Model 3 adds the coefficient of the color-coded conflict variable, which appears to be negatively significant, suggesting that large-scale conflicts between pro- and anti-Thaksin groups are associated with a reduction of insurgent violence in the Deep South. The significant positive association of regime type, measured by Polity-4 scores, and conflicts in models 2 and 3 appears to capture the fact that the junta imposed tougher security policies in the Deep South after toppling the democratically-elected government. Periods of color-coded conflict do not appear to be statistically influenced by Metternich et al.'s network measurement or the intensity of the Southern conflict as shown by models 4 and 5.

In summary, the main explanatory variable – the lowest eigenvalue – does not turn out to be significant with alternative conflict measurements. In fact, periods of extraordinary conflict between pro- and anti-Thaksin groups are negatively associated with violence in the Deep South, contradicting Metternich et al.'s (2013) argument that groups have a stronger tendency to fight the government on their own when ideologically dissimilar groups challenge the government. In addition, other studies on the color-coded conflict show that red and yellow shirts invested in their organizational capacity to integrate heterogeneous groups into their movements. Movement unity rather than network fragmentation was key to effectively challenge the government in the color-coded conflict (Naruemon and McCargo 2011; Sinpeng and Kuhonta 2012; Sinpeng 2014).

4. Conclusion and Recommendation

While Schrodt (2012, 556) maintains that "event data is perfectly suited for network analysis," my results suggest that missing observations in ICEWS make it, at best, an incomplete source for studying the networks of insurgents and political parties.[10] Furthermore, the high rate of incorrect data entries indicates that the inclusion of irrelevant stories due to name confusion remains a major problem for event data. While more reliable machine-coding programs might reduce the amount of incorrect entries in the future, the analysis of ICEWS and a hand-coded alternative suggests that event data tend to draw inaccurate network pictures due to systematic biases of media reports. Insurgent networks are too secretive to be accurately depicted, while

[10] O'Brien (2010) claims that ICEWS is useful for predicting conflicts. However, predicting conflicts by measuring how often journalists report on mass protests, tensions, threats, fights, or killings provides only limited scientific insight, because the explanatory variable is only a weaker variation of the conflict variable that one wants to predict (Jäger 2016).

alliances among regular political parties are often too well-known and constant over time to be considered newsworthy. Using alternative conflict measurements, the empirical analysis above draws the findings of Metternich el al. (2013) into question. Their network measurement was not associated with more insurgent violence, and my analysis shows that periods of extraordinary conflict between pro- and anti-Thaksin groups were negatively associated with violence in the Deep South.

Automated content analysis methods show promise for advancing research, but they "will not […] eliminate the need for careful thought by researchers nor remove the necessity of reading texts" (Grimmer and Stewart 2013, 270). Researchers using automated- or hand-coded event data should be aware that an association of network activity with a dependent variable of interest, such as violent incidents, could simply reflect the tendency of journalists to report more on groups when these groups make headlines by causing conflicts. Connections between groups could be constant over time, but reporters are likely to discuss these connections more often when the involved groups are in the spotlight.

A possible solution to the risk of spurious correlation is to conduct robustness tests by changing the weight of networks ties. For instance, researchers might discount reports published during extraordinary periods, or average the network variable over a longer period. As connections between groups with strong organizational overlap are likely to be underreported, it might be useful to increase the leverage of these ties, or to treat such groups as one organizational unit. In any event, manual modifications of the network data should be justified with case-specific considerations, highlighting the need for empirical researchers to "engage in *Verstehen* before we engage in reductionism" (Friedman 2012, 438). A qualitative understanding of a country, its major political players, and key actors' intentions is required if we are to estimate the soundness of large-scale datasets.

References

Althaus, Scott L., Nathaniel Swigger, Svitlana Chernykh, David J. Hendry, Sergio C. Wals, and Christopher Tiwald. 2011. "Assumed transmission in political science: A call for bringing description back in." *Journal of Politics* 73 (4): 1065-80.

Askew, Marc. 2010. "Fighting with Ghosts: Querying Thailand's "Southern Fire." *Contemporary Southeast Asia: A Journal of International and Strategic Affairs* 32 (2): 117-55.

Balcells, Laia, and Patricia Justino. 2014. "Bridging Micro and Macro Approaches on Civil Wars and Political Violence Issues, Challenges, and the Way Forward." *Journal of Conflict Resolution* 58 (8): 1343-59.

Bangkok Post. 2006. "South militants reportedly planning to intensify attacks September to December." *Bangkok Post*, August 27.

Baum, Matthew A., and Tim Groeling. 2010. *War stories: The Causes and Consequences of Citizen Views of War.* Princeton: Princeton University Press.

Baum, Matthew A., and Yuri M. Zhukov. 2015. "Filtering Revolution: Reporting bias in international newspaper coverage of the Libyan civil war." *Journal of Peace Research* 52 (3): 384-400.

Boschee, Elizabeth, Jennifer Lautenschlager, Sean O'Brien, Steve Shellman, James Starz, and Michael Ward. 2015. "ICEWS Coded Event Data." Harvard Dataverse, V5.

Cederman, Lars-Erik, and Kristian Skrede Gleditsch. 2009. “Introduction to Special Issue on Disaggregating Civil War.” *Journal of Conflict Resolution* 53 (4): 487–95.

Croissant, Aurel. 2007. "Muslim Insurgency, Political Violence, and Democracy in Thailand." *Terrorism and Political Violence* 19 (1): 1-18.

Croissant, Aurel, and Philip Völkel. 2012. "Party System Types and Party System Institutionalization Comparing New Democracies in East and Southeast Asia." *Party Politics* 18 (2): 235-65.

Dalpino, Catharin. 2011. "Thailand in 2010." *Asian Survey* 51 (1): 155-62.

Earl, Jennifer, Andrew Martin, John D. McCarthy, and Sarah A. Soule. 2004. "The Use of Newspaper Data in the Study of Collective Action." *Annual Review of Sociology* 30: 65-80.

Friedman, Jeffrey. 2012. "Beyond Cues and Political Elites: The Forgotten Zaller." *Critical Review* 24 (4): 417-61.

Gleditsch, Kristian Skrede, Nils W. Metternich, and Andrea Ruggeri. 2014. "Data and Progress in Peace and Conflict Research." *Journal of Peace Research* 51 (2): 301-14.

Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21 (3): 267-97.

Groeling, Tim, and Matthew A. Baum. 2008. "Crossing the Water's Edge: Elite Rhetoric, Media Coverage, and the Rally-round-the-flag Phenomenon." *Journal of Politics* 70 (4): 1065-85.

Hammond, Jesse, and Nils B. Weidmann. 2014. "Using Machine-coded Event Data for the Micro-level Study of Political Violence." *Research & Politics* 1 (2): 1-8.

Hookway, James. 2010. "Deadly Grenade Blasts Shake Bangkok." *The Wall Street Journal*, April 23.

Hoff, Peter D., Adrian E. Raftery, and Mark S. Handcock. 2002. "Latent space approaches to social network analysis." *Journal of the American Statistical Association* 97 (460): 1090-98.

Jäger, Kai. 2012. "Why did Thailand's Middle Class Turn against a Democratically Elected Fovernment? The Information-gap Hypothesis." *Democratization* 19 (6): 1138-65.

Jäger, Kai. 2016. "Not a New Gold Standard: Even Big Data Cannot Predict the Future." *Critical Review* 28 (3-4): 335-55.

Jäger, Kai. 2017. "The Potential of Online Sampling for Studying Political Activists around the World and across Time." *Political Analysis* 25 (3): 329-43.

Japan Economic Newswire. 2008. "Thai opposition party leader Abhisit becomes new premier." *Japan Economic Newswire*, December 15.

Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War*. Cambridge: Cambridge University Press.

Kalyvas, Stathis N. 2012. "Micro-level Studies of Violence in Civil War: Refining and Extending the Control-collaboration Model." *Terrorism and Political Violence* 24 (4): 658-68.

Kitti Prasirtsuk. 2009. "Thailand in 2008: Crises Continued." *Asian Survey* 49 (1): 174-84.

Kitti Prasirtsuk. 2010. "Thailand in 2009: Colored by Turbulence." *Asian Survey* 50 (1): 203-10.

McCargo, Duncan, 2000. *Politics and the Press in Thailand*. London: Routledge.

McCargo, Duncan. 2005. "Network Monarchy and Legitimacy Crises in Thailand." *Pacific Review* 18 (4): 499-519.

McCargo, Duncan. 2008. *Tearing apart the Land: Islam and legitimacy in Southern Thailand*. Ithaca: Cornell University Press.

McCargo, Duncan. 2010. "Thailand's Twin Fires." *Survival* 52 (4): 5-12.

Metternich, Nils W., Cassy Dorff, Max Gallop, Simon Weschle, and Michael D. Ward. 2013. "Antigovernment Networks in Civil Conflicts: How Network Structures Affect Conflictual Behavior." *American Journal of Political Science* 57 (4): 892-911.

Mills, Elizabeth. 2004. "Talks Between Thai Government, Southern Separatists Set for July." *IHS Global Insight*, May 24.

Naruemon, Thabchumpon, and Duncan McCargo. 2011. "Urbanized Villagers in the 2010 Thai Redshirt Protests." *Asian Survey* 51 (6): 993-1018.

Narut Wasinpiyamongkhon. 2013. *People's politics in Thailand: a critical study of the Assembly of the Poor, 2001-2010*. Leeds: University of Leeds.

O'Brien, Sean P. 2010. "Crisis Early Warning and Decision Support: Contemporary Approaches and Thoughts on Future Research." *International Studies Review* 12 (1): 87-104.

Ockey, James. 2007. "Thailand in 2006: Retreat to Military Rule." *Asian Survey* 47 (1): 133-40.

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

Sinpeng, Aim. 2014. "Party-Social Movement Coalition in Thailand's Political Conflict (2005–2011)." In *Contemporary Socio-Cultural and Political Perspectives in Thailand*, edited by Pranee Liamputtong, 157-68. New York: Springer.

Sinpeng, Aim, and Erik Martinez Kuhonta. 2012. "From the Street to the Ballot Box: The July 2011 Elections and the Rise of Social Movements in Thailand." *Contemporary Southeast Asia* 34 (3): 389-415.

Srisompob Jitpiromsri. 2011. *The Protracted Violence amidst the Unstable Political Situation after 2011 Elections*. Pattani: Deep South Watch Center for Conflict Studies and Cultural Diversity, Prince of Songkla University.

Srisompob Jitpiromsri, and Duncan McCargo. 2010. "The Southern Thai Conflict Six Years On: Insurgency, Not Just Crime." *Contemporary Southeast Asia* 32 (2): 156-83.

Thai News Service. 2006. "Thailand: Anti-Thaksin demonstrations set to continue." *Thai News Service*, April 6.

Thai News Service. 2008. "Thailand: Commentary - Third party can help resolve political conflict." *Thai News Service*, November 13.

Thai News Service. 2009. "Thailand: Pheu Thai Party to vote in new executive board." *Thai News Service*, March 23.

The Nation. 2009. "Today's red rally puts Mr Flip-Flop in a quandary." *The Nation*, December 10.

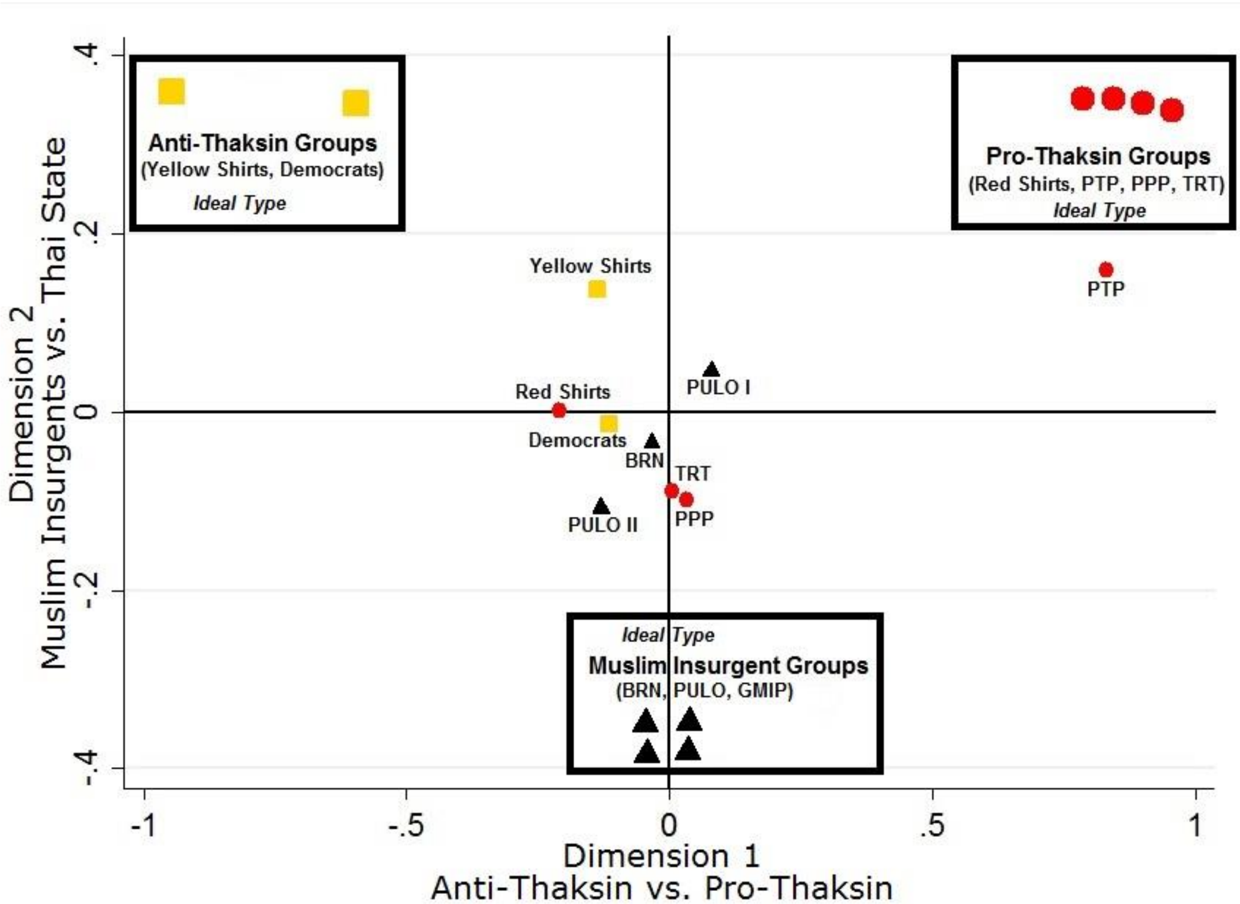
Thompson, Allan (ed.). 2007. *The Media and the Rwanda Genocide*. London: Pluto Press.

Weidmann, Nils B. 2015. "On the Accuracy of Media-based Conflict Event Data." *Journal of Conflict Resolution* 59 (6): 1129-49.

Weidmann, Nils B. 2016. "A Closer Look at Reporting Bias in Conflict Event Data." *American Journal of Political Science* 60 (1): 206-18.

Xinhua General News Service. 2004. "Thai Muslim separatists raid arsenal in south: Army chief." *Xinhua General News Service*, January 4.

Figure 1. The Ideal-type Cleavages of the Political Conflicts in Thailand Compared to the Average ICEWS Placement Based on the Defined Minimum Necessary Benchmarks of Chapter 3.1



Note: The labeling of the dimensions is exemplary. There are potentially eight different ways to depict the two dimensions.

Figure 2. Incorrect Entries According to the Evaluation Criteria

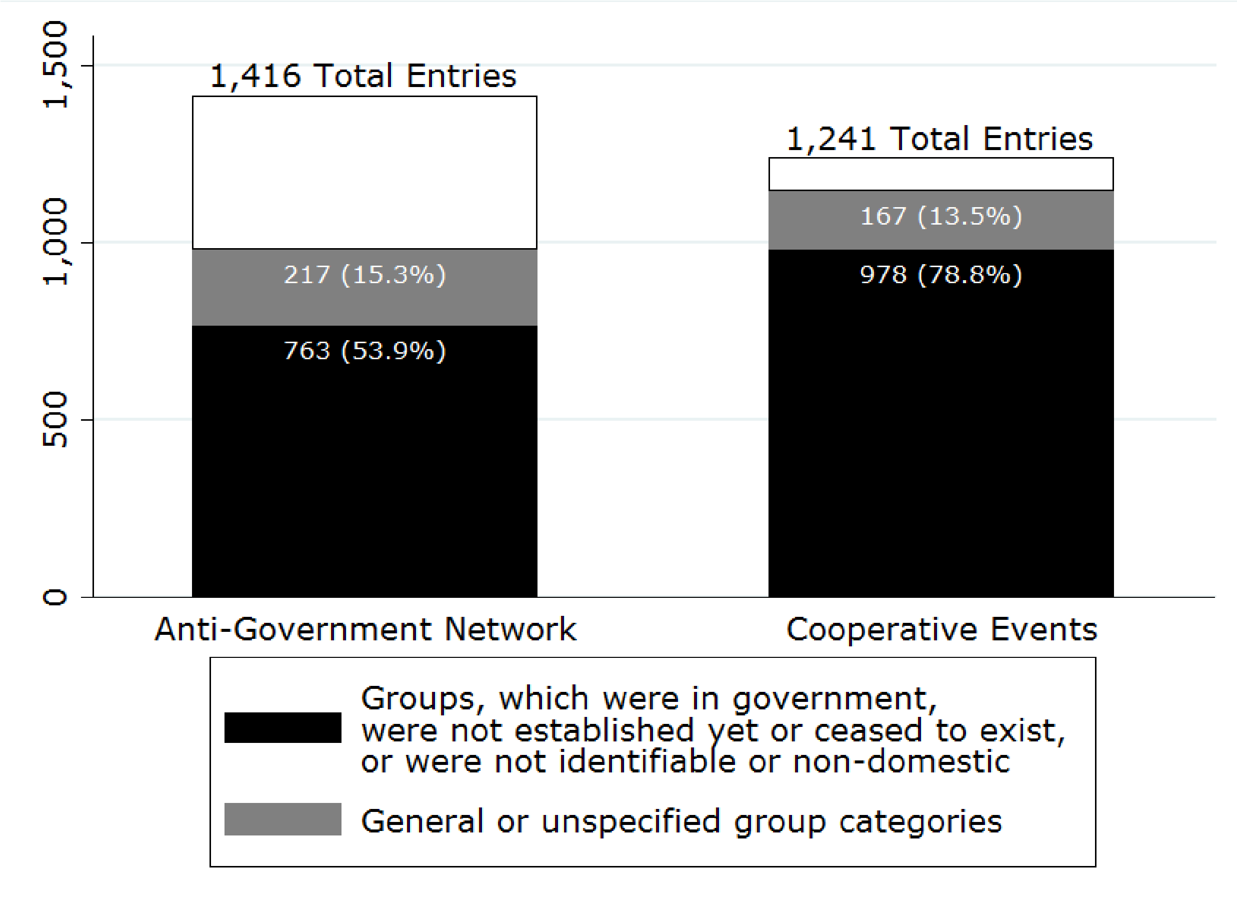


Figure 3. Reports and Cooperative Events for Yellow Shirts

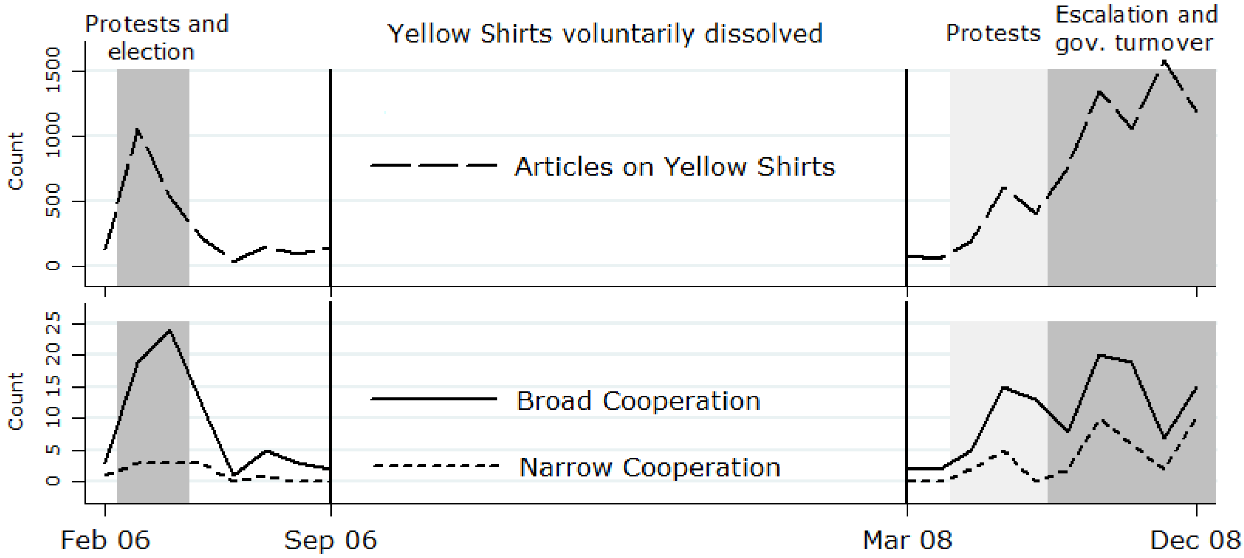


Figure 4. Reports and Cooperative Events for Red Shirts

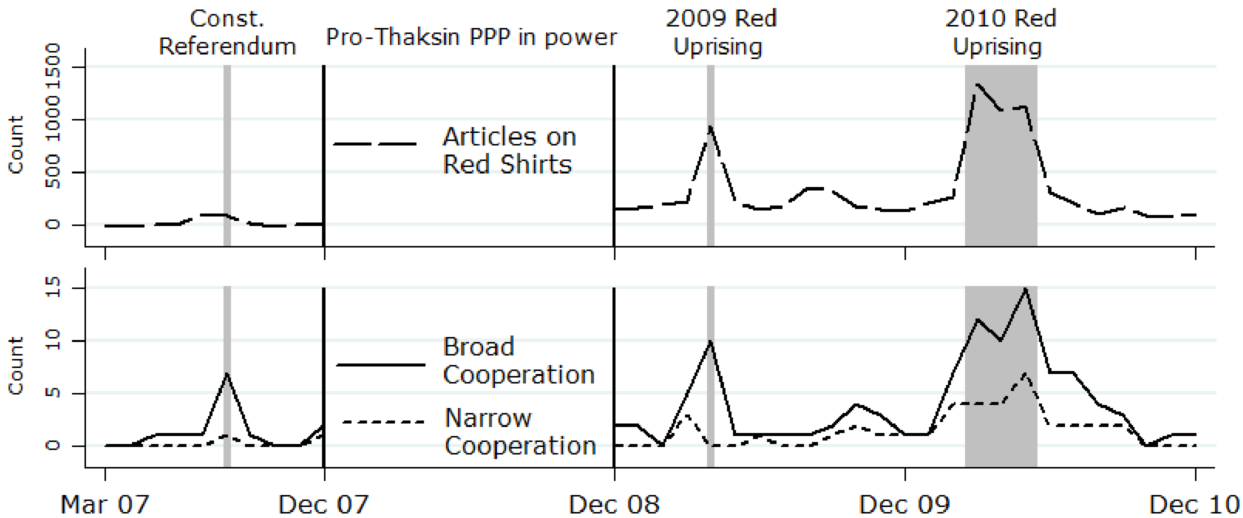


Figure 5. Comparing Metternich et al.'s (2013) ICEWS Dataset with the DSW Dataset and Periods of Extraordinary Color-coded Conflicts

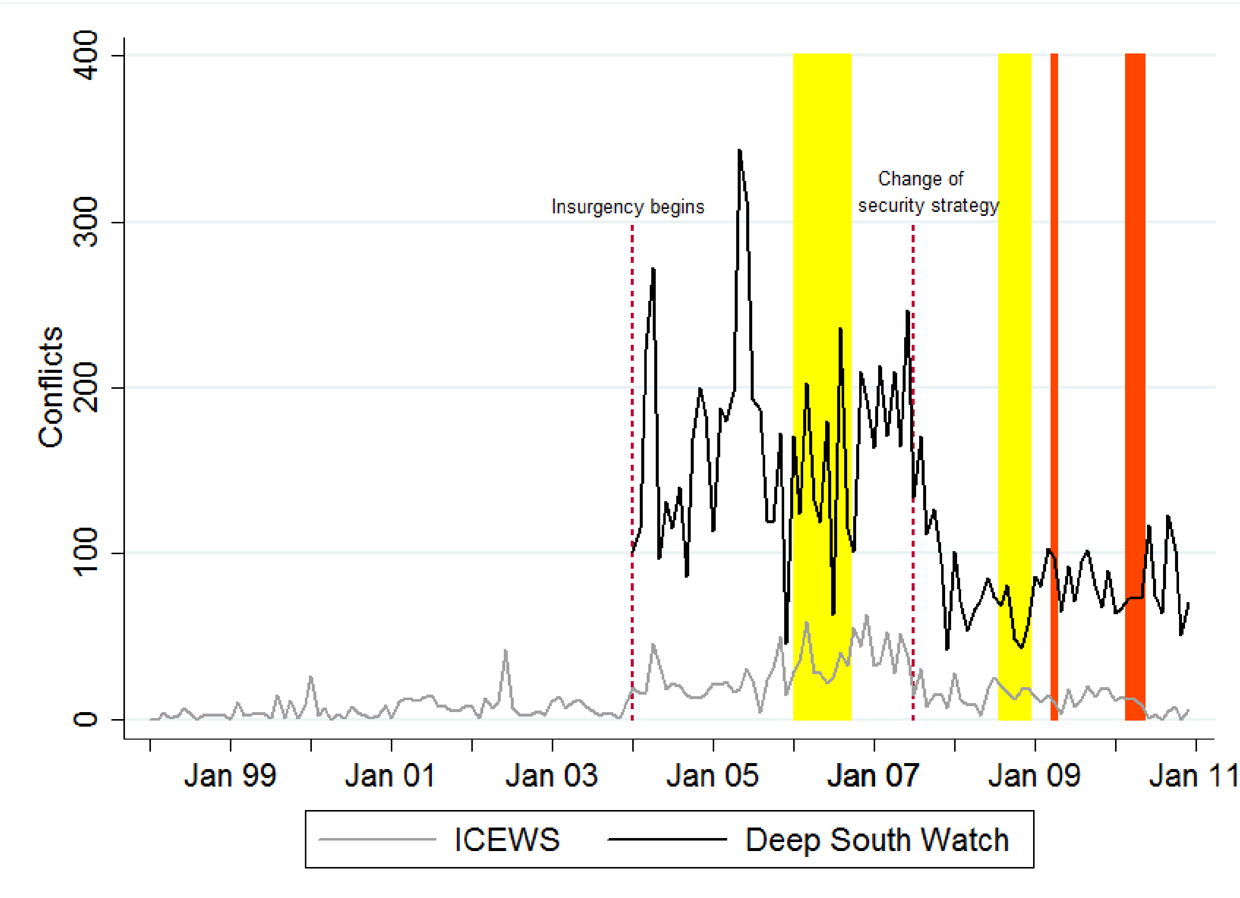


Table 1. Summary of the Expected Problems of News-based Event Datasets for Studying Networks

Issue	Predicted Problem	Type of Problem	Cost of Solution
Insurgent networks are too secretive to be revealed by media reports.	Interaction between insurgents are underreported: Allied insurgent groups are not placed relatively closer to each other in the network. (**Media description bias**)	False negative	High
Political relations are too constant over time to be considered newsworthy.	Stable interaction between political groups are underreported if they are not in the spotlight: Allies in the party system are not placed relatively closer to each other in the network. (**Media selection bias**)	False negative	High
Irrelevant stories / name confusion	Groups appear too often in the dataset. (**Machine-coding bias**)	False positive	Low

Table 2. Replicating Metternich et al.'s (2013) Results (DV: ICEWS) with the Two Alternative Dependent Variables DSW and CCC.

	Observed network model				
	1	**2**	**3**	**4**	**5**
Dependent Variable:	**ICEWS**	**DSW**	**DSW**	**CCC**	**CCC**
Lowest eigenvalue	-0.21***	-0.04	-0.05	-0.08	-0.11
(Network variable)	(0.07)	(0.05)	(0.05)	(0.63)	(0.63)
Number of attacking members	0.05	-0.03	0.03	1.08**	0.95*
	(0.07)	(0.05)	(0.05)	(0.54)	(0.55)
Number of attacked members	0.05	-0.03	-0.00	0.68	0.76
	(0.07)	(0.05)	(0.05)	(0.52)	(0.53)
Size of network	0.12	-0.04	-0.05	1.65*	1.66
	(0.10)	(0.10)	(0.09)	(0.98)	(1.03)
Growth	0.17**	0.07	0.02	-0.37	-0.34
	(0.07)	(0.09)	(0.09)	(0.88)	(0.98)
Proximity to elections	-0.14*	-0.13*	-0.15**	-0.85	-1.13
	(0.07)	(0.07)	(0.06)	(0.71)	(0.77)
Polity	0.19**	0.30***	0.29***	-0.49	-0.24
	(0.08)	(0.06)	(0.06)	(0.61)	(0.70)
Rebellion in profile similar countries	-0.18**	-0.01	-0.02	-0.85	-0.57
	(0.08)	(0.08)	(0.07)	(0.86)	(0.90)
Other Political Conflict			-0.11***		-0.71
			(0.04)		(0.55)
AIC	746.99	756.42	751.78	64.91	64.94

Note: * p ≤ 0.10; ** p ≤ 0.05; *** p ≤ 0.01. Standard errors are in parentheses. Model specifications are based on Figure 9 of Metternich et al. (2013: 906). The models with CCC as dependent variable are binomial logit regressions.