Appendix for Subject to Change: Quantifying Transformation in Armed Conflict Actors At Scale Using Text

June 7, 2024

Contents

1 Positioning Relative to Existing Approaches

Previous attempts to model transformation within opaque militant groups have generally followed one of two strategies. Scholars working qualitatively have developed rich scholarship on the dynamics and trajectories of specific organizations. This approach is invaluable for "why" questions and for insight into dynamics of specific movements and their larger political and cultural contexts (Giustozzi 2019; Morrison 2013; Mosinger 2019; Sinno 2011; Woldemariam 2018; Wood 2008). However, this approach relies on detailed case knowledge, limiting the universe of cases that any individual scholar can draw from. As well, qualitative case scholarship requires access, which may not be possible.

Scholars working in a large-N setting have often looked for indirect and proxy strategies to measure group evolution and transitions. Previous work has operationalized change points via the emergence of new militant groups within a conflict setting (Lounsbery and Cook 2011) or transitions one form of mobilization into another, including between violence and non-violence (Dudouet 2013), violence and political engagement (Acosta 2014; Mampilly 2012), and insurgency and terrorism (Cronin 2009; Findley and Young 2012; Fortna, Lotito, and Rubin 2018).

The operationalization captures changes in third-party representation of rebel group operations. It differs from existing data on organizational attributes such as the Foundations of Rebel Group Emergence (FORGE) dataset, which focuses on lineage and organizational endowments (Braithwaite and Cunningham 2020), or REVMOD, which combs data sources to code group organizational traits (Acosta 2019). These and similar data sources such as Nonviolent and Violent Campaigns and Outcomes (NAVCO), All Minorities at Risk (AMAR), and Big Allied and Dangerous (BAAD) all build data from expert assessments of group attributes and structure (Asal and Rethemeyer 2008; Birnir et al. 2018; Chenoweth and Lewis 2013). This strategy produces data that is more specifically about organizations, but requires that information be accessible to research team(s). Moreover, the level of dynamism varies for each dataset. I model media coverage of activities and only indirectly capture organizational dynamics, effectively exchanging directness for coverage. Both data sources could be easily combined in future work, which would produce both depth from the existing organizational data and breadth from the media framing measure. A second extension could incorporate government behavior, which this analysis has largely elided.

2 Organized Armed Actor Identification and Selection

To build a list of conflicts and actors, I used the UCDP Georeferenced Event Database and the UCDP dataset of armed conflict actors. I cross-referenced the dataset with armed non-state actors featured in the UCPD/PRIO Armed Conflict Dataset (Pettersson et al. 2021). I selected the UCDP GED dataset as it has both the finest granularity of data and the underlying source articles that I use for classification. However, the GED's aperture is too wide for the research question, which is how to model the evolution of organized armed conflict actors. In order to narrow the scope, I derived the list of 244 extrasystemic, intrastate, and internationalized intrastate conflicts from the Armed Conflict Data. To identify unique non-state actors in these conflicts, I used the UCDP PRIO Actor Dataset (Pettersson et al. 2021). The Actor Dataset provides information about the 1661 unique actors of which 1016 are distinct non-state armed groups; I subset that to all non-state actors that are also in the list of 244 conflicts identified from the ACD.

Using an existing database means that my results are sensitive to a number of sources of potential error: measurement and specification errors of my approach but also errors arising from the general process of producing data from news reporting on conflicts, selection effects, and any errors specific to the UCDP (Demarest and Langer 2019; Eck 2012). At the same time, these sources of error are counter-balanced by several advantages associated with the UCDP GED: First, the UCDP has an extensive selection and validation process to minimize gaps and errors in the data collection process (Högbladh 2021; Sundberg and Melander 2013; Pettersson et al. 2021). Second, the UCDP publishes an encyclopedia of conflicts and organized conflict actors. Their Conflict Encyclopedia provides a point of entry to check the validity of the computational approach: periods of transformation and volatility implied by the computational analysis should also be reflected in researcher-generated profiles.[1]

3 Modeling

In general, text methods can be classified as *supervised* or *unsupervised*, depending on whether the researcher provides a coding scheme before the analysis (Benoit 2020; Wilkerson and Casas 2017). An unsupervised model is appropriate for the goal of this project because it can discover actor-specific patterns of representation at scale. A supervised model would require the researcher to identify dictionary terms for each of the candidate actors. This type of approach could model specific theoretical changes, such as threat perception over time or shifts between insurgent and terror tactics. However, using a supervised approach to identify group-specific theoretical shifts requires specifying the dimensions to find for each candidate group or restricting the potential dimensions to only those that fit a

[1] Ideally, the narrative profiles in the Encyclopedia would always echo periods of change in the computational model. Indeed, the computational analysis closely matches references to instability and transition in the corresponding UCDP encyclopedia profile for several of the conflict actors in the validation set. However, it is important to note that the encyclopedia profiles were created as an overview of each group and do not specifically focus on changes. The difference in goals is also the reason why I focus the analysis on modeling the underlying source data rather than on modeling descriptions of transitions in the UCDP Encyclopedia or similar projects, such as the Mapping Militant Groups project. As with most cases of validation of text analysis, the gold standard remains validation via the judgement of subject matter experts (*e.g.*, Atteveldt, Velden, and Boukes 2021).

general trajectory.

The GED's source articles range in length from 1 to 998 words (from 1-6,228 characters), with a median of 14 words. Texts in the UCDP GED tend to be well under 50 words and there does not appear to be a systematic difference in article lengths across each quartile. The 10-word threshold impacted a handful of groups for whom, for whatever reason, the Georeferenced Event Database source "article" constitutes metadata. These entries are often single-word records sourced from proprietary databases such as NigeriaWatch and the South Asian Terrorism Portal (SATP).

On the other extreme, the longest articles are often aggregations of event notifications and often record jihadi conflicts. The Syrian Civil War and the activities of Al Shabaab and the Islamic State are particularly over-represented in the list of articles over 250 words, with, respectively, 21, 9, and 7 of the 44 articles. Likewise, of the six articles over 500 words, three are associated with "Syrian insurgents," two are associated with the Islamic State, and one reports on the activities of the Caucasus Emirate.

The most active actors —Al Shabaab, the LTTE, the PKK, Kashmir insurgents, the Islamic State, the Taliban, and Syrian insurgents— have tens of thousands of violent events in the data. Indeed, the top three— the Islamic State (15,011 events), the Taliban (31,590 events), and "Syrian Insurgents" (60,814 events)— make up over 107,000 of the dataset's 191,252 records of violence. Fortunately, as I model candidate nonstate actors separately, the imbalance in the data does not infect the results on the whole.

Figure 1 shows the distribution of article lengths for each of the 251 "armed actor" (aka rebel) groups over the 10-event threshold. The Figure presents actors according to quartiles of recorded activity: the first quartile contains groups associated with $[10, 25]$ violent event records, the second quartile comprises groups with $(25, 75]$ violent events, the third quartile covers groups with $(75, 219]$ events, and the fourth quartile contains groups with more than 219 violent events.

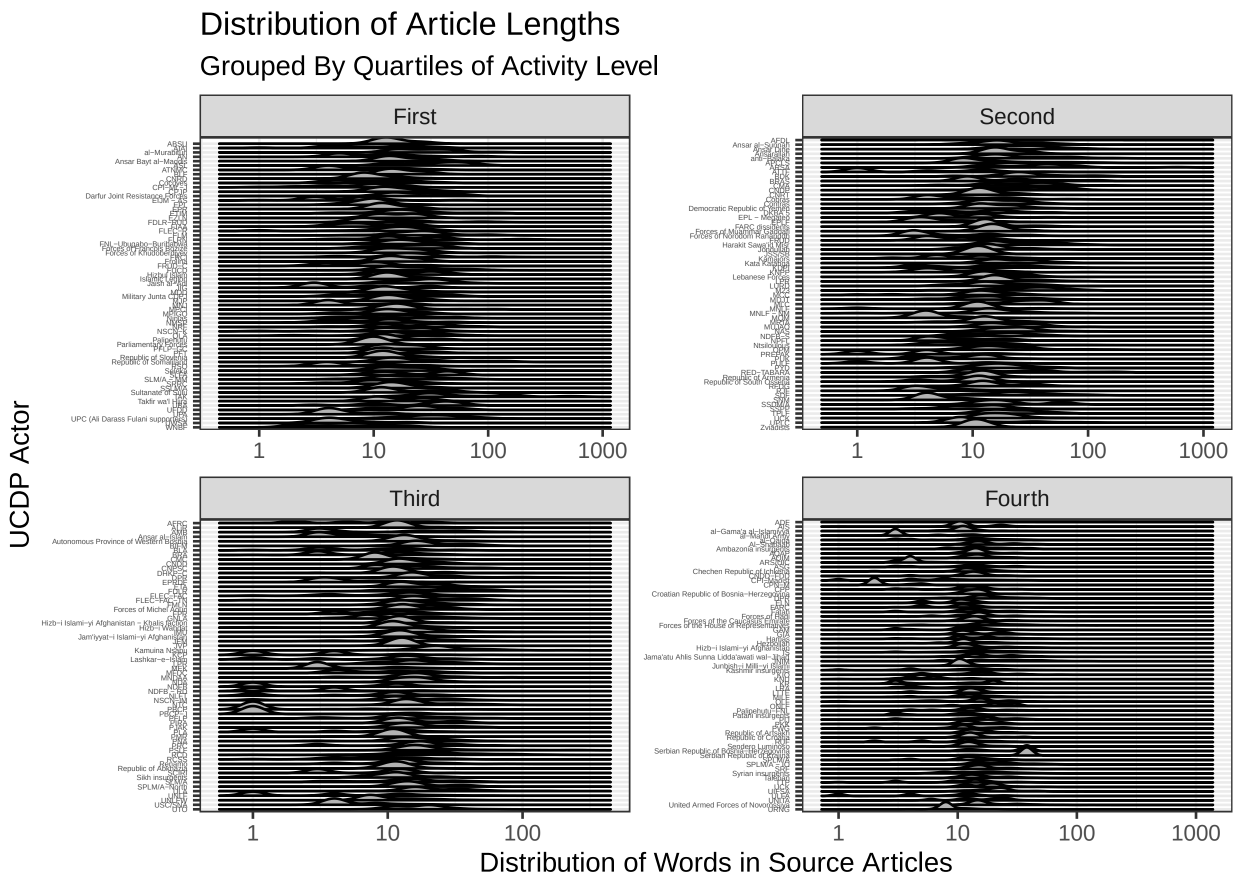

Figure 1: Distribution of number of words in articles associated with specific armed groups. For ease of interpretability, the plot groups UCDP Actors into quartiles of activity level.

3.1 Preprocessing

As with all applications of text-as-data methods, preprocessing steps fundamentally shape the outcome. In this case, significant choices include which words and symbols to drop from the underlying data; whether or not to stem the data; and what, if any, word frequency thresholds to use. The results reported here were generated by using preprocessing native to version 3 of the quanteda package (Benoit et al. 2018). Using quanteda's built-in stemming and symbol, url, and punctuation removal tools produced the most stable results. Numbers were removed using regular expressions.

The first decision point, and one of the most subtle, is which words and symbols to drop. We want to remove words that reflect the production of the news document rather than the featured event. After that, the researcher must decide whether to remove symbols and other potentially-spurious content. Removing non-text characters and single-letter words reduces the risk of inadvertently modeling substantively irrelevant features of the texts, such as metadata and stray markup. However, doing so risks distorting the corpus by pushing many articles under the ten document and ten word modeling threshold. I chose to error on the side of minimal processing.

I processed each actor's news corpus to remove words that are more likely to be part of the reporting process than of the events being reported. This list, which is available below, focused on words that are linked to international and regional news organizations, as well as dates and cardinal directions. Removing these words increases the likelihood that the topic model will pick up on changes in the presentation and description of group activities and minimizes the risk that the clustering will, instead, sort along an orthogonal dimension such as changes in information collection and reporting processes.[2]

The second preprocessing choice is whether to stem words in the documents. As with list-wise word removal, the choice to stem influences the outcome of the modeling (Denny and Spirling 2018). I chose to stem the texts because not stemming produced widespread duplication in both FREX and high-probability word summaries of each model (e.g., kill/kills/killed, Peru/Peruvian, soldier/soldiers). This made post-analysis interpretation more difficult by obscuring substantive content in the group-specific topics.

Ultimately, these specific choices are only one path through a complicated menu of potential preprocessing decisions. One of the most difficult elements of text analysis work is to decide on a specific preprocessing and analytical specification. I opted for a pragmatic approach of evaluating competing preprocessing approaches based on whether or they successfully captured known trends in a spotlight case. The specific criteria was whether the resulting models "found" the tactical oscillation in Abu Sayyaf.

[2]The list of "reporting" words includes directions and months because they often represent a dispatch signaling their Bureau location or situating the location of a report.

4 Proof of Concept Via Case Illustrations

A skeptical reader may wonder whether we can expect the analysis of short news texts to meaningfully capture changes in the operation of armed conflict actors. In this section, I present a proof of concept in the results for four cases chosen to highlight a range of challenges for the measurement strategy.

The first validity case, al-Qaeda in the Arabian Peninsula (AQAP), is a transnational insurgency that appears to have become localized. The second highlighted case is Abu Sayyaf (ASG) in the Philippines, selected because the group is known to have operated in phases dominated by ideological or profit-oriented behaviors. A third case, the United Liberation Front of Asom (ULFA) focuses on the substantive results and estimated change period for a group whose UCDP Encyclopedia summary indicates tactical and strategic shifts, but does not provide a specific benchmark year. In this case, the measure captured ULFA's tactical evolution and suggests three specific years of transition. The fourth case, spotlighting the Lord's Resistance Army (LRA), analyzes the measurement strategy in the case of a conflict actor with a relatively stable tactical and strategic profile.

A summary of the group-level representation and estimated change periods for the organizations follows. Each case is accompanied by a plot in which years of operation (within UCDP records) are presented along the X-axis with article-level topic proportion on the Y-axis. Each point is a single GED "source article" entry documenting group behavior. Articles (points) positioned closer to -1 are more associated with that group's Topic One while articles positioned closer to 1 are more associated with the group's Topic Two. To help interpret the scales, each plot highlights selected text from representative articles. Finally, the dotted line summarizes trends in the yearly proportion of articles associated with each topic. When all articles are assigned to the group's Topic One, the summary line registers -1 for that year, conversely when all articles for that year are assigned to the group's Topic Two the line registers 1 for the year. By extension, whenever the line crosses 0, the dominant representation has changed between one year and the next.

As the first illustrative example, consider the case of al-Qaeda in the Arabian Peninsula (AQAP). The Yemeni-based jihadi organization operates in a complex security environment: since AQAP emerged as a local offshoot of the transnational jihadi movement lead by al-Qaeda, the country has undergone a political revolution that removed Ali Abdullah Saleh, the country's leader for over two decades, and entered into a complex and internationalized civil war. This civil war pitted AQAP against the Iranian-backed Houthi militia movement, and has involved not only religious tension but also within-Yemen geographic grievances and regional power rivalries. Acting in conjunction with local tribal interests, AQAP has gained and lost significant territory and lost many of their prominent leaders. Within this context, AQAP has undergone a number of shifts, originating as a vanguard with leaders thought to be particularly close to the internationalist core of al-Qaeda's central leadership to deeply-rooted Yemeni group made up of tribal cadres (Barfi 2010).

To obtain a high-level view of the trajectory of AQAP, I modeled 916 articles in the UCDP Database, spanning 2009 through 2020. Figure 2 presents time-trends in topic estimates for the AQAP articles. Each point on the plot represents a UCDP source article about AQAP's activities, with the placement capturing that article's location relative to a (hypothetical) purely Topic 1 assignment, positioned at -1 on the Y-axis, or a (hypothetical) purely Topic 2 assignment, 1 on the Y-axis.[3]

The plot highlights excerpts of several of the modeled articles to allow for a better sense of what the articles record. Topic summary words printed at the top and bottom of the plot provide a tool to identify and summarize the substantive meaning of each side of the scale. From the topic words and the highlighted articles, the first topic appears to capture a regional political mode, describing conflict away from the capital and operations against non-state targets, such as local power-brokers and rival military groups. Conversely, the second topic appears to be oriented more towards AQAP's conflicts against the Yemeni state and state forces.

[3]Points outside of the range are due to the jitter added for visualization.

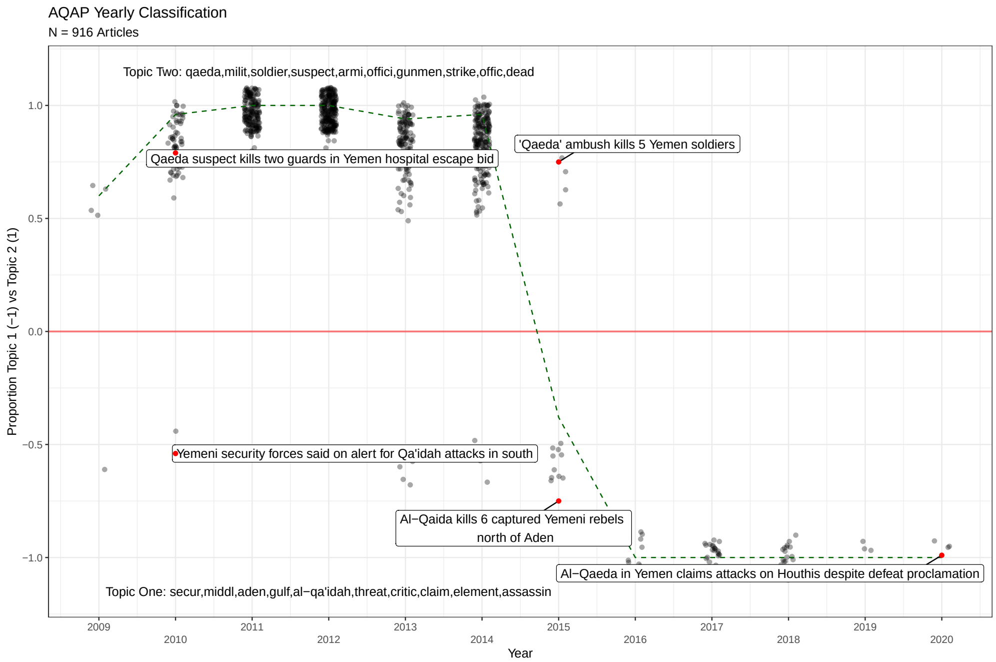

Figure 2: Example of an identified change period, with a representational change point between 2014-2015. Each point represents a story from the Georeferenced Event Database reporting on violent events associated with al-Qaeda in the Arabian Peninsula while the dotted line tracks the yearly average topic proportion assignments in the AQAP sub-corpus. The caption boxes feature text from selected stories before and after the 2014-2015 change point to highlight the trends in story themes.

To elaborate, in 2009 there were five articles about AQAP as the violent actor; after the text analysis, four of those articles were assigned to Topic Two and one was assigned to Topic One. The four articles assigned to the Topic Two comprise two reports of a Yemeni raids on AQAP targets, AQAP killing state security officials, and an AQAP claim of responsibility for an operation. The fifth article reports that AQAP kidnapped and killed a detective in the Ma'rib, a strategically important territory in the center of the country. Thus, the yearly third-party representation in the UCDP article collection is 80% Topic Two (The AQAP-Yemeni state theme). By the following year, there were 61 articles about the group, 60 of which were in Topic Two. Conversely, in 2015, of the 16 articles about the group, 11 were in Topic One (the rebel competition/regional conflict theme). That conversion gives AQAP's estimated "change" year, with 2015 as the year when third-party representation of AQAP shifted from describing an organization primarily involved in conflict with the state—

consistent with their self-presentation as a vanguard jihadi group seeking to overthrow the political system of Yemen— to a group heavily involved in a regional confrontation via the prism of hyper-local Yemeni rivalries.

The second featured illustration is Abu Sayyaf (ASG), an Islamist-motivated separatist group based in the southern Philippines. ASG is a difficult case because the group is widely described as a small, closed, group that operates in both a militant separatist mode and and an organized crime and piracy dominant mode (*e.g.* Bowden 2007). One may expect that identifying transition points for ASG would be challenging because banditry and crime may be less present in the news record and could plausibly be less likely to be collected for a dataset covering politically motivated violence.

The UCDP encyclopedia describes the erratic nature of Abu Sayyaf; particularly noting that observers perceived Abu Sayyaf as "oscillating between ideology-driven and profit-driven motivations" between 1998 and 2005 (*ASG* 2021). After the deaths of interim leaders, Khadaffy Janjalani attempted to reconsolidate ASG around its religious origins. Janjalani continued this effort until his 2006 death. In the intervening years, Abu Sayyaf has continued to engage in both insurgency and hostage-taking and piracy (*ASG* 2021). Figure 3 captures this period of oscillation and stabilization into an insurgency-framed pattern after about 2010.

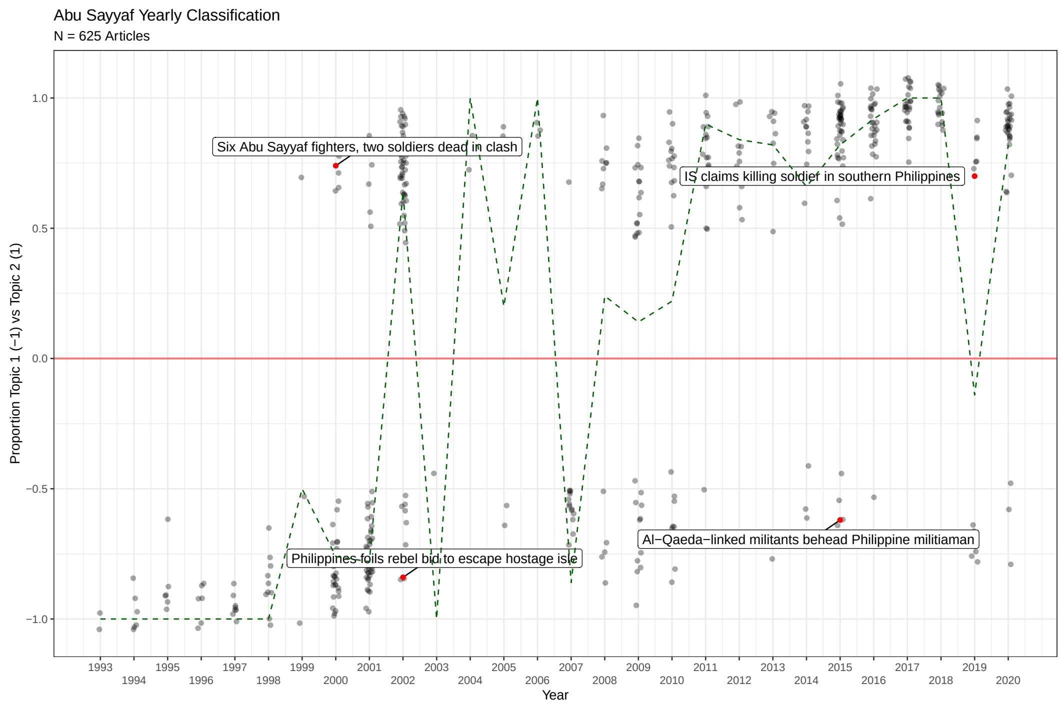

Figure 3: Illustration of Modeling Output, Abu Sayyaf

The third featured case is the United Liberation Front of Asom (ULFA), a separatist group fighting for the independence of Assam from the rest of India. This case represents a group that has shifted tactics across time and which is out of the jihadi space. Founded as a guerilla movement in 1979, the ULFA began insurgent operations in 1990 with the years between their establishment and the start of the insurgency used to recruit and develop cadres (*United Liberation Front of Asom (ULFA)* 2021; *ULFA* 2021). The ULFA's trajectory moved from a rural insurgent network with widespread popularity to an organization that, under military and personnel pressure, steadily became more extractive and predatory of the local population throughout the 1990s (*ULFA* 2021).

The ULFA modeling, depicted in Figure 4, is suggestive of an organization that started the 1990s as a militant group in conflict with state forces before transitioning to one using terror tactics which reverted to a militant thematic representation around 2010.

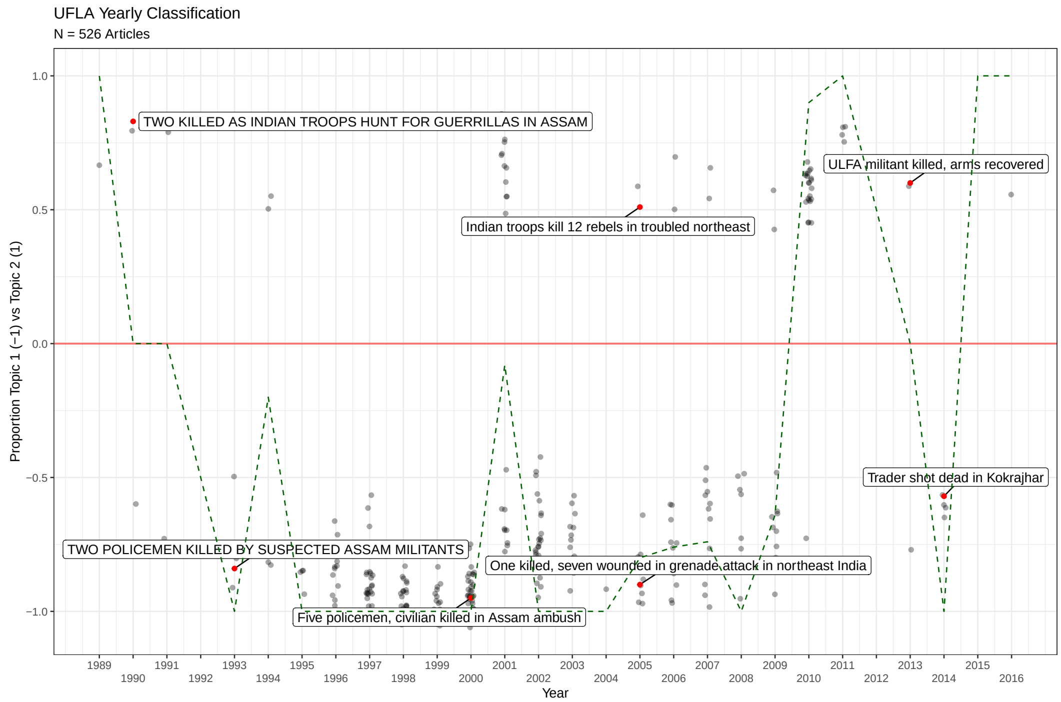

Figure 4: Illustration of Modeling Output, ULFA

Finally, Figure 5 shows the yearly estimated topic proportions for the fourth highlighted case: the Lord's Resistance Army (LRA). Despite ebbing strength by the 2010s, the LRA's "historical patterns...suggest consistent strategies and tactics used by the LRA" extending from the 1980s to the 2010s (Lancaster, Lacaille, and Cakaj 2011, pp. 4, 21). Their tactical profile has remained consistent: abductions and ruthless predation of civilian communities; rural operations (especially ambush and hit-and-run) across a large region; and violence conducted by small groups operating under varying levels of central coordination and control (Beber and Blattman 2013; Lancaster, Lacaille, and Cakaj 2011).

The stable characterization of LRA is reflected in the model results. The plot for the LRA is based on the classifications of 1,038 articles from 1989-2020, covering 31 of the group's 34 years of operation to date. The model consistently scales the LRA-associated news articles around -0.5 on the [-1, 1] scale, which substantively corresponds to most of the articles being classified as about 75% Topic One. Digging more deeply into the FREX words related to the topic further underscores the consistency of presentation, as the two topics share an unusual number of overlapping FREX terms and articles focus heavily on massacres and one-sided violence.

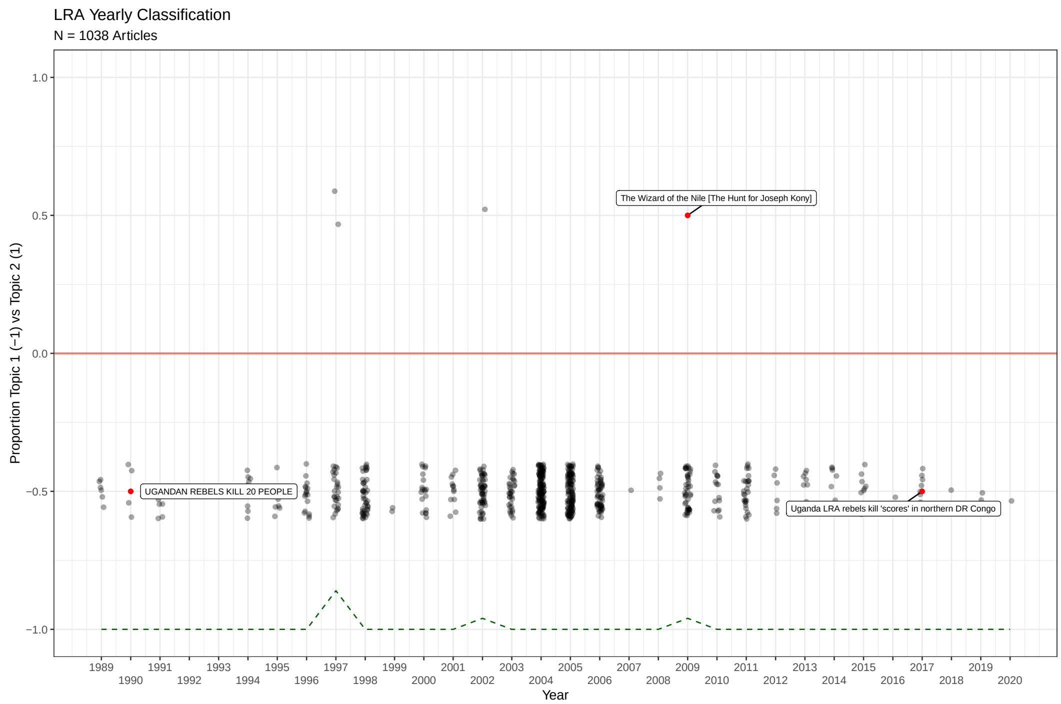

Figure 5: Illustration of Modeling Output, LRA

5 Challenges

In this section, I focus on four points of potential concern: selection effects, media access, model specification, and face validity. As well, I describe in general how a researcher could analyze the results for specific groups of interest. I evaluate the face validity of the measure for four groups chosen to cover a range of operational and situational environments.

The most foundational question is whether to expect the modeling strategy to work at all. Several validity tests address this question. First, I compare temporal trends and estimated change years in cases selected to cover an array of contexts and priors. An indication that the measure is meaningfully capturing an underlying process comes via the application: the transformation measure substantively and statistically follows theoretical expectations of how uncertainty impacts conflict termination.

5.1 Selection and Media Access Effects

Readers may be concerned about selection effects. Actor events missed by the data aggregation process will not be modeled. This would be particularly worrying if certain types of activities are systematically not recorded by local reporters, not included in the data collection process, or not described within the news corpus. The data used in this application was aggregated via a focus on violent events; thus, a major strategic shift to non-violent or politically focused activities would not be captured without incorporating additional data sources.

The sensitivity of event datasets to missing data and the processes by which the records can be distorted in news coverage has been carefully studied (*e.g.* Baum and Zhukov 2015; Eck 2012; Dawkins 2021; Weidmann 2015; Weidmann 2016). However, analysis has found that, despite known limitations in the data gathering processes, machine-coded and human-coded event data provide comparable predictions and inferences (Bagozzi et al. 2019).[4]

A related question is whether there are media access effects. Research has shown that restrictions on media freedom shape reporting of conflicts. If access determines narrative, then it may be the case that reporting on militant groups is decoupled from their activities. There are several pathways through which the media environment can artificially shape what the text models finds. I examine whether the media environment is associated with length of time between model-identified group change periods. If lack of access tends to freeze media discourse, then less precision should be associated with longer periods between-group thematic change points. The inverse may also be at play: if access leads to more variation in reporting on group activities, a more open media environment should correspond to shorter intervals between identified change years. If a restrictive media environment encourages journalists to reuse frames and established tropes, militant groups operating in difficult media environments should go longer stretches between identified changes. Conversely, if a

[4] An exception may be fatality estimates, which are notoriously difficult to verify and extremely sensitive to source choice (Davenport and Ball 2002; Landman and Gohdes n.d.; Manrique-Vallier, Price, and Gohdes 2013)

more accessible media environment produces more vibrant media narratives of conflicts the text model may incorrectly identify this as a period of change. In that case, a more free media environment should be associated with more changes in representational description of the group.

I address the media access concern via two proxies of the media environment. The first are the variables that the UCDP reports to characterize the precision of reporting about violent events in the database. The GED precision variables allow us to check whether low/high precision reporting years and are associated with low/high change group-years. To capture media restrictions, I use the Media System Freedom (MSF) index developed by Solis and Waggoner 2021. The results are presented in full in the Location section below. In both cases, country-year media freedom and group-country-year reporting precision averages are generally uncorrelated with the identified change years.[5]

5.2 Location Precision

I evaluate the measure's sensitivity to location precision by using the Media System Freedom (MSF) Index and variables native to the UCPD's Georeferenced Event Database. The first set of precision variables comes from UCDP event metadata. These variables are *event clarity*, *location precision*, and *date precision*. Event clarity indicates whether the report was detailed enough to allow the UCDP coder fully identify the event or whether the event listed in the dataset was part of an aggregation of information. In this case, it proxies for possible media access conditions. Date precision ranges from 1 to 5, spanning the exact date of the event known to the coders (level 1) to the date only known at a level of precision between one month and a year (level 5). Likewise, the location precision variable ranges from 1 (the exact location of the event is known and coded) to 6 (only the country where the event took place is known).[6] For each of the clarity variables, higher values denote less

[5]I use group-country-year to allow for contexts in which a militant group operates transnationally and there is a meaningful difference in media conditions across the countries of operation.

[6]The UCDP's coding includes an additional category for events in international waters or airspace. This category corresponded to a handful of naval piracy events occurring near Somalia and Sri Lanka in my

precision. I converted the event-level precision metrics into a group-country-year summary to allow transnational militant groups to operate in different media environments. I then use the group-country-year average of each variable as a proxy for ease of media access and reporting on the group's activities.

Figure 6 presents a linear regression model with year-fixed effects and standard errors clustered at the group level. The model takes the absolute value of the lagged year-on-year representation change as the outcome variable. The model fails to find statistically or substantively significant relationships between the media precision variables and change years.

For this, I use the Media System Freedom (MSF) index developed by Solis and Waggoner 2021 to proxy media access. This measure uses a Bayesian Item Response Theory model to synthesize the findings of ten existing indicators. The MSF Index provides coverage from 1948 to 2017, allowing me to incorporate a measure of media freedom for nearly the entire range of my transformation data. I capture the overall media environment using this index, which ranges from 0 (absence of media freedom) to 1 (complete media freedom).

Figure 7 uses the precision and media freedom variables as independent variables, again with standard errors clustered at the group level with year-fixed effects.

data. I omit naval events because they form a small component of the data and because the coding of the category breaks the coding pattern in which lower values indicate more precision and higher values indicate less precision.

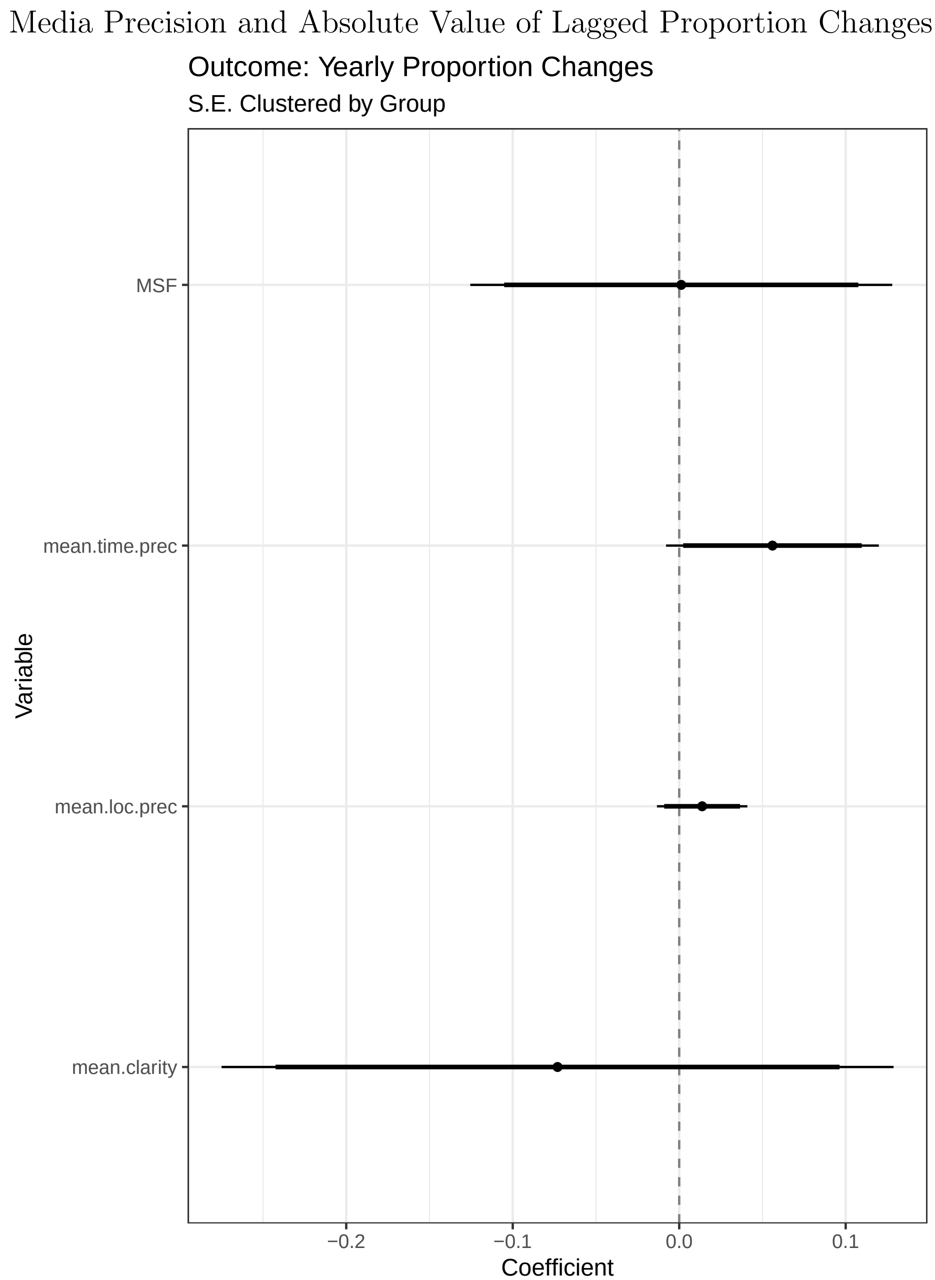

Figure 6: Relationship between Media System Freedom (MSF) measure of a country's media environment on the absolute value of lagged topic proportion changes.

Effect of Media Environment And Change Records

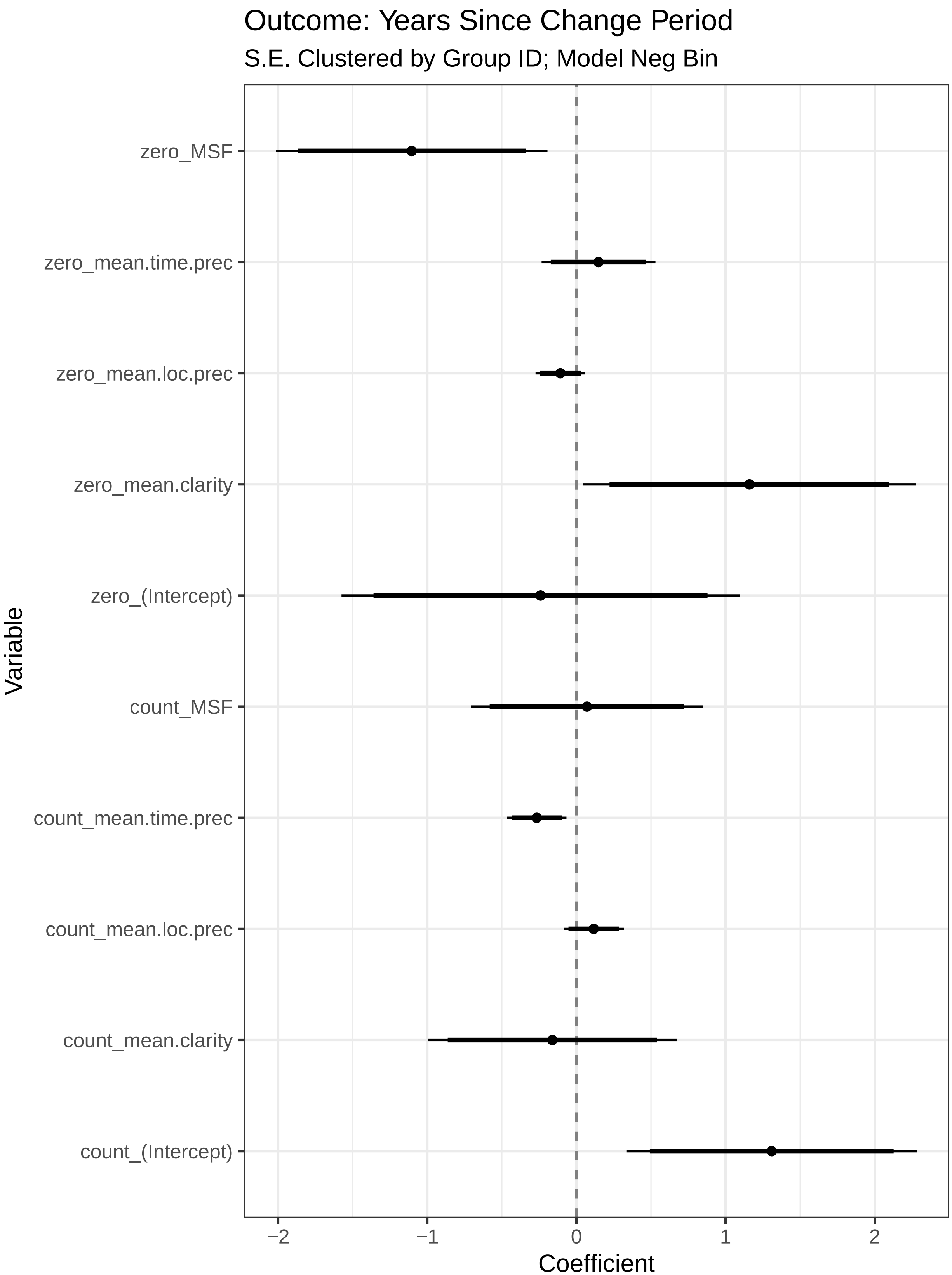

Figure 7: Zero-inflated negative binomial model with the interval between change year(s) and termination as the dependent variable and media environment as the main independent variable.

Figure 7 presents years since change as the outcome variable and media environment variables as the independent variables. The analysis uses a zero-inflated negative binomial (ZINB) model, which models count data in which there are both event and structural zeros. In this context, that approach allows the model to account for zeros that indicate that no change has occurred for a group (structural zeros) as well as zeros that indicate the year of a representational change (the event zeros). The outcome variable is a count variable recording the number of years since a change occurred in the reporting of a group. As with the previous analysis, if media access drives reporting of changes, we should expect to see a relationship between the variable that counts years since a change and the UCDP precision variables. In this case, that relationship should be positive because the UCDP precision variables increase in value as article precision declines. In the results reported in Figure 7 militant groups are less likely to have no changes at all (structural zeros) when they operate in countries where the MSF score ranks as less free. Thus, there may be a slight media effect in that groups operating in countries with less media freedom are significantly less likely to be described as having any change periods at all. However the other predictions are not borne out.

5.3 Document Sparsity

Readers may be concerned that the strategy of modeling reported activity means that years with very few documents may create a distorting effect on the data. If there are groups with relatively few events, the topics on which the representation measurement is base may be estimated on the basis of small numbers of potentially very short documents. In order to address this concern, the following section presents results of the termination analysis using a progressively stricter threshold on minimum number of documents appearing in a given year.

The re-estimated the change point analysis and replication uses subsets of groups that meet a set of thresholds intended to ensure a baseline amount of data across years. The thresholds are: at least one article/year for 75%, 90%, and 100% of the group-years in the data; at least 5 articles/year for 75%, 90%, and 100% of group-years; and 10 articles/year for 75%, 90%, and 100% of group-years. The thresholds can have a large influence on the data that is included. For example, the subset of the data that imposes a threshold of at least one document (N=1) for every year that the group was included (T=100) models 299 subjects for 1229 group years with 398 conflict terminations. The most restrictive version, which requires ten articles (N=10) for every year that the group is in the data (T=100) models outcomes for 105 unique non-state groups versus 285 unique non-state groups in the version that enforces one article in every year of activity. At all thresholds, these criteria are stricter than the application in the paper, which requires at least 10 words across all years.

Figure 8 presents results for each of the inclusion thresholds. The analysis recreates the replication in Figure 3 of the main text, but with differing criteria for how groups are included in the model that estimates change. The results are largely unchanged in both point estimate and precision of estimation until the N=10, T=100 point, at which the effects are substantively and statistically equal to zero. This result should not be surprising, as the subset with the most restrictive 10 articlesyear threshold omits many of the longest, most complex, conflicts in the dataset – unsurprisingly, as the threshold omits years with sparse information as well as years in which a conflict has ebbed. Notable missing rebel groups include Abu Sayyaf, the FARC, the LRA, the LTTE, and the PKK.

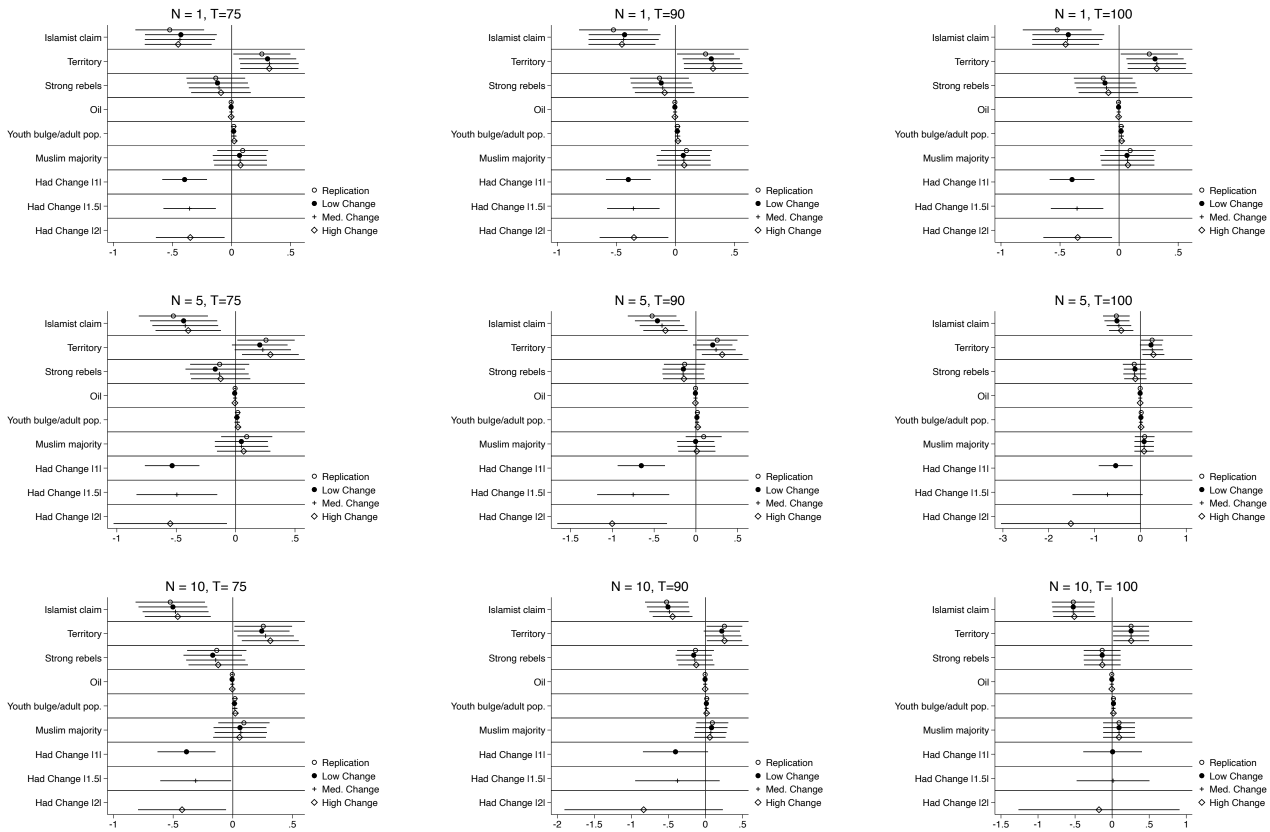

Figure 8: Effect of minimum document requirements on regression output

6 Topic Model Specification

Readers might wonder about the choice of a single model specification across all groups. Can we expect an unsupervised topic model to generate meaningful scales of group behavior absent targeted researcher feedback and validation? In particular, setting K=2 for all cases supposes two and only two different modes of thematic representation. In theory, restricting all groups to only two topics should pose a significant limitation. This could result in overestimating change periods if there is one extremely dominant frame that the model is forced to fit two dimensions. On the other extreme, the uniform two-topic model may force a true multi-dimensional representation into an artificially unidimensional space. This could result in the model underestimating change periods, if multiple thematic periods are compressed onto a two-dimensional axis.

There are several ways to evaluate these theoretical and practical concerns about model specification. The first is via the source articles themselves. It is reasonable to expect that applying a two-topic model regardless of group should flatten important variation across contexts. However, in practice, articles are thematically circumscribed by their genre, generally short length, and the UCDP's selection and vetting process. Removing stopwords relating to the reporting process further consolidates the themes. Together, these features reduce the complexity of each group's corpus.

Secondly, we can inspect the processed group scales for coherence. An advantage of using a text model to derive the group-specific scales is that researchers can easily inspect the substantive meaning associated with each group-specific scale. These are accessed via the high probability and FREX words associated with each group-specific topic. The FREX words– terms that are relatively common within a given topic and relatively rare outside of the topic– are reported in the yearly summary data for each group. The terms summarize the topics, which produce each end of the thematic scale. In doing so, the terms provide an entry point to assess the substantive implications of each group-specific thematic cluster.[7]

Finally, in cases where researchers suspect that a two-topic model is misidentifying change periods to a degree that challenges their research design, it is straightforward to adjust the algorithm to identify a larger number of topics. However, doing so adds considerable overhead, so this avenue is most useful for scholars looking in depth at relatively few cases.

Ideally, one could quantify the degree to which the two-topic model deviates from a ground-truth ideal text-based model. At that point, one could quantify the gains in scalability from using the two-topic model relative to the loss of precision and accuracy of using a model with fewer topics. However, in the case of optimizing topic selection, "ground truth" is fundamentally unknowable.

Moreover, identifying a group-specific optimal number of topics is not the only challenge inherent in increasing the number of topics above K=2. As the ultimate goal is to identify

[7]The associated terms for each group's scale are available from the code output and also featured in the replication dataset.

periods in which reporting suggests a change in activity, adding more topics creates a decision point as to how to aggregate similar modes of activity described by the different topics. For some group-model sets, this may be straightforward. However, for others it may require considerable researcher oversight and judgement.

As a middle ground, I examine a handful of cases for which prior knowledge would suggest that a K=2 model is either too many or too few underlying topics relative to that group. For each, I look at what the STM diagnostics would suggest as an optimal number of topics. To get an idea of the information being left behind by the K=2 specification, I then compare the conclusions about change years suggested by models with more topics with those derived from the K=2 model. In particular, I look for whether the results of the different models produce different estimated change points.

Given the lack of scalability in comparing alternative specifications, I highlight four armed groups selected to cover an array of failure modes for the approach. Each of these organizations was chosen because they are relatively long lasted armed groups, often with several modalities. The length of operations creates more opportunities to discover that the K=2 model is insufficient. These groups cover expectations of different eras, and are coherent organizations (versus more amorphous actors, such as "Syrian Insurgents".)

The first two, the Ogaden National Liberation Front (ONLF) and the Lord's Resistance Army (LRA), are armed organizations with no representational changes in the K=2 model. Running a range of topics for these can help to uncover whether the K=2 modeling approach presented in the Note missed change points that could have otherwise been uncovered by a more fine-grained approach. The third case, al-Qaeda in the Arabian Peninsula (AQAP), is selected as an example of an organization in which substantive researcher knowledge might suggest a specific number of topics other than K=2 (K=3 in this case). The final category represents the failure model in which an organization is not well-represented by a two-topic topic model, and increasing the number of topics would imply different change years than does the K=2 model. This case is modeled with the Kurdistan Workers' Party (PKK).

6.1 Case One: No Change in K=2

The first set of groups to spotlight are representatives of the category of militant groups that have no thematic changes in the K=2 model. The central question is whether the apparent stability captures the underlying representation, or if this is an artifact of the topic model specification. This is an important category to open with because it challenges the underlying approach: if having more than two topics would give change points that the two-topic model misses, then the which conclusions (and follow-on analysis) may be wrong because the categorization of groups is incorrect. For this category, I look at the Ogaden National Liberation Front (ONLF) and the Lord's Resistance Army (LRA). These groups are chosen for being among the longest-running of the groups without a change year in the K=2 specification.

Lacking a strong prior on the number of topics to find for either the LRA or the ONLF, I first investigated whether the built-in STM diagnostics suggest a number of topics that better fits the data. The results of these diagnostics are presented in Figure 9. The oscillation of benchmarks across the range of 5–20 topics suggests that there is not an obvious number of topics to aim for. I thus selected four topics: a low number that reflects the narrow range of topics in the corpus of local news reporting on violent events associated with each group.

Looking naively at a four-topic model for the ONLF might give the impression of 8 years of change, as different topics predominate. A more conservative coding could cluster Topics 1 and 2, which would produce an estimate of 4 change years. However, within the four-topic model, the topics are extremely similar. Indeed, two of the discovered topics are observationally identical. Thus, one can force periods of "change" by adding more topics, the "changes" capture very similar (and sometimes identical) topics. The texts from the LRA behave similarly: among the four topics, two seem to be so closely related that they have near complete overlap in most likely terms.

6.2 Case Two: Prior About K

The second category to spotlight is the case of a group where substantive knowledge would suggest a specific number of topics other than two. In this case, I look at the results for al-Qaeda in the Arabian Peninsula. I chose this group to highlight because knowledge of the case could lead to conclude that AQAP was operating in at least three frames: as a transnational jihadi group active in the global War on Terror, as a regional militant group battling the Yemeni state, and as a participant in the internationalized civil war centered around the Houthi insurgency.

Based on this background, one might expect AQAP to have at least three distinct periods of operations— pre-Arab Spring, then AQAP-Houthi civil war, then the AQAP-Forces of Saleh conflict. So it makes sense to assume that at two-topic model under-specifies changes. Thus, we want to see (1) what model would look better and (2) how many and which change points are lost by putting AQAP into a two-topic model. In this case, adjusting AQAP to a three-topic topic model does find the three frames that one might expect: a topic about AQAP in the War on Terror, a topic about AQAP as a militant group opposing state forces, and a topic about AQAP's involvement in the internationalized civil war. Despite picking up a War on Terror topic, the majority topic does not change for any given year and so the transition points and themes estimated by the three-topic model are the same as those for the two-topic model. This can be seen in Table 1, which presents the yearly paragraph counts and topic assignments for AQAP from 2009-2020.

Thus far, this appendix has analyzed the effect of changing the baseline number of topics in two cases in which the two-topic default does not find major shifts in framing (ONLF and LRA) and for a case (AQAP) in which there are substantive reasons to believe that different number of topics would more accurately model the group's evolution. Exploring different specifications for each of these groups ultimately did not change the number and year of change periods from that identified by the two-topic model.

Year	CountT1	CountT2	CountT3	PropT1	PropT2	PropT3	maxTopic
2009	0	5	0	0	1	0	2
2010	0	61	0	0	1	0	2
2011	0	174	0	0	1	0	2
2012	19	228	0	0.08	0.92	0	2
2013	25	101	0	0.2	0.8	0	2
2014	4	211	0	0.02	0.98	0	2
2015	0	11	5	0	0.69	0.31	2
2016	0	0	10	0	0	1	3
2017	0	0	35	0	0	1	3
2018	0	0	18	0	0	1	3
2019	0	0	4	0	0	1	3
2020	0	0	5	0	0	1	3

Table 1: Yearly Summary for AQAP Three Topic Model

6.3 Case Three: Different Change Points

The rest of the appendix explores a third type of contingency, in which change periods are sensitive to the model specification. This is explored via the case of the Kurdistan Workers' Party (PKK). Since being founded in 1978, the PKK has fought an insurgency against Turkey to create a Kurdish state using territory in Iraq, Syria, and Turkey.

PKK is a long-running and complex organization. The group has undergone several strategic and ideological shifts during their operation, at times fighting for independence and at times for autonomy. Likewise, the degree to which the group has remained close to their secular leftist ideology has waxed and waned, as they sought support from a more Islamist-leaning rural community. Thus, as with the AQAP case, we can expect the K=2 to overlook possible change periods and therefore to lose information. Looking at the results, we may be concerned that the two-topic model is inadequate to capture the richness of the group's trajectory. Indeed, the K=2 model essentially compresses the PKK's operation in the GED into pre-and post Syria.

STM features a tool to check model fit, which suggests that the most conservative number of topics for the PKK would be around 5. Modeling an alternative topic model with 5 topics gives a more nuanced perspective: one topic appears to be about Turkish offensives against the PKK[8]; the second describes PKK insurgent activities[9]; the third appears to relate to the PKK's efforts to establish an independent or autonomous territory of Rojava[10]; the fourth discusses militant activities[11]; and the fifth discusses the PKK's involvement in the Syrian civil war.[12]. The five-topic model records eleven changes in dominant topic between 1989 and 2020, although with a general trajectory that from 1989-2006 the dominant topics record the ebb and flow of PKK and Turkish offensives against each other. These are topics 1,2, and 4. Then from 2018 onwards, the dominant topics have an underlying theme of battles over PKK territorial bases in eastern Syria and Iraq. After 2006, Topics 3 and 5 dominate, with a single interjection of Topic 4 as the dominant frame for 2011.

Figure 10 compares the change plots for the base two-topic model and a five-topic model with the topics clustered according to the thematic overlap in the topics. The trajectory of dominant frames in a 5-topic PKK model (on the right) suggests 2006 as the primary transition year. This differs from the conclusions of the 2-topic model on the left, in which 2013 and 2018 are transition years between the PKK-Turkish (topic 1) and Rojava (topic 2) thematic clusters. Thus, a scholar interested in a more precise model of PKK change trajectory and year of change may opt to customize the model's number of topics for the PKK corpus.

The modeling approach is flexible and provides tools for analysts to customize the degree to which they want to tune the number of underlying topics of activity. The approach is most straightforward when there are only two dimensions being sought, but using the STM and Georeferenced Event Database an analyst could base their decisions on a different number of dimensions, and, if resources allowed, tailor the number of dimensions to the groups. he most precise way to use this approach to measuring change via text would be to adapt it on

[8] Characteristic (FREX) words: ankara, terrorist, f, anatolia, jun, _f, apr[ril]

[9] Characteristic words: violenc, insurg, kurdish-rel, kurdish-link, bus, die, kurdish

[10] Characteristic words: peopl, defens, etkisiz [neutralize], hale [make], terörist, getirildi [brought], roj[ava], ozgur [free]

[11] Characteristic terms: conflict, eur, sweep, militiamen, x, rebel, kurd

[12] Characteristic terms: milit, european, iraq, militari, strike, air, iraqi

a case-by-case basis. A researcher could use their substantive knowledge of a given case to customize the number of topics, hand cluster the topics into thematic groups of topics, and then count change years between these thematic clusters. This would produce an estimate that richly captures the specific trajectories of an individual group. However, this increase in precision comes with a significant trade-off in reducing scalability. It may be worthwhile for research questions for which capturing the exact trajectory of change is important. Yet, as long as they record the specification of the model and their decision criteria, the cutpoints are transparent and replicable. They can thus be built upon in future work.

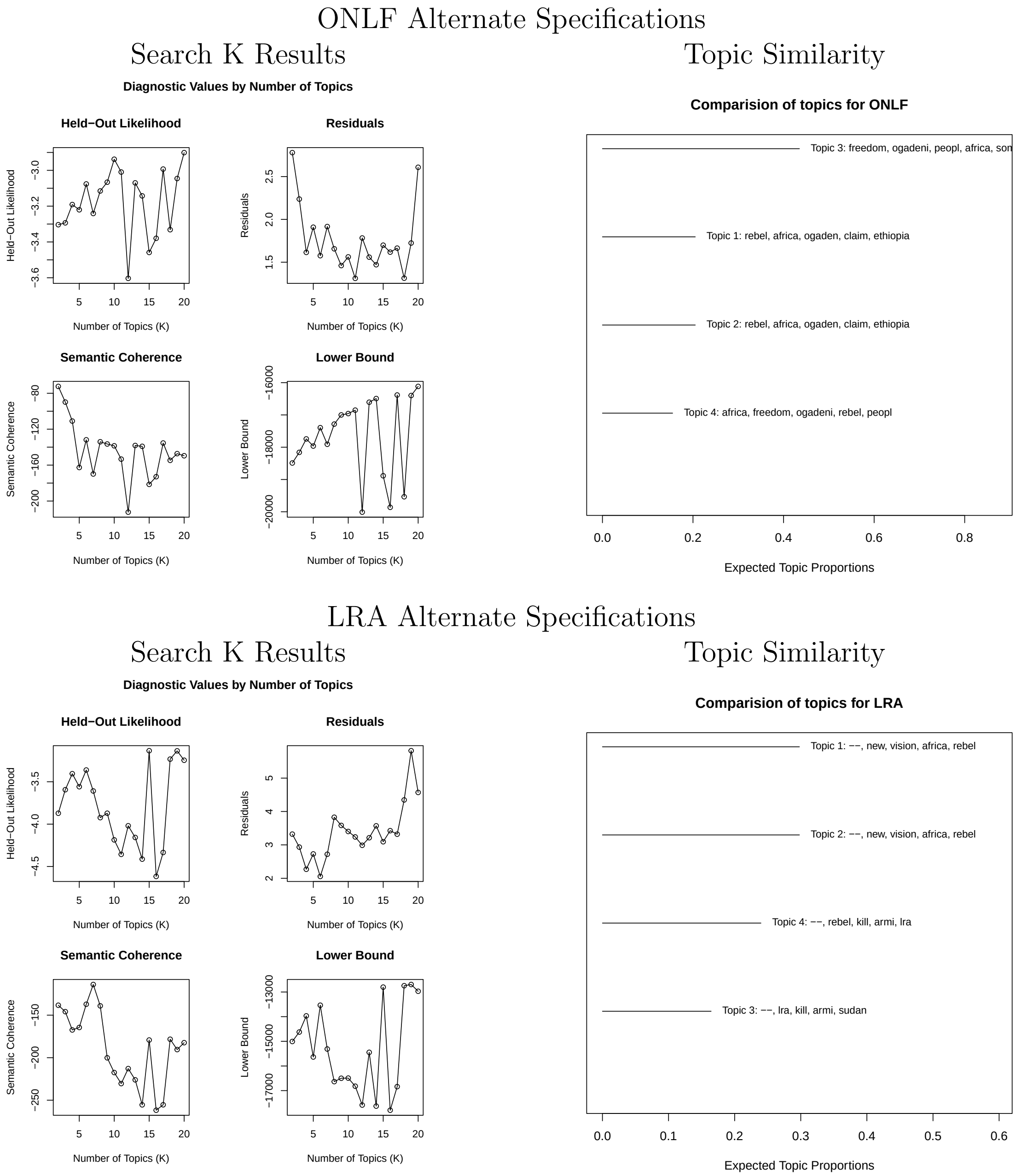

Figure 9: Comparison of Alternate Specifications for ONLF and LRA

PKK Change Graph with K=2

PKK Change Graph with K=5
(Clustered themes)

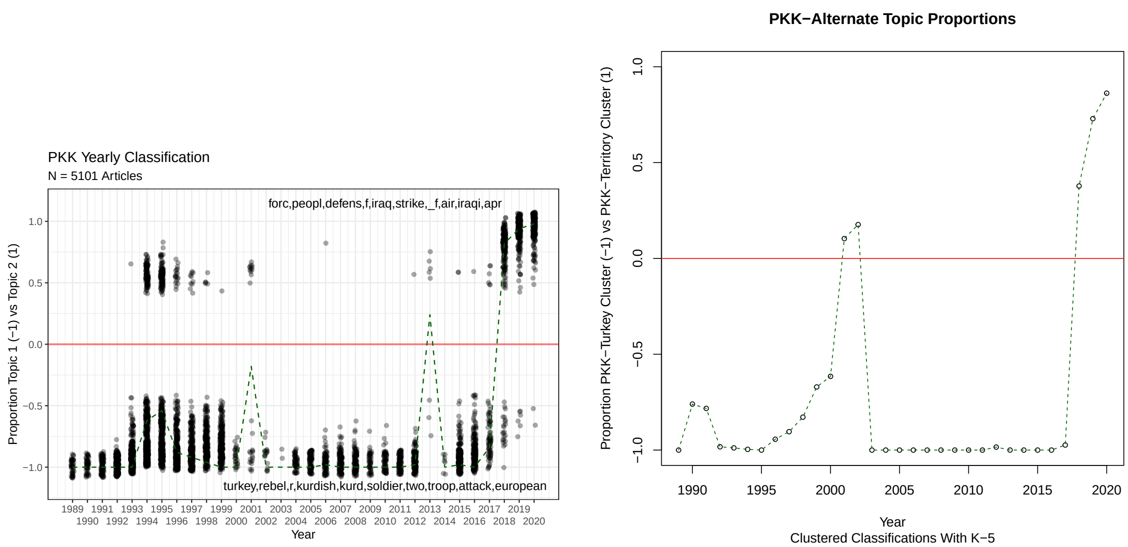

Figure 10: Change plots for a two-topic and five-topic model for the PKK. The plots illustrate a case in which using a different number of topics than K=2 may have the same overall trajectory but with different specific change years.

7 Application

7.1 Variable Correlations

Figure 11 shows the bivariate correlations between the change ("delta") and ambiguity variables and the variables in Nilsson and Svensson 2021's Termination dataset. The change variables are all highly correlated with each other—unsurprising, as they are a progressively increasing cutoff point for the same measure— but generally uncorrelated with other attributes gathered for civil conflicts. As in the specific termination and recurrence replications, this suggests that the change and ambiguity variables are adding value rather than indirectly capturing some latent attribute that is already measured in the data. The correlation plot omits location and identity variables as well as the variable for *outcome*. The outcome variable had an unusually high degree of missingness (more than 800 observations), leading to concerns that including it would require dropping enough observations to change the relationships among the data.

7.2 Additional Termination Models

The replication presented in the main text covers the most conservative application of the change measurement, namely a binary variable for whether the algorithm has identified any change at all. A more precise conceptualization may capture dynamics specific to the start or end of a conflict, such as conflating representational instability at the start of a conflict as both the group and observers develop expertise, with an impact on the duration of a conflict. If this is the case, we might expect to see an association between the number of years since a given group experienced change and the termination of the conflict that the group is part of. We would also expect to see many of the identified change points occurring relatively early in a group's active years. The "Years From Change" variable depicted in Figure 14 regresses the number of years since a change measure with the likelihood of conflict termination. The effect is both substantively close to zero and statistically insignificant, which suggests that

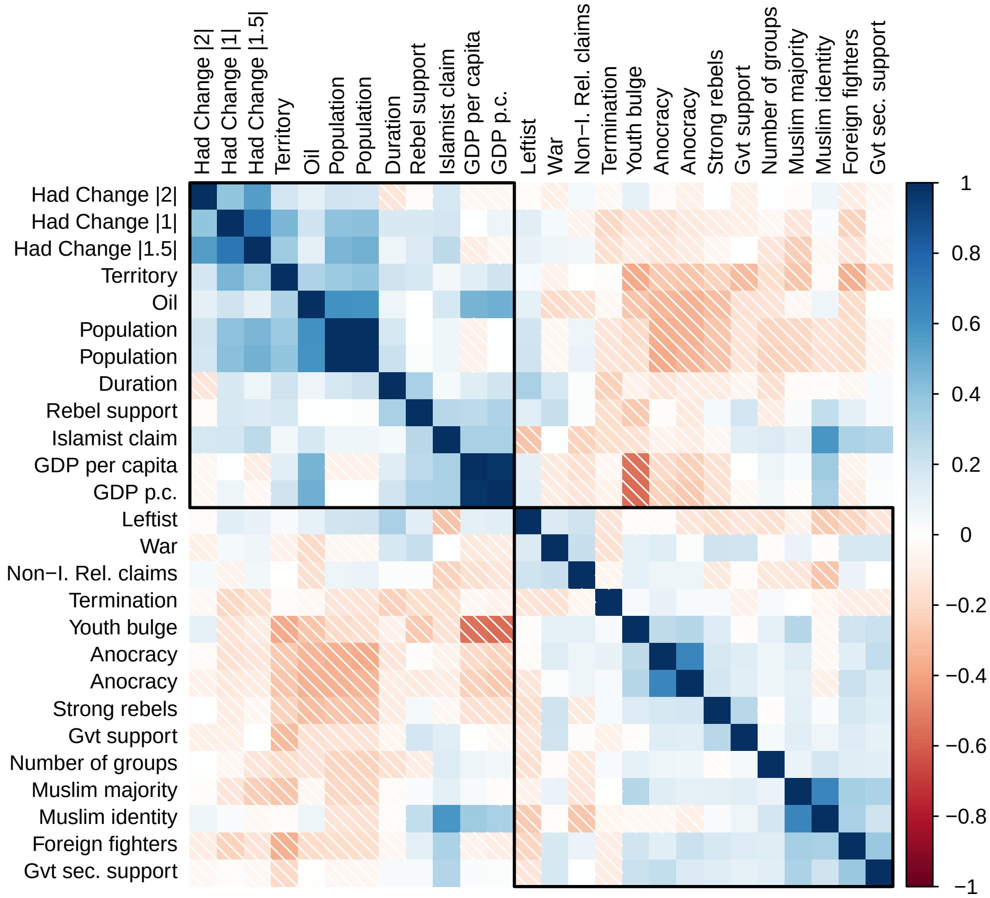

Figure 11: Correlation of change and ambiguity variables with variables in Nilsson and Svensson termination dataset.

the early stabilization story outlined above is driving the results.

One might also imagine changes in the data generating process (*i.e.* the conflict reporting) when long, intractable, conflicts grind towards a close. In this case, we should see a correlation between a recent representational change and the end of the conflict. The model depicted in Figure 14 introduces the "Change in Prev.2 Years" variable which records whether the group underwent a change in the previous two years. The data has 224 group-years in which there was a representational change in the previous two years, and 894 group-years in which there was no representational change in the previous two years. This is compared to

126 group-years with a group change and 992 group-years without any changes. The point estimates suggests that conflicts are slightly more likely to terminate within two years of a militant group change, thus lending credence to the perspective that the representation measure emphasizes changes around the end of conflicts. However, the lack of statistical significance indicates that the effect cannot be precisely separated from what might occur by chance.

Additionally, we can continue the logic that representation changes represent an underlying uncertainty that make conflicts more difficult to conclude by observing that if a single change is associated with more intractable conflict, the effect should hold for multiple representation changes. The "Change Frequency" variable counts the number of change periods recorded for a given group. The expectation is born out: more group change periods is associated with a decreased risk of conflict termination. Qualitatively, the effect is underscored by outliers in the changes variable. The two groups with the most change years are Abu Sayyaf (ASG) and the FARC in Colombia, each of which underwent four representation change periods. Moreover, the number of representational changes are not simply an artifact of both ASG and FARC being long-running militant groups and thus having more opportunities for observers to change how they describe the groups. The data has several examples of similarly long-running militant groups with few or even no representational changes. One such example is the Lord's Resistance Army, which was in operation for the entire 1989-2020 data window, but which had an extremely consistent yearly representation and thus had no periods of change.

Figure 12 summarizes the new count variables for change frequency as well as years between change periods.

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Number of Changes	1,114	1.061	1.166	0.000	0.000	2.000	4.000
Years Since Change	1,118	0.830	2.057	0.000	0.000	0.000	16.000

Figure 12: Distribution of new count variables

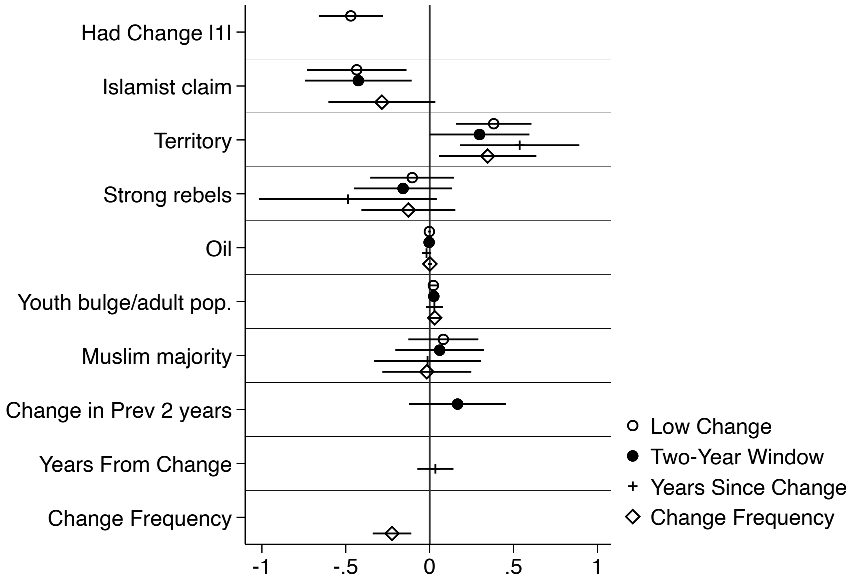

Figure 13: Replication of Termination Model, taking advantage of the extra information that the change measure provides by modeling whether conflict termination is associated with a recent change (change in previous two years), the years since a change point, and the number of representational shifts for each group.

7.3 Recurrence Replication

Replicating the recurrence dataset is complicated by the fundamental difference in data collection and inclusion. In particular, conflicts can endure despite years-long pauses in violence. This can create conflict years for which there is no corresponding event (and therefore no representation) data. For any given year, a low intensity conflict might remain in a dataset of ongoing conflicts but not appear in an event dataset.

A major difference between Nilsson and Svensson's data and my own is the presence of long-running conflicts with low intensity after 1989. In fact, of the 157 groups for which I do not have topic information, 129 were in operation for more than 20 years and 95 operated for 25 or more years. This means that many of the longest-lasting conflicts in their data are censored from my data. These conflicts include NSF [Lebanon], PLO-Israel [Levant], and Polisario - Morocco. The latter experienced more than half a century of conflict, but is associated with fewer than ten violent events in the GED and thus is not included in my

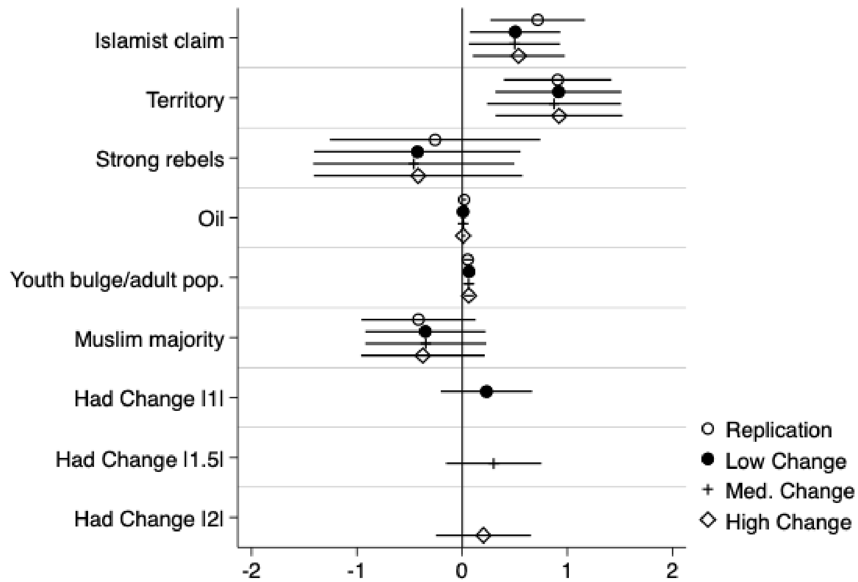

Figure 14: Replication of recurrence model, with censored time window and proposed change variable

primary analysis. Likewise, "Serbian irregulars" were associated with 41 conflict-years, but only 7 events in the GED.

Figure **??** shows the result of the censored recurrence model. As with the primary Termination model, the recurrence model replicates Nilsson and Svenson's original design using a binary for whether the non-state side was associated with an Islamist claim at the outset of the conflict. The model then adds three binary variables: whether the non-state actor had any one-year change of magnitude $\geq |1|$ on a $[-1, 1]$ scale, whether the group had any one-year topic summary changes of $\geq |1.5|$ or equal to $|2|$ on the same scale. In this case, we fail to reject the hypothesis that conflicts are more likely to recur when the violent nonstate actor experiences periods of change.

References

Acosta, Benjamin (2014). "From bombs to ballots: When militant organizations transition to political parties". In: *The Journal of Politics* 76.3, pp. 666–683.

Acosta, Benjamin (2019). "Reconceptualizing resistance organizations and outcomes: Introducing the Revolutionary and Militant Organizations dataset (REVMOD)". In: *Journal of peace research* 56.5, pp. 724–734.

Asal, Victor and R Karl Rethemeyer (2008). "The nature of the beast: Organizational structures and the lethality of terrorist attacks". In: *The Journal of Politics* 70.2, pp. 437–449.

ASG (2021). Date of retrieval: November 10, 2021. URL: https://ucdp.uu.se/actor/277.

Atteveldt, Wouter van, Mariken ACG van der Velden, and Mark Boukes (2021). "The Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary Approaches, and Machine Learning Algorithms". In: *Communication Methods and Measures* 15.2, pp. 121–140.

Bagozzi, Benjamin E et al. (2019). "The prevalence and severity of underreporting bias in machine-and human-coded data". In: *Political Science Research and Methods* 7.3, pp. 641–649.

Barfi, Barak (2010). "Yemen on the Brink?" In: *New America Foundation.*

Baum, Matthew A and Yuri M Zhukov (2015). "Filtering revolution: Reporting bias in international newspaper coverage of the Libyan civil war". In: *Journal of Peace Research* 52.3, pp. 384–400.

Beber, Bernd and Christopher Blattman (2013). "The logic of child soldiering and coercion". In: *International Organization* 67.1, pp. 65–104.

Benoit, Kenneth (2020). "Text as data: An overview". In: *The SAGE Handbook of Research Methods in Political Science and International Relations, SAGE Publishing, London (forthcoming).*

Benoit, Kenneth et al. (2018). "quanteda: An R package for the quantitative analysis of textual data". In: *Journal of Open Source Software* 3.30, p. 774. DOI: 10.21105/joss.00774. URL: https://quanteda.io.

Birnir, Jóhanna K et al. (2018). "Introducing the AMAR (All Minorities at Risk) Data". In: *Journal of Conflict Resolution* 62.1, pp. 203–226.

Bowden, Mark (2007). "Jihadists in paradise". In: *Atlantic Monthly*, p. 54.

Braithwaite, Jessica Maves and Kathleen Gallagher Cunningham (2020). "When organizations rebel: introducing the foundations of rebel group emergence (FORGE) dataset". In: *International Studies Quarterly* 64.1, pp. 183–193.

Chenoweth, Erica and Orion A Lewis (2013). "Unpacking nonviolent campaigns: Introducing the NAVCO 2.0 dataset". In: *Journal of Peace Research* 50.3, pp. 415–423.

Cronin, Audrey Kurth (2009). *How terrorism ends*. Princeton University Press.

Davenport, Christian and Patrick Ball (2002). "Views to a kill: Exploring the implications of source selection in the case of Guatemalan state terror, 1977-1995". In: *Journal of conflict resolution* 46.3, pp. 427–450.

Dawkins, Sophia (2021). "The problem of the missing dead". In: *Journal of Peace Research*, pp. 1098–1116.

Demarest, Leila and Arnim Langer (2019). "How Events Enter (or Not) Data Sets: The Pitfalls and Guidelines of Using Newspapers in the Study of Conflict". In: *Sociological Methods & Research*, p. 0049124119882453.

Denny, Matthew J and Arthur Spirling (2018). "Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it". In: *Political Analysis* 26.2, pp. 168–189.

Dudouet, Véronique (2013). "Dynamics and factors of transition from armed struggle to nonviolent resistance". In: *Journal of Peace Research* 50.3, pp. 401–413.

Eck, Kristine (2012). "In data we trust? A comparison of UCDP GED and ACLED conflict events datasets". In: *Cooperation and Conflict* 47.1, pp. 124–141.

Findley, Michael G and Joseph K Young (2012). "Terrorism and civil war: A spatial and temporal approach to a conceptual problem". In: *Perspectives on Politics* 10.2, pp. 285–305.

Fortna, Virginia Page, Nicholas J Lotito, and Michael A Rubin (2018). "Don't bite the hand that feeds: rebel funding sources and the use of terrorism in civil wars". In: *International Studies Quarterly* 62.4, pp. 782–794.

Giustozzi, Antonio (2019). *The Taliban at War: 2001-2018*. Oxford University Press.

Högbladh, Stina (2021). "UCDP GED Codebook version 21.1". In: *Department of Peace and Conflict Research, Uppsala University*.

Lancaster, Philip, Guillaume Lacaille, and Ledio Cakaj (2011). "Diagnostic study of the Lord's Resistance Army". In: *Washington, DC: World Bank* 3.8, p. 9.

Landman, Todd and Anita Gohdes (n.d.). "Challenges of non-random data in analyzing human rights violations during conflicts in Peru and Sierra Leone". In: ().

Lounsbery, Marie Olson and Alethia H Cook (2011). "Rebellion, mediation, and group change: An empirical investigation of competing hypotheses". In: *Journal of Peace Research* 48.1, pp. 73–84.

Mampilly, Zachariah Cherian (2012). *Rebel rulers*. Cornell University Press.

Manrique-Vallier, Daniel, Megan E Price, and Anita Gohdes (2013). "Multiple systems estimation techniques for estimating casualties in armed conflicts". In: *Counting civilian casualties: An introduction to recording and estimating nonmilitary deaths in conflict* 165.

Morrison, John F (2013). *The Origins and Rise of Dissident Irish Republicanism: The Role and Impact of Organizational Splits*. Bloomsbury Publishing USA.

Mosinger, Eric S (2019). "Balance of loyalties: Explaining rebel factional struggles in the Nicaraguan Revolution". In: *Security Studies* 28.5, pp. 935–975.

Nilsson, Desirée and Isak Svensson (2021). "The Intractability of Islamist Insurgencies: Islamist Rebels and the Recurrence of Civil War". In: *International Studies Quarterly* 65.3, pp. 620–632.

Pettersson, Therése et al. (2021). "Organized violence 1989–2020, with a special emphasis on Syria". In: *Journal of Peace Research* 58.4, pp. 809–825.

Sinno, Abdulkader H (2011). *Organizations at war in Afghanistan and beyond*. Cornell University Press.

Solis, Jonathan A and Philip D Waggoner (2021). "Measuring media freedom: An item response theory analysis of existing indicators". In: *British Journal of Political Science* 51.4, pp. 1685–1704.

Sundberg, Ralph and Erik Melander (2013). "Introducing the UCDP georeferenced event dataset". In: *Journal of Peace Research* 50.4, pp. 523–532.

ULFA (2021). Date of retrieval: November 10, 2021. URL: `https://ucdp.uu.se/actor/326`.

United Liberation Front of Asom (ULFA) (2021). Date of retrieval: November 10, 2021. URL: `https://www.satp.org/satporgtp/countries/india/states/assam/terrorist_outfits/ulfa.htm`.

Weidmann, Nils B (2015). "On the accuracy of media-based conflict event data". In: *Journal of Conflict Resolution* 59.6, pp. 1129–1149.

— (2016). "A closer look at reporting bias in conflict event data". In: *American Journal of Political Science* 60.1, pp. 206–218.

Wilkerson, John and Andreu Casas (2017). "Large-scale computerized text analysis in political science: Opportunities and challenges". In: *Annual Review of Political Science* 20, pp. 529–544.

Woldemariam, Michael (2018). *Insurgent Fragmentation in the Horn of Africa: Rebellion and Its Discontents*. Cambridge University Press.

Wood, Elisabeth Jean (2008). "The social processes of civil war: The wartime transformation of social networks". In: *Annu. Rev. Polit. Sci.* 11, pp. 539–561.