

Over Pressure: Grassroots-Driven Transformation of (Militant) Organizations

Margaret J. Foster*

April 21, 2018

Abstract Why might it not be beneficial for a clandestine or militant organization take advantage of an upswell of potential recruits who can increase group strength and resilience? I argue that when resource constrained organizations grow quickly, their leaders face the choice of balancing short-term organizational survival with long-term mission focus. If leaders are unable to quickly socialize their recruits, bottom-up internal pressures can transformation priorities and operational focus. I illustrate the logic and outcome via the example of al-Qaeda in Yemen (AQAP). Changes AQAP's activity patterns and self-presentation, suggest that rapid introduction of personnel may come at the cost of fundamentally transforming the character of an organization.

*PhD Candidate, Department of Political Science, Duke University. Mailing address: Gross Hall, 140 Science Dr, Durham, NC 27708. E-mail: margaret.foster@duke.edu, Web: margaretjfoster.net. Special thanks to David Siegel, Kyle Beardsley, Laia Balcells, Scott DeMarchi, Peter Feaver, David Banks, Ben Jones, Mia Bloom, Victor Asal, seminars at Duke University, and participants from Peace Science 2017 and ISA 2018 for their insightful comments, guidance, and support throughout the development of this project. I am also grateful to my former colleagues at the SITE Intelligence Group for providing data used in this project and for the generosity with which they have shared their expertise. All remaining errors are my own.

Introduction

When might improving the strength and resilience of an organization profoundly undermine the leader’s long-term success? This paper argues that taking advantage of some opportunities for recruiting can strengthen a group in the short run but introduce upwards internal pressure on leaders to satisfy the priorities of their new recruits. This pressure encourages leaders to adjust the strategies and goals of their movements to become more palatable to the new membership base—even if these strategies and goals were not originally desirable to the leader. This process of transformation suggests an under-explored trade-off that organizations face if they try to increase their resilience by deepening ties to a desirable recruitment demographic.

For a issue-motivated organization, such as a revolutionary militant group, this process risks undermining their very *raison d’être*. As their interests and identity appear driven by concerns of the new members, the group becomes less appealing to their former constituencies and more appealing to a membership base mobilized by issues salient to the new members.¹ A bottom-up transformation represents a particular trap for militant leaders, for whom integration with local communities provides resources and protection but can leave them beholden to the parochial interests of local recruits.

In highlighting a bottom-up mechanism of organizational transformation and presenting evidence consistent with this process in al-Qaeda in the Arabian Peninsula (AQAP), I contribute to the literatures on the organizational dynamics of rebel organizations and the social processes of civil war (Staniland, 2014; Weinstein, 2006; Wood, 2008) and terrorism (Asal and Rethemeyer, 2008; Crenshaw, 1987; Gupta, 2008; Oots, 1989). I make a methodological contribution by adopting an empirical strategy that leverages two distinct sources of text data—news reports and official communiqués—to generate separate but complementary avenues

¹I assume a strategic interplay between the leaders, who intend to quickly socialize the new members and thereby minimize the organizational cost and the members, who wish to harness the group’s capacities to advance their personal agendas. Additional work formalizes this dynamic and identifies the contexts in which each outcome can be expected.

of insight into internal dynamics of an opaque organization operating in a relatively information poor context. In aggregating multiple avenues of suggestive evidence, I strengthen the conclusions that can be drawn from any individual component of the empirical approach.

This paper proceeds in three parts: I first outline a theory of bottom-up organizational transformation in militant revolutionary organizations and position the theory within the existing literatures of civil war and terrorism. Secondly, I exploit a series of exogenous political and military events in Yemen to suggest two points at which a bottom-up transformation should be visible to outside observers. In each of these cases, events outside of AQAP's control influenced tribal reception of AQAP's recruitment efforts, thereby encouraging an influx of recruits with a particular set of priorities. As AQAP became more attractive to local Sunni communities, AQAP's self-presentation should have also become increasingly parochial.

The first large scale influx of tribal recruits was spurred by collateral damage from international counter-terror operations, which generated resentment against the Yemeni state and international actors. The combination of physical threats and resentment made local Sunni communities more open to the jihadi group's overtures (Human Rights Watch, 2013; Salama, 2014). The second recruitment shock followed rising domestic polarization along sectarian lines. After 2014, Sunni tribes mobilized around AQAP to counteract the local influence of a domestic Zaidi militia. I contrast these strategic events with a political development, the 2011 Yemeni Revolution. For reasons outlined below, the Yemeni Revolution is expected to have a modest impact on the receptiveness of Yemeni tribes to the Sunni insurgent organization.

To gather evidence that recruitment shocks were followed by shifting organizational priorities, I carry out a qualitative reading of translated texts from al-Qaeda in Yemen. I then use supervised clustering algorithms to analyze nearly 600 stories published by Yemeni and international news sources between 2002 and 2014. These reports describe activities of AQAP; a local spin-off, Ansar al-Shariah; and a local Shia militia also active in Yemen

during these years. I supplement the qualitative reading and news analysis with a statistical analysis of English translations of over 800 pieces of propaganda released by al-Qaeda in the Arabian Peninsula and direct precursors from 2004 through 2016.²

Theoretical Contribution: Organizations Are Who They Recruit

My central theoretical claim is that organizations are subject to upward-driving pressure from their grassroots. This paper addresses a significant empirical implication of the theoretical claim, namely that embedding into local communities and social networks increases organizational resilience but entails costs that have been largely overlooked in the academic literature. Unless socialization is comprehensive, incorporating a large number of recruits with established preferences can introduce a new constituency into the organization. In the most dramatic outcome, rapidly expanding recruitment from among a new social or political base can cost the movement's leaders control of their movement by facilitating a bottom-up transformation of the organization.

The tension between resilience and mission purity occurs because failing to accommodate the interests of new recruits can be catastrophic. First, recruits are logistically necessary: without their manpower, resources, and networks, the organization will cease to operate. Thus, some accommodation can be expected in order to keep the recruits satisfied.³ On the other, if the new recruits feel that their goals are not being met, they can form an internal faction within the group. Strong factions are dangerous for leaders, as they can lead to internal fissures, thereby reducing operational efficacy and increasing the risk of schism and collapse. For example, the stress of factional politics can lead militant organizations to

²The text analysis was conducted using English translations as the morphology of Arabic presents challenges for topic modeling. Although the development of text-as-data methods for Arabic is an active research area (see, for example, Abbas, Smaili and Berkani 2011; Brahmi, Ech-Cherif and Benyettou 2012; Salloum et al. 2018), the existence of an accessible corpus of professional translations for these texts allowed for a more straightforward research design.

³Some organizations attempt to restrict the exit options of their members, such as by encouraging them to engage in illegal activity.

splinter (Bakke, Cunningham and Seymour, 2012; Pearlman and Cunningham, 2012), while competition between factions can have lethal implications for the conflict theater (Bloom, 2004; De Mesquita et al., 2008).

Recruitment shocks may generate a self-reinforcing cycle of internal transformation. First, the production of violence is extremely sensitive to the private interests of combatants and, where internal institutions of socialization and control break down, local actors use the conflict as an opportunity to settle private grievances or enrich themselves, as Kalyvas (2003) describes theoretically and Peters (2012) analyzes in the context of the Haqqani network in Afghanistan).

Second, changes in rebel practices can generate downstream consequences for perception. For example, Weinstein (2006) traces how recruitment strategies can influence how organizations interact with civilian communities via a self-reinforcing virtuous or vicious cycle of restraint or predation. Likewise, Beardsley and McQuinn (2009) demonstrate an avenue through which group actions influence perception, shape relationships with the population, and can influence conflict termination. They identify how resource availability can influence rebel organizational structures and behavioral patterns, which then characterize the conflict environment. Indeed, they propose that “many of the ostensibly random behaviors of insurgents will begin to produce patterns of behavior by certain types of groups” (Beardsley and McQuinn, 2009, 628).

In the absence of strong socializing or monitoring institutions, the preexisting motivations that the new recruits bring into the organization are likely to inform the actions carried out under the auspices of the group. In this case, mobilization around an existing social cleavage is likely to generate actions that enhance the relative position of the recruits’ social base. This can result in violence and military operations that are intended to advance the interests of that internal constituency rather than the organization as a whole. Such interests may relate to the current conflict dynamics or to expectations about relative positions after the

conflict has concluded (Balcells, 2010). Either way, when rank and file members pursue their existing interests, these actions are likely to have follow-on effects as they are the behaviors that the community will ascribe to the group. In this way, the actions of recruits inform the expectations that other actors will use to shape their own involvement in the conflict.

Two examples from the previous century illustrate the transformative internal stresses that can result from recruitment shocks. In each of the vignettes presented below, the organization appeared to be able to strengthen their short-term position due to an influx of recruits. However, for both groups, incorporating the new members triggered internal strains by precipitating a bottom-up shift in the actions and priorities of the organization.

The first example comes from the trajectory of the Provisional Irish Republican Army during and after the 1970s. After heavy-handed British repression, the Provisional Irish Republican Army (PIRA) appeared to be in a position to benefit from an immense recruitment windfall. Indeed, the so-called “Bloody Sunday” on January 30, 1972 resulted in a massive influx of recruits for the PIRA. However, the gains in human capital came at a cost as “the new members were not traditional Republicans, as the men who founded the PIRA had been; rather, they were motivated by their everyday experience of British tanks and soldiers on the streets, and attacks by Loyalist militias” (Kenny, 2010, 554). While the increased manpower boosted their military success, in the long run the changing recruitment base instigated nearly catastrophic schisms between the PIRA’s original leaders and the Northern-driven factions. Moreover, the influx of volunteers who were motivated by revenge and anger rather than political ideals drove the PIRA to take on a “very different character” and the level of collateral damage that their attacks generated “rose precipitously after the early 1970s” (Kenny, 2010, 551, 554). Bowyer-Bell (2000) summarized the bottom-up evolution of the IRA, writing that: “as the new [members] were being absorbed the balance in the IRA shifted, became more representative of the emerging Provisional IRA rather than of the organizing fathers.... The leadership became younger and those from Northern Ireland grew in importance. The shift at the very top was slow but inevitable because of the focus

of the campaign. Further down as the Provisionals expanded rapidly local leadership went directly to the competent volunteer. And increasingly the Northern units became central to the focus of the movement” (Bowyer-Bell, 2000, 153).

Another recruitment-driven organizational change is alluded to in primary-source documents from radical British environmentalists. A 2003 periodical released by and for “ecological resistance” activists bitterly traced what the editor(s) viewed as erosion of the seriousness of the movement. The editor(s) of the tenth issue of a violent “ecological resistance” publication, titled *Do or Die*, complained that previous changes to recruitment strategies intended to encourage wider social involvement resulted in a class of members who were “attracted to ‘campaign’ jobs...[and] inclined to paper pushing rather than physical action” (Voices from the Ecological Resistance, 2003, 3). As well, they blamed attempts to increase the popular appeal of radical environmentalism with instigating “a terrible internal pressure crushing the radical content and practical usefulness of the groups” (Voices from the Ecological Resistance, 2003, 3). As an example of this process, the document cited Greenpeace ejecting a director, Paul Watson, for failing to moderate his activities. To the *Do or Die* editors, former leaders were concerned that that illegal actions would alienate new and prospective members, and thus the need to appease new members resulted in curtailing activists.

As the theory and vignettes suggest, access to a new recruitment base can be both vital and dangerous. The influx of members and social connections has the potential to make the group stronger, more resilient, and bring a new set of resources to exploit. However, recruits are the future of an organization and changing the membership base may result in transforming the organization’s own priorities as well as reshaping how other conflict actors perceive the group.

Case Study: Local and Transnational Tensions For AQAP In Yemen

I illustrate the trade-off between short-term strength and resilience and long-term mission creep via the example of al-Qaeda in the Arabian Peninsula (AQAP), a hybrid terror-insurgent organization based in Yemen. Three factors make AQAP a particularly valuable illustrative case. First, AQAP's association with al-Qaeda heightens international interest in the group, and so sufficient data exist to build a narrative of AQAP's recruitment goals and successes over a decade. As well, al-Qaeda affiliation means AQAP faces the challenge of balancing local-transnational push forces that run in countervailing directions. This balancing act presents an extreme challenge, as “[constituents] interested in nationalist or parochial concerns often have little interest in pursuing transnational religious violence, while the international ambitions of globally oriented jihadists frequently alienate them from broad-based local support” (Koehler-Derrick (ed), 2011, 11). Due to AQAP's participation in the al-Qaeda network, there is documentation of strategic instruction indicating that al-Qaeda Central (AQC) leadership strongly encouraged AQAP to avoid becoming entangled in local conflicts.⁴ These documents indicate that any localizing evolution was likely directed by either the local or the transnational al-Qaeda leadership.

In one such letter, believed to be written by al-Qaeda strategist Atiyah Abd al-Rahman, directed the Yemeni branch away from local operations, writing that in the past “...many jihadist groups did not succeed in gaining their objective because they concentrated on their internal enemy.” (al-Qaeda strategist, 2006-2011, 11) AQAP was correspondingly chastised for their domestic activities and urged to train their new members to prioritize the international struggle ahead of the local one:

“...it is very important to remind all of our brothers about it with a note to the new generation, who joined the jihad road and were not advised about this

⁴I referring to al-Qaeda's Central Asian-based leadership as “al-Qaeda Central” to distinguish leaders such as Ayman al-Zawahiri and Atiyah Abd al-Rahman from the Yemeni al-Qaeda leadership

issue. Thus, they conduct separate operations rather than concentrating on the main objective [the external enemy, i.e.,: the US] as we heard in the news about operations in Marib and Ataq against the government forces. I hope these operations were important for the mujahidins self-defense only.” (al-Qaeda strategist, 2006-2011, 6)

Second, AQAP’s unsuccessful attempts to recruit within the tribal communities suggests that they may quickly take in new recruits once that channel opened to them. Within Yemen, domestic instability and international military engagements created the opportunity for events outside of AQAP’s control to shape the recruitment environment. These events can generate local shocks that create opportunities to recruit from previously-reluctant communities. Thus, evidence suggests that AQAP initially had difficulty establishing significant recruitment networks within Yemen’s Sunni tribal communities, high casualty counter-terror operations mobilized the tribes and made these communities more amenable to AQAP’s recruitment appeals (Abdul-Ahad, 2015).

Third, AQAP created a comparison case by founding a locally-focused spin off organization, Ansar al-Shariah (Supporters of the Shariah). Ansar al-Shariah was established in 2011 as an arms-length local wing that could focus on domestic grievances and administration rather than AQAP’s transnational mission and which would be free of negative local sentiment associated with the al-Qaeda brand (International Crisis Group, 2017). Although quickly identified as an alias for AQAP, the two different brands provides a convenient referent point. A transnationally motivated, transnationally behaving, AQAP should operate in ways that are readily distinguishable from the Ansar al-Shariah local insurgency. In turn, both should be distinguishable from other domestic insurgent groups.

Trajectory of AQAP Recruitment

Although officially created only in 2009, the roots of AQAP can be traced to al-Qaeda offshoots in Yemen and Saudi Arabia active in the late 1990s and early 2000s.⁵ In late 2008, a Saudi crackdown on Al-Qaeda in Saudi Arabia sent the Saudi wing of al-Qaeda to Yemen. The two branches merged shortly thereafter, making AQAP a child organization of al-Qaeda in Saudi Arabia. In January 2009, the movement's Yemeni and Saudi wings announce their merger into the current form of the organization (Cutler, 2010).

An assessment built on interviews among Yemen's Sunni tribes conducted from 2008 through the fall of 2009 asserts that at the time of their official establishment, AQAP's recruitment was mainly drawn from urban centers, particularly from specific neighborhoods of Sanaa and Taizz (Koehler-Derrick (ed), 2011, 138). The assessment added that two common strategies of building local support were being rebuffed as their attempts to recruit from the tribal areas of Marib and al-Jawf did not result in extensive marriage links with the local tribes. Likewise, their attempts to generate support by solving tribal disputes or otherwise providing public services also failed (Koehler-Derrick (ed), 2011). At the time of the 2009 merger, the United States State Department estimated that AQAP's membership was approximately 200-300 (Johnsen, 2012).

After their failures in 2008 and 2009, AQAP shifted their recruitment strategy to re-market themselves with a more local face. In 2011, AQAP organization created a subsidiary, Ansar al-Sharia, that could more credibly promote local concerns. Under the Ansar al-Sharia name, AQAP could strike a more parochial message, exploit local grievances, and avoid the encumbrances of the al-Qaeda brand (Swift, 2012). Although given a different name for

⁵Within Yemen, AQAP has had at least three predecessor organizations: the Islamic Jihad in Yemen from 1990-1994, the Army of Aden Abyan from 1994-1998, and al-Qaeda in Yemen from 1998-2003. Reflecting the strong influence of precursor movements, histories of the group's activities often extend several years before the formal establishment of AQAP. Thus, reports of AQAP's activities often include actions carried out by al-Qaeda in Saudi Arabia (see, for example, BBC News 2015) likewise a February 2006 jailbreak that allowed senior jihadi leaders to escape a high security prison in Sanaa is often identified as the defining moment in establishing the current incarnation of AQAP (Counsel on Foreign Relations, 2015).

the purposes of local recruiting, Ansar al-Sharia is widely regarded as being an alias for AQAP (United States Department of State, 2012).

Along with a more local message and a campaign of public service provision, Ansar al-Shariah reportedly recruited new members with promises of salaries and equipment (Swift, 2012). The result was a greater successes in making inroads not only among the underemployed and impoverished Sunni communities, but also access to tribal elders who could reportedly draw on Ansar al-Shariah fighters to provide manpower for public works projects (Swift, 2012). In keeping with the expectation that local recruits would be primarily invested in the local conflict, many of these fighters “have deployed exclusively for an insurgency against the Yemeni government” (Human Rights Watch, 2013, 14).

In 2011 and 2012 Yemeni military and political institutions imploded. The subsequent political and military distractions allowed AQAP to take and hold territory in the southern Abyan and Shabwa governorates, engaging with government force, and even approaching Yemen’s capital, Sanaa.⁶ In 2012 the White House counter-terror advisor John Brennan estimated AQAP’s membership base as over thousand fighters (Cruickshank, 2013; Johnsen, 2013).

By the end of 2013, the United States State Department’s estimate of AQAP’s strength had increased threefold relative to 2009, and the group was credited with nearly a thousand members. AQAP was not the only local militia to capitalize on the weakness of the Yemeni state: the Houthi movement, an association of militias that champion the cause of the Zaidi Shia minority, consolidated their power and influence. The movement subsequently took over the capital of Sanaa in 2014 to pressure—and then replace—the transitional government (BBC News, 2016). The rise of a Shia militia widely perceived to be backed by Iran triggered Saudi Arabia and other Sunni Arab to mobilize to restore the Yemeni government under President Abdrabbuh Mansour Hadi.

⁶The Yemeni state reclaimed the lost urban areas during the following year, but AQAP has continued fighting in southern provinces such as Hadramawt (Sharp, 2015, 9)

Yemen's civil war officially began in March, 2015, the following month AQAP captured al-Mukalla, Yemen's fifth-largest city. In an action seen as an indicator of their new domestic strategy, AQAP quickly appointed local forces to administer the city (Hubbard, 2015). In doing so, the organization deliberately rebranded itself under the name "Sons of Hadhramaut," to underscore their local connections. According to a self-described al-Qaeda member interviewed by *The New York Times*, the decision to use a name with local references was a deliberate attempt to undercut the ability of American-backed Yemeni forces to push them out of newly seized territory. In the aftermath of their advance on al-Mukalla, a self-described Qaeda member reported that the name selection was a deliberate attempt to give the impression of local connections (Hubbard, 2015). By 2016, international media organizations were crediting AQAP with a "mini-state" in Yemen's southeast (Yara Bayoumy and Ghobari, 2016).

Two additional trends accelerated AQAP's newfound ability to rally local supporters: revenge for civilian casualties of counter-terrorism operations and internal balancing against the Houthi militias. By alienating the population, drone operations are reported to have had the effect of drawing otherwise-pragmatic tribes closer to the jihadi militant group and seeding widespread hatred of the United States. Thus, without considering a resilience-autonomy trade-off, we should expect to see AQAP capitalize anti-American rhetoric to capitalize on ill-will towards the United States, particularly with spikes just after major drone incidents, by pushing an anti-American jihadi ideology.

Fighters that join AQAP for revenge may have no particular ideological affinity for the movement, but are responding to perceived threats to their local interests and identities (Johnsen, 2013). This dynamic underlies the broader logic described in Kilcullen (2009), who traced how transnational revolutionary movements embedded themselves within local communities. Some of these fighters, who saw AQAP membership as an avenue to avenge relatives killed by drone strikes, were reportedly so motivated by local concerns that, by late 2013, an AQAP commander in the south-east of the country was complaining that his fighters were so in-

sufficiently ideologically motivated that they neglected basic religious obligations (Muslimi, 2014).

As well, the rise of a Shia Zaidi insurgency associated with the Houthi movement allowed al-Qaeda to better integrate with the local tribes by making sectarian identity increasingly salient, driving the Sunni tribes closer to the Sunni jihadi AQAP (Campbell, 2015; Hubbard, 2015; Worth, 2015). By mid-2015, reporting from Yemen indicated that AQAP was able to use the Houthi threat to Sunni interests in order to forge the tribal alliances that eluded them in 2009 (Al-Batati and Fahim, 2015; Hubbard, 2015). Summarizing the new accessibility to the tribes that AQAP enjoyed, a Sunni militiamen observed: “Even if al-Qaeda and I have disagreements, if we are fighting in the same trench against the Houthis, he is my brother” (Worth, 2015). At the same time, Yemeni and Saudi military preoccupation with the Houthi uprising deflected state resources thereby allowing AQAP to expand their territorial reach (Yara Bayoumy and Ghobari, 2016). In these areas, AQAP has sought to publicize social service provision and pragmatic governance to reinforce support among the communities that they control (Yara Bayoumy and Ghobari, 2016).

Empirical Implications and Quantitative Analysis

The theory of grassroots driven transformation presented above suggests that increased recruitment from among Yemeni tribes leads to shifting organizational priorities. After exogenous events boosted the group’s ability to leverage local grievances to recruit among, and become more deeply tied to, Sunni tribes, AQAP should develop a large internal base for whom security and political developments in Yemen are more relevant than abstract global jihadi revolution.

An ideal quantitative test of the theory would draw on micro-level recruitment and operations data that is sufficiently rich as to be able to identify actors, target priorities, and tactics. However, existing conflict data sets emphasize accuracy of event counts rather than richness

of features for each event. One strategy to generate data that can suggest a transformation of focus is to treat news texts as a source of feature-rich data that encode a combination of observed behavior of local actors as well as the views of regional experts.

Under the assumption that reporting tracks behavior, clustering techniques should be able to highlight on-the-ground changes in group behavior and perception. In particular, a growing interest in local conflicts within AQAP's base should induce the group to behave more as a local Yemeni Sunni militia than as a transnational jihadi vanguard.

For example, a June 28, 2011 story from *Xinhua* news describing an al-Qaeda operation highlights the target of an al-Qaeda offensive (a military hospital), remarks on a distinctive weapon choice (heavy machine guns).

“Al-Qaida militants carried out offensive attacks targeting the 25th Mechanized Brigade in the east of Zinjibar city, which aroused heavy clashes, leaving three soldiers and eight al-Qaida militants killed in addition to injuring dozens of others from both sides, the official told Xinhua, who asked to remain anonymous. Fierce battles are still ongoing around the military brigade, which was surrounded by the militants, the official said. Al-Qaida group was trying to bring down the military brigade by using heavy machine guns, he added. Meanwhile, a local medic at the Basuhib military hospital in Aden said that dozens of injured soldiers were receiving treatment from the clashes...”

Similarly, a September 22, 2013 article from Agence France Presse about activities attributed to Ansar al-Shariah likewise identified the target and tactics of the attack, and indicated that the method of assassination has been a frequent strategy of Ansar al-Shariah.

“ADEN: Gunmen have shot dead a Yemeni man in the south of the country because they suspected he was homosexual, police said yesterday, in the sixth such killing this year. One of two men on a motorbike opened fire at the man

in his 20s outside his house in Huta, the capital of Lahij province, killing him on the spot. Police said the attackers, presumed Islamist militants, escaped. Similar murders in the nation’s provinces of Abyan and Aden have been blamed on an al-Qa’ida-affiliated group, Ansar al-Sharia. For the past year, Ansar al-Shariah has imposed Islamic law on areas of Abyan where it still holds sway. Its so-called courts have condemned to death several people. Others have had hands amputated after being ‘convicted’ of theft. After an army offensive in May last year ousted the militants from areas they controlled, they holed up in mountainous regions of south and southeast Yemen.”

Notably, in the two articles excerpted above—which were selected randomly from the corpus—the two Sunni groups are presented as operating in different spheres. AQAP as a strong military threat, using heavy weaponry against a hard military target, with Ansar al-Shariah trying to impose local administration and social behaviors via a campaign of individual-level violence and assassination.

The expectation is that if AQAP’s local membership induces a local focus, reports on the activities of AQAP and Ansar al-Shariah should be sufficiently similar as to make news stories difficult to distinguish from each other. Houthi militants are included as a point of comparison: if inferring targets, tactics, and priorities via news reporting is viable, the Sunni groups should be distinguishable from the Shia movement.

Data and Analysis: News Corpus

To assess similarities of behavior attributed to AQAP and Ansar al-Shariah, I pulled all media reports of violent activity from Yemen for 2009-mid 2015 from the ICEWS database. This generated 10818 stories, covering November 1993- January 2015. For tractability, I randomly

sampled 1,772 stories of violent activities, which were then hand-coded for primary actor.⁷ To tighten the focus on whether AQAP is distinguishable from Ansar al-Shariah, the stories were filtered to 566 stories that reported activities attributed to al-Qaeda in the Arabian Peninsula, Ansar al-Sharia, and the Houthi militia.⁸

The stories were randomly divided into training and test sets at a .67-33% ratio yielding a training set of 432 news stories and a test set of 144 new stories with publication dates ranging from October 30, 2002 through January 3, 2015. The temporal distribution of the news articles and actor labels can be seen in 1.

In order to convert the news articles into data for classification, I tokenized each story and removed numbers, standard English stopwords, whitespace, and stray HTML markup. I additionally removed a custom list of stopwords that strongly signal the group, such as variations on the group name and signifiers of sectarian identity. Word frequency was normalized via term frequency-inverse document frequency (tf-idf), producing a pair of tf-idf matrices, from which I took the intersection of features (*i.e.* words).⁹ This generated a set of 2,222 “features” for classification in the texts.

An initial visualization of these news articles can be accomplished using t-Distributed Stochastic Neighbor Embedding (tSNE). Essentially, tSNE allows visualization of high dimensional data into two or three dimensions by converting distance between data points (here: stories) into conditional probabilities of similarity based clustering under an abstract Gaussian distribution of all data points (Maaten and Hinton, 2008). Since being introduced in 2008, tSNE a widespread tool for visualizing similarities in high-dimensional data. Figure 2 presents tSNE clustering for all stories from 2011 through 2014, after Ansar al-Shariah was established.

⁷Approximately 15% of the data was coded by workers on Mechanical Turk, the rest was coded by the author.

⁸Activities attributed to the Yemeni government would introduce considerable noise and, as the classification is focused on predicting militant authorship, were removed as the noise would outweigh the insight of the extra information.

⁹As Denny and Spirling (2018) observe, decisions in processing can have significant effects on the outcome. In this case, I minimally process the texts as I am looking for potentially subtle differences in the presentation of the groups, rather than thematic similarities among the texts.

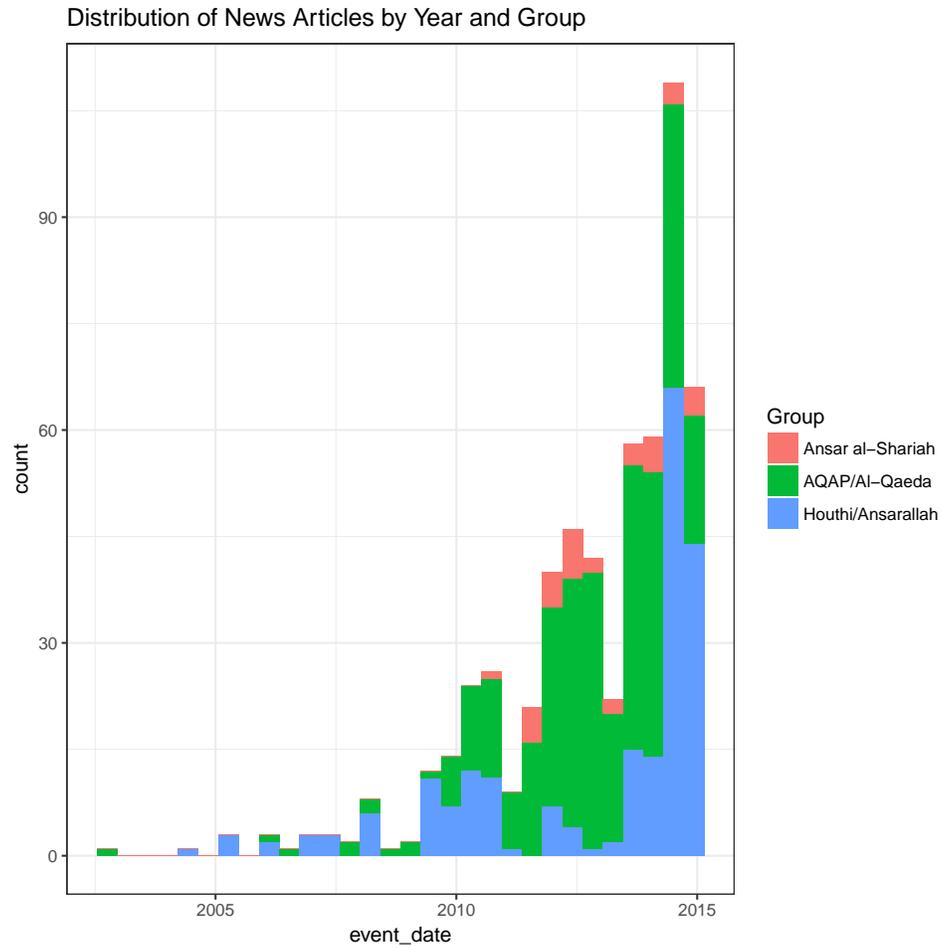


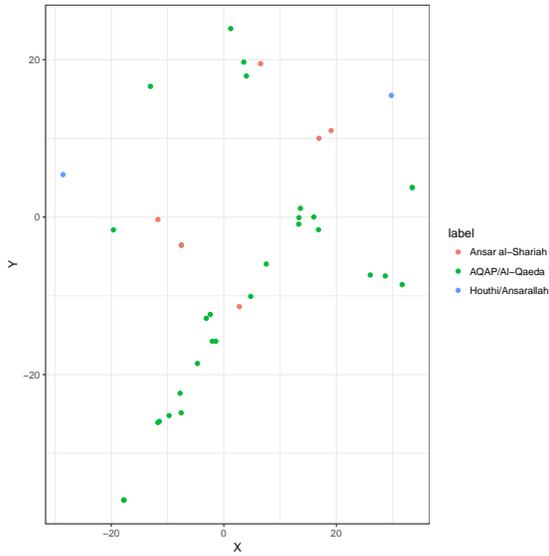
Figure 1: Distribution of News Articles

These indicate a consistent separation between the Sunni and Shia groups and moderate separation between AQAP and Ansar al-Shariah.¹⁰

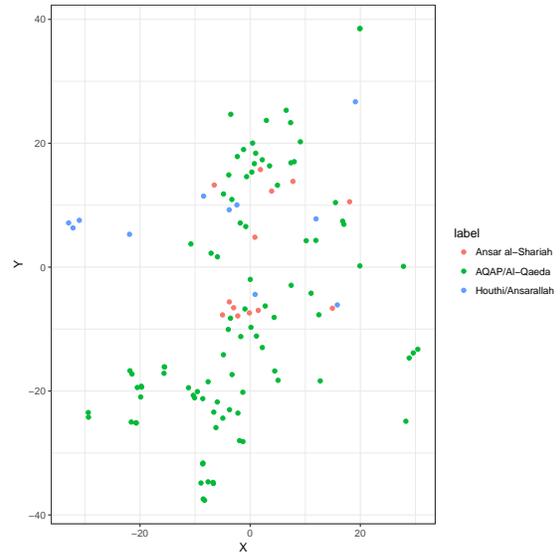
To more precisely analyze the distinctiveness of reporting about the local Ansar al-Shariah branch and AQAP, I classified news articles using a random forest classifier, using the `randomForest()` method from the R package `randomForest`, and a support vector machine using the `ksvm()` method from the `kernlab` R package. In each case, the classifier’s predictions were assessed against hand-coded labels. The results of these supervised approaches reinforce broad outline of the tSNE visualization: within the test set, the Sunni and Shia groups are relatively separate and Ansar al-Shariah stories are consistently predicted as being associated with AQAP. I use three approaches, all of which produce generally the same findings, to increase confidence in the validity of the results.

The first classification algorithm was a random forest, an approach that is particularly effective for high-dimensional data and which has been widely used to analyze text data (Beauchamp, 2017; Jones and Linder, 2016; Muchlinski et al., 2015; Siroky et al., 2009; Spirling, 2012). The random forest analysis is consistent with the tSNE visualization in that articles about AQAP and the Houthi insurgency separate relatively cleanly. This separation is shown in 3, which uses a principle component analysis implementation from the *edarf* package for R to plot proximity of stories, as measured by the proportion of times that individual stories are in the same terminal node (Jones and Linder, 2016). However, the random forest model consistently misclassifies Ansar al-Shariah stories as belonging to al-Qaeda in the Arabian Peninsula. 1 shows the confusion matrix for the random forest model, indicating that any differences in reporting on the activities of al-Qaeda in the Arabian Peninsula and Ansar

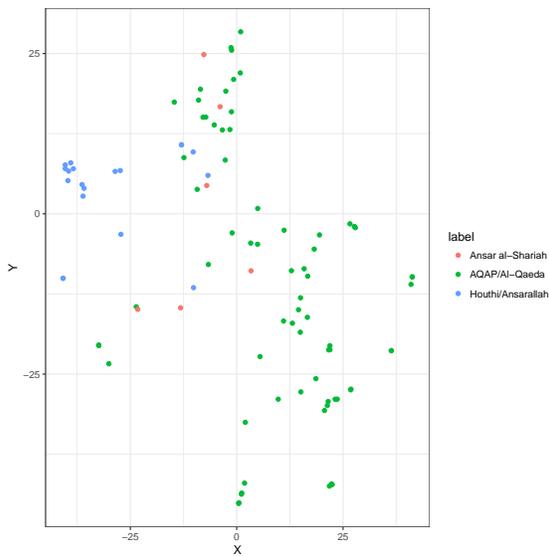
¹⁰The visualization presented here was generated by running the tSNE algorithm on for 5,000 iterations on the pooled data. The perplexity hyperparameter presented below was selected after grid sweeping from 5-50, at intervals of 5. To address concerns that the observed clustering is random noise or driven by a specific initialization, the clustering was carried out with the same specifications on different machines. While adjusting the hyperparameters changes the exact outcome, as expected from a probabilistic approach to summarizing structure in complex high dimensional data, the conclusions are broadly consistent across specifications. Additional discussion of parameterization choices for the tSNE, random forest, SVM, and STM models are discussed in an online Technical Appendix.



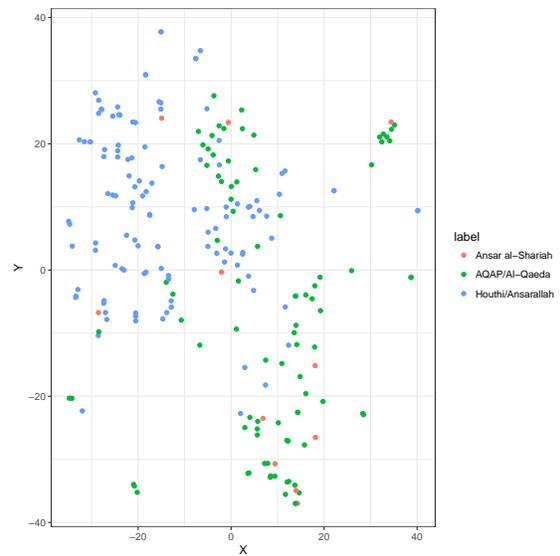
2011



2012



2013



2014

Figure 2: tSNE Clustering, 2011-2014

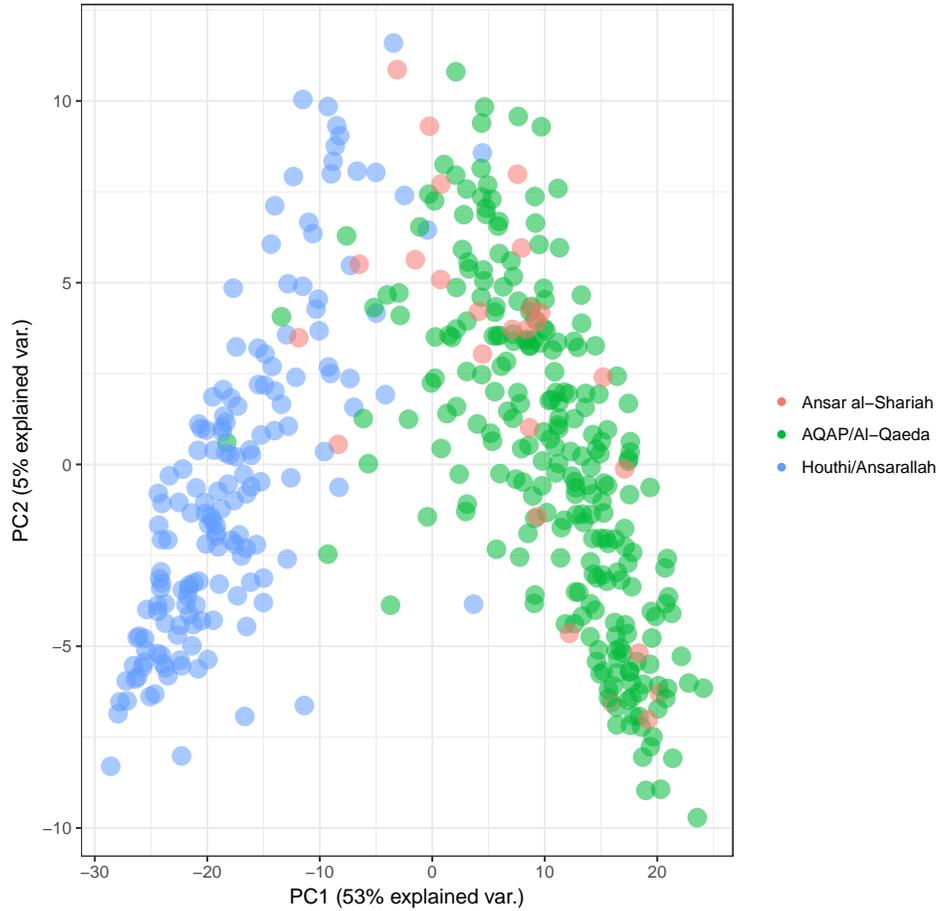


Figure 3: PCA Visualization of Group Classification

al-Shariah are overwhelmed by the differences between the two Sunni groups and the Houthi group.

	Ansar al-Shariah	AQAP/Al-Qaeda	Houthi/Ansarallah	class.error
Ansar al-Shariah	0.00	24.00	2.00	1.00
AQAP/Al-Qaeda	4.00	240.00	3.00	0.03
Houthi/Ansarallah	0.00	7.00	152.00	0.04

Table 1: Random Forest Confusion Matrix

One advantage of the random forest approach for text classification is that as the features are also terms in the document, evaluating the most important variables may produce readily-interpretable insights into the data. The fifteen most important words for the random forest

15 Most Important Words For Story Classification

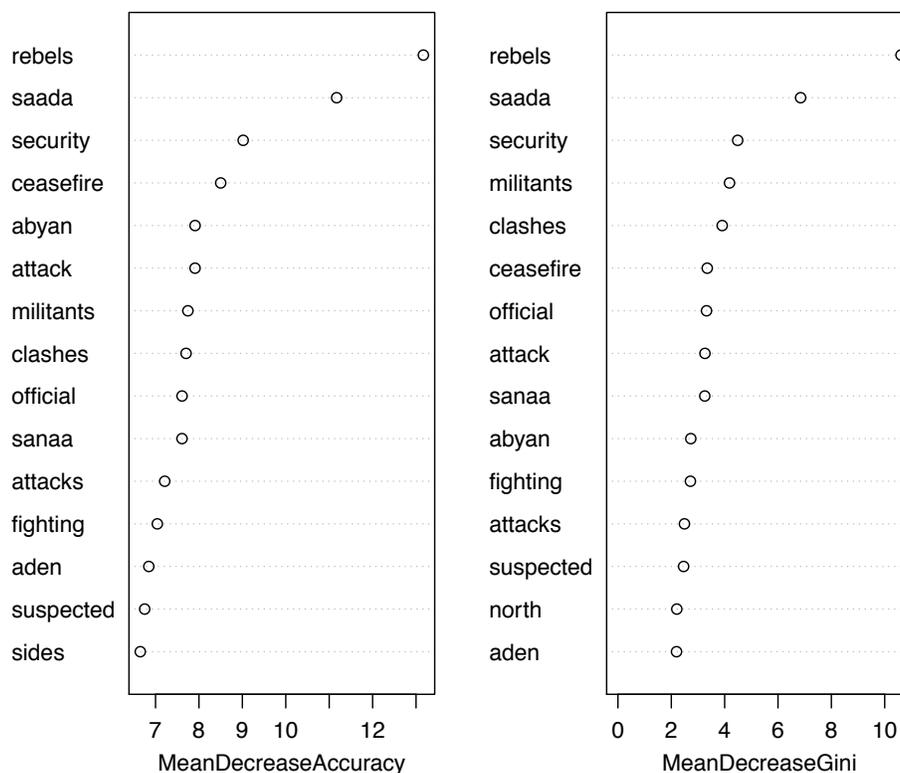


Figure 4: Important Words for Classification

classification—after removing stopwords that describe or name the active group— suggest that reason for the relatively clean split across the Sunni and Shia movements is simply in journalist framing. Houthis are consistently described as “rebels” while the Sunni fighters are frequently presented as “militants.” From there, words that describe the location of operations and military occupations are, unsurprisingly, important classifiers: Aden, Saada and Sanaa are regions associated with Houthi territorial gains, while Abyan is more closely linked to AQAP and Ansar al-Shariah activities.

A second classification, using a support vector machine, generated similar results. As in the random forest classification, Ansar al-Shariah stories are consistently identified with AQAP.

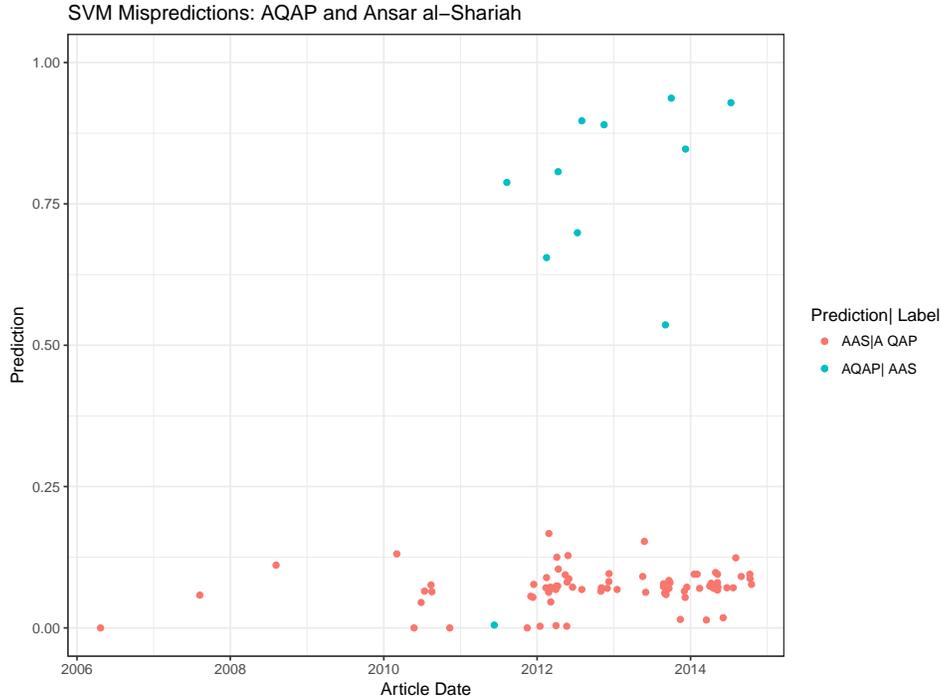


Figure 5: SVM (Mis)Predictions over time

Indeed, as with the random forest classification, the SVM fails to predict any stories as being Ansar al-Shariah. Figure 5 provides a closer look at whether the SVM’s confidence in whether to assign stories to Ansar al-Shariah or AQAP change over time. Ideally, support for the transformation theory would indicate the SVM assigning more and more weight to the Ansar al-Shariah label for news stories about AQAP activities. The plot depicts the predicted probability that an article would be assigned to AQAP given that the story was labeled as being an Ansar al-Shariah action (blue) or that a story would be labeled as an Ansar al-Shariah action given that the label was AQAP (red). Throughout the corpus, the SVM model consistently assigns a low probability, to Ansar al-Shariah, remaining relatively stable at a predicted probability of the Ansar al-Shariah label at about $p = .015$.

Although the results of the clustering indicate convergence between coverage of the activities of al-Qaeda in the Arabian Peninsula and Ansar al-Sharia, these methods don’t reveal whether AQAP is like Ansar al-Shariah, whether Ansar al-Shariah acts like AQAP, or

whether there are simply so few articles that directly attribute activities to Ansar al-Shariah that the differences between the Sunni and Shia groups overwhelm within-Sunni differences. The two most straightforward avenues to keeping these members are to socialize the members into alignment with with AQAP's goals or to adjust AQAP to fit the interests of the new members. Both predict a convergence between AQAP and their local offshoot, Ansar al-Sharia, but imply different directions of movement: if the former dominates, the local branch should gain a greater international focus as local actors are socialized into the transnational jihadi ideology. If the later avenue dominates, AQAP itself should gravitate towards parochial Yemeni concerns over transnational jihadi concepts.

As an illustration, consider interaction between AQAP and the Houthi militia. A narrowly Yemeni political and religious movement, the Houthis have little resonance with a transnational jihadi cause beyond the general contours of the Sunni-Shia sectarian conflict. If the local Sunni community is converging towards AQAP's transnational agenda, AQAP could present their combat with the militias within an established, transnational, anti-Shia framing. Moreover AQAP could link the conflict to other hotspots in jihadi sectarian battlegrounds, particularly in Syria and Pakistan. Conversely, if the influence of the new tribal grassroots is driving AQAP to adopt more local preferences, the group could be expected to move away from a transnational Sunni-Shia framing and towards specifically Yemeni facets of the Houthi movement.

Moreover, the news classification approach does not directly measure activity, but rather what is covered and how it is attributed. Third-party discussions about local events are sensitive to idiosyncrasies introduced at the descriptive level. So, for example, we should expect over-attribution to AQAP, given that the al-Qaeda name is of high salience. Additionally, we should expect under attribution to Ansar al-Shariah, as many journalists have associated AQAP and Ansar al-Shariah.¹¹

¹¹For example, 11 of the stories in the corpus carried language explicitly blur the distinction between AQAP and Ansar al-Shariah, such as the following quote from Xinhua News stating: Al-Qaida in the

Finally, the consistent increase in uncertainty about story classification across the years of the news corpus is likewise ambiguous. The uncertainty largely derives from mispredicting AQAP-Houthi pairs. This may be a result of journalists increasingly including discussion of the security situation in the country, thus dedicating article space to similar terminology. Alternately, the conflation of AQAP and Houthi stories may pick up a tendency of writers to frame AQAP mobilization using the rebellion frame that dominates coverage of Houthi activities.

In order to identify the origins of the convergence and to evaluate differences in messaging in a more controlled domain, I turn to communiques and self-presentation as a second data source.

Topic Modeling To Identify Direction of Convergence

A second approach to using text data as a window into opaque organizations leverages topic models, applied to materials issued by the groups themselves. These methods have gained wide adoption, and notably in other difficult-to-reach domains. organizations, such as understanding how military elites influenced Russian foreign and defense policy under Putin’s leadership (Stewart and Zhukov, 2009) and how signals from leaders influence Russian policy makers (Baturu and Mikhaylov, 2013). Within scholarship on the transnational jihadi movement, Nielsen (2014, 2013) has used statistical text modeling techniques to investigate how career and educational networks influence the adoption of jihadi rhetoric.

In order to assess the apparent convergence in AQAP and Ansar al-Shariah activity, I use the structural topic model, to characterize propaganda messages by uncovering latent topics within the document. I interpret a rise in in Yemeni specific topics and decreases in transnational and pan-jihadi rhetoric as suggestive of increased influence of a parochial interest base.

The structural topic model allows researchers to incorporate document-level information as

Arabian Peninsula, also known locally as Ansar al-Sharia.

a covariate that is statistically related to topic prevalence (Roberts, Stewart and Tingley, 2014), which permits modeling group-level changes in attention over time. Existing work has applied the STM to a variety of corpora comprised of short and moderate-length documents (Roberts et al., 2014), such as open-ended surveys (Tingley, 2017), social media messages (Bail, 2016), and deepweb forum posts (Munksgaard and Demant, 2016).

Communique Analysis

I present three models: the first of which illustrates a general localizing trend in the self-presentation of AQAP and Ansar al-Shariah. The second highlights changes in AQAP messaging following events, such as the introduction of drone warfare to Yemen and the deaths of Usama bin Laden and Anwar al-Awlaki, that can be expected to have an influence on AQAP's self-presentation. In order to account for the possibility that observed shifts in messaging were driven by top-down directions from al-Qaeda leadership or general trends in the jihadi environment, the third model contrasts the rhetoric of AQAP with materials released by as-Sahab.¹²

I focus on media released online to jihadi media platforms and outlets. Officially-released propaganda is an illustrative area to demonstrate changes in organizational preferences because we can expect these documents to be an area in which the group can choose, within ideological and alliance constraints, how to best represent their movement. Moreover, since 2011, online platforms have been a "major means of communication" within Yemen (Carpico, 2014, 33), so issuing releases online means that the same document can be consumed domestically as well as internationally. The potential for their messages to be consumed by both a domestic and international audience can be expected to discourage AQAP from differentiating the priorities that they signal in propaganda media from their local priorities. The documents included in this study were aggregated and translated by an American-based

¹²As-Sahab is closely associated with the core of al-Qaeda's "central" leadership, which is why their releases were chosen for the comparison.

research organization, the SITE Intelligence Group, and represent a large sample of media released by AQAP.¹³ In analyzing AQAP’s releases, I use a corpus of 875 documents associated with al-Qaeda in the Arabian Peninsula’s propaganda output from June 18, 2004 through September 18, 2016.¹⁴ The corpus is predominantly comprised of English translations of documents originally released in Arabic.¹⁵

Preprocessing removed words that occurred in fewer than two or more 70% of the documents in the corpus.¹⁶

The three models are based on a corpus of 1353 documents, spanning October 25, 2005 through September 21, 2016. Approximately 500 documents are associated with as-Sahab. A histogram of the distribution of can be seen in 6.

Model One: Comparison of Ansar al-Shariah and AQAP Branding

To provide more specific indicators of how AQAP’s messaging differed from communiqués released under the Ansar al-Shariah brand, I estimated I estimated a 18-topic model on 875

¹³The SITE Intelligence Group collects and translates official and unofficial jihadi media releases, with near real-time coverage of prominent online distribution sites. As no official universal archive of AQAP’s releases exists in the public domain, the corpus is necessarily a sample of the releases. However, as the SITE Intelligence Group is an internationally-focused monitoring organization that specializes collecting al-Qaeda and jihadi propaganda, we can expect that any selection effect would prioritize AQAP documents aimed at an international audience and with an internationalist message. Thus, any selection effect would be expected to increase the percentage of documents aimed at an international audience, and which could be expected to prioritize an internationalist message. Likewise, readers may worry that the relatively low internet penetration rate of Yemen may imply a largely international audience for their propaganda. However, if this the case, it should bias my results against finding a drive towards local self-presentation.

¹⁴The corpus includes content from both the current—*i.e.* post-2006— AQAP and a predecessor organization of the same name that is occasionally referred to as al-Qaeda in Saudi Arabia. As I have argued in the above section, the post-2009 iteration of AQAP grew directly from the latter, and the post-2009 AQAP leadership actively sought to present the organizations as a linked entity. Therefore I allow the AQAP corpus to accommodate both organizations as this permits examination of the impact of the upswing in drone casualties in 2009 and 2010.

¹⁵A small subset of the documents, such as selected articles from *Inspire Magazine*, were originally released in English. However, these make up a very small portion of the corpus. The documents were all translated by the SITE Intelligence group, which has internal procedures to ensure consistent translation in style and tone.

¹⁶In the AQAP corpus, there was no change to the number of tokens in the corpus for an upper bound threshold between 70-95%. I evaluated coherence and exclusivity at an upper threshold of 50%, but did not find results that would suggest either a coherence or exclusivity benefit from the additional reduction in corpus size.

AQAP statements released from 2004-2016. The primary covariate was a binary variable for whether the communique was branded by AQAP or Ansar al-Shariah. A second covariate located the document in time, by counting elapsed days between the date of the first document in the corpus and the release of the specific document's translation. Expected topic proportions can be seen in Figure 10 and the substantively interesting topics are summarized in Figure 11.¹⁷

Each topic is summarized by its FREX words, which are words or tokens that are highly associated with the topic but relatively unlikely elsewhere in the corpus. Finally, as a post-processing step to enhance interpretability of the topic presentation, I clustered the topics into four thematic groupings: locally-focused war reports, discussions about and threads of clandestine operations, topics with a transnational jihadi sentiment and goals, and jihadi-associated descriptors. After estimating the model, I grouped topics into thematically similar clusters.¹⁸ Topics are generally similar those estimated in the first STM model, although the re-estimation with covariates generated a few notable differences. For example, although both models identify a topic linked to individual attacks, the FREX words for the second model places a greater emphasis on recruitment for emigration to Yemen than in the corresponding topic in the first model.

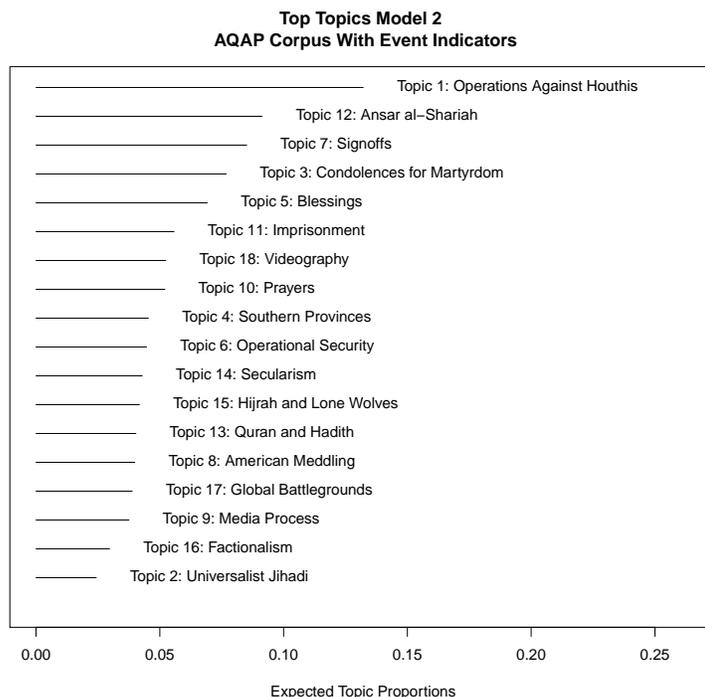
Figure 9 shows the change in topic prevalence for each of the 18 topics for material issued under the Ansar al-Shariah name. The results underscore the attempt to position Ansar al-Shariah, as topics related to local battleground significantly grow in prevalence. Correspondingly, several of the broadly transnationalist topics become significantly less common.

Among these less common themes are a number of telling topics, including references imprisoned jihadists and incarcerated Muslim women. The latter of which is often a jihadi rallying cry to mobilize far-flung audiences to become emotionally invested in otherwise-local man-

¹⁷The figure omits three topics that are associated with the editing and production of the releases.

¹⁸The three remaining topics relate to the actual communique document and are less interesting as a reflection of AQAP's self-presentation. Words associated with this topic relate to videography, habitual sign-off terms, and transcript production.

Figure 7: Expected Proportion of Topics in Model 2



ifestations of the American War on Terror, such as the American occupation of Iraq and the incarceration of Aafia Siddiqui.¹⁹ Interestingly, the group also decreases their emphasis on religious themes, which would be consistent with downplaying their identity as a revolutionary religious organization in favor of self-presentation as an actor in the domestic social conflict. These results are all the more striking, given the directions featured in the al-Qaeda strategic guidance highlighted in the case introduction but the lack of separation in the news corpus. Thus, while AQAP was explicitly warned against focusing on local conflicts, and appears to have made some effort to separate their al-Qaeda brand from local conflicts, this separation did not extend to local perception of their activities.

¹⁹For a discussion of Siddiqui’s position in jihadi propaganda, see Di Giovanni (2014).

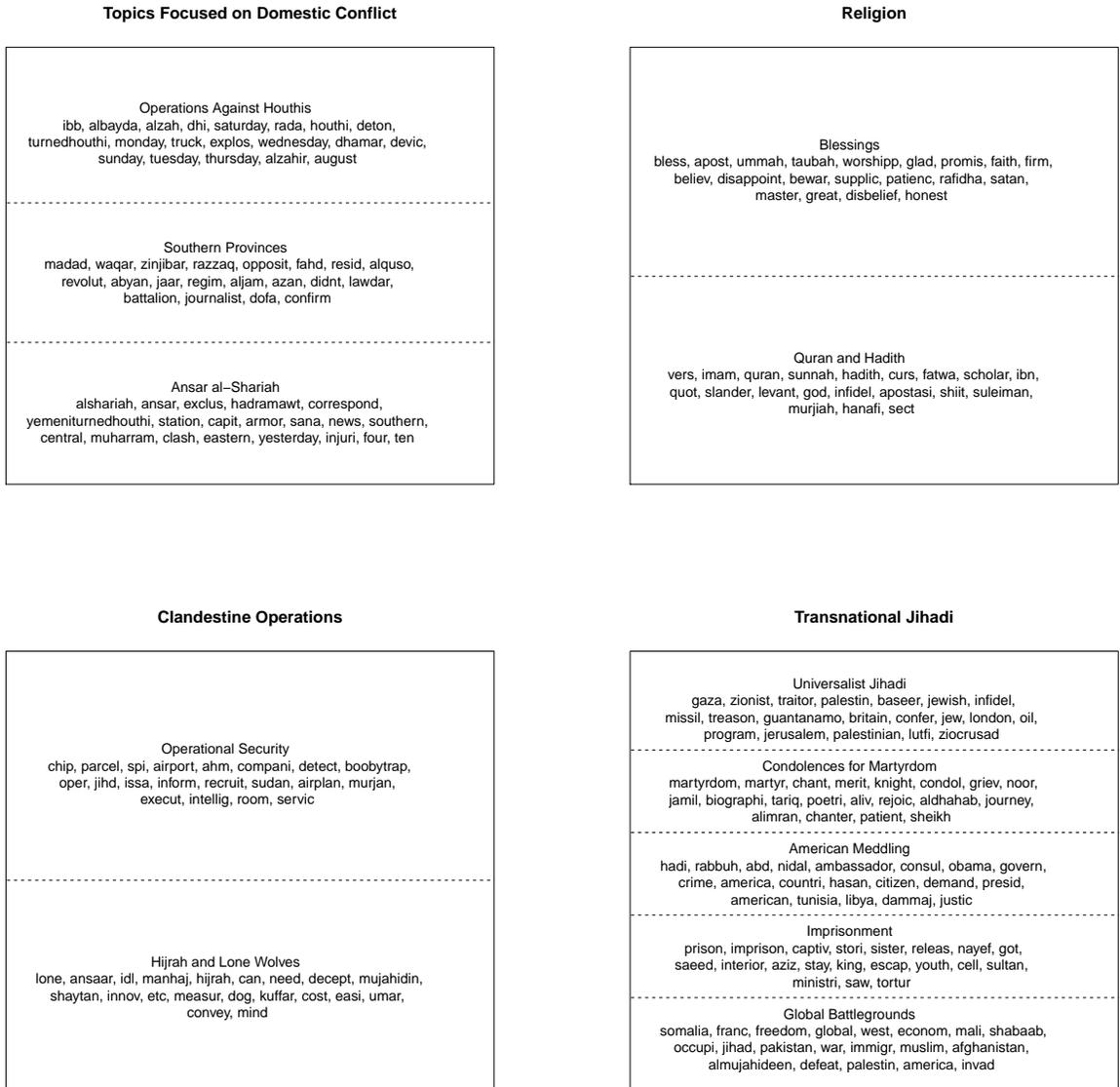


Figure 8: Groupings of Substantive Topics in Event Covariate Model

Figure 9: Covariate Influences on Topic Prevalence

STM Model Two: Evolution of AQAP Messaging

The first model presented the estimated thematic divergence of communiqués associated with AQAP or Ansar al-Shariah, which indicated that although event reporting does not differentiate between the groups, the jihadi group itself did present each wing differently. The second model evaluates trends in the thematic content of AQAP’s messaging. Although general trends alone are unable to directly test the predictions of my theory, changes in topic proportions can indicate the general plausibility of my argument—that increased access to local recruits should drive a localized agenda—in the case of Yemen and suggest a driver of why AQAP and Ansar al-Shariah activities are difficult to correctly categorize.

The bottom-up transformation theory predicts that an influx of local recruits should encourage AQAP to adopt increasingly local priorities, and thus that convergence between AQAP and Ansar al-Shariah should be driven by changes in AQAP messaging. This should manifest in the topic model identifying an increase in AQAP attention paid to domestic fissures, such as the Houthi-Sunni civil war, and a decline of pan-jihadi topics.

As this is a completely separate model from the one presented above, the topics differ slightly in composition from the topics presented in the first model. The topics, summarized in 10 and 11 are generally consistent between the first and second models.

Figure 12 depicts the expected document-level topic proportion for three topics from the “local conflict” cluster. The FREX words for each of these three topics include tactical terms that refer to specific local geography and operations, and thus likely to not be motivated by incitement for (trans)national revolution. I identify these topics according to the content or place names that are emphasized by the FREX words, and name them as “Houthis,” and “Local Targets.” Each of the selected topics appears to suggest a close local focus, as one might find in battlefield reports and communiqués issued to an audience with an interest in local territorial control. Notably, the emphasis on the bellwether Houthi topic has dramatically increased in recent years.

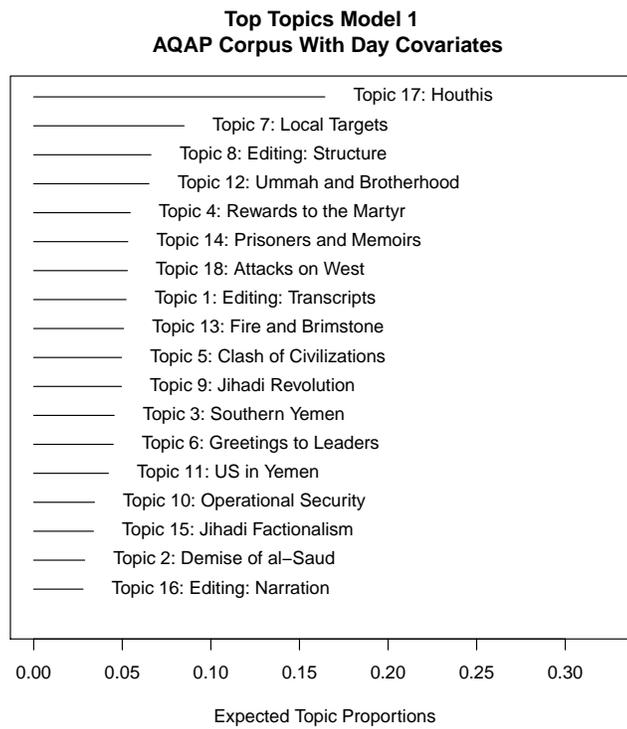


Figure 10: Estimated Topic Proportions in al-Qaeda in the Arabian Peninsula Corpus

Figure 11: Groupings of Substantive Topics in General Trends Model

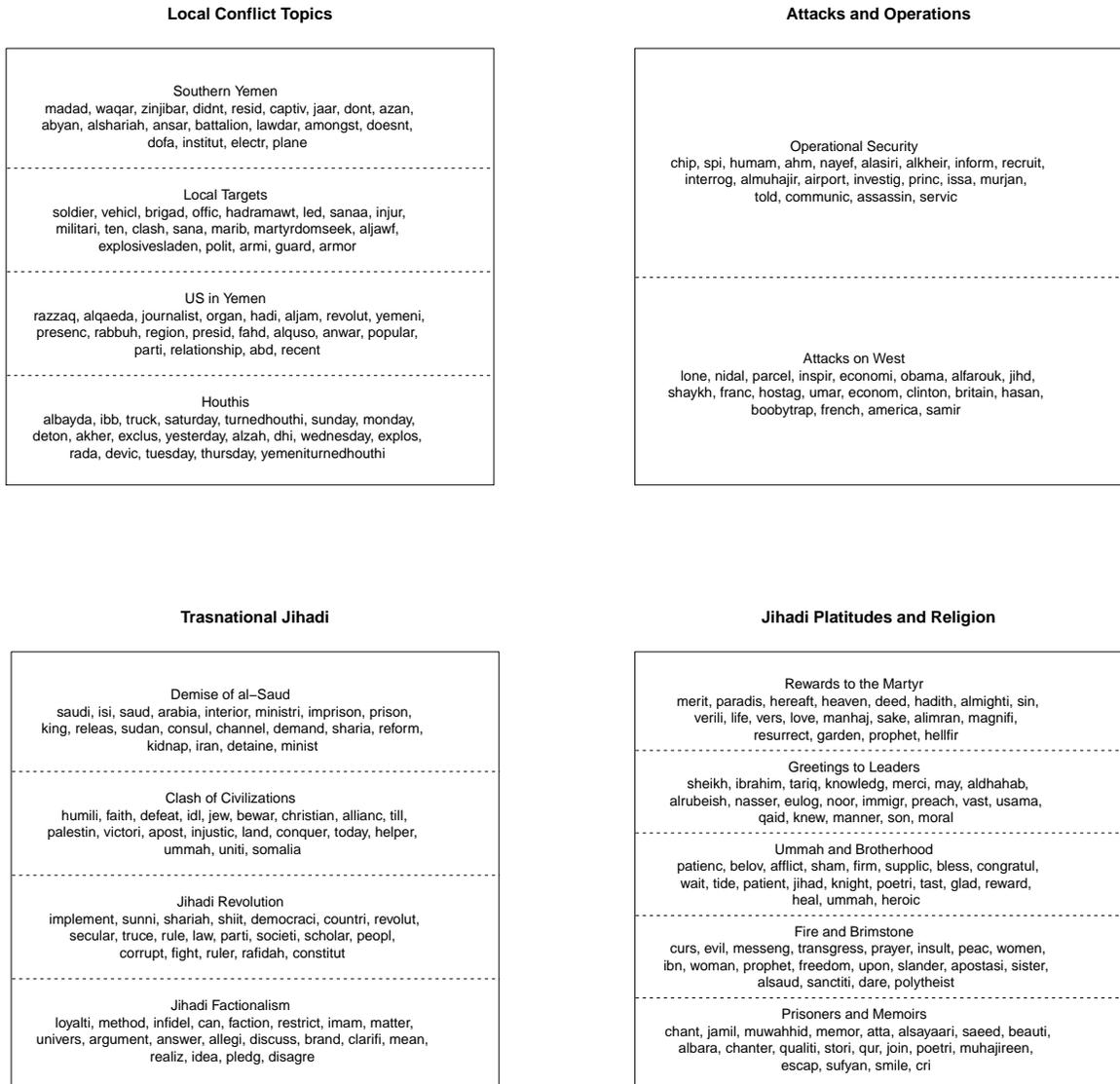
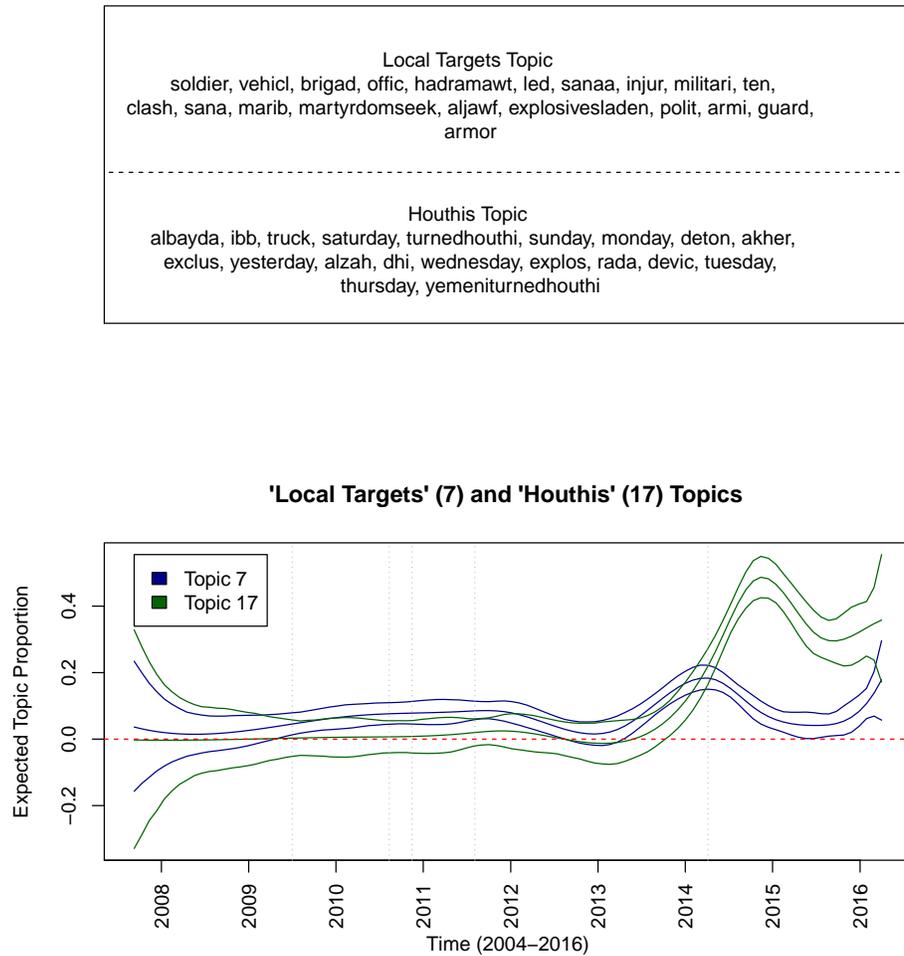


Figure 12: Changes over time to attention dedicated to local topics



The two remaining Yemen-centric topics are topics 3 and 11. Topic 3, labeled as the “Southern Yemen” topic strongly relates to AQAP’s 2011 seizure of territory in the south of the country during the 2011 Revolution. The topic rises in prominence while the organization controlled territory in 2011 and 2012, and then quickly recedes as an expected thematic topic. As such, the topic is a useful indicator that the STM model identified are reflective of underlying events. The third Yemen-related topic, loosely named the “US in Yemen” topic, does not has as clearly interpretable thematic structure as do the other three Yemen-centric topics, and so is not presented here.

A localizing trend is underscored by looking at changes in the expected prevalence of two of the four topics that speak to a transnational jihadi sentiment. Figure 13 depicts the expected proportion of the transnational jihadi topics presented according to time.²⁰ The y-axis represents the expected proportion of each document dedicated to each topic. For context, I added four vertical lines marking important dates identified above. From left to right, the lines represent: the al-Majalah airstrike on December 17, 2009; the start of the Yemeni Revolution on January 27, 2011; the death of Usama bin Laden on May 2, 2011; the end of the first Obama Administration on January 19, 2012; and the Houthi takeover of Sanaa on September 21, 2014.

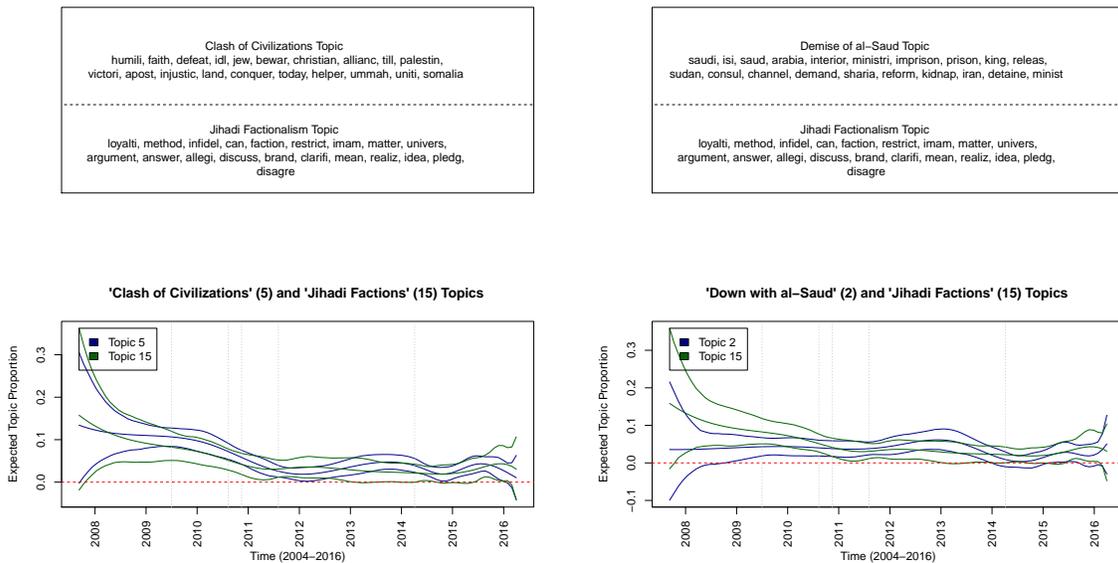
The first topic is dedicated to naming groups that the jihadi worldview considers global enemies of Islam. I term this the “Clash of Civilization” topic, as the FREX words reflect a pervasive jihadi doctrinal focus on fighting a perceived global alliance of Jews and Christians who are attempting to subjugate Muslims. The “Clash of Civilizations” topic begin to decline after about 2009, an important benchmark, as the year saw a number of high-profile drone strikes that caused widespread resentment. This decline may speak to the bottom-up transformation of interest described above: if localizing rhetoric was driven by top-down marketing decisions, and the group turned away from global jihadi branding to capitalize on domestic frustrations, we might expect to see a sharper decline in the topic during 2009.

The second transnational jihadi topic, topic 15, I label as the “Jihadi Factionalism” topic, as the topic is focused on terms that indicate an attempt to navigate conflicts between jihadi communities and other Muslim groups and among jihadi communities themselves. The FREX words for the topic feature words used when presenting the jihadi cause and in attempts to recapture ideological legitimacy. The topic declines throughout 2009 and 2010,

The third topic presented in Figure 13, the “Jihadi Revolution” topic, is centered around concepts used to incite for overthrow of secular governments and implementation of an

²⁰For interpretability, the x-axis is labeled by year. The model was estimated according to the number of days from the start of the document corpus.

Figure 13: Changes over time to attention given to global jihadi topics



Islamic theocracy. Such revolutionary rhetoric is central to the transnational jihadi view of themselves as a vanguard of social and political revolution. The topic declines briefly after the first inflection point, then rises from mid-2010 through mid-2013 before taking a more dramatic downturn at the second point. One reason why the “Jihadi Revolution” topic may not behave as expected of the transnational jihadi topics may be that the 2011 Yemeni Revolution increased the group’s interest in presenting itself as a viable alternative.

The fourth transnational topic in Figure 13 is topic 2, which I have labeled “Demise of al-Saud.” The topic is focused on Saudi-centric themes, with references to Saudi repression of jihadi dissidents along with references to Saudi government officials. The topic is largely stable at an expected prevalence of approximately 5% throughout the time period.

Comparison to As-Sahab

Thus far, the analysis has suggested that local observers generally characterized AQAP and Ansar al-Shariah using similar terminology and framing, but that the Sunni group attempted

to differentiate their own transnational and local brands. However, despite the apparent investment in a local alias, AQAP's communiques become progressively more thematically local. Although these results are generally consistent with the theory's expectation that incorporating an influx of local fighters has generated internal pressure on AQAP to focus on local issues, an alternate explanation could point to broader forces among the transnational community. In particular, 2011, the year of the Yemen Revolution, also saw a series of shockwaves through the jihadi world after Usama bin Laden was killed on May 2, 2011. His second in command, Ayman al-Zawahiri, subsequently assumed leadership of the global al-Qaeda network. In September of that year, an American drone strike killed Anwar al-Awlaki, an American-Yemeni cleric who had been a strident internationalist face for AQAP.

The deaths of bin Laden and Awlaki, and Ayman al-Zawahiri's subsequent ascendancy to the leadership of al-Qaeda, are challenging for the theory developed above. al-Zawahiri has long been associated with an internal al-Qaeda faction that prioritizes focusing on implementing country-specific social and political revolutions in the home countries of al-Qaeda's operatives, rather than engaging in civilizational struggle against the American-lead global system (Miller, 2015, 147). Correspondingly, Awlaki was instrumental in AQAP's efforts to radicalize and mobilize fighters internationally, particularly from the West.

Thus, one possible counter-narrative to the bottom-up transformation argument maintains that the change in al-Qaeda's leadership may have triggered a wider ideological shift that filtered to local branches and that the loss of Awlaki amplified the effect in Yemen. Indeed, if AQAP was following the lead of as-Sahab after 2011, the effect may directly contravene the predictions of the theory as it would suggest that the group's messaging demonstrates a closer adherence to al-Qaeda's direction as their local membership increased.

To differentiate a locally-driven change in Yemeni rhetoric from a centrally-motivated change prompted by changing al-Qaeda strategy, this section presents selected results from a structural topic model that estimated topic prevalence in a corpus of propaganda released by

both AQAP and as-Sahab, al-Qaeda’s central media wing. AQAP’s changing rhetorical style is presented alongside that of the central al-Qaeda media wing to establish that observed changes in AQAP rhetoric are not driven by an underlying pan-jihadi trend. As a separate model estimated on a different set of data—a corpus of AQAP communiqués and releases from as-Sahab—the resulting topics differ from those presented in models one and two.

Figure 14 provides an overview of twenty-nine substantively interesting topics identified by the model.²¹ For ease of interpretation, the topics are clustered into four general categories: those primarily relating to Yemen, topics with strong religious overtones, topics that suggest engagement with global jihadi issues, and topics that address specific countries and battlefields other than Yemen.

Within each cluster, topics are summarized according to the top FREX words. The point estimates of the difference in topic prevalence are presented along with 95% confidence intervals. Positive values indicate a stronger association with AQAP’s corpus, while negative values indicate topics more strongly associated with as-Sahab.

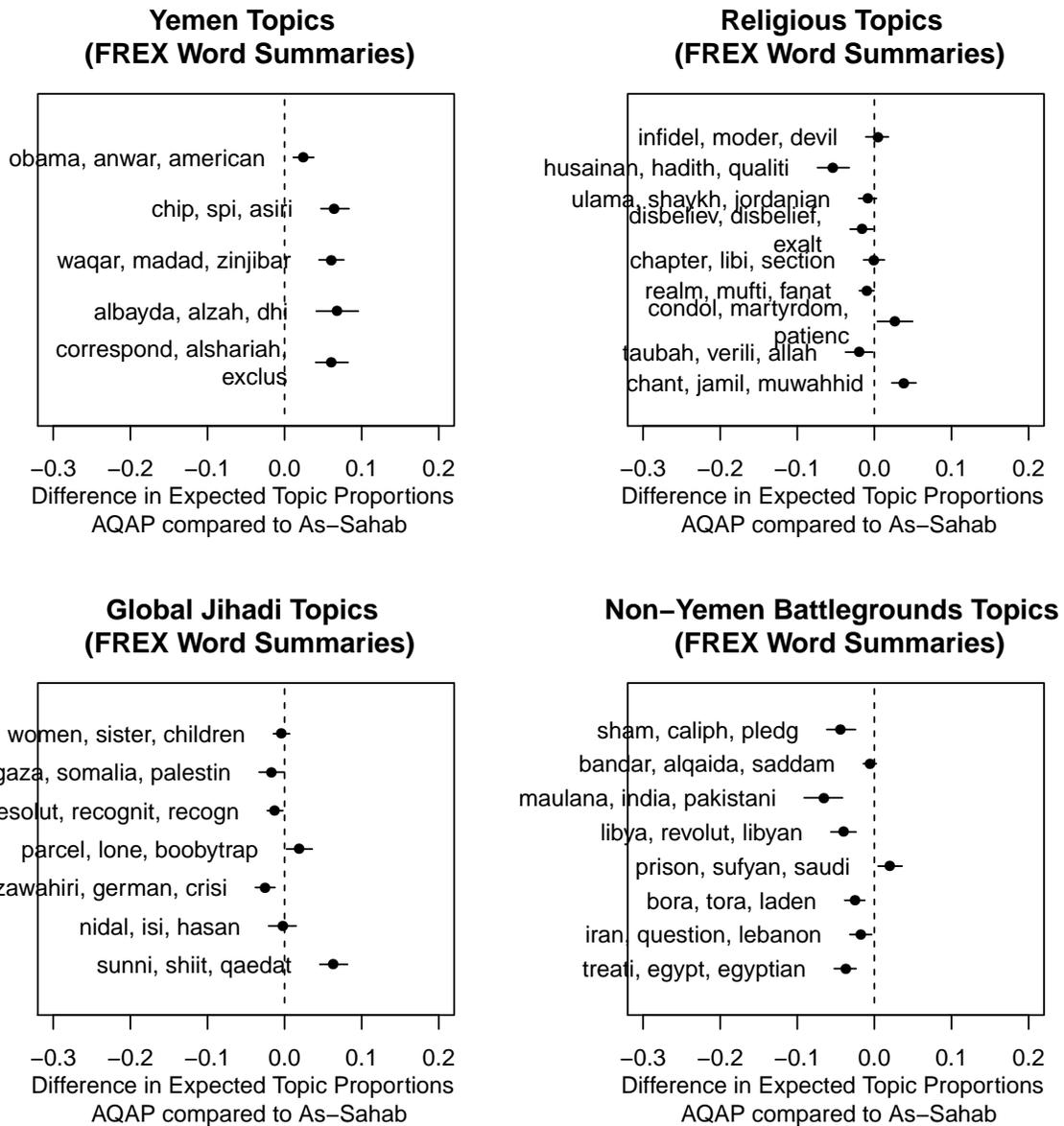
The patterns are generally intuitive: overall AQAP is more likely to talk about Yemen-related topics while topics that discuss other battlegrounds and targets for revolution are more associated with as-Sahab.

Figure 15 highlights three outcomes from the model, which predicted topic prevalence in the corpus as a linear function of time interacted with issuing group.

The first is a topic analogous to both of the “Clash of Civilizations” topics above which I call “Crusaders and Zionists.” This topic is notable for having a particularly transnational jihadi focus; exactly the type of subject that AQAP should cease to discuss if their domestic recruitment is driving a local focus. Indeed we see that although al-Qaeda Central’s rhetoric

²¹Of the five topics excluded from the summary, four were associated with editing and document preparation and the fifth consists of declarations of defiance and intention. FREX words for this last topic include: “know”, “think”, “can”, “see”, “thing”, “happen”, and “now”.

Figure 14: Summary of 34-Topic AQAP and As-Sahab Model



does increasingly feature this topic, AQAP become progressively less and less inclined to use words associated with the topic. Interestingly, this move away from the “Crusaders and Zionists” happens even though at the outset of the dates covered in the corpus, AQAP was more likely to refer to the topic than was as-Sahab.

Similarly, after 2011, as-Sahab documents become more likely to discuss a topic addressing alleged injustices against vulnerable populations such as women and children. The topic, termed, “Defending the Weak,” expresses indignation about alleged crimes against Muslim women and children and is a pervasive component of general jihadi rhetoric. Throughout the time period covered by the corpus, the topic becomes increasingly less prevalent in AQAP documents and progressively more prevalent in As-Sahab releases.

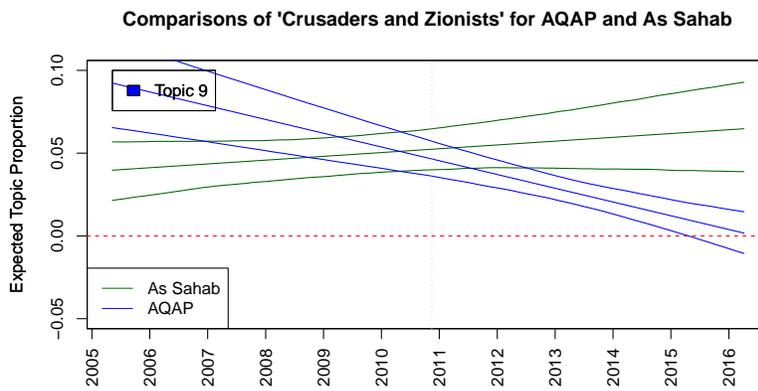
For the third highlighted topic— termed “Strike the US” as the FREX words make specific references to attempted and successful attacks within the United States— the general pattern is again the same. Although the expected prevalence of the topic remains statistically indistinguishable across the two groups, the predicted prevalence of the topic drops slightly for AQAP documents, while trending upwards for content from as-Sahab.

Limitations

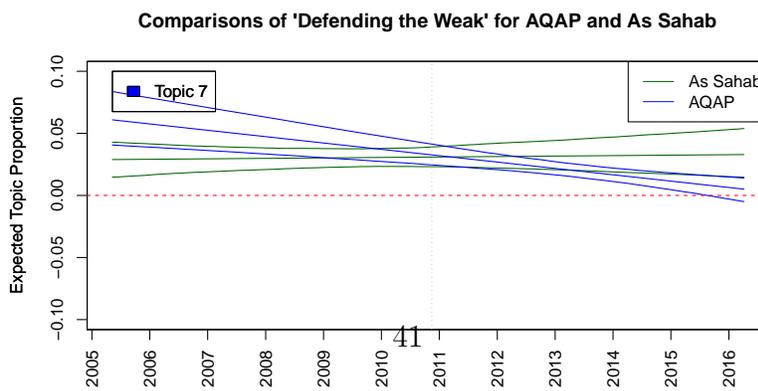
The theory of organizational change that I present above is agnostic about the direction through which group changes influence rhetoric. One possibility is that the organization changes their presentation to appeal to a local population. In this case, tone presages recruitment drives and the ensuing internal changes. Alternatively, rhetorical changes may follow the incorporation of local recruits, as their interests begin to dominate. Although the textual evidence presented above cannot adjudicate between these two alternatives, results can remain agnostic about the exact sequencing of the shifts while still supporting the theory of bottom-up transformation.

Figure 15: AQAP and As-Sahab Divergence

Topic 9:
 gaza, somalia, palestin, everywhere, jerusalem, jew, levant, crusad, traitor,
 invad, defeat, liber, allianc, zionist, campaign, alaqsa, maghreb, agent, land,
 lion



Topic 7:
 women, sister, children, condemn, saud, injustic, crimin, crime, woman,
 corrupt, hurt, sanctiti, dare, univers, free, ruler, document, digniti, gang,
 hous



Topic 25:
 nidal, isi, hasan, barack, corpor, jeremi, samir, takfir, signific, scahil,
 awlaki, khan, qaeda, hood, jesus, anwar, alawlaki, sharia, cruiz, umar

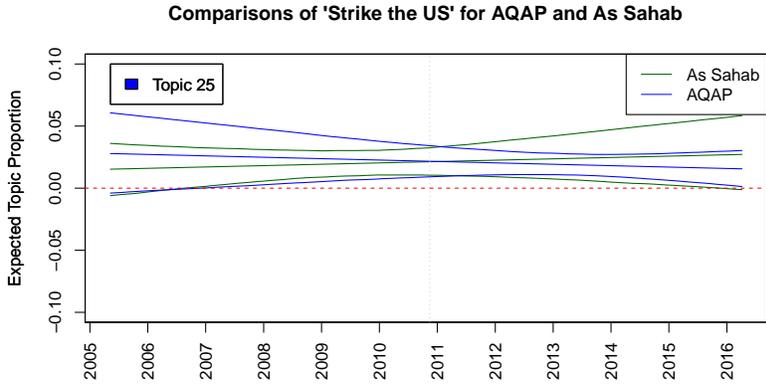


Figure 16: AQAP and As-Sahab Divergence

A second limitation of the analysis presented above is that the text data from which I draw may not be representative of the full scope of al-Qaeda in the Arabian Peninsula's local outreach efforts. First, documents released online may be disproportionately aimed at an international audience. Secondly, as an international research organization, the SITE Intelligence Group might systematically feature internationally-focused media from AQAP. If non-sampled local messages differ significantly from those directed online, then AQAP may have more opportunity to tailor different messages for different audiences. Given the poverty of Yemen and relatively low internet penetration in the country—reported as 15% in 2011, rising to 25% by 2016 and with a 42% smartphone penetration according to a 2016 industry publication (GSMA Intelligence, 2016)—AQAP may view their online audience as largely international. Thus, while scholars have reported that AQAP targets media output at a domestic audience—for example, the *Sada al Malahem* magazine is identified by Yemen analysts such as Phillips (2011) as being an avenue through which AQAP directed specific messages at local tribes—the text-based approach may not capture the full spectrum of their messaging.

However, this bias would make for a harder test for my theory, as the direction of the bias would encourage over-representation of international themes. If AQAP is able to separate their transnational messaging from their local messages and my data disproportionately selects transnational messages, the resulting media should heavily target the transnational jihadi community. As AQAP is clearly identified as the al-Qaeda wing in Yemen, the group has little opportunity to disassociate from the transnational jihadi movement. By adopting the name of the organization, they have committed to the association. Thus, the group is incentivized to speak to an audience that is friendly to al-Qaeda. If this is the case, we should expect to see the topics to which the group is attentive either remain transnational or otherwise closely track the messaging from al-Qaeda Central.

Conclusion And Larger Research Agenda

In this paper, I have presented a theory of grassroots-driven, bottom-up organizational transformation with the specific context of a militant organization faced with the opportunity to incorporate a coveted local population. Future work extends this framework to other types of organizations operating in a contested domain or under threats to organizational survival and models the identified trade-off to analyze when organizational cohesion will diminish and thus suggest when interventions may be most efficacious. Refinement of the theory will evaluate the effect of institutional design on organizational vulnerability to upwards pressure.

The paper makes a methodological contribution, by applying both unsupervised (tSNE and the structural topic model) and supervised (random forest and support vector machine) machine learning tools to two distinct text data sets and in treating the text of event reporting as a source of feature-rich data about the tactical and strategic goals of a local actor.

When there are tight control mechanisms and powerful internal institutions, leaders can ensure that their new recruits implement the actions that the group's leadership wants. Training and indoctrination procedures can reorient the preferences of the membership base towards those of the leadership. However, in the absence of strength in either, or both, areas, leaders can easily lose control over their movements. The process is especially dangerous if leaders find themselves recruiting from already-cohesive populations or recruiting under heavy repression. The risk of losing their mission is one reason why an ambitious movement might not want to seize the opportunity to become locally embedded. This observation has consequences for domains beyond the militant context presented here: it suggests that what appears to be a process that brings in much-needed strength and resources can create internal strains and transformative pressure.

The upwards drive of the theory suggests that we should see AQAP and Ansar al-Shariah acting similarly, despite their official (and top-level) strategy of remaining at arms-length from each other. Due to lack of fine-grained event data from Yemen, a strategy to get

around this is to take newspaper reporting of violence in Yemen as being indicative of how close observers of the country attribute expectations about the behaviors of armed groups in the country. Due to lack of fine-grained event data from Yemen, a strategy to get around this is to take newspaper reporting of violence in Yemen as being indicative of how close observers of the country attribute expectations about the behaviors of armed groups in the country. The goal is to identify whether activities attributed to AQAP and Ansar al-Shariah are easily distinguished from each other— as they should be if the al-Qaeda branch was following al-Qaedas strategic guidelines and their own attempts to generate an arms-length local organization or if AQAP and Ansar al-Shariahs activities are difficult for a machine-lead clustering approach to distinguish. Introducing stories about the Houthi uprising provides a counter-point to assess whether the clustering techniques are able to capture operational or descriptive differences that signal different insurgents in the country. In order to to assess the degree that AQAP and Ansar al-Shariah were perceived by close observers to maintain a similar profile of activities, I treat the news articles as data points to be classified, using hand-coded tags of actor authorship.

References

- Abbas, Mourad, Kamel Smaïli and Daoud Berkani. 2011. "Evaluation of Topic Identification Methods on Arabic Corpora." *JDIM* 9(5):185–192.
- Abdul-Ahad, Ghaith. 2015. "Yemeni implosion pushes southern Sunnis into arms of al-Qaida and Isis."
- Al-Batati, Saeed and Kareem Fahim. 2015. "War in Yemen is Allowing Qaeda Group To Expand."
- al-Qaeda strategist. 2006-2011. "Letter to Abu Basir." *Abbottabad Documents* .
- Asal, Victor and R Karl Rethemeyer. 2008. "The nature of the beast: Organizational structures and the lethality of terrorist attacks." *The Journal of Politics* 70(2):437–449.
- Bail, Christopher A. 2016. "Cultural carrying capacity: Organ donation advocacy, discursive framing, and social media engagement." *Social Science & Medicine* .
- Bakke, Kristin M, Kathleen Gallagher Cunningham and Lee JM Seymour. 2012. "A plague of initials: Fragmentation, cohesion, and infighting in civil wars." *Perspectives on Politics* 10(02):265–283.
- Balcells, Laia. 2010. "Rivalry and Revenge: Violence against Civilians in Conventional Civil Wars." *International Studies Quarterly* 54(2):291–313.
- Baturo, Alexander and Slava Mikhaylov. 2013. "Life of Brian revisited: Assessing informational and non-informational leadership tools." *Political Science Research and Methods* 1(01):139–157.
- BBC News. 2015. "Profile: Al-Qaeda in the Arabian Peninsula."
- BBC News. 2016. "Yemen crisis: Who is fighting whom?"

- Beardsley, Kyle and Brian McQuinn. 2009. "Rebel Groups as Predatory Organizations The Political Effects of the 2004 Tsunami in Indonesia and Sri Lanka." *Journal of Conflict Resolution* 53(4):624–645.
- Beauchamp, Nicholas. 2017. "Predicting and interpolating state-level polls using Twitter textual data." *American Journal of Political Science* 61(2):490–503.
- Bloom, Mia M. 2004. "Palestinian suicide bombing: Public support, market share, and outbidding." *Political Science Quarterly* 119(1):61–88.
- Bowyer-Bell, J. 2000. *The IRA, 1968-2000: analysis of a secret army*. Vol. 7 Taylor & Francis.
- Brahmi, Abderrezak, Ahmed Ech-Cherif and Abdelkader Benyettou. 2012. "Arabic texts analysis for topic modeling evaluation." *Information retrieval* 15(1):33–53.
- Campbell, Leslie. 2015. Yemen: The Tribal Islamists. Technical report Wilson Center.
- Carapico, Sheila. 2014. "Yemen between revolution and counter-terrorism." *Why Yemen Matters: A Society in Transition* .
- Counsel on Foreign Relations. 2015. Al-Qaeda in the Arabian Peninsula (AQAP). Technical report Council on Foreign Relations.
- Crenshaw, Martha. 1987. "Theories of terrorism: Instrumental and organizational approaches." *The Journal of strategic studies* 10(4):13–31.
- Cruickshank, Paul. 2013. "Brennan on bin Laden raid, and 'Dangerous' Yemen." *CNN* .
- Cutler, David. 2010. "TIMELINE: Al Qaeda activity in Yemen." *Reuters* .
- De Mesquita, Ethan Bueno et al. 2008. "Terrorist factions." *Quarterly Journal of Political Science* 3(4):399–418.

- Denny, Matthew J and Arthur Spirling. 2018. "Text preprocessing for unsupervised learning: why it matters, when it misleads, and what to do about it." *Political Analysis* pp. 1–22.
- Di Giovanni, Jannine. 2014. "Aafia Siddiqui: The Woman ISIS Wanted to Trade for Foley, Then Sotloff." *Newsweek* .
- GSMA Intelligence. 2016. "The Mobile Economy: Middle East and North Africa 2016."
- Gupta, Dipak K. 2008. *Understanding terrorism and political violence: The life cycle of birth, growth, transformation, and demise*. Routledge.
- Hubbard, Ben. 2015. "Al Qaeda Tries a New Tactic to Keep Power: Sharing It." *The New York Times* .
- Human Rights Watch. 2013. "'Between a Drone and Al-Qaeda': The Civilian Cost of US Targeted Killings in Yemen."
- International Crisis Group. 2017. "Yemen's al-Qaeda: Expanding the Base." *Middle East Report* .
- Johnsen, Gregory D. 2012. "A profile of AQAP's upper eschalon." *CTC Sentinel* .
- Johnsen, Gregory D. 2013. "How We Lost Yemen." *Foreign Policy* .
- Jones, Zachary and Fridolin Linder. 2016. *edarf:Exploratory data analysis using random forests*. Vol. 1.
- Kalyvas, Stathis N. 2003. "The ontology of political violence: action and identity in civil wars." *Perspectives on Politics* 1(03):475–494.
- Kenny, Paul D. 2010. "Structural integrity and cohesion in insurgent organizations: Evidence from protracted conflicts in ireland and burma1." *International Studies Review* 12(4):533–555.

- Kilcullen, David. 2009. *The accidental guerrilla: Fighting small wars in the midst of a big one*. Taylor & Francis.
- Koehler-Derrick (ed), Gabriel. 2011. A False Foundation? AQAP, Tribes and ungoverned spaces in Yemen. Technical report The Combatting Terrorism Center at West Point.
- Maaten, Laurens van der and Geoffrey Hinton. 2008. "Visualizing data using t-SNE." *Journal of machine learning research* 9(Nov):2579–2605.
- Miller, Flagg. 2015. *The Audacious Ascetic: What the Bin Laden Tapes Reveal about Al-Qa'ida*. Oxford University Press.
- Muchlinski, David, David Siroky, Jingrui He and Matthew Kocher. 2015. "Comparing random forest with logistic regression for predicting class-imbalanced civil war onset data." *Political Analysis* 24(1):87–103.
- Munksgaard, Rasmus and Jakob Demant. 2016. "Mixing politics and crime—The prevalence and decline of political discourse on the cryptomarket." *International Journal of Drug Policy* 35:77–83.
- Muslimi, Farea Al. 2014. "US drone attacks in Yemen protect no one but Al Qaeda." *The National* .
- Nielsen, Richard. 2014. Networks, Careers, and the Jihadi Radicalization of Muslim Clerics PhD thesis Harvard University.
- Nielsen, Richard Alexander. 2013. The lonely Jihadist: Weak networks and the radicalization of Muslim clerics PhD thesis.
- Oots, Kent Layne. 1989. "Organizational perspectives on the formation and disintegration of terrorist groups." *Studies in Conflict & Terrorism* 12(3):139–152.
- Pearlman, Wendy and Kathleen Gallagher Cunningham. 2012. "Nonstate actors, fragmentation, and conflict processes." *Journal of Conflict Resolution* 56(1):3–15.

- Peters, Gretchen. 2012. Haqqani network financing: the evolution of an industry. Technical report MILITARY ACADEMY WEST POINT NY COMBATING TERRORISM CENTER.
- Phillips, Sarah. 2011. “Al-Qaeda and the Struggle for Yemen.” *Survival* 53(1):95–120.
- Roberts, Margaret E, Brandon M Stewart and Dustin Tingley. 2014. “stm: R package for structural topic models.” *R package* 1:12.
- Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson and David G Rand. 2014. “Structural Topic Models for Open-Ended Survey Responses.” *American Journal of Political Science* 58(4):1064–1082.
- Salama, Vivian. 2014. “Death From Above: How American Drone Strikes Are Devastating Yemen.” *Rolling Stone* .
- Salloum, Said A, Ahmad Qasim AlHamad, Mostafa Al-Emran and Khaled Shaalan. 2018. A Survey of Arabic Text Mining. In *Intelligent Natural Language Processing: Trends and Applications*. Springer pp. 417–431.
- Sharp, Jeremy M. 2015. Yemen: Background and U.S. Relations. Technical report Congressional Research Service.
- Siroky, David S et al. 2009. “Navigating random forests and related advances in algorithmic modeling.” *Statistics Surveys* 3:147–163.
- Spirling, Arthur. 2012. “US treaty making with American Indians: Institutional change and relative power, 1784–1911.” *American Journal of Political Science* 56(1):84–97.
- Staniland, Paul. 2014. *Networks of rebellion: Explaining insurgent cohesion and collapse*. Cornell University Press.

- Stewart, Brandon M and Yuri M Zhukov. 2009. "Use of force and civil–military relations in Russia: an automated content analysis." *Small Wars & Insurgencies* 20(2):319–343.
- Swift, Christopher. 2012. "Arc of Convergence: AQAP, Ansar AL-Shari'a and the Struggle For Yemen." *CTC Sentinel* .
- Tingley, Dustin. 2017. "Rising power on the mind." *International Organization* 71(S1):S165–S188.
- United States Department of State. 2012. "Terrorist Designations of Ansar al-Sharia as an Alias for Al-Qaida in the Arabian Peninsula."
- Voices from the Ecological Resistance. 2003. "The Rise of Environmentalism." *Do or Die* .
- Weinstein, Jeremy M. 2006. *Inside rebellion: The politics of insurgent violence*. Cambridge University Press.
- Wood, Elisabeth Jean. 2008. "The social processes of civil war: The wartime transformation of social networks." *Annu. Rev. Polit. Sci.* 11:539–561.
- Worth, Robert F. 2015. "Yemen: The Houthi Enigma." *The New York Review of Books* .
- Yara Bayoumy, Noah Browning and Mohammed Ghobari. 2016. "How Saudi Arabia's war in Yemen has made al Qaeda stronger and richer." *Reuters* .