

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

Margaret J. Foster*

May 5, 2018

Abstract Why might it not be beneficial for an 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 the Arabian Peninsula (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 leadership. This process of transformation suggests an under-explored but widespread trade-off that organizations face if they try to increase their resilience and capacity 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: integration with local communities provides resources and protection but can undermine their broader goals by leaving them beholden to the parochial interests of local recruits. In this paper, I use the example of al-Qaeda in the Arabian Peninsula (AQAP) to highlight the bottom-up mechanism of organizational transformation within an organization caught in a particularly acute tradeoff in satisfying their local and global constituencies.

However, the internal dynamics of militant organizations are often opaque and micro-level event data is not always available from conflict zones. I therefore make a methodological contribution by adopting an empirical strategy that leverages two distinct sources of text data—

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.

news reports and official communiqués— to generate separate but complementary avenues 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, with particular focus on how the process may influence militant revolutionary organizations. I position the theory within the existing literatures of substate conflict and organizational dynamics of rebel organizations and the social processes of civil war and terrorism. I then apply this theory to the case of al-Qaeda in the Arabian Peninsula, widely identified to have gained strength due to an influx of members from local communities.

To gather evidence that recruitment shocks were followed by shifting organizational priorities, I examine AQAP through the lenses of their reported activities and through their self-presentation in communiqués and propaganda material. 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 then carry out a qualitative reading of translated texts from al-Qaeda in Yemen, and supplement the conclusions 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 to adjust to reflect the preferences of the grassroots. Three circumstances make a leader more

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, Smaïli 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, although at the cost of nuance and depth present in the original Arabic documents.

likely to feel pressure: labor mobility among the rank-and-file; differences in priorities of the leadership and new base, and a situation in which the leader has a restricted ability or desire to titrate the inflow of new members. Labor mobility is necessary for the organization's rank and file to be credibly able to exit if the organization does not accommodate their priorities, while difference in priorities and preferences are required for there to be an internal tension between leaders and the grassroots.³ The third assumption, restricted ability or desire to limit inflow, is an avenue through which leaders are prevented from effectively socializing the new entrants. Finally, the theory is agnostic as to whether the leader is able to foresee the effects of an influx of members: either the leader fails to foresee the potential for tension or, more likely, circumstance or optimism cause the leader to discount the effects of the internal tension.⁴

The tension between resilience and mission purity occurs because unless leaders can reduce labor mobility or ensure that they comprehensively socialize their new members,⁵ failing to accommodate the interests of new recruits can be catastrophic. Without the manpower, resources, and networks of grassroots members, the organization will cease to operate. On the other, if the new recruits feel that their goals are not being met, they can form an internal faction within the group or simply leave. 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 splinter (Bakke, Cunningham and Seymour, 2012; Pearlman and Cunningham, 2012), while competition between factions can have lethal implications for the conflict

³Another way to conceptualize the absence of interest divergence between the leaders and the members of their organizations is to consider an organization without a strong attachment to a particular set of goals. In this case, we should expect a grassroots-driven transformation to occur quickly, as the leadership does not undergo internal pushback against modifying their organizations to capitalize on the preferences of a desirable pool of recruits.

⁴Survival threats provide one powerful reason for a leader to discount the effects of tension between their goals and the goals of their new membership. A leader who wants their group to continue to exist in the short-run may be tempted to discount the long-term effects of admitting new members whose preferences are similar to, but not exactly in alignment with, their own.

⁵Some organizations can thus attempt to restrict the exit options of their members, such as by encouraging them to engage in illegal activity or by creating club goods contingent on membership.

theater (Bloom, 2004; De Mesquita et al., 2008).

Recruitment shocks may also generate a self-reinforcing cycle of internal transformation. 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. 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 organization's issue space.

Changes in group practices can generate also 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 context of militant groups, 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. 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).

Two militant 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. Four 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.⁶ 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 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,

⁶I 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

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. Under the Ansar al-Shariah name, AQAP could strike a more parochial message, exploit local grievances, and avoid the encumbrances of the al-Qaeda brand (Swift, 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).

Each of the three previous factors are magnified by Yemeni state weakness, which means that AQAP's trajectory is less likely to be driven by strategic interaction with a strong state and security apparatus (Bureau of Counterterrorism, 2017). In particular, 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.⁷ In 2012 the White House counter-terror advisor John Brennan estimated AQAP's membership base as over thousand fighters (Cruickshank, 2013; Johnsen, 2013).

Finally, AQAP is widely reported to have experienced a dramatic inflow of members from among the Sunni tribal communities. Moreover, the inflow of recruits is from among a population that AQAP is reported to have failed to make inroads into, thus suggesting that it is more likely that the group would seize the opportunity to expand their membership, even at risk of being unable to incompletely socialize.

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).

The speed of AQAP's membership growth after 2009 was attested to in the yearly *Country Reports on Terrorism* produced and released by the United State Department of State. In 2009, the publication estimated that AQAP's membership was approximately 200-300 (Johnsen, 2012), with a similar estimate of a "few hundred" members the following year (Bureau of

⁷The 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)

Counterterrorism, 2011). From 2011 onward the estimated group membership spiked dramatically, from an estimated strength of a “few thousand members” (Bureau of Counterterrorism, 2012) to “up to four thousand members” in 2015 and 2016 (Bureau of Counterterrorism, 2016, 395).⁸

Two trends widely held responsible for accelerating AQAP’s ability to recruit local supporters may provide vectors for new preoccupations.

The first major trend driving recruitment is desire for revenge against the United States and Yemeni government for collateral damage of American drone strikes (Bayoumy, 2013). By alienating the population, drone operations are reported to have had the effect of drawing otherwise-pragmatic tribes closer to the jihadi militant group. A Yemeni journalist with ties to AQAP likewise noted that revenge drove Yemenis closer to AQAP, “Hundreds of families are seeking revenge from the U.S. so they deal with that by joining al Qaeda” (Bayoumy, 2013).

For example, the brother of a man killed in a strike described how drone strikes quickly changed local receptiveness to AQAP among communities with little previous engagement or affinity for the jihadi group’s appeals. He noted that “In our area there was never anyone linked to al Qaeda. After the strike, everyone in the area started listening to al Qaeda types, exchanging videos on mobile phones” (Bayoumy, 2013). This dynamic underlies the broader logic described in Kilcullen (2009), who traced how transnational revolutionary movements embedded themselves within local communities. Yet, while fighters motivated to join AQAP for revenge add needed local strength, they may also join the group in response to threats to their local interests and identities with no particular ideological affinity for the movement (Johnsen, 2013). Once recruited, AQAP must then undertake the process of socializing the new members to adopt the jihadi ideology. Although Yemeni state failure may suggest that AQAP should benefit from territory in which to train and indoctrinate, anecdotes from Yemen suggest that the group has difficulty indoctrinating new members. One explanation

⁸Note that the 2013 and 2014 Country reports revised the strength estimate to about a thousand members

may be that pressure from the American drone campaign has limited AQAP's ability to move trainers around the country (Bayoumy, 2013) at the same time that local communities are joining for revenge. The effect of a base expansion and restricted ability to indoctrinate has led to situations such as one reported in 2013 in which an AQAP commander in the south-east of the country was complaining that his fighters were so insufficiently ideologically motivated that they neglected basic religious obligations (Muslimi, 2014).

The second trend accelerating AQAP's ability to recruit in Yemen is the rise of a Shia Zaidi insurgency associated with the Houthi movement. As with drone strikes, the Houthi insurgency allowed al-Qaeda to better integrate with the local tribes. By making sectarian identity increasingly salient, the Shia insurgency drove 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 militiaman 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). As with the recruitment influx driven by anger and resentment over drone strikes, recruitment into AQAP due to sectarian polarization has the effect of encouraging membership from large numbers of grassroots fighters without a strong preexisting commitment to the jihadi ideological cause.

Empirical Implications and Quantitative Analysis

The combination of strategic direction to avoid local entanglements and creation of a separate spin-off as a front for involvement in the local insurgency provides a series of natural counterfactuals for the development of AQAP's trajectory. Without internal or external countervailing pressure, we should expect to see AQAP's activities reflecting the priorities that advance the transnational jihadi agenda that they were instructed to implement. On the whole, AQAP's behavior should reflect this focus and center on attacks against the Yemeni military, offensives against the central state, and external initiatives against transnational enemies, notably the US and Saudi Arabia. Likewise, Ansar al-Shariah's behavior should indicate a greater involvement in capturing and administering local territories, such as assassinations of local leaders, score-settling attacks, and intimidation. 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.

Conversely, the theory of grassroots driven transformation presented above suggests that increased recruitment from among Yemeni tribes should result in local conflicts and priorities being more strongly reflected in AQAP. Thus, in expectation, as AQAP became more deeply tied to Sunni tribes, their new membership base is expected to create an internal constituency 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. Thus, 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. Conversely, if AQAP is able to carry out the bipartite local presence that they attempted to design, activities attributed to AQAP and Ansar al-Shariah should be readily distinguishable from each other. Finally, if inferring targets, tactics, and priorities via news reporting is viable, the Sunni groups should be distinguishable from the Houthi Shia insurgency, which has little in common with either Sunni group other than shared activity in Yemen.

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.⁹ To tighten the focus on whether AQAP is distinguishable from Ansar al-Shariah, corpus was filtered to 566 articles that reported on events attributed to al-Qaeda in the Arabian Peninsula, Ansar al-Sharia, and the Houthi militia.¹⁰ 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 news stories with publication dates ranging from October 30, 2002 through January 3, 2015. In order to convert the news articles into data for classification, I converted each tagged story into a tokenized bag of words, normalized via term frequency-

⁹Approximately 15% of the data was coded by workers on Mechanical Turk, the rest was coded by the author.

¹⁰Activities 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.

inverse document frequency (tf-idf).¹¹ For the training set, this produced a 432 x 2,222 matrix of those words common to both the training and test sets of articles about the violent activities of AQAP, Ansar al-Shariah, and the Houthis. The clustering analysis presented below evaluates behavioral similarity among the three groups by analyzing the similarity of the rows of this matrix.

The following section presents analysis derived by training random forest and support vector machine classifiers of news articles about their activities. The results are underscored via a tSNE visualization of stories about each group. In each case, the classifier's predictions were assessed against hand-coded labels of the story's primary actor. I use the two classifiers and the tSNE visualization to increase confidence in the validity of the results; as the methods all of which produce generally the same findings, the results are unlikely to be driven by an idiosyncrasy specific to a particular algorithm.

The classification strategy allows for a systematic analysis of group behavior in an information-poor setting. In order to be meaningful, the classification must pass a basic plausibility test: stories about the Sunni groups should be distinguishable from other domestic insurgent groups. In the analysis that follows, news reporting about the Houthi Shia insurgency stand in for other domestic actors, as the organization is sufficiently active to provide enough stories to compare with AQAP and Ansar al-Shariah. Then, if AQAP is constrained by a growing local preference within their base, this interest should push the al-Qaeda branch to act in ways that make it hard to separate from Ansar al-Shariah, their local spin-off. Conversely, without a pressure to accommodate local preferences and interests, AQAP should be expected to follow the leadership's apparent preference to separate the activities of the globally-branded AQAP from the locally-branded Ansar al-Shariah. This should result in news coverage of the two groups being different from each other, and from other local actors.

The results of these supervised approaches reinforce broad outline of the tSNE visualization:

¹¹Examples of the stories and full details about preprocessing are available in the Technical Appendix, which is available at margaretjfoster.net/over-pressure-technical-appendix.html

within the test set, the Sunni and Shia groups separate, and thus fulfill the basic plausibility criteria. However, Ansar al-Shariah stories are consistently predicted as being associated with AQAP, suggesting that AQAP has been unable to separate the activities of their local and global fronts. In the following section, I summarize the results and focus on attributes of each classifier that can produce insight how the stories systematically differ. The random forest analysis focuses on the most important features that separate Houthi from AQAP stories, while the analysis of the support vector machine looks closely at the predictions of attribution to AQAP and Ansar al-Shariah over time.¹²

Random forests are particularly effective for high-dimensional data and have 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 expectations from the plausibility test and the convergence expectation. Articles about AQAP and the Houthi insurgency separate cleanly. However, the random forest model consistently misclassifies Ansar al-Shariah stories as belonging to al-Qaeda in the Arabian Peninsula. Table 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 al-Shariah are overwhelmed by the differences between the two Sunni groups and the Houthi movement.

| | 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 is that as the features used for classification are also terms in the document, which provide interpretable insights into what words drive separation among stories in the data. The fifteen most important words for the random forest

¹²The results are more fully characterized in the Technical Appendix, along with the results of classification on just the Ansar al-Shariah and AQAP stories.

15 Most Important Words For Story Classification

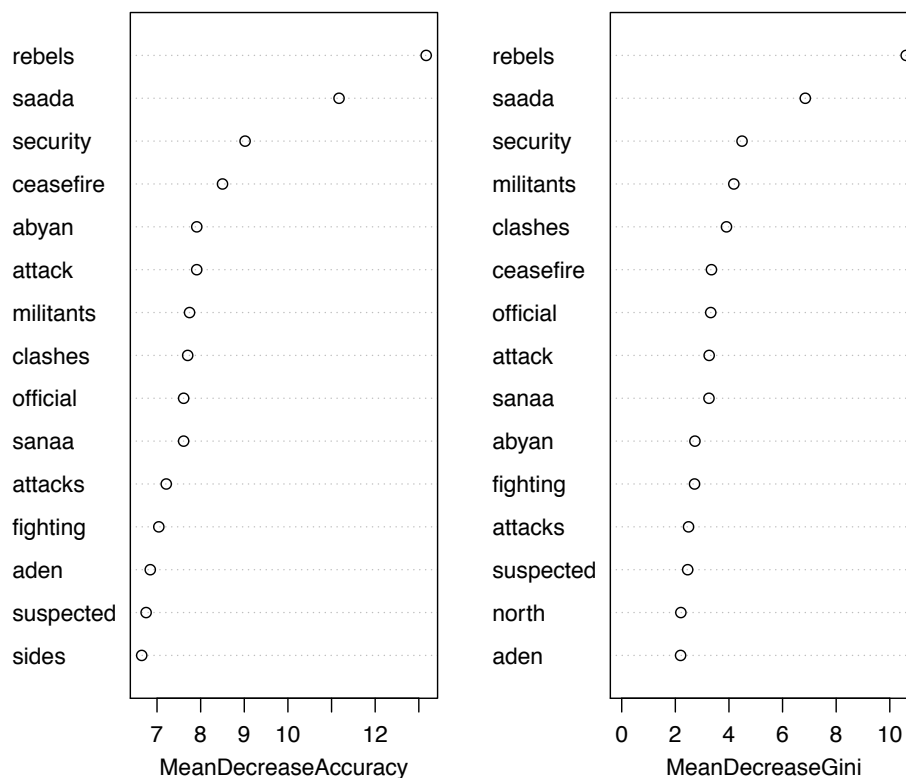


Figure 1: Important Words for Classification

classification—after removing stopwords that describe or name the active group— suggest that reason for the clean split across the Sunni and Shia movements lies in framing. Houthis are consistently described as “rebels” while the Sunni fighters are frequently presented as “militants.” After the rebel/militant split, 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.¹³

A second classification, using a support vector machine (SVM), generated similar results as

¹³This list has group-specific terms removed, as described in the Technical Appendix.

the random forest classification. As in the random forest, the support vector machine consistently passes the basic plausibility test by correctly distinguishing news stories describing activities of the Shia insurgency from those describing activities of the Sunni insurgents. However, as with the results of the random forest, Ansar al-Shariah stories are consistently identified with AQAP, and similarly fails to predict any stories as being Ansar al-Shariah. The support vector machine provides a natural way to assess whether similarity among AQAP and Ansar al-Shariah stories become more similar over time, as would be expected if the apparent convergence is driven by a growing internal constituency pushing for local activities.

Figure 2 provides a closer look at whether the support vector machine’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 increasing weight to the Ansar al-Shariah label for news stories about AQAP activities as AQAP members push the group to engage with local concerns. The plot focuses on stories with a true label of AQAP and depicts the predicted probability that an article would be assigned to the label of Ansar al-Shariah given that it is a story about AQAP actions (red) as well as the SVM’s confidence in the classification for AQAP (blue) for the 144 stories in the test set. As the SVM consistently predicts all AQAP and Ansar al-Shariah stories for AQAP, the expected probability of assignment to Ansar al-Shariah remains constant at approximately $p = .15$. The classification predictions provide mixed support for the expectation that an influx of local members should increase the difficulty of assigning group labels: although the SVM’s confidence in predicting the AQAP label becomes more variable over time, the predicted probability of assignment to Ansar al-Shariah remains constant over the time period.

The support vector machine’s assigned probabilities of the Ansar al-Shariah label contains information about both the differences between AQAP and Ansar al-Shariah stories as well as the relative frequencies of each label. Using a t-Distributed Stochastic Neighbor Embedding (tSNE) allows a way to focus on just the content of each article by visualizing the differences

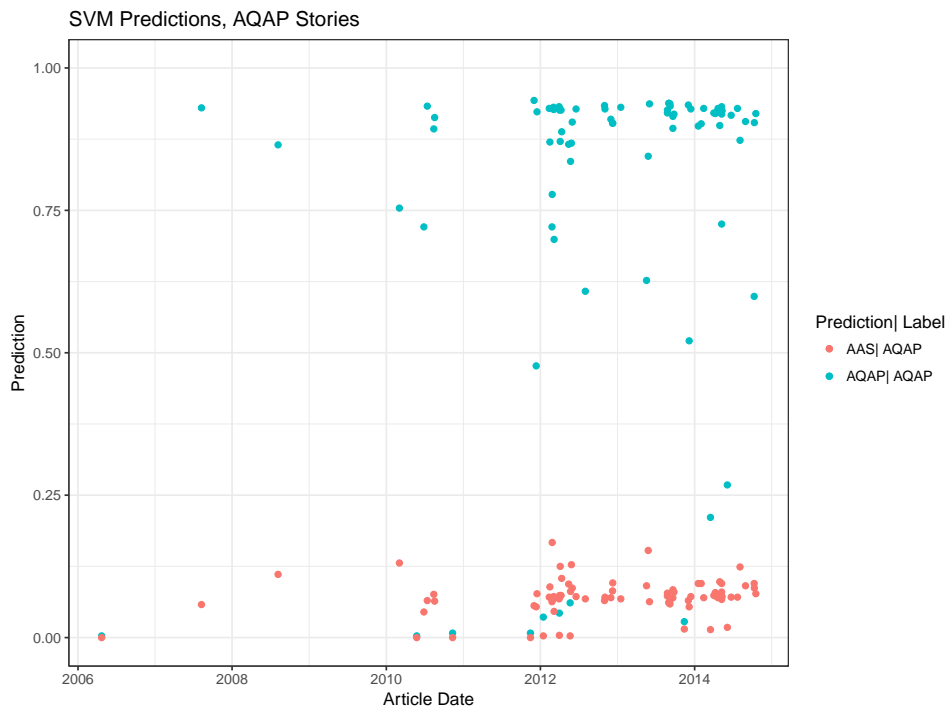


Figure 2: SVM Predictions Over Time For AQAP Stories

in article words in a two-dimensional projection.¹⁴ If news stories about the two Sunni groups are similar, as might be expected if there is little difference between texts about their activities, the dimensionality reduction would be expected to show consistent intermingling of Ansar al-Shariah and AQAP stories across the time period. However, if descriptions of the activities of each group become more or less similar over time, this separation should also be reflected in the tSNE visualization. Figure 2 shows tSNE clustering of stories tagged as describing activities of AQAP (blue) and Ansar al-Shariah (red), presented by year.¹⁵ Although distances are not necessarily interpretable in tSNE, the plots indicate that stories of Ansar al-Shariah’s activities and Al-Qaeda in the Arabian Peninsula’s activities become

¹⁴tSNE allows for the reduction 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. As a stochastic process, tSNE is sensitive to initial starting values and parameter specification. The Technical Appendix details parameter selection.

¹⁵The dimensionality reduction did not use time as a covariate, and thus the changing clustering patterns are not a result of applying different iterations of tSNE on different data, nor of random changes from iterated runs.

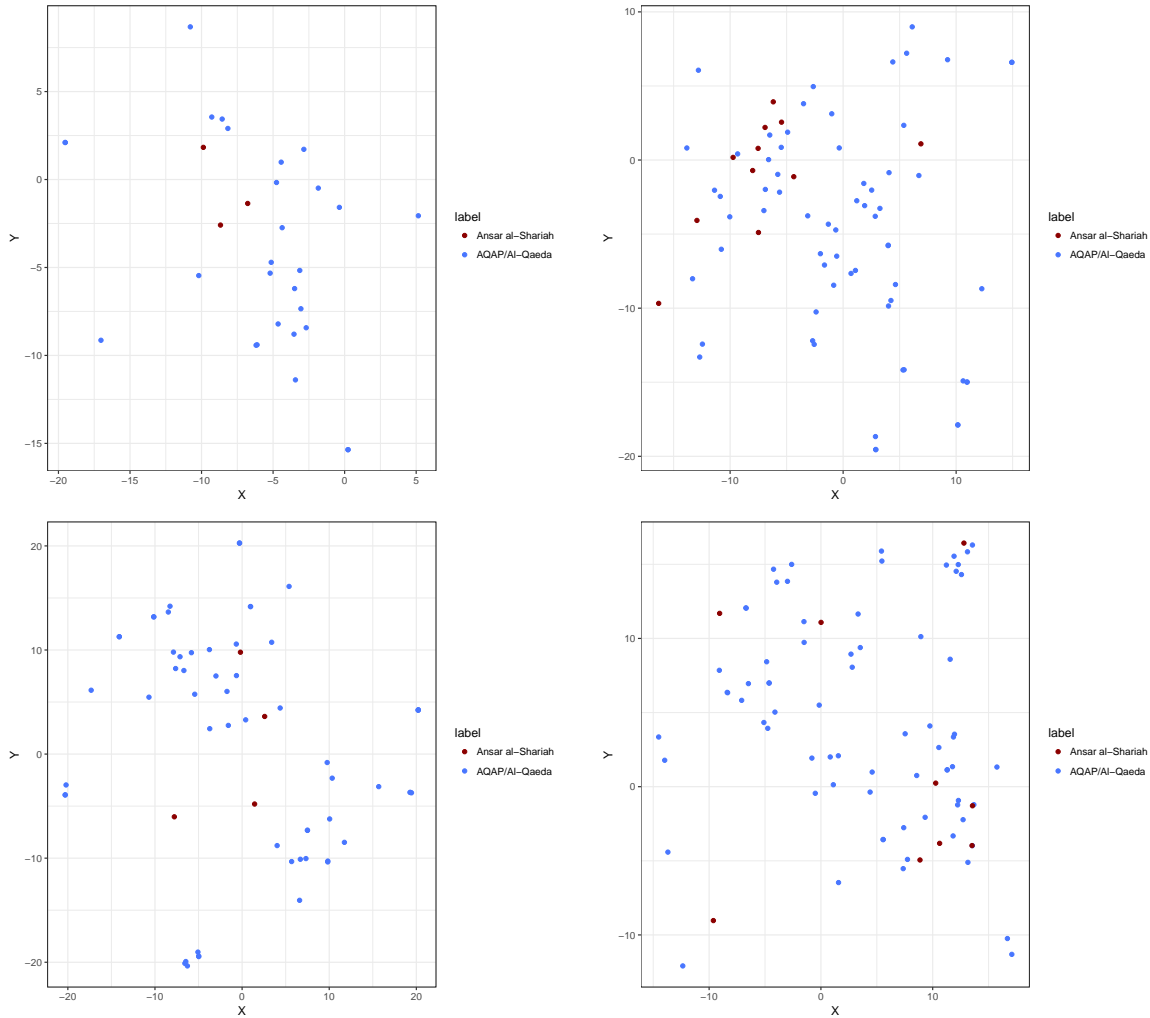


Table 2: tSNE Visualization Sunni Groups, 2011-2014

progressively less distinct as time passes.

Although the results of the clustering indicate convergence between coverage of the activities of al-Qaeda in the Arabian Peninsula and Ansar al-Shariah, these methods don't adjudicate between competing situations in which AQAP acts like Ansar al-Shariah; Ansar al-Shariah acts like AQAP; or the results of each clustering algorithm are driven by some feature other than underlying behavioral similarities and differences.

The supervised machine learning techniques provide an angle to adjudicate between the

counterfactuals introduced in the previous section. In particular, the clustering analysis above suggests that AQAP has been unable to maintain a local spin-off with a distinctive operational profile. However, these techniques are unable to distinguish between AQAP acting like Ansar al-Shariah or Ansar al-Shariah acting like AQAP, an important difference for the theoretical expectation that an inflow of recruits should pressure AQAP's leadership to adopt a local emphasis. One natural counterfactual in which an influx of local fighters is followed by behavioral convergence of AQAP and Ansar al-Shariah could be that AQAP's socialization has been so successful that the group has changed the preferences of the communities in which they operate. In this scenario, the local Ansar al-Shariah should gain a greater international focus as local actors are socialized into the transnational jihadi ideology.

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 uprising.

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, there is likely to be over-attribution to AQAP, given that the al-Qaeda name is of high salience. Additionally, journalists associating AQAP and Ansar al-Shariah should drive under-attribution of to Ansar al-Shariah.¹⁶ Similarly, the consistent increase in uncertainty over the SVM

¹⁶For example, 11 of the stories in the corpus carried language explicitly blur the distinction between

story classification for AQAP-articles is likewise ambiguous. The uncertainty largely derives from mispredicting AQAP-Houthi pairs.¹⁷ 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, 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 adjudicate the two possible directions driving convergence in descriptions of 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 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

AQAP and Ansar al-Shariah, such as the following quote from Xinhua News stating: “Al-Qaida in the Arabian Peninsula, also known locally as Ansar al-Sharia.”

¹⁷Results are similar for a classification carried out on a subset of the corpus comprising just AQAP and Ansar al-Shariah stories.

researchers to incorporate document-level information as 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

To address the question the direction of the apparent convergence of AQAP and Ansar al-Shariah behavior, the following section presents selected results from a Structural Topic Model that shows a general localizing trend in AQAP’s self-presentation.¹⁸ This model addresses the question of direction of convergence between AQAP and Ansar al-Shariah, as the scenario predicted by the bottom-up theory of organizational change expects that an influx of fighters was followed by AQAP adopting a more local, Yemeni-centric, self-presentation. In the counterfactual presented at the end of the previous section, growing AQAP strength could have been followed by increasing adoption of the globalist jihadi ideology. This case would not require AQAP to dramatically change their messaging. Moreover, the timing and content of any change in self-presentation can shed light on the question of whether the supervised clustering responds to changes in ground truth: although communiques and speeches do not necessarily perfectly map onto ground truth, finding that themes in AQAP self-presentation move along with the convergence in stories about AQAP and Ansar al-Shariah should bolster confidence that each technique is picking up a real change.

The second and third models address the counterfactual that regional and global developments may account for changes in AQAP messaging, independent of any changes in membership base. The second model highlights changes in AQAP messaging following events,

¹⁸The Technical Appendix features information about text processing and a fuller characterization of the results

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.¹⁹

STM Model One: Evolution of AQAP Messaging

The first model presented the estimated thematic divergence of communiques 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.

Each topic is summarized by its FREX words, which are words or tokens that are 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 promoting transnational jihadi sentiments and goals, and

¹⁹As-Sahab is closely associated with the core of al-Qaeda’s “central” leadership, which is why their releases were chosen for the comparison.

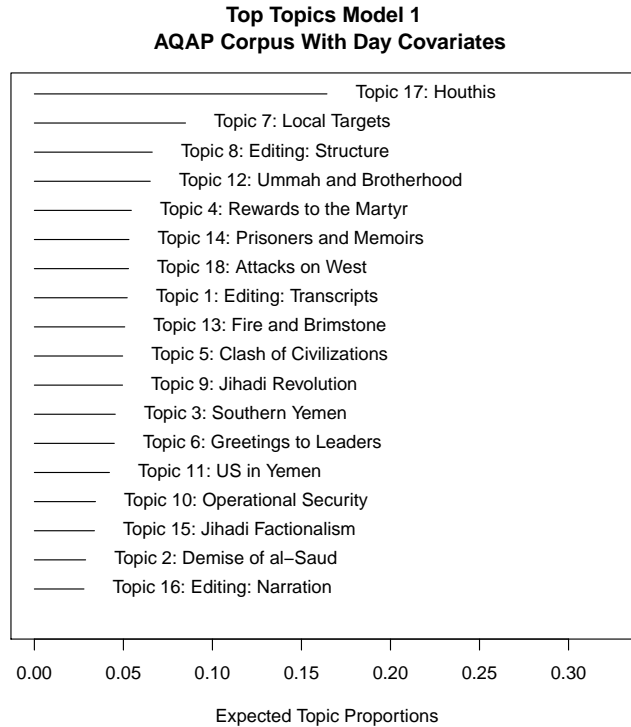


Figure 3: Estimated Topic Proportions in al-Qaeda in the Arabian Peninsula Corpus

jihadi-associated descriptors. After estimating the model, I grouped topics into thematically similar clusters.²⁰

Figure 5 depicts the expected document-level topic proportion for two 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 communiques issued to an audience with an interest in local territorial control. Notably, the emphasis on the bellwether Houthi topic

²⁰The three remaining topics relate to the documents themselves 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 4: Groupings of Substantive Topics in General Trends Model

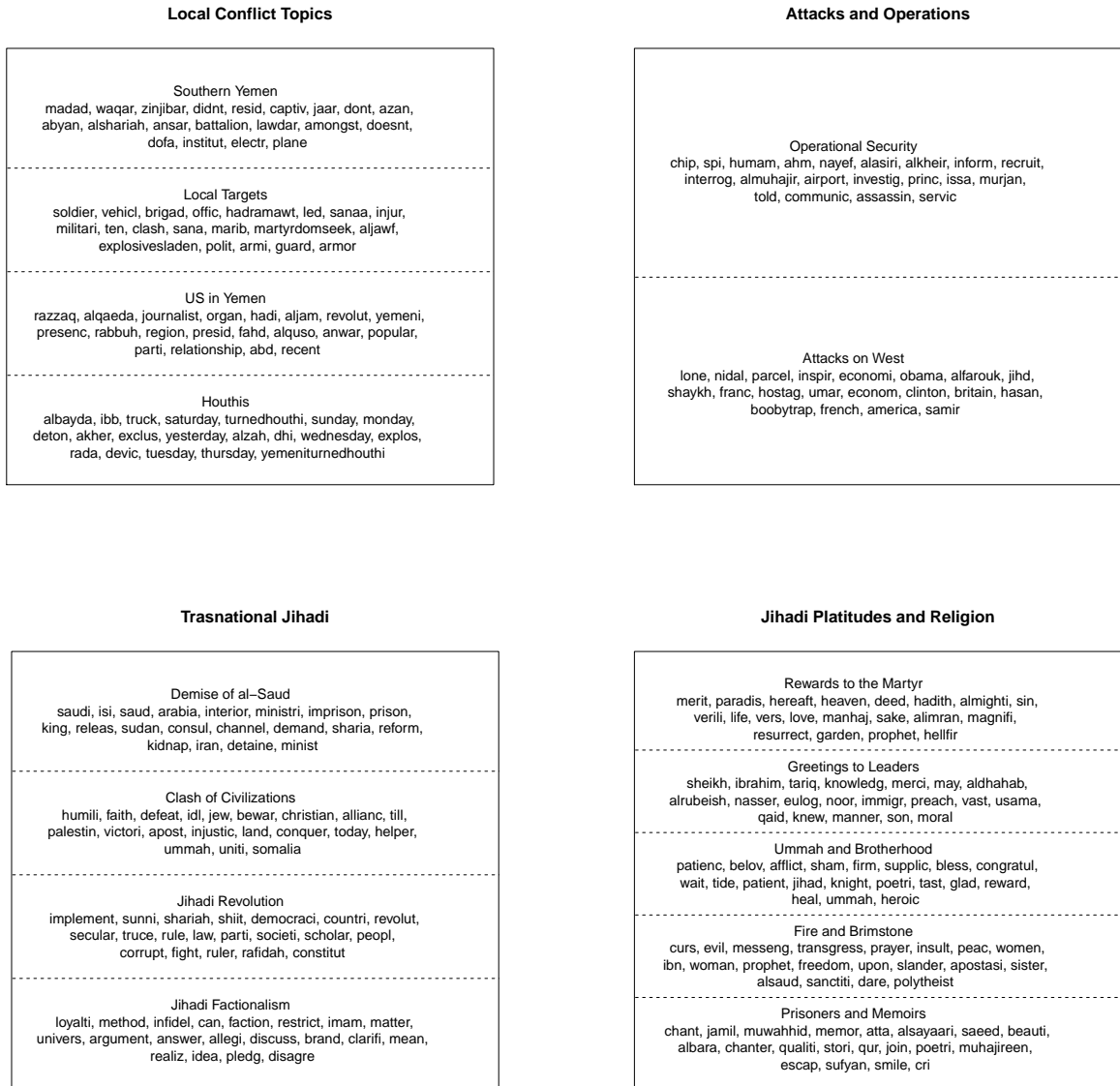
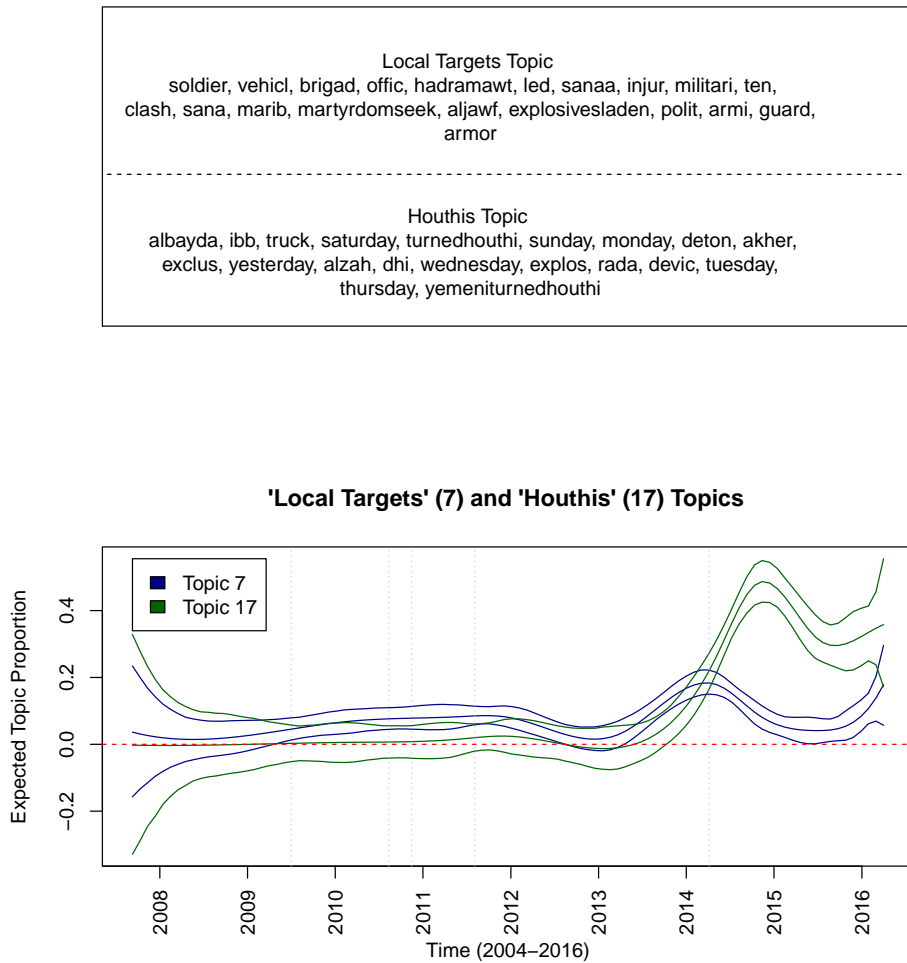


Figure 5: Changes over time to attention dedicated to local topics



has dramatically increased in recent years.

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 6 depicts the expected proportion of the transnational jihadi topics presented according to time.²¹ 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

²¹For 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.

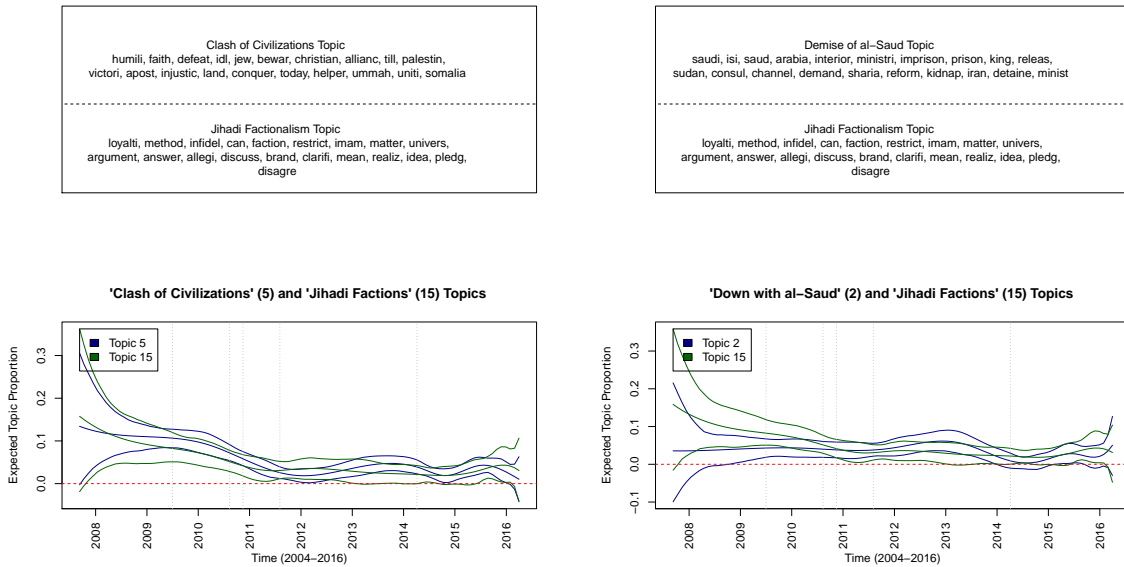
right, the lines represent: the al-Majalah airstrike on December 17, 2009; the start of the Yemeni Revolution on January 27, 2011; the death of Usama bin Laden on May 2, 2011; the end of the first Obama Administration on January 19, 2012; and the Houthi takeover of Sanaa on September 21, 2014.

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 6, the “Jihadi Revolution” topic, is centered around concepts used to incite for overthrow of secular governments and implementation of an 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.

Figure 6: Changes over time to attention given to global jihadi topics



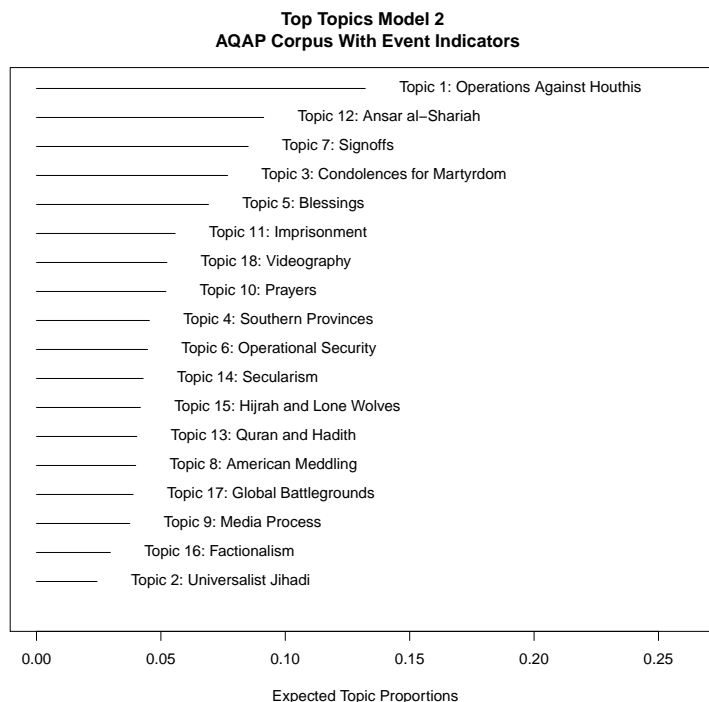
The fourth transnational topic in Figure 6 is topic 2, which I have labeled “Demise of al-Saud.” The topic is focused on Saudi-centric themes, with references to Saudi repression of jihadi dissidents along with references to Saudi government officials. The topic is largely stable at an expected prevalence of approximately 5% throughout the time period.

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

To provide more specific indicators of how AQAP’s messaging differed from communiques released under the Ansar al-Shariah brand, I estimated an 18-topic model on 875 pooled AQAP and Ansar al-Shariah 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 7.²²

²²The figure omits three topics that are associated with the editing and production of the releases. A thematic clustering of the topics, with their associated FREX words can be seen in the Technical Appendix.

Figure 7: Expected Proportion of Topics in Model 2



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 8 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 manifestations of the American War on Terror, such as the American occupation of Iraq and the

incarceration of Aafia Siddiqui.²³

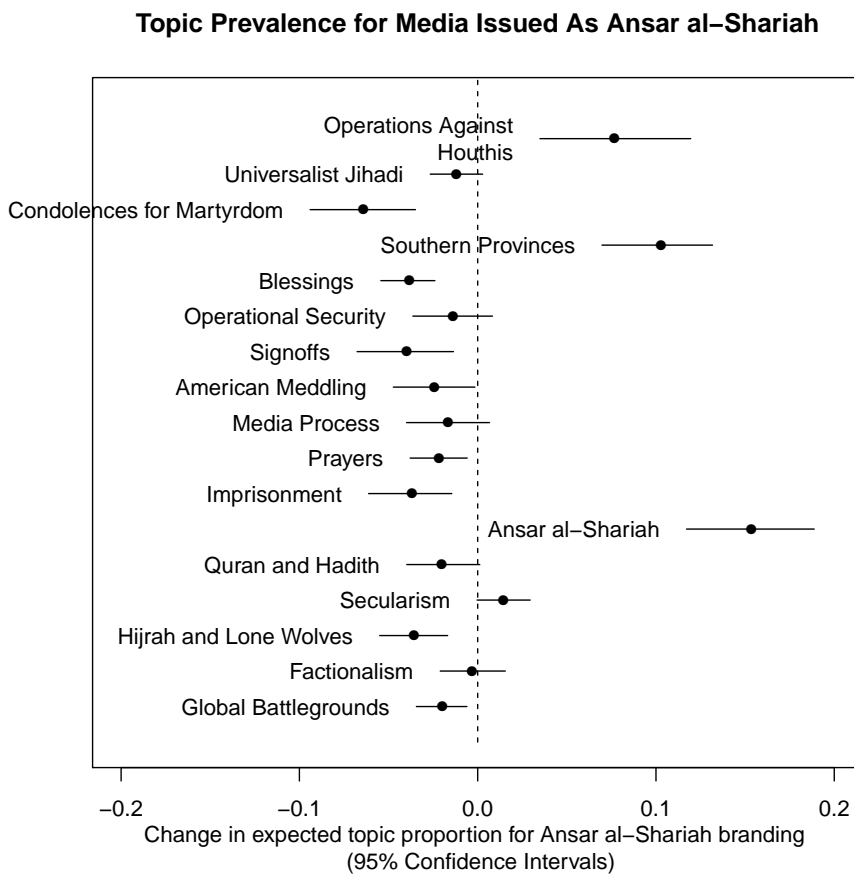
Interestingly, media released under the Ansar al-Shariah name has less emphasis on religious themes, which would be consistent with downplaying their identity as a revolutionary religious organization in favor of self-presentation as an actor in the domestic social conflict. These results are all the 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. Evidently, although 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 as indicated in the analysis of AQAP and Ansar al-Shariah communiques, this separation did not extend to local perception of their activities as indicated by news articles about their activities.

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 local in theme. 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,

²³For a discussion of Siddiqui's position in jihadi propaganda, see Di Giovanni (2014).

Figure 8: Covariate Influences on Topic Prevalence



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.

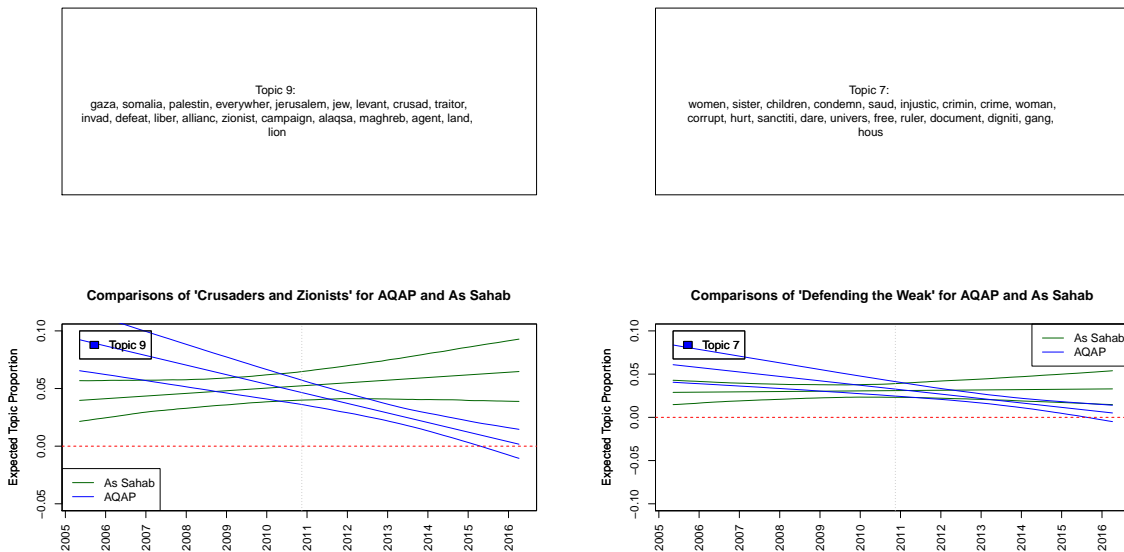
Thus, one possible counter-narrative to the bottom-up transformation argument maintains that the change in al-Qaeda's leadership may have triggered a wider ideological shift that filtered to local branches and that the loss of Awlaki amplified the effect in Yemen. Indeed, if AQAP was following the lead 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 communiques and releases from as-Sahab—the resulting topics differ from those presented in models one and two.

Figure 9 highlights two topic outcomes from the model, which predicted topic prevalence in the corpus as a function of time interacted with issuing group.²⁴

²⁴Time is included in this model as a linear function for computational tractability.

Figure 9: AQAP and As-Sahab Divergence



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 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.

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.²⁵ However, the divergence in thematic prevalence of pan-jihadi topics between releases issued by As-Sahab and AQAP indicate that AQAP's increasing Yemen focus was not indicative of a localizing turn lead or directed by al-Qaeda Central, via their As-Sahab mouthpiece.

Concern: Biased Communique Data

Although approaching each component of the analysis from multiple perspectives, significant limitations remain for the empirical strategy to be able to conclusively claim to have demonstrated an upwards-driving pressure causing AQAP's leadership to evolve from a transnational jihadi vanguard into an increasingly local insurgency. The following section discusses the implications of one limitation: systematically biased text data in the communique analysis.

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

²⁵A full analysis is discussed in the Technical Appendix

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 the 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 the 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

The paper has 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. Additionally, the paper makes a methodological contribution for research in information-poor settings, by suggesting a method for treating the text of news reports as a source of feature-rich data about the tactical and strategic goals of a local actor and in using multiple methods and data sources to triangulate a view onto the behavior of actors that would otherwise be difficult to systematically characterize. In taking overlapping methods, both unsupervised and supervised machine learning, and distinct text data sets, the empirical approach builds confidence in the absence of a ground truth against which to compare the findings.

Substantively, this project contributes to the literature on organizational behavior and evolution. 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.

However, the empirical strategy followed above is agnostic about the direction through which group changes influence rhetoric. One possibility is that leaders change their self-presentation in order to appeal to a local population, who then join and begin to change the behavior of the group. In this case, changes in tone presage 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 qualitative evidence about AQAP's strategic instructions and attempts to create a local spin-off suggest an organization trying to avoid a top-down localizing shift, the text evidence cannot adjudicate between these two alternatives. Indeed, the results are agnostic about the exact sequencing of the shifts while still supporting the theory of bottom-up transformation.

Future work extends the theoretical framework to other organizations operating in a contested domain or under survival threats, particularly seeking to apply the theory of grassroots-driven organizational change to contexts in which microlevel data on recruitment and management decisions is more accessible. Additionally, modeling the identified trade-off can 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.

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* .
- 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.
- Bayoumy, Yara. 2013. Insight: In Yemen, al Qaeda gains sympathy amid U.S. drone strikes. Technical report.
- 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.
- Bureau of Counterterrorism. 2011. Country Reports on Terrorism 2010. Technical report United States State Department.
- Bureau of Counterterrorism. 2012. Country Reports on Terrorism 2011. Technical report United States State Department.
- Bureau of Counterterrorism. 2016. Country Reports on Terrorism 2016. Technical report United States State Department.
- Bureau of Counterterrorism. 2017. Country Reports on Terrorism 2016. Technical report United States State Department.
- Campbell, Leslie. 2015. Yemen: The Tribal Islamists. Technical report Wilson Center.
- Cruickshank, Paul. 2013. "Brennan on bin Laden raid, and 'Dangerous' Yemen." *CNN* .
- De Mesquita, Ethan Bueno et al. 2008. "Terrorist factions." *Quarterly Journal of Political Science* 3(4):399–418.
- 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."

- 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.
- 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 Combating Terrorism Center at West Point.
- 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.

- 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.
- 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.
- 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.
- 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* .