Bureaucracy and Control in Terrorist Organizations*

Jacob N. Shapiro†

This version August 8, 2008.

*I thank Scott Ashworth, Jon Bendor, Ethan Bueno de Mesquita, Dan Butler, Jim Fearon, Kris Ramsay, Scott Sagan, Dave Siegel, Jeremy Weinstein and the participants in the Columbia University International Politics Seminar and UT Dallas Terrorism and Policy Conference for provided outstanding feedback on various aspects of this paper and related research. Joe Felter and Brian Fishman, and Assaf Moghadam at the Combating Terrorism Center offered many helpful comments on the al-Qa'ida in Iraq case study and facilitated access to documents from the Harmony database. All errors are my own.

†Assistant Professor, Woodrow Wilson School and Department of Politics, Princeton University, Princeton, NJ 08544-1013; E-mail: jns@princeton.edu; Phone: 609-258-2256; Fax 609-258-0482.

Abstract

Terrorists face a hard organizational challenge. Political and ideological leaders, the principals, have to delegate certain duties—planning attacks, soliciting funds, recruiting, and the like—to middlemen or low-level operatives, their agents. Such delegation is problematic as operational elements often see the world differently than their leaders do and tend to disagree with them on how best to serve the cause. Strategies to control the resulting agency problems all entail security costs. Thus terrorists face a fundamental tradeoff between security and control. I model this tradeoff as an agency problem where principals are better informed than agents about the mapping from policies to outcomes, but sharing information to help agents achieve better outcomes is costly. This approach explains why terrorist groups employ many of the same mechanisms of control as traditional hierarchical organizations and suggests a number of specific principals for counterterrorism policy.

Introduction

Terrorist organizations face a difficult task in a hostile operational setting. First, they must achieve the controlled use of violence as a means of advancing their specified political ends. Doing too much, or hitting the wrong targets, can be just as damaging to the cause as doing too little. Second, they must maintain this calibrated use of force in an environment where becoming known to government leads to operational failure.

These twin challenges lead to several recurring themes in terrorists' organizational writings from the 1880s through the present day. We see a consistent focus on how to achieve the appropriate use of violence when the rank and file often clamor for more violence than is useful from the leader's perspective or seek to enrich themselves in the course of their duties.[1] Groups repeatedly struggle with the problem of maintaining situational awareness while staying covert, so that members understand which actions will support the political goal, and which will be counterproductive.[2] Finally, there is regular concern with balancing the need to control operational elements with the need to evade government attention and limit the consequences of any compromise.[3]

One surprising fact about terrorist organizations is the frequency with which standard bureaucratic mechanisms—organization charts, expense reports, and the like—emerge as the

[1]Russian Marxist groups such as the Party of the Socialist Revolutionaries (PSR) and the Russian Social Democratic Labor Party (RSDLP) had regular problems with local cells conducting revenge attacks that could not be justified by Marxist theory (Geifman, 1992).

[2]The problem of becoming detached from the masses because of the exigencies of maintaining security features prominently in both jihadi descriptions of their fight against the Assad regime in Syria from 1976 to 1982 (AFGP-2002-600080) and in discussions by Russian revolutionaries about why Russia's first prominent terrorist organization—*Narodnya Volya* (The People's Will)—failed (Naimark, 1983).

[3]See for example the 1977 "Staff Report" for the Provisional Irish Republican Army (PIRA) General Headquarters (GHQ) which details reorganization plans explicitly intended to minimize security vulnerabilities while maintaining sufficient operational control (Twomey, 1977).

solution to the challenge of balancing security and control. Groups as diverse as the Red Brigades, Euskadi Ta Askatasuna (ETA), Aum Shinrikyo, Fatah, al-Qaida, and al-Qa'ida in Iraq generate organizational minutiae and paperwork that, were it not for their violent subject matter, could have come from any traditional organization.

Using relatively simple formal models to carefully analyze the tasks terrorist organizations must achieve in light of their unusual operational environment can help us understand this puzzling pattern.[4] In the context of these models, bureaucracy serve to communicate critical facts up and down the chain of command thereby reducing agency losses. Viewing expense reports and the like as solely providing information to leaders is a mistake. Reporting requirements imposed by leaders in any organization provide subordinates with valuable information about their superiors' priorities and preferences.

To see why such communication is valuable, we can break the terrorist's challenge down into three distinct internal problems. The first problem is that operational terrorists often have different preferences over targets than do terrorist leaders, either because they have different political goals or because they disagree on how best to advance those goals. Leaders can mitigate the impact of this preference divergence by sanctioning operatives who misbehave, but doing so is problematic in terrorist organizations where operatives wield both implicit and explicit threats over their leaders (Shapiro, 2007). The second problem is that terrorist leaders often have different information about the relative political value of targets than do operational elements. Communicating this information has security costs. This information problem means that even if leaders' preferences over targets are perfectly aligned with those of operators, groups may face a tradeoff between security and control when they operate in an uncertain environment.[5] Of course, both problems can be present at the same

[4]The collective purpose of the models in this paper is to probe key organizational dilemmas and communicate some basic hypotheses about how the equilibrium solutions to these dilemmas will vary.

[5]Many groups have noted that different information creates problems, including leftist terrorists operating in Europe in the 1900s (McDaniel, 1976) and foreign insurgents operating today in Iraq (IZ-060316-02).

time. When they are, leaders face a third problem as the conflict of interests over targets makes it harder to credibly communicate about which targets should be hit.

Taken together, these three problems create the fundamental tradeoff between security and control that has troubled terrorist leaders from Lenin to Zarqawi. To better understand how terrorist organizations resolve this tradeoff, the remainder of this paper proceeds as follows. Section 1 explains why preference divergence is so pervasive in terrorist organizations. Section 2 uses two nested models to study terrorists' organizational challenges. Section 2.1 presents a static model of the situation where leaders and operators have different preferences over targets and receive different information about what are the politically optimal targets. Section 2.2 extends the model to a dynamic setting in which conditioning agents is costly for leaders. This extension allows us to explore the relationship between internal discipline and organization. Section 2.3 summarizes the models' predictions about communication and control in terrorist organizations. Section 3 uses internal documents from a portion of al-Qa'ida in Iraq (AQI) to conduct an out-of-sample test of some of these predictions. Section 4 concludes by summarizing the key results and drawing out a few policy implications.

Before proceeding, a short note on sources. Throughout this paper I refer to documents captured from Al-Qa'ida and other groups. These documents were originally contained in the United States Department of Defense's Harmony Database. The database contains more than one million documents captured during operations in Afghanistan, Iraq, and elsewhere. Roughly a quarter of these documents have been fully translated. The documents run the gamut from strategic policy studies, to accounting reports, to membership lists, to technical training manuals, to draft ideological screeds, to letters between family members. A number of organizations have released documents from the Harmony database including the Institute for Defense Analysis, the National Defense University, and the Combating Terrorism Center at the United States Military Academy. Originals and translations of the documents I refer to and many others are available at: http://www.ctc.usma.edu/harmony_docs.asp. Documents

from the Harmony database are referenced by their document number.

1 Why Preference Divergence?

Substantial evidence indicates that members of terrorist groups are not uniformly motivated by the cause, are not equally willing to sacrifice for the cause, often disagree on what the cause is, and rarely see eye-to-eye on the best tactics to achieve their strategic end. Scholars who have done extensive interview work with terrorists report their organizations are torn by strife and disagreement (Bell, 1989; Crawford, 2003). Supporting this view, the Harmony documents are full of sometimes vitriolic letters flying back and forth as members of al-Qa'ida debate ideology, strategy, and tactics.[6] Even when there is no conflict within groups, leaders often engage in costly efforts to monitor their agents, suggesting the potential for disagreement exists.[7] Historically, terrorist groups have repeatedly splintered because of differences of opinion about how to conduct the struggle. The Irish Republican Army, for example, has spawned at least five splinter groups since the mid-1970s. Evidence from Indonesia suggests that the cohesion of Islamist terrorist groups is similarly tenuous, although conflict at the upper levels does not always preclude low-level cooperation (International Crisis Group, 2004, 2005). The Harmony documents reveal a surprising level of infighting and conflict over strategic focus and arcane points of doctrine within al-Qa'ida (Brachman

[6]Hassan al-Tajiki's Third Letter to the Africa Corps is typical. Hassan writes "Here once again I remind you of one of your fatal mistakes, which is the quick changing of strategic targets, whereby now every action is tactical and improvised." AFGP-2002-600053, 25. The best analysis of leadership schisms in al-Qa'ida is Brown (2007).

[7]Some leaders in Jemaah Islamiyah (JI), for example, required members to report their travel expenses in order to know if there was any corruption. That they never had problems with corruption could mean agents did not have different preference from the leaders, or it could mean the monitoring deterred corruption. Author interview, Jakarta, February 20, 2007. Al-Qa'ida units operating in Africa in the early 1990s also had to file detailed expense reports. AFGP-2002-800573.

and McCants, 2006).

To understand why there is so much preference divergence in terrorist groups, it helps to differentiate between divergence in underlying preferences and divergence in induced preferences.Induced preferences are a function of three factors: underlying preferences, the information individuals receive about the world, and beliefs about the actions implied by specific information. People who have exactly the same underlying preferences can have very different induced preferences if they get different information or have different beliefs about how to interpret that information. Focusing on induced preferences suggests several reasons why the very nature of terrorist operations drive preference divergence.

The first reason is that people who are good at violence, those who make ideal recruits as far as their ability to conduct operations, often have underlying preferences for violence which lead them to seek more violence than is politically desirable. Early Marxist militants were the first to document problems caused by the positive correlation between skill at violent action and the amount of violence one finds appropriate. The Russian Social Democratic Labor Party had regular problems in the 1890s and 1900s with local cells conducting revenge attacks that could not be justified by Marxist theory (Geifman, 1993, 1992). Lenin and others noted repeatedly that those recruited for their ability to conduct military operations often pushed for such activities even when not politically advantageous (Newell, 1981). In like fashion, the PIRA suffered repeated problems with Active Service Units (ASU), made up of combat specialists, pushing for violence when the organization as a whole wanted to limit attacks (Bishop and Mallie, 1994; Smith, 1995). Protestant paramilitaries suffered similar problems of control. As a former Ulster Volunteer Force (UVF) bomb-maker put it, "we had some very heinous and counterproductive activities being carried out that the leadership didn't punish because they had to maintain the hearts and minds within the organization."[8] Islamist groups are not immune from these phenomena. JI suffered both

[8]Author interview, David Ervine, March 8, 2006.

excessive violence and criminality because of its recruitment of *preman*, career criminals, to fill out paramilitary units (International Crisis Group, 2003). Problems of excessive violence also created headaches for leaders among the foreign elements of the Iraqi insurgency. In 2005 senior leaders in AQI instructed a cell operating in Ramadi to be more careful in whom they kill or else "...[the] people will start fighting us in the streets" (IZ-060316-02). Even Ayman al-Zawahiri famously noted that AQI's campaign of beheadings in 2004 significantly damaged the foreign insurgents' reputation among Iraqi Sunnis.[9]

The second reason is that those operating underground often receive different information about the political impact of their actions than do their leaders. In larger groups, where the leadership is organizationally isolated from operational cells, or where it is geographically separated from them, these informational differences can be quite problematic. Islamist groups suffered deeply because of this problem during the Syrian jihad from 1979-82, where local cells repeatedly made attacks that the outside leadership opposed given their knowledge about the larger political situation (AFGP-2002-600080). Of course, information differences do not always mean operatives push for more violence. Sometimes the opposite is true as when Hamas' Amman-based leaders authorized attacks by the group's military wing when the Gaza-based political leadership wanted to lay low (Kristiansen, 1999).

The third reason is that the cognitive dynamics of underground organization—isolation from the outside, negative physical incentives to external contacts, excessively strong affective ties, and the like—often lead operational cells to lose touch with their original goals and come to value action in its own right, leading cells to seek more violence than those removed from the situation would like (Bell, 1990). This problem plagued both the Party of the Socialist Revolutionaries (PSR) in the early 1900s (Knight, 1979), the Malaysian Communist Party (MCP) in the early 1950s (Pye, 1956),[10] and the Italian Red Brigades (Della Porta, 1995).

[9]Harmony, Zawahiri letter.

[10]The MCP also had problems enforcing a decision to use more discriminate violence because members who had become skilled at violence naturally wanted to focus on activities they were good at.

Such preference divergence, and the counterproductive violence it engenders, appears to be a serious problem in foreign elements of the Iraqi insurgency. One recently captured letter to a local cell leader admonishes the cell to "Stop the killing of people unless they are spying, military, or police officers. We have to find a secure method because if we continue using the same method, people will start fighting us in the streets."[11] The recipient of this letter was ordered to attend a personal meeting two days later to discuss this and other matters of contention between himself and the leadership.

Two disputes within JI cleanly illustrate these concepts. In early 2004, differences in underlying preferences played a key role in preventing cooperation between two factions in the group. Noordin Mohammad Top, the leader of the break-away faction within JI that has conducted all of the group's major attacks since the first Bali bombing, approached Sarwo Edi Nughroho, a mid-level leader in JI's mainline military wing, looking for support.[12] Noordin asked Sarwo to provide operatives, funds, and explosives to support jihad, which Noordin defined as "...fighting America, and its interests, assets, citizens, and allies, wherever they may be."[13] Sarwo refused because "I have a different understanding of jihad than he does." Essentially, Sarwo did not support Noordin because they had different underlying preferences about what JI should do with its military forces.[14]

Differences in beliefs created conflict in JI over whether the group should get involved in communal violence in Ambon in 1999. In the middle of that year, JI leader Abdullah Sungkar ordered one of JI's regional commands, *Mantiqi II*, to send fighters to Ambon. Members

[11]IZ-060316-02.

[12]For a thorough discussion of Noordin's organization see International Crisis Group (2006).

[13]This account and quotes are taken from Budi Setyarso, Imron Rosyid, and Ivansyah, "The Teacher's New Group," *Tempo*, April 10-16, 2007.

[14]After the first Bali bombing a split emerged within JI between those who felt the time was right for terrorist attacks and those who felt violence was proving counterproductive for the group's long-term goals. The split grew worse over time. Author interviews with journalists, JI members, and government officials, Jakarta, April 10-18, 2007.

of the *mantiqi* resisted doing so for several months because they felt the conflict was not a religious one, but was a political conflict between indigenous Ambonese and newly-arrived migrants from Java.[15] Here, both leaders and operatives saw the same information, but their beliefs about what it meant led to divergent preferences over getting involved.

We thus see three internal dynamics leading to preference divergence over tactics: (1) individuals recruited because of their skills in violence will tend to have an underlying preference for more action, or different actions, than leaders would prefer; (2) leaders and their covert operatives receive different information about the world; and (3) the cognitive dynamics of underground organizations lead operational units to see the world differently than their leaders, typically interpreting the same information as implying more violence. All three result in agency problems. Strategies to solve these problems create security costs, leading to an inescapable tradeoff between security and control.

2 Models of Terrorist Organization

Understanding how terrorist groups will organize under different conditions is easiest if we break their organizational challenge down into two discrete questions. First, how much communication should we expect in groups when leaders and operatives have different preferences over targets and receive different information about the political impact of their actions? Second, how should these expectations depend on the ability of leaders to discipline operatives? This section addresses both questions by modeling the situation where a terrorist leader, B, has to delegate attacks to an operational terrorist, T.

We begin with a static model in which nature generates the ideal target, $\theta \in \mathbb{R}$, from some known distribution of potential targets. B observes θ and T does not. B can send a message, m to T about θ, but doing so increases the risk of being identified by the government and so reduces the probability that the attack succeeds. After receiving B's message, or noting the

[15] Author interviews, Jakarta, February 21 and April 13, 2007.

lack thereof, T decides where to attack. B and T can have different preferences over targets, so that while B's utility is maximized when T attacks θ, T would rather attack $\theta - d$. In this setting, problems of credible signaling are central and lead to a number of interesting predictions about when leaders will accept the security costs entailed in communicating with their operatives.

Of course, the story does not end with communications. Terrorist leaders face the second problem that even if operatives know what leaders want, they may not do it. Historically, efforts to exercise control over terrorist operatives have often entailed sanctioning those who do not behave as leaders would like.[16] Turning to a two-period model that cleanly separates communications from sanctioning allows us to explore how leaders' credibility and reputational concerns impact group organization. Incorporating these elements also helps to illuminate the important distinction between efforts to exercise control—by providing information and possibly sanctioning operatives—and the realized level of control leaders achieve.[17] As we will see, the model makes an explicit argument about how the amount of control leaders need to exercise to get a given closeness of outcomes depends on the extent of preference divergence in a group.This fuller model also provides insight into why actually exercising control can be so problematic in terrorist organizations.

Both models are unusual from an agency theory perspective in that the information advantage is the reverse of that in standard models of delegation.[18] Here the principals are better informed than the agents about the mapping from policies into outcomes.[19] Leaders

[16]For many specic examples, see the outstanding discussion of the mechanics of discipline in the PIRA in McIntyre (2003).

[17]By making predictions about both, this model differ from Baccara and Bar-Isaac (2006). Their models deal only with groups' information structures and do not make predictions about outcomes.

[18]For a review of models of delegation, see Bendor et al. (2001). These models do not fall into either of their two canonical categories.

[19]This will generally be the pattern in covert political groups where the exigencies of underground organization mean agents are ill-informed about the political impact of their actions (Bell, 1989, 1990).

can share information to help their agents achieve better outcomes, but doing so comes at a cost. This information structure is actually quite common in more mundane organizations. Whether by dint of experience, or because they have more developed social networks, senior leaders in many organizations often know more than their subordinates do about how to respond to the environment. This is especially true in organizations like the military where managers work their way up from the lower ranks and so can drawn on more contacts and a much larger experiential base when deciding what to do given ambiguous information. Delegation is not always about who can do a better job, often it is about the exigencies of the situation, or the limited work capacity of bosses. The models thus have theoretical relevance beyond the immediate setting.

As the proofs of key results build on well-estabished results or simple algebra, I relegate them to an online appendix.[20]

2.1 Different Preferences and Different Information

Play in the the static game proceeds through three steps. First nature chooses the B's optimal target θ from a symmetric, unimodal distribution $g(\theta)$ with $var(\theta) = b$. B then observes θ and can send a message m to T. If B sends a message, the probability of the attack succeeding drops from l to h. Taking B's decision into account, T attacks a target a in the support of $g(\theta)$. Utilities are as follows:

$$\begin{aligned} U_B(\theta, a, nc) &= -u(\beta|a-\theta|)l - f(1-l) \\ U_B(\theta, a, c) &= -u(\beta|a-\theta|)h - f(1-h) \\ U_T(\theta, a, nc) &= -v(|a-(\theta-d)|)l - f(1-l) \\ U_T(\theta, a, c) &= -v(|a-(\theta-d)|)h - f(1-h) \end{aligned}$$

Here $u(\cdot)$ and $v(\cdot)$ are increasing functions so that B's expected utility is maximized at

[20] Anonymized appendix provided to reviewers.

$a = \theta$ and T's is maximized at $a = \theta - d$. The intensity of government counterterrorism is captured in the reduced probability of success when B communicates with h being the probability of success when B communicates and l being the probability when B does not. Assume $h < l$ so that telling T about θ makes it more likely government will disrupt the plot.[21] Let f be the cost of failure which may vary in response to political conditions or in response to leaders' risk aversion. Uncertainty over targets is captured by $var(\theta)$. The value of discrimination is reflected in $\beta \in [0, 1]$ which scales the costs to B from hitting the wrong target relative to the costs to T. The greater is β, the more costly it is for B to hit the wrong target. Preference divergence is captured by d.

This simple setting allows us to address three basic questions. First, how will the level of communication vary in response to changes in basic environmental parameters such as the discrimination in violence required by the group's political goals? Second, how will the level of communication vary in response to government counterterrorism? Third, how will the precision of communication vary in response to changes in the amount of preference divergence within a group?[22]

2.1.1 Level of Communication

We can begin by analyzing how the level of communication varies in response to changes in key environmental variables.[23] To simplify the exposition, we'll consider absolute value loss

[21] We can also think of h and l as high and low levels of control. When B exercises low levels of control, T can only rely on her prior beliefs about what targets should be hit but the plot is less vulnerable to government disruption.

[22] All of these variables could be endogenized, though doing so would detract from our central focus. Recent studies of how the strategic interaction between a terrorist group conducting attacks and a government deciding how much to invest in counterterrorism determines the level of government counterterrorism include Bueno de Mesquita and Dickson (2007) and Siqueira and Sandler (2006).

[23] Technically the results in this model address the probability of communication, not the level of communication. The statement that comparative statics about the probability of communication in a single

utilities and $\theta \sim unif[0, b]$ so that larger b indicate greater uncertainty about what targets will serve the political goals.[24]

A Perfect Bayesian Equilibrium in this interaction is a messaging strategy for B, $m(\theta)$, an attack strategy for T, $\alpha(m)$ and off-the-path beliefs for T. As in any such game, there are a wide variety of possible equilibrium.[25] For $d > 0$ there can be no equilibrium in which B signals honestly. Suppose B did signal honestly, setting $m(\theta) = \theta$. Then T's best response would be to play $\alpha(m) = m - d$. But if T played $\alpha(m) = m - d$, then B's best response would be to set $m(\theta) = \theta + d$, and so on. Drawing on the classic Crawford and Sobel (1982) model of cheap talk, however, we can construct an equilibrium based on cutpoints. B can communicate that θ is in a given range and the size of these ranges will depend on the extent of the preference divergence.[26]

Consider the following two cut-point messaging strategy:

$$m(\theta) = \begin{cases} r_1 & \text{if } \theta < \theta_1 \\ nc & \text{if } \theta_1 \leq \theta \leq \theta_2 \\ r_2 & \text{if } \theta_2 < \theta \end{cases}$$

As θ is drawn from a uniform distribution, T's best response function is to hit d below

interaction tell us something about the level of communication expected from a group is a common shorthand in the literature. It implicitly assumes that the interactions modeled here occur repeatedly in a world with stochastic parameters.

[24]Quadratic loss utilities yield the same results but do not admit illustrative algebraic simplifications. For $\theta \sim unif[a, b]$, $var(\theta) = \frac{(b-a)^2}{12}$ which is increasing in b.

[25]For a recent discussion of this issue and one method of solving it, see Olszewski (2006).

[26]A key finding from the Crawford and Sobel paper is that as preference divergence goes to zero, the size of these ranges shrinks to a point and communication becomes completely informative.

the middle of each segment:[27]

$$\alpha(m) = \begin{cases} \frac{\theta_1}{2} - d & \text{if } m = r_1 \\ \frac{\theta_1+\theta_2}{2} - d & \text{if } m = nc \\ \frac{\theta_2+b}{2} - d & \text{if } m = r_2 \end{cases}$$

If T sees any $m \notin \{r_1, nc, r_2\}$ then it believes B is one of the three types, which one does not matter. Since our equilibrium concept places no restrictions on off-the-path beliefs this weak restriction is fine.

At θ_1, the terrorist boss must be indifferent between: (1) paying the security cost of signaling and having T hit d below the mid-point of the lower section; and (2) paying no security cost and having T hit d below the mid-point of the no-communications region. Likewise at θ_2 B must be indifferent between: (1) paying the security cost of signaling and having T hit d below the mid-point of the upper section; and (2) paying no security cost and having T hit d below the mid-point of the no-communications region. Figure 1 illustrates B's utility as a function of θ and the cutpoints.

These two indifference conditions describe a system of two equations in two unknowns which we can solve for θ_1 and θ_2, providing an explicit expression of the distance between them.[28] Noting that the probability of communication is decreasing in this distance leads to the following propositions:

Proposition 1: There are at least two cutpoints whenever preference divergence is sufficiently small given the level of uncertainty about the optimal targets. The range of parameters for which a two-cutpoint equilibrium exists increases as the value of discrimination increases and as uncertainty increases. The range shrinks as preference divergence decreases and as the relative reduction in success probability for communicating gets smaller.

[27]For a generic distribution and symmetric loss function T's best response would be to hit d below the expected value of each segment.

[28]This approach essentially modifies that presented in Osborne (2003).

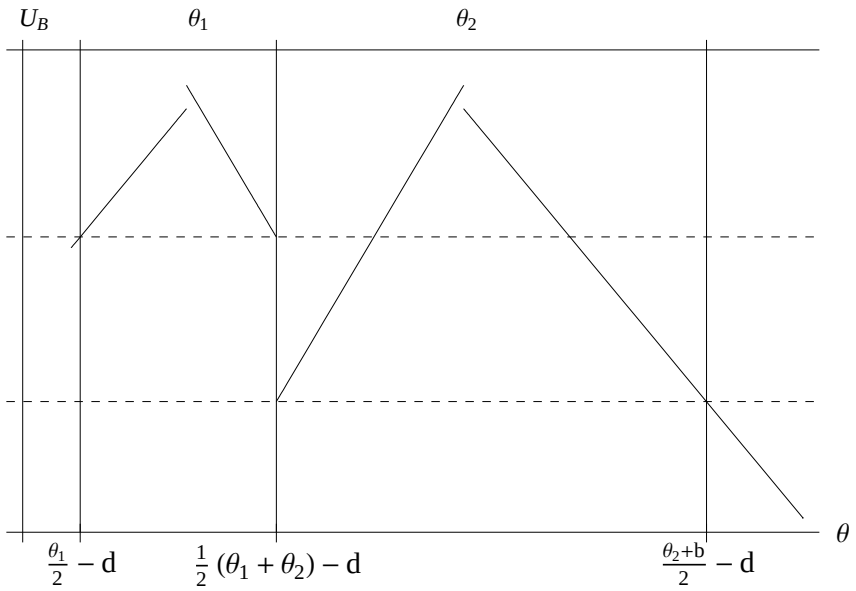

Figure 1: Cutpoint Equilibrium

Proposition 2: When a two-cutpoint equilibrium exists, the probability of communication is decreasing in the cost of failure and in the relative reduction in the success probability when B communicates.[29] The probability of communication is increasing in the level of discrimination required by political goals. The probability of communication is increasing in the uncertainty of the operational environment provided the costs of failure are not too low. The level of preference divergence has no effect on the distance between the cut-points in a two-cutpoint equilibrium.

Two characteristics of the static game merit greater attention: the impact of counterterrorism and the relationship between preference divergence and the precision of communications.

[29] Another way to state this is that the size of the no communication region is increasing.

2.1.2 The impact of government counterterrorism

The marginal impact of an increase in government counterterrorism on the probability an attack succeeds depends on both the specific nature of that increase and on whether B communicates. Target hardening, for example, can be expect to have a similar impact on both l and h while increased spending on communications intercepts would have a greater impact on h. Whether such changes in government strategy increase or decrease communication depends on the relationship between the costs of uncertainty and the costs of failure, on whether success is more important than precision. If it is, then leaders would rather have a high probability of success against the wrong target ($b < \frac{6f}{\beta}$) Otherwise a lower probability of success against the right target ($b > \frac{6f}{\beta}$) is preferred. Proposition 3 formalizes the relationship between government counterterrorism, the level of uncertainty, and the costs of failure for a two-cutpoint equilibrium.

Proposition 3: When $\frac{h'}{h} < \frac{l'}{l}$ so that the proportional decrease in the success probability when B communicates is larger than the proportional decrease in the no-communication success probability, then the probability of communications is decreasing in counterterrorism if $b < \frac{6f}{\beta}$ and increasing otherwise. When $\frac{h'}{h} > \frac{l'}{l}$ the comparative statics are reversed.

These results demonstrate an interesting relationship between uncertainty over targets – here the range of possible targets – and the impact of government counterterrorism. The intuition behind them is that when b gets larger for a fixed number of cutpoints, the sizes of both the no-communication and communication regions increase. Both changes mean that the relative gains to communicating shrink. $b = \frac{6f}{\beta}$ effectively marks the point at which the world is sufficiently certain that the value of communicating outweighs the cost of failure. Once b drops below this point, we find the intuitive relationship that government actions which lower T's success probability proportionally more when T communicates than when he does not make T less likely to communicate. The key point to emphasize is that when the effects of increased government counterterrorism depend on whether B communicates,

increased counterterrorism can actually lead to increased communications.

2.1.3 Preference Divergence and the Precision of Communication

Despite the fact that there is no relationship between d and the size of the no-communication region in a two cut-point equilibrium, there is a sense in which the level of control gets larger as d gets smaller. If d is sufficiently small, then B can partition the signaling spaces below θ_1 and above θ_2 into more than one region, leading to attacks that are closer to θ in expectation. Doing so will also decrease the size of the no-communication region. The following proposition formalizes this insight.

Proposition 4: As d gets small, the number of feasible partitions of the communication regions is weakly increasing and the size of the no-communication region is weakly decreasing. Both the probability of communication and the precision of communication are thus weakly increasing as d gets small.

This result leads naturally to the following remark.

Remark 1: B's *ex ante* expected utility from the *ex ante* pareto-optimal equilibrium is increasing as d gets small.[30]

There are two reasons for this result. First, B does better for a fixed number of partitions with smaller d because T's attacks will be closer to what B would like. Second, B's utility is increasing in the number of partitions of the communications regions. The intuition for this result is that B pays no additional cost for sending more distinct messages in the signaling regions and receives strictly greater payoffs in those regions by doing so.

This result on precision has strong implications for the two-period game. There B's ability to credible threaten T, and so sustain an honest equilibrium in the first period, depends critically on B's expected utility from the second period.

[30]Crawford and Sobel (1982) prove this result for any fixed interval. See appendix for a proof in this context.

2.2 Discipline, Communication, and Control

So far we have discussed how preference divergence between B and T limits B's ability to communicate, thereby lowering the expected level of communication. While the static model provided a rich set of predictions on when and how effectively terrorist leaders will communicate, it left two key questions unanswered: (1) Under what conditions will operatives be motivated by threats; and (2) Under what conditions will leaders' threats be credible?

Terrorists' own observations about command and control suggest there is an interesting interaction between B's ability to punish T for not taking the desired action, the extent to which B will communicate, and the level of control leaders exercise. When asked to explain why the leadership of his group did not do more to restrain wayward operatives, former UVF bomb-maker David Ervine responded "in a military organization the admiral doesn't have to worry about the sailor getting off watch and shooting him. My admiral did have that concern."[31] Ervine's observation suggests that if we want to understand how terrorist organize, it is critical to understand their disciplinary challenge.

To do so, consider a two-period extension of the model presented in section 2.1. After observing T's attack, B can sanction T for poor performance in the first period and thereby motivate better performance. For B's decision about whether to sanction B to depend on T's actions in the first period—and thereby make it possible for the threat of being sanctioned to motivate T—B must be uncertain as to the level of preference divergence. To keep things tractable restrict T to three types: 'good' ($d_g = 0$), 'bad' ($d_b > 0$), and 'acceptable' ($0 < d_a < d_b$). The players' utility functions for each period are the same as in the previous model with second-period payoffs discounted by a common discount factor δ.

Play begins when nature chooses T_1's type. Both players share the belief that T_1 is bad with probability q and good with probability $(1-q)$. The stage game then proceeds as before. Having observed T_1's choice of attacks, B decides whether to sanction T_1, paying cost s if

[31] Author interview, David Ervine, March 8, 2006.

it does. T_1 suffers disutility w from being sanctioned and is not rehired. If B sanctions it provides information to potential recruits about its preferences, so that B is guaranteed an acceptable agent in the second round.[32] In the next round B's preferences remain the same, and the stage game is repeated.

A Subgame Perfect Bayesian Equilibrium in this repeated interaction has four components: (1) a messaging strategy for B in every period conditional on B's prior (q) and posterior ($q_1(\cdot)$) beliefs about T, $m(\theta^1, \theta^2, s(a_1), q, q_1(a_1))$, (2) an attack strategy for every T in every period conditional on B's sanctioning strategy, $\alpha_t^{T_i}(m_1, m_2, s(a_1))$; (3) a sanctioning strategy for B after seeing T_1's attack in the first period, $s(a_1)$; (4) and off-the-path beliefs for B and T at both stages if applicable. Importantly, the core results from the previous section carry through when we introduce uncertainty about T_1's type. B simply takes this uncertainty into account when generating it's messaging strategy by integrating over the distribution of possible types.[33]

In this setting B's communication strategy clearly depends on its ability to force bad T to play as though they are good in the first period. B's ability to do so depends on the set of credible sanctioning strategies, which in turn depend on B's expected outcome in the final period. Given this complexity, one way to get at the relationship between the level of communication and B's disciplinary capacity is to focus on the conditions under which an honest equilibrium can be sustained in the first period.

Because sending messages is costly, an honest equilibrium is one in which B sends an honest message, $m_1(\theta, q) = \theta$ when θ is sufficiently far from $E(\theta)$ that the reduced probability

[32]We can think of this as B establishing a reputation for toughness, or as the pool of T's learning something about B and so those who don't agree leave, rendering the remaining potential recruits more likely to share B's preferences.

[33]B's utility from any message is a strictly concave function of T's type. Because any linear combination of strictly concave functions is a strictly concave function, the method of establishing n cutpoints by solving a system of n indifference conditions in n unkowns still establishes B's equilibrium messaging strategy.

of success from communicating is less costly than having T rely on its prior beliefs. T then attacks exactly where B instructs or at $E(\theta)$ if no message is received—$\alpha(m_1) = m_1$ if T gets a message and $\alpha(nc) = E(\theta)$ otherwise. An honest equilibrium in the first round thus achieves the maximal level of communication in that the no-communications region shrinks to what it would be if B and T had identical preferences over targets. Formally an honest equilibrium entails the following messaging strategy for B in the first period:

$$m_1(\theta) = \begin{cases} \theta & \text{if } \theta < E(\theta) - \frac{(l-h)f}{\beta l} \\ \theta & \text{if } \theta > E(\theta) + \frac{(l-h)f}{\beta l} \\ nc & \text{otherwise} \end{cases}$$

For T's best response function in sustain an honest equilibrium B's sanctioning strategy must guarantee there is no profitable deviation for T. Since the bad T would clearly prefer to deviate from this strategy in any one-shot interaction, two constraints must be met to sustain an honest equilibrium in the first round:

1. Motivational Constraint (MC): T_1 must prefer the honest equilibrium over deviating, being sanctioned, and losing the discounted value of second-period game.
2. Credibility Constraint (CC): T_1 must believe B will sanction if he deviates from the honest equilibrium.

We can rewrite the constraints more formally using some additional notation. $\alpha(m, d)$ indicates the optimal attack strategy given d. $m_t(\theta, q)$ is the optimal messaging strategy in period t given B's prior beliefs over T's type. $m_t(\theta, d_i)$ is the optimal messaging strategy if B believes T to be of type $i \in \{g, b, a\}$. $m_t(\theta, q_1(\alpha_1, m_1))$ is B's optimal messaging strategy given posterior beliefs about T's type. The motivational and credibility constraints amount to the following two equations which respectively define w^*, the smallest w for which MC holds and s^*, the largest s for which CC holds:

$$\begin{aligned} w^* &\geq EU_1^{T_1}(\alpha_1(m_1(\theta, d_g), d_b)) - EU_1^{T_1}(\alpha_1(m_1(\theta, d_g), d_g)) - (1-\delta)EU_2^{T_1}(\alpha(m_2(\theta, q)), d_b), \\ \frac{s^*}{1-\delta} &\leq EU_2^{B}(m_2(\theta, d_a)) - EU_2^{B}(m_2(\theta, q_1(m_1, \alpha_1))). \end{aligned}$$

Carefully inspecting these equations makes it immediately clear that the constraints interact to produce the outcomes we care about. We will therefore work through the logic of the results before discussing how changes in the core parameters of the model impact the likelihood of getting different outcomes.

The motivational constraint depends on three factors: B's messaging strategy, the signal T receives in the first round, and B's sanctioning strategy. The intuition for this dependence is clear from considering the situation where B plays as though he faces a good agent in the first round. If B signals, then T knows θ exactly and so gets the maximum gain in the present period from defecting. If B does not signal—because θ is close to $E(\theta)$—then defecting from an honest equilibrium in the present period is not as good for T as he cannot perfectly absorb the impact of uncertainty over targets. This illustrates a more general principle, the bad T's value from defecting is greater when B's message is more precise. This principle means that the motivational constraint will be met for lower w in the no-communication region. We can therefore identify three motivational regimes: (1) the motivational constraint is met in both regions; (2) the motivational constraint is only met in the no-communication region; and (3) the motivational constraint is never met. Because the last term in the MC equation captures the value to T of being rehired, we need to understand both B's sanctioning strategy and his second-period messaging strategy to fully analyze how the MC works.

The CC is a bit simpler. It is essentially the value to B of guaranteeing himself an acceptable agent in the next round and so clearly depends on whether T's actions are informative as to her type. If the Ts separate, then it is easier to meet the credibility constraint as the expected value of the one-shot game with a bad T is lower than the expected value of the game where T faces a lottery between a bad T and a good T. Lemma 1 formalizes this intuition.

Lemma 1: The highest cost of sanctioning, s^*, for which the credibility constraint is met is greater when the T play true to type the when the Ts pool.

We can thus identify three credibility regimes: (1) the credibility constraint is never met; (2) the credibility constraint is met only when B knows it is facing a bad T; and (3) the credibility constraint is weak enough the B will fire if it remains uncertain as to T's type.

The three regimes each for two constraints lead to 9 possible equilibrium outcomes in pure strategies. Rather than discuss each regime, we can focus on three key qualitative features of the interaction between MC and CC.[34].The first is that whether or not the MC is met depends critically on the stochastic signal B receives about the world. The second is that communication is maximized—in the sense of being most likely in the first period—when B can effectively threaten T, but when doing so is sufficiently costly that B will only fire if he is certain he is facing a bad T. If firing is too inexpensive for B, then the incentives for bad T to play honestly in the first round disappear and an honest equilibrium cannot be sustained. The third is that realized discipline occurs as the interaction of a random event—nature's choice of the ideal target—and the players' strategic incentives. This means that observing disciplinary actions in group X, but not in group Y, provides limited evidence of greater preference divergence in group X. It may simply be that leaders in group Y wield a sufficient threat that their agents toe the line.

With these results in mind, we turn to closer inspection of the two constraints to see how the chances of being in the different regimes vary with the key independent variables.

2.2.1 Motivation

The first condition that must be met for bad T to be willing to follow an honest equilibrium is that deviation not be profitable. This motivational constraint is met when the combined costs of being sanctioned and losing the benefits of the second-round outweigh the present

[34]Table A1 in the appendix describes the equilibrium outcomes for each constraint combination under the assumption that B has off-the-path beliefs which ensure that if the Ts pool, they do so on $\alpha(m, d_g)$. This is not a strong restriction. Good T would always benefit by deviating from pooling on $\alpha(m, d_b)$ as long as B's off the path beliefs do not lead B to fire T upon getting an unexpectedly good outcome in the first period.

gains from defection. T's first period payoff to defecting from an honest equilibrium clearly depends on whether or not B signals. If B sends an honest message, T can attack its own ideal point. If B sends no message, T's utility is maximized by attacking d below the expected value of the no-communication region. We will consider each possibility in turn.

If B communicates honestly, T knows θ and the MC reduces to

$$w^* \geq ld - (1-\delta)EU_2^{T_1}(\alpha(m_2(\theta, q)), d_b).$$

For all the parameters except for the level of preference divergence, taking comparative statics on this amounts to taking comparative statics on the bad T's continuation value in a pooling equilibrium. Doing so leads to the following proposition.

Proposition 5: If B communicates in the first period in a two-cutpoint equilibrium, w^* is increasing in the cost of failure, the uncertainty of the operational environment, and the level of preference divergence for the bad T. w^* is decreasing in the level of discrimination required by political goals.[35]

If B does not communicate, then the situation is a bit more complicated. Because B's messaging strategy in an honest equilibrium neatly bounds the no-communication region we can explicitly solve for T's expected utility from defection. Doing so shows that T's utility from defecting in the first period responds differently to changes in the parameters than its utility in the second period. For example, the first-period value to the bad T of playing true to type is decreasing in f while the second period expected utility given the T pooled is increasing in f. Proposition 6 formalizes how the consequences of changes in the cost of failure on w^* depend on other parameters.

Proposition 6: If B does not communicate in the first period in a two-cutpoint equilibrium, w^* is increasing in the uncertainty of the operational environment and the level of

[35] w^* is insensitive to changes in the probability of getting a bad T because with linear utility functions changes in q do not impact T's second-period utility.

preference divergence. w^* is increasing in the cost of failure as long as the discount factor and level of preference divergence are not too large. w^* is decreasing in the level of discrimination required by political goals as long as the discount factor and level of preference divergence are not too large.

Regardless of whether B communicates, the MC is easier to meet as preference divergence and uncertainty drop.

2.2.2 Credibility

The second condition that must be met for T_1 to be willing to follow the honest equilibrium is that he must believe that B will sanction any first-period deviation. Such a threat is only credible if the discounted gains from getting an acceptable agent in the second round exceed the present cost. Notice that first period outcomes have no impact on this constraint. Because no sanctioning is credible in the second round, the discounted gains amount to the difference in B's expected utility if he gets the better agent in the second round. The credibility constraint thus depends on whether the T pool. We can thus identify two constraints

$$\begin{aligned} \frac{s^*_{sep}}{1-\delta} &\leq EU_2^B(m_2(\theta, d_a)) - EU_2^B(m_2(\theta, d_b)), \\ \frac{s^*_{pool}}{1-\delta} &\leq EU_2^B(m_2(\theta, d_a)) - EU_2^B(m_2(\theta, q)). \end{aligned}$$

Using the implicit solutions for a two-cutpoint equilibrium we can identify how the probability of meeting the constraint varies with the parameters and thereby identify the minimum reputational gain, $d_b - d_a$, for which the credibility constraint is met in a two-cutpoint equilibrium. Lemma 1 shows that $s^*_{sep} < s^*_{pool}$ and so we focus on the former. Because B's signaling strategy depends on the type it is facing, the exact difference in expected utilities is quite complex. What we can do is sign the cross-partial derivatives of B's expected utility with respect to preference divergence and the parameters. Where the cross partial is positive, the gains from getting an acceptable agent are increasing in the parameter, meaning

the minimum reputation gain required to meet CC is decreasing in the parameter. Taking this approach yields the following proposition.

Proposition 7: For a two-cutpoint equilibrium, the minimum reputational gain for which B can credibly threaten to sanction T for deviating from an honest equilibrium in the first round is in decreasing in uncertainty. The minimum reputational gain is also decreasing in the level of discrimination required to meet political goals as long as uncertainty is sufficiently large. The greater the level of preference divergence in first period, the smaller the reputational gain required to make sanctioning credible.

These results make intuitive sense. The value to B of having T attack close to θ is obviously increasing in uncertainty and discrimination. This means the value of getting an acceptable agent in the second period is increasing in uncertainty and discrimination. Getting a slightly better agent yields higher returns the worse the first agent because B's expected utility for the second period is strictly concave in d.

Before moving on, we should not that changes in the precision of communication can have outsized impacts on the expected utility of the second round. This means that if d_b is close to the level that would allow a finer partition of the communication regions, then reputational gains that would be too small to meet the CC in a two-cutpoint equilibrium may suffice to make sanctioning credible. As a practical matter, this finding once again suggests the link between latent preference divergence and observed discipline and is quite complex.

2.2.3 Outcomes of Two-Period Model

Combining the results from the previous two sub-sections shows how the outcomes of the game in a two-cutpoint equilibrium vary depending on which combination of constraints hold. Figure 2 illustrates how the probability of communication varies by regime. Figure 3 shows how the *ex ante* expected level of control in the first period, $-E(|a^* - \theta|)$, does so. In

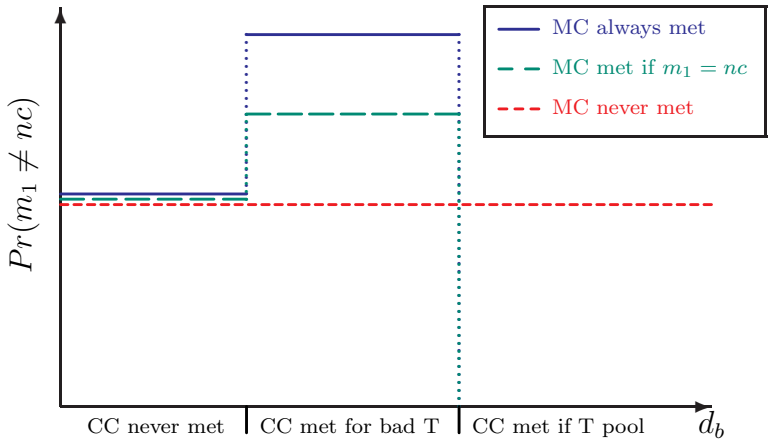

Figure 2: Probability of Communication

both figures the x-axis measures preference divergence between B and the bad T, d_b. Moving right on the x-axis thus equates to increased reputational gain, $d_b - d_a$, for B from firing a bad T. As proposition 7 shows, the larger this gain, the more likely it is that the credibility constraint is met. Each figure has three lines representing the three possible regimes on the MC.[36]

Two features of the outcomes with respect to the level of communication merit further explanation. First, the impact of preference divergence on outcomes is distinctly non-monotonic. Starting from a point at which preferences are sufficiently aligned that CC is never met, the probability of communication is insensitive to the level of preference divergence until it becomes sufficiently large that B would actually want to fire a bad T.[37] At this point B's threat of sanctioning T becomes relevant and the MC can begin to bite. If the MC is met, the probability of communication increases suddenly as B is able to motivate

[36]It is important to keep in mind that an increase in d_b can make it harder to meet the MC. To focus in on the interaction between the constraints, these figures assume the changes in d_b over the range of the figure are insufficient to change the MC regime. Instead the different MC regimes pictured here depend on B's ability to sanction, w.

[37]Recall that both cutpoints shift by the same amount as d changes and so the probability that $m_1 \neq nc$ does not change.

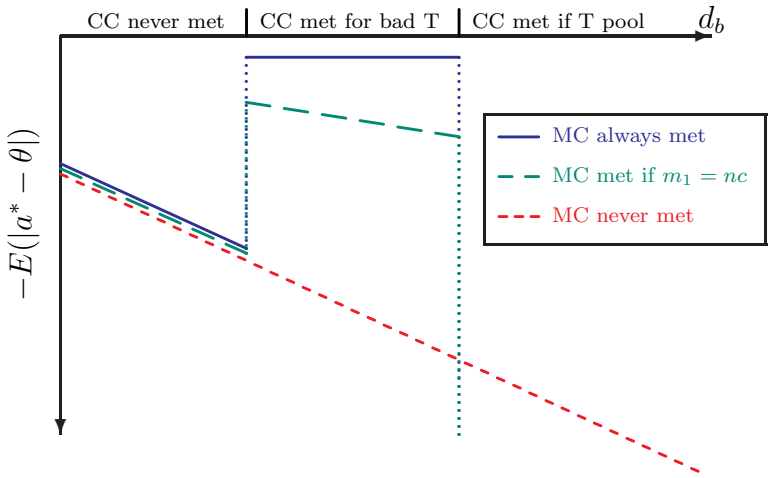

Figure 3: Expected Level of Control

either a pooling equilibrium—if the MC is met regardless of B's message—or a semi-pooling equilibirum—when MC is met only if $m_1 = nc$. As we should expect, communication is maximized when the MC is met regardless of the message from B.

Second, when it becomes too easy for B to punish T, when d_b is large enough that the CC is met even if the T pool, then the there is no informative pure-strategy equilibrium. This is a standard result. If the T pool on $\alpha(m, d_g)$ they are fired, so the bad T would play true to type. But knowing this, B would not fire on seeing a T play $\alpha(m, d_g)$. But then the bad T would want to dissemble if the MC bites. And so on. One way to interpret this result is that being able to threaten helps T, but when B wields too great a threat credible communication becomes impossible and internal communication breaks down.

Turning to the outcomes with respect to control, the key results mirror those for the probability of communication. Starting from low levels of preference divergence, the expected level of control is decreasing until preference divergence becomes large enough that the CC is met for bad T. At that point if the MC is met, B is able to motivate bad T to play as good T, increasing the level of control. If the MC is met regardless of the signal B sends then the level of control becomes insensitive to preference divergence for some time as B can motivate all T to play $\alpha(m, d_g)$ in the first period. However, if the MC is only met when $m_1 = nc$,

then the level of control will be decreasing in d_b in this region, albeit at a slower pace than in the case where MC is never met. Once the level of preference divergence with the bad T becomes so large that the CC is met when the T pool, cooperation within the group breaks down if the MC is met. If the MC is never met, then the CC simply does not matter for a linear utility function.

Pulling these results together suggests an interesting paradox: groups that suffer such high levels of preference divergence that the CC will be met even when the T pool may be better off with limited disciplinary capacity. At least when the MC is never met, informative equilibria are possible with such high levels of preference divergence. As in many organizations, leaders in terrorist groups appear to do best when they can punish their agents but cannot do so too easily.

2.3 Common Results

Certain common threads run through both models. The first is that the costliness of hitting the wrong targets is a key factor. The greater the discrimination required by a group's political goals, the more control leaders will exert. In the models in sections 2.1 and 2.2 this meant that the terrorist boss was more likely to communicate when hitting the wrong targets was costly. This suggests that giving groups a stake in the political process so that they become concerned with maintaining popular support can have two beneficial effects. First, it can cause them to be more discriminate in their attacks; perhaps by focusing on legitimate military targets, thereby sparing civilians lives. Second, it can push them to accept lower levels of security, rendering government's counterterrorism job easier.

The second is that leaders will exercise more control when there is more uncertainty about how specific targets or methods of attack will support political goals. This result should not be surprising, but put more generally it has real bite. The greater the information advantage

of the principals, the more likely they are to exert control.[38]

Finally, we saw that greater preference divergence leads to decreased communication for two reasons. In the static model greater preference divergence makes it harder to credibly signal about the nature of the world. The less B and T agree, the more B will try to mislead T when he does signal. This means the amount of information T can take from the signal is relatively small. In the dynamic model, greater preference divergence meant B needed to wield a greater threat to compel T to behave as he would like. In general, greater preference divergence can be thought of as raising the price of exerting control. Taken together, these results mean that greater preference divergence will lead to less communication and control.

Testing these hypotheses presents something of a challenge in that each group operates in a unique context and there is no obvious reference point from which one could measure relative levels of the independent variables across conflicts.[39] One way to deal with this problem is to employ a within-case design that takes advantage of clear variation over time on one independent variable. The next section conducts such a test using evidence unavailable when these models were developed.

3 The 'Sinjar Organization': An Out-of-Sample Test

This section draws on a sample of 109 recently-released documents that provide a striking insiders view of the management challenges facing al-Qa'ida in Iraq's (AQI) successor organization, the Islamic State of Iraq (ISI). The documents from what I call the 'Sinjar organization' were captured by coalition forces in October 2007 in a raid near Sinjar, along

[38] Additionally, in the face of divergent preference, greater uncertainty can facilitate more precise communication. Greater uncertainty does so discontinuously by allowing B to squeeze in more partitions of the type space.

[39] Another approach is to test these comparative hypotheses by studying conflicts with multiple groups that are similar on two of three group-specific variables. Author (2007) conducts such a test.

Iraq's Syrian border.[40] The majority of the documents (70) are typed or were found on computer files and just over half of the documents (61) are either standardized forms or are handwritten according to a common format. A number of the documents are actually blank standardized forms for everything from tracking weapons (NMEC-2007-657916), to receiving group funds (NMEC-2007-657782), to swearing allegiance to the ISI (NMEC-2007-637854).

The combined documents cover a wide variety of institutional minutiae including: 44 signed pledges by fighters agreeing to conditions for different activities; 43 managerial reports covering personnel, equipment, and finances; 10 internal receipts for funds disbursed to group members; 9 internal policy memos; and 4 press releases. While some of the documents are undated and the exact dates on many others remain obscure, the production dates for 37 can be identified with relative certainty. These documents were produced during the 11-month period from September 2006 through August 2007, and contain information about the Sinjar organization's finances going back to March 2006.[41]

This time span is critical for hypothesis testing. By many accounts AQI's relationship with the local population in Western Iraq began to sour in the summer of 2006 (Kagan, 2007; Roggio, 2007). For the ISI, carefully-targeted violence and the appearance of financial efficiency became critical elements in a struggle for legitimacy in the eyes of the locals.[42]

[40]For a more wide-ranging analysis of these documents see Fishman (2008). Fourteen of the documents explicitly refer to the "Border Emirate" (Imarat al-Hudud) of the ISI. It is not clear when this organizational distinction became official within the group, nor whether what would become the Border Emirate was recognized as a distinct geographical command within AQI. What is clear is that these documents were produced by a cohesive organization with shared personnel across 'official' names, institutional memory, embedded management practices, and permanent salaried employees. I therefore refer to it as the Sinjar organization.

[41]The vast majority of the documents are from the "Border Emirate" of the Islamic State of Iraq (ISI), though the earliest documents were completed shortly before AQI and its allies publicly declared the independent ISI across a swath of Western Iraq on Oct. 15, 2006.

[42]Fishman (2007) analyzes the ISIs pitch to Iraqis and the wider Arab world and Kohlmann (2008)

This meant the ISI had to both control operations and convince the locals it was doing so. In this light, seemingly obscure bureaucratic actions—having the group's financial administrator, Shahin, sign as a witness on the outgoing leader's summary of the group's activities in September 2006, for example—contributed to the ISI's ability to present itself as a disciplined, well-run, financially circumspect organization (NMEC-2007-658016).

If the broad causal claims made in section 2 are correct, several patterns should be clear in these data. First, there should be evidence that ISI leaders are aware of both the political value of exercising control and of the fact that the communications required to do so create security risks. Second, the Sinjar organization should exhibit a non-trivial level of bureaucracy because: (1) controlling violence is critical to its political mission; and (2) it operates with a very diverse membership. Third, under the identifying assumption that the ISI's loss of public support led to an increase in the level of discrimination required by the ISI's political goals, the models predict that documents which can be placed in temporal order relative to each other should show a steady trend towards greater bureaucratization.[43]

The first expected pattern is clearly evident. The group's leaders repeatedly note that they must control their subordinates because doing too much violence and engaging in inappropriate fundraising efforts, can damage their cause (IZ-060316-02; NMEC-2007-637813). It is instructive that the ISI's mid-level managers, and those of AQI in earlier years, had repeated problems with personnel engaging in politically-problematic fundraising that hurt their relationship with the locals.[44] A December 2004 "Threat Announcement" from AQI

discusses the divisions and competition between Sunni insurgent groups.

[43]Recall the models predicted communication would be increasing in the demand for discrimination.

[44]Press reporting on coercive fund raising by foreign fighters include: David Loyn, "Iraq's migrants from violence," *BBC News*, January 10, 2008; Michael R. Gordon, "G.I.'s Forge Sunni Tie in Bid to Squeeze Militants," *New York Times*, July 6, 2007; Vanessa Arrington, "Most likely kidnap victims in Iraq are Iraqis," *Associated Press*, March 30, 2006; and Hamza Hendawi, "Kidnappings fund insurgent attacks," *Houston Chronicle*, February 8, 2006.

highlights the political impact of such actions when it lists the names of people who took part in "mugging and looting" and disavows them in the following harsh terms:

> "Due to the increase of illegal activities from theft and looting from those who claim to be part of Jihadi organizations and have committed such act in the name of their organization, let it be known [that] we at Al Qaeda organization in the land of two river are free from all their wrong doing and their acts we have sworn to expel those whom allow themselves to exploit the blood or the money of the Muslims" (NMEC-2007-637813).

The discussion of 'expelling those who exploit' suggests that the appearance of financial impropriety hurt AQI politically. Indeed, concerns with public perceptions are implicit in a series of later ISI memos that lay down administrative rules which provide leaders the ability to monitor how their operatives spend money.[45] Two requirements from these memos stand out:

1. "For every amount paid out of the Muslim people funds, the recipient is required to provide two signatures, in his own handwriting, one for receiving the money and another one to show how the money was spent."
2. "All properties, small and large, will be inventoried, a report will be kept. We will keep a copy of the report, and all changes will be annotated from the previous report."

These are sensible requirements for most any organization, except for one that is ostensibly covert. For a covert organization, such record keeping presents dramatic security risks and only makes sense if the group's leaders were either: (1) deeply concerned with the costs of operatives engaging in unauthorized behavior; or (2) willing to give up a measure

[45]The list of rules appears in a series of three slightly amended memos, two of which have standardized expense forms tacked on at the end, presumably to ensure that ISI operatives do their paperwork correctly (NMEC-2007-657945, NMEC-2007-657926 and NMEC-2007-657877).

of security in order to be able to demonstrate financial propriety to a skeptical public. Both possibilities reflect a deep concern with the political impact of inappropriate actions.

The leaders of the Sinjar organization are also keenly aware that they operate in an environment where becoming known to Iraqi or American government forces leads to operational failure.[46] One Sinjar organization memo detailing standard reporting procedures specifically admonishes local leaders that:

> "All information is to be uploaded on a [flash USB], and sent every week to the administrator of the Emirate of Borders. Due to the security risk involved, do not keep any information' (NMEC-2007-657926).

Despite such concerns the Sinjar organization persists in using standardized forms,[47] keeping lists of operatives and equipment,[48] demanding regular reports from lower-level units,[49] and sending intra-organizational memos on everything from recommendations for a movie memorializing Abu Musab al-Zarqawi to a meeting agenda with a proposed organizational chart to be discussed.[50] Put simply, these documents show that for at least a one-year period, the Sinjar organization repeatedly sacrificed operational security for managerial control. In doing so, the group's leaders appear to have been motivated by a deep concern with managing public perceptions of their organization.

The second expected pattern, that the Sinjar organization would use bureaucratic tools, is especially prominent in the group's "Fighter Registry Report[s]." These reports provided

[46]This tradeoff is especially stark for organizations such as the ISI whose enemy, the Coalition Forces, can strike anywhere without the constraints security forces face in traditional counter-terrorism operations. The British, for example, could not use indirect fire weapons against Provisional Irish Republican Army (P-IRA) meetings in Belfast in the 1990s.

[47]NMEC-2007-657694, NMEC-2007-658008, NMEC-2007-657694, NMEC-2007-658008.

[48]NMEC-2007-658070, NMEC-2007-657674, NMEC-2007-657911, NMEC-2007-657921.

[49]NMEC-2006-657777, NMEC-2007-657926, NMEC-2007-657945.

[50]NMEC-2007-637123 and NMEC-2007-657738 respectively.

a clearly useful leadership tool.[51] Each details the groups personnel in three categories: "incoming fighters," "permanent Emirate fighters," and "exiting brothers." For incoming fighters the reports provide names, dates of arrival, the work they will do (martyr or fighter), and the assets they brought with them. For permanent fighters, the reports list individuals' names and gives information on "salary" and "work" for a number of individuals. For "exiting brothers" the report gives the individuals' "reason for leaving," who gave permission for them to leave, and their dates of departure. Any human resources manager would want to capture such information. For an organization whose members depend on anonymity for survival, however, such record-keeping is quite literally a disaster waiting to happen.[52]

In many ways this kind of bureaucracy is exactly what we should expect to see in the ISI. As section 1 discussed in detail, people who are good at violence, those who make ideal recruits as far as their eagerness and ability to conduct operations, often have an underlying preference for violence which leads them to seek more violence than is politically desirable. The ISI surely faced this dynamic. Most of the ISIs fighters were foreigners who had traveled long distances to kill and die for a cause they deeply believed in. These fighters came from a wide variety of different countries and different socio-economic backgrounds (Felter and Fishman, 2007). They were exactly the kind of diverse, highly-motivated group of operatives that should be expected to seek out violence and combat for its own sake. Outside of any other considerations, ISI leaders likely had a great need to exercise control simply due to the nature of their operatives.

The documents also provide strong support for the third expected pattern in that docu-

[51]Chronologically these are NMEC-2007-657921, NMEC-2007-657680, NMEC-2007-657775, and NMEC-2007-657927.

[52]Other Islamist terrorist organizations use personnel tracking spreadsheets to monitor and control operatives. At least one manager in Al-Qa'ida (Abu Huthayfa) in the late-1990s called for the establishment of a database on al-Qa'ida members and programs, the goal being to guide the organization and the broader jihadi movement by the study of its people (AFGP-2002-003251).

ments serving similar functions become much more formal and detailed over time. There is an illustrative contrast between two documents that report monetary income, expenses, personnel entering and departing, and equipment status for defined periods. The first document covers the two month period from October 29 to December 21, 2006 (NMEC-2007-657676). The second covers the 15-day period from January 27 to February 10, 2007 (NMEC-2007-657731). The first difference between these two documents is the time period they cover. The later document is one of two bi-monthly reports in the sample, suggesting the reporting frequency required of the Sinjar organization increased over time.[53] The second difference of note is that the while the first document breaks expenses down into large categories—"Guest houses," "Returnable debts," and "The brother's salaries", for example—the second document records each disbursement individually, noting the amount in dinars, recipient, and purpose for each while summarizing the total amount disbursed to each individual in dollars.

A similar increase in bureaucratization can be seen in how the organization tracks its personnel. In October and November 2006 the Sinjar organization reports its personnel activities in relatively informal notes (NMEC-2007-657696; NMEC-2007-657973). In December the personnel memos are typed but still appear ad hoc (NMEC-2007-657941). In January the first "Foreign Fighter Registry Report" appears with a standard format (NMEC-2007-657921). These reports appear with the same format in February (NMEC-2007-657850) and March (NMEC-2007-657775; NMEC-2007-657927). Finally in May 2007 the "Islamic State of Iraq Salah al Deen Province Media Office" issues a memo that nicely illustrates the peak of the Sinjar organization's bureaucratization by direction all "sections and media personnel" to:

> "Requesting [sic] an application for everyone of your soldiers in your district, if they have not done that before, please send them via postal service, in order

[53]The other such report is NMEC-2007-657680.

> for us to have all those who pledge of allegiance. If you did not have enough applications and if you don't have the capabilities for copying please inform our office to secure your needs" (NMEC-2007-637927).

From the perspective of the models presented in section 2, this increase in bureaucratization at a time when the group was under intense security pressure makes sense since the level of discrimination required by the group's goals was increasing.

Two notes of caution must be sounded when deciding how much confidence these documents should provide in my theory. First, these documents do not represent a random sample of all ISI documents captured in Western Iraq from September 2006 on, much less a random sample of all ISI correspondence produced during this time period. Thus we cannot be certain they are representative of the group's internal correspondence. Second, some of the staid managerial tone in the translated documents may be due more to the translators' experiences in highly-bureaucratic government organizations than to the original authors' affinities for Western business jargon. These problems aside, the Sinjar documents examined above provide the best public evidence on the financing and management of one important terrorist organization.

4 Conclusion and Policy Recommendations

I have presented a series of theoretical arguments that terrorist groups should look strikingly familiar from an organizational perspective and provided some detailed evidence that at least one prominent group actually does. It should no longer be surprising that the regulations for an al-Qa'ida guest house in Afghanistan stated that "The guest can't exceed his visit for more than a day and a half unless a special permission from the dispatcher was submitted."[54] The very normality of this kind of bureaucratic minutiae taps into a deeper reality: terrorist

[54] AFGP-2002-600148.

groups have the same strengths as other human organizations and they suffer the same problems. Once we take the unusual characteristics of their production task and operational environment into account, our existing tools of organizational analysis are more than up to the task of understanding these groups.

As a first step in developing such an understanding, I modeled the interaction between terrorist leaders and their subordinates as an agency problem with an unusual information structure; one where the principals were better informed about the impact of actions than their agents. The models addressed three specific phenomena that lead to the security-control tradeoff. First, leaders often have different preferences over targets than do operational elements. Second, leaders often have different information about the relative political value of targets than operational elements. Third, the conflict of interests over targets makes it difficult for leaders to credibly communicate about which targets should be hit.

A common set of hypotheses about the relationship between control and four independent variables emerged from the models. First, the impact of government counterterrorism depends on how it impacts the cost of different actions. When government counterterrorism increases the cost of exercising control relative to the cost of not doing so, it reduces the level of control leaders exert. Otherwise, greater counterterrorism can actually lead to more control. This finding helps explain the interesting historical pattern that many groups respond to government crackdowns by centralizing authority and increasing control over all aspects of organizational activity. Second, the greater the discrimination required by a group's political goals, the more control leaders exert. Third, the more uncertainty there is about how specific targets or methods of attack will support political goals, the more control leaders exert. Fourth, greater preference divergence leads to less control.

With these results in hand, I presented an out-of-sample test using documents from a subset of al-Qa'ida in Iraq, the 'Sinjar Organization', that were not available when the theory was developed. The case study offered strong evidence that terrorist leaders recognize the

tradeoffs embodied in my models and that bureaucracy emerges, at least in part, as the chosen solution to those tradeoffs. Moreover, using the identifying assumption that secular variation in the baseline level of support for the group increased the level of discrimination required to achieve its political goals, the case study supported the prediction that this change would lead to increased efforts by the group's leaders to exercise control over their operatives. These results should give us confidence that the patterns identified are fundamental to any human organization facing the terrorists' challenge.

Government efforts to degrade terrorists' operational capacities can take advantage of these patterns in a number of ways. Examining these in more detail highlights the important fact that the appropriate intervention depends on the endgame policy-makers have in mind. For example, interventions which lead groups to reduce control can actually make ending the conflict harder. Even if the increased pressure makes terrorist leaders want to stop using violence and switch to a purely political strategy, they may no longer be able to force their organization to go along with such a switch. Taking these subtleties into account, here are a few suggestions that follow from this analysis.

Information Operations can Drive Intelligence

Information operations (IO) can generate intelligence in a manner similar to traditional combat operations.[55] One implication of the Sinjar organization's extensive financial record keeping is that publicizing financial improprieties by insurgents can increase their need for record-keeping. This will, on average, increase the intelligence gained from traditional combat operations, driving further operations and providing material to be used in further information campaigns.

[55]The U.S. Army counterinsurgency field manual notes that counterinsurgent "operations produce intelligence that drives subsequent operations." FM 3-24, 1-23.

Create uncertainty about operationally relevant political information

One fundamental challenge of underground organization is staying sufficiently in touch with the community to understand the political impact of operations. If terrorist cells are allowed to choose their own targets to enhance security, the leadership does not necessarily face an unpalatable tradeoff. Problems arise only when cells attack targets that arouse public ire. Thus there is a greater premium on control when operational units are uncertain about the political environment. One of the clear weaknesses of decentralized terrorist organizations is that those who form cells inspired by an ideology such as al-Qa'ida's are often quite ignorant about how their actions will be perceived. The plot to attack New York's John F. Kennedy International Airport that was disrupted in June 2007 provides a striking example on this score. The plot's alleged leader, Russell Defreitas, explained his choice of targets by arguing that "[Americans] love John F. Kennedy. If you hit that, the whole country will be in mourning. It's like you can kill the man twice."[56]

There are three policy implications here. First, information operations that damage jihadi consensus about politics in target states, or about how their actions are perceived in their supporting populations, will worsen agency problems. Such operations can include engaging in discussions on jihadi websites in ways that provoke uncertainty about politics in targeted nations. Second, officials in the U.S. can take advantage of disrupted plots to sow uncertainty. Through September 2007 U.S. officials hyped every terrorist arrest as a major victory, thereby increasing terrorists' beliefs about the political impact of successful attacks, making such attacks more likely. A better strategy would be for officials to sometimes hype arrests and sometimes announce them with a discrete web-only press release. Doing so can

[56]Cara Buckley and William K. Rashbaum, "4 Men Accused of Plot to Blow Up Kennedy Airport Terminals and Fuel Lines," *New York Times*, June 3, 2007. The alleged plotters also believed that by setting the airport fuel farm on fire they could trigger a series of explosion in an underground pipeline. The plotters were apparently unaware that combustion requires both fuel *and* oxygen.

increase political uncertainty, pushing leaders to exert greater levels of control.[57] Third, creating a sufficient level of uncertainty can prevent groups from engaging in terrorism. Greater uncertainty reduces the utility of terrorism. When the utility of terrorism becomes too low relative the value of other activities, groups may choose to use other methods to achieve their political goals.[58]

Take the Endgame Into Account

The interventions identified above are likely to induce changes in terrorists' organizations. The most important consequence of such changes from a policy perspective relate to how much control terrorist leaders can exercise over their operatives. Whether a government should engage in interventions that reduce control depends critically on a number of factors.

The first such factor is the kinds of violence the government is willing to accept. If a government cracks down hard on a group whose leaders valued discrimination and were playing a restraining role, for example, the expected result would be a short term spike in violence. Such spikes may prove politically damaging to the terrorists, as they often have for the groups we have studied, but it is not obvious that political leaders should accept greater levels of indiscriminate violence in exchange for less political impact. Moreover, a strategy that puts so much pressure on a group that leaders loses the ability to control their operatives leaves the government with no way to end the conflict except to kill or imprison everyone in the group.

[57]Of course, such a strategy would be deny federal officials certain political benefits. This is well-studies agency problem between the citizenry, the principal, and government officials, the agents.

[58]Hamas' decision in 1987 to break with the Muslim Brotherhood's long-standing focus on non-violent action, for example, was based exactly on the calculation that the political gains from participating in the violence exceeded those from staying focused on ideological outreach and social service provision (Abu-Amr, 1994, 68-9). Similarly, in 1995 a leading Hamas figure in Gaza, Mahmud al-Zahar argued his group "...must calculate the benefit and cost of continued armed operations. If we can fulfill a goal without violence, we will do so" (Mishal, 2002, 93).

This leads us to the second factor leaders should consider: the endgame. If a government decides that it cannot win militarily and needs a negotiating partner, as the British government did in Northern Ireland, then it may want to take actions that allow leaders to increase control over their operatives.[59] There is a hard balancing act here as reducing security pressure also reduces the incentives for a group to negotiate. The key thing for decision makers to keep in mind is that short-term policies to reduce terrorists' political impact can foreclose long-term options for ending the conflict.

[59]One contributing factor to the failure of the Oslo peace process was that Hamas's political leaders were never able to enforce their preferences over the use of violence. Another factor was the failure of successive Israeli prime ministers to follow through on commitments regarding halting settlement activity because doing so would have meant alienating key members of their ruling coalitions. It is not just the Palestinians' agency problems that doomed Oslo.

References

Ziad Abu-Amr. *Islamic Fundamentalism in the West Bank and Gaza: Muslim Brotherhood and Islamic Jihad*. Bloomington: Indiana University Press, 1994.

Mariagiovanna Baccara and Heski Bar-Isaac. How to Organize Crime. Working paper, New York University, 2006.

J. Bowyer Bell. Aspects of the Dragonworld: Covert Communications and the Rebel Ecosystem. *International Journal of Intelligence and Counterintelligence*, 3(1):15–43, 1989.

J. Bowyer Bell. Revolutionary Dynamics: The Inherent Inefficiency of the Underground. *Terrorism and Political Violence*, 2:193–211, 1990.

Jonathan Bendor, Amihai Glazer, and Thomas H. Hammond. Theories of Delegation. *Annual Review of Political Science*, 4:235–269, 2001.

Patrick Bishop and Eamonn Mallie. *The Provisional IRA*. London: Corgi, 1994.

Jarret Brachman and William McCants. Stealing Al-Qa'ida's Playbook. *Studies in Conflict and Terrorism*, 29(4):309–321, 2006.

Vahid Brown. *Cracks in the Foundation: Leadership Schisms in al-Qa'ida from 1989-2006*. West Point: Combating Terrorism Center, 2007.

Ethan Bueno de Mesquita and Eric S. Dickson. The Propaganda of the Deed: Terrorism, Counterterrorism, and Mobilization. *American Journal of Political Science*, 51(2):364–381, 2007.

Colin Crawford. *Inside the UDA: Volunteers and Violence*. University of Michigan Press, 2003.

Vincent P. Crawford and Joel Sobel. Strategic Information Transmission. *Econometrica*, 50 (6):1431–1451, 1982.

Donatella Della Porta. *Social Movements, Political Violence, and the State: A comparative analysis of Italy and Germany*. New York: Cambridge University Press, 1995.

Joseph H. Felter and Brian Fishman. *Al-Qa'ida's Foreign Fighters in Iraq: A First Look at the Sinjar Records*. Combating Terrorism Center, 2007.

Brian Fishman. Fourth Generation Governance: Sheikh Tamimi defends the Islamic State of Iraq. Technical report, Combating Terrorism Center, March 2007.

Brian Fishman, editor. *Sinjar Two: al-Qa'ida in Iraq's Foreign Recruiting, Finances, and Future*. West Point: Combating Terrorism Center, 2008.

Anna Geifman. Aspects of Early Twentieth-Century Russian Terrorism: The Socialist-Revolutionary Combat Organization. *Terrorism and Political Violence*, 4(2):23–46, 1992.

Anna Geifman. *Thou Shalt Kill: Revolutionary Terrorism in Russia, 1894-1917*. Princeton University Press, 1993.

International Crisis Group. Jemaah Islamiyah in South East Asia: Damaged but Still Dangerous. *ICG Asia Report*, 63, 2003.

International Crisis Group. Indonesia Backgrounder: Why Salafism and Terrorism Mostly Don't Mix. *ICG Asia Report*, 83, 2004.

International Crisis Group. Recycling Militants in Indonesia: Darul Islam and the Australian Embassy Bombing. *ICG Asia Report*, 92, 2005.

International Crisis Group. Terrorism in Indonesia: Noordin's Networks. *Asia Report*, 114, 2006.

Kimberly Kagan. The Anbar Awakening: Displacing al Qaeda from its Stronghold in Western Iraq. *Iraq Report (The Weekly Standard)*, 2007.

Amy Knight. Female Terrorists in the Russian Socialist Revolutionary Party. *Russian Review*, 38(2):139–159, 1979.

Evan F. Kohlmann. State of the Sunni Insurgency in Iraq. Technical report, NEFA Foundation, 2008.

Wendy Kristiansen. Challenge and Counterchallenge: Hamas's Response to Oslo. *Journal of Palestine Studies*, 3:19–36, 1999.

James Frank McDaniel. *Political Assassination and Mass Execution: Terrorism in Revolutionary Russia, 1878-1938*. PhD thesis, University of Michigan, 1976.

Anthony McIntyre. Provisional Republicanism: Internal Politics, Inequities and Modes of Repression. In Fearghal McGarry, editor, *Republicanism in Modern Ireland*, pages 178–198. Dublin: University College Dublin Press, 2003.

Shaul Mishal. How Hamas Thinks. In Joseph Ginat, Edward J. Perkins, and Edwin G. Corr, editors, *The Middle East Peace Process: Vision Versus Reality*, pages 91–103. Norman, O.K.: University of Oklahoma Press, 2002.

Norman M. Naimark. *Terrorists and Social Democrats: The Russian Revolutionary Movement Under Alexander III*. Harvard University Press, 1983.

David Allen Newell. *The Russian Marxist Response to Terrorism: 1878-1917*. PhD thesis, Stanford University, 1981.

Wojciech Olszewski. Rich language and refinements of cheap-talk equilibria. *Journal of Economic Theory*, 128:164–186, 2006.

Martin J. Osborne. *An Introduction to Game Theory*. New York: Oxford University Press, 2003.

Lucian W. Pye. *Guerilla Communism in Malaya*. Princeton: Princeton University Press, 1956.

Bill Roggio. Securing Eastern Anbar Province, June 2007.

Jacob N. Shapiro. The Greedy Terrorist: A Rational Choice Perspective on Terrorist Organizations' Inefficiencies and Vulnerabilities. In Harold Trinkunas and Jeanne K. Giraldo, editors, *Terrorist Financing in Comparative Perspective*. Stanford, C.A.: Stanford University Press, 2007.

Kevin Siqueira and Todd Sandler. Terrorists versus the Government: Strategic Interaction, Support, and Sponsorship. *Journal of Conflict Resolution*, 50(6):878–898, 2006.

M.L.R. Smith. *Fighting for Ireland? The Military Strategy of the Irish Republican Movement*. Routledge, 1995.

Sean Twomey. Staff Report. Technical report, Irish Republican Army, September 1977.