Vanguards in Revolution: Sacrifice, Radicalism, and Coercion[1]

Mehdi Shadmehr[2] Dan Bernhardt[3]

[1]We thank Raphael Boleslavsky, Odilon Camara, Mark Harrison, Patrick Kuhn, Salvatore Nunnari (our 2014 SPSA discussant), Michael Peress, Manuel Santos, seminar participants at the Harris School of Public Policy, University of Chicago, and participants at the 2012 MPSA, Spring 2013 Midwest Economic Theory conferences, and SPSA 2014 for helpful comments.

[2]Department of Economics, University of Miami, 5250 University Dr., Coral Gables, FL 33146. E-mail: shad@miami.edu

[3]Department of Economics, University of Illinois, 1407 W. Gregory Dr., Urbana, IL 61801, and Department of Economics, University of Warwick, Coventry CV4 7AL. E-mail: danber@illinois.edu

Abstract

We analyze strategic interactions between a revolutionary vanguard and a citizen in a setting where vanguard and follower have private information about payoffs from successful revolution. A vanguard sacrifices by initiating a revolt that may not win support, mitigating a follower's risk of revolting alone. We show: (1) The regime may prefer a radical vanguard to a moderately conservative vanguard because even though a radical vanguard initiates more revolt, it generates less following; (2) the vanguard appears more conservative than the follower when a disloyal follower is punished harshly (by the vanguard in a civil war or guerrilla movement settings) or rewarded generously (by the regime); and (3) even when the vanguard can freely punish a non-compliant follower, its incentive to exploit a follower's information limits its use of coercion; but (4) in contrast, the follower may prefer a radical vanguard that always revolts.

1 Introduction

We study the strategic interactions between revolutionary vanguards[1] and citizens in a setting where the merits of changing the status quo are uncertain, and both vanguards and citizens have private information about those merits. Censorship and restrictions on assemblies together with the fear of identification and punishment by the regime sharply limit direct communication between vanguards and citizens resulting in both parties having substantial private, uncommunicated information about the merits of revolution.[2] We characterize the coordination and information aggregation considerations that arise. We study how the vanguards' radicalism affects the likelihood of successful revolution, the subtleties that enter the state's decisions about how to maintain the status quo, and we distinguish between conservative vanguards and those who only *appear* to be conservative. Finally, we investigate the optimal use of coercion by vanguards in settings where the vanguards can punish disloyal citizens.

Our model features a representative vanguard and a representative citizen (the "follower"). The vanguard first decides whether to revolt.[3] The follower observes the vanguard's action and decides whether or not to revolt. For the revolution to succeed, both the vanguard and follower must revolt. Otherwise, the status quo prevails. This framework captures two key features: (1) without the vanguard's tactical and organizational skills, a "spontaneous" protest cannot sustain to change the status quo; and (2) the general population typically plays a more passive role—they may join a

[1] The word vanguard is from the Middle English *vauntgard*, and from the Anglo-French word *avantgarde*, where *avant-* means "front" and *garde* means "guard". By vanguards we mean activists with skills and experience in organizing anti-regime activities. This notion of vanguard resonates with the "professional revolutionaries" advocated by Lenin in his treatise, "What is to be Done?" (Ch. 4), as well as the notions of "revolutionary entrepreneurs" in formal models of revolution (Bueno de Mesquita 2010; Roemer 1985), and "brokers", "social entrepreneurs" and "entrepreneurs of violence" in the social movements literature (Della Porta 1995, p. 108, 195-201; Diani 2003; Tilly and Tarrow 2007, p. 29-31). However, our vanguards need not be committed revolutionaries. What distinguishes them from other citizens is that they have the know-hows of "protest technology"; that is, they are have access to organizational resources (e.g., religious funds or networks) or are endowed with skills that enable them to initiate anti-regime activities such as protests or armed attacks.

[2] An alternative interpretation is that only some citizens can communicate with vanguards, and we consider the interaction between the vanguards and the remaining citizens.

[3] We use the term revolt broadly, including actions to overthrow the current regime as well as actions directed toward changing a status quo policy.

protest, but they do not initiate it. Moreover, whoever revolts alone, gets punished by the regime. Thus, by moving first, the vanguard incurs the risk of being the sole revolter who is punished, but this sacrifice mitigates that risk for the follower.

The vanguard's strategic considerations are subtle: (1) with a more compliant follower, the vanguard faces less risk of a failed revolt; (2) but a more compliant follower also supports change even when his private information indicates that the vanguard is "mistaken", i.e., that a successful protest would be worse than the status quo. This reduces the vanguard's ability to use the citizen's information to protect against outcomes that could be far worse than the costs of revolting alone, for example, protecting the vanguard against a successful revolution that "devours its own", as happened to the Jacobins in the French Revolution. We establish the strategic substitute/complement structure of strategies and show that a unique equilibrium exists.

We then investigate what happens when the vanguard is more (less) radical than the follower. By definition, a radical vanguard derives a higher payoff from a successful revolution than the follower, and hence has more incentive to revolt. A more radical vanguard revolts after worse information about the successful revolution payoff, making the follower less willing to support it: for the follower to revolt, his information about the successful revolution payoff must be better. Thus, a more radical vanguard revolts more, but is less likely to generate a following.

This simple result has important implications for a state that seeks to minimize the probability of a *successful* revolution. Obviously, the state would like to have a very conservative vanguard, one that is reluctant to overturn the status quo. So, too, the state would like to be able to catch and harshly punish any vanguard in a failed revolt—as then a vanguard is very unlikely to initiate a revolt. But what should the state do if it can neither anoint a puppet as a leader, nor punish so harshly? We show that the state may be better off with an extremely radical vanguard than with a moderate vanguard; and the state may be better off not punishing leaders of a failed revolt than punishing moderately. While more radical and less harshly-punished vanguards revolt more, they are also less likely to be followed; and what matters to the state is the probability of *joint* action, and not *unilateral* action. We find that the probability of a successful revolution may be *maximized* by a slightly *conservative* vanguard, or by moderate punishments for a failed revolt. This provides a rationale for why a regime might "forgive" leaders of a failed revolt, or why Assad "has funded and co-operated with al-Qaeda in a complex double game even as the terrorists fight Damascus (*Telegraph*, 20 Jan 2014)." So, too, a state must be careful when deciding how to combine a reward that it gives a potential revolutionary follower for defying a vanguard together with anointing a more conservative vanguard. We illustrate how, when the reward to a disloyal follower is raised, the probability of successful revolt is maximized by a more *conservative* vanguard, i.e., by a vanguard that is *more reluctant* to overturn the status quo.

We then highlight a distinction between a vanguard that is truly radical and one that only *appears* to be radical. A vanguard appears to be more radical than her follower when she would revolt following signals that would not cause her follower to do so. In contrast, a vanguard appears to be more conservative when she sets a higher threshold for revolt than her follower. When the punishment for leading a failed revolt is high enough, a vanguard always appears to be more conservative than her follower. However, as long as this punishment is not too high, a vanguard appears to be more conservative *only* if the follower's payoff for non-support is either (a) high (e.g., the state generously rewards a disloyal follower), or (b) low (e.g., the state punishes indiscriminately).

The intuition for why a vanguard appears more conservative when a follower is harshly punished for failing to revolt is information-based: since the follower fears the consequences of not revolting, he revolts even after very bad signals, causing the vanguard to revolt more selectively in order to avoid "successful" revolutions that turn out badly. The intuition for why a vanguard appears more conservative when a follower is generously rewarded by the regime for disloyalty reflects the risks of miscoordination: the vanguard's fear of punishment is so high that she revolts less than the follower.

The level of a vanguard's radicalism directly affects the follower's payoffs; and the follower's payoff from not supporting the vanguard affects the vanguard. These observations give rise to natural questions: If the follower could choose the vanguard, would he select a radical vanguard that is likely to initiate revolt or a conservative one that seldom does? And, if a vanguard can employ coercion to reduce a follower's payoffs if he does not provide support, how much coercion should she employ?

We first identify the follower's preferred level of radicalism in a vanguard, and study how it varies in different environments. This level of radicalism induces the optimal choice by the vanguard from the follower's perspective of when to revolt. Thus, if their status quo payoffs are equal, the follower prefers a radical vanguard if and only if his reward for not supporting the vanguard exceeds the vanguard's payoff from a failed revolt. Further, the higher is the vanguard's status quo payoff or the higher is the vanguard's punishment, the more reluctant she is to revolt—to unwind these effects, the follower prefers a more radical vanguard. So, too, the follower prefers a more radical vanguard when his status quo payoffs are lower.

The revolutionary vanguard's design of the coercion to employ on her follower reflects different considerations. In settings such as civil wars or guerrilla movements, a vanguard can punish a follower who does not follow her lead. For example, a clandestine armed organization (the vanguard) contemplating an attack on government forces can punish villagers (the follower) who have information about the strength of those forces when they do not cooperate (e.g., revealing the hiding locations of guerrillas to government forces) by setting crops on fire or killing villagers. When choosing coercion, the vanguard internalizes that her follower may have information suggesting that the outcome of their joint action is likely to be very bad; so bad, in fact, that the vanguard would be better off if the follower does not provide support, causing the joint action to fail. This means that the vanguard does not want to coerce her follower too harshly. We show that if the vanguard can freely choose the extent of coercion, she always punishes a "disloyal" follower who does not follow her by just enough that he faces the same payoffs as the vanguard. The logic driving the limited use of coercion is strategic in nature: a vanguard wants the follower to take whatever action the vanguard would were she in his shoes, i.e., were the vanguard taking the follower's decision given his information, which means that the vanguard wants the follower to face the same payoffs. Thus, the more a vanguard sacrifices, the more coercion she wants to use. Ironically, a follower can *benefit* from a vanguard's use of coercion, as it makes the vanguard more willing to initiate revolt.

Extending this reasoning, we show that if, rather than having common beliefs about the informational structure, a vanguard has more confidence in her knowledge than does her follower, then she wants to use harsher coercion. If a follower has less faith in the vanguard's information than does the vanguard itself, he weighs the vanguard's information by less. The vanguard then wants to punish more harshly in order to adjust for her follower's "ignorance". These findings highlight that the source of extreme coercive measures by vanguards is not that the vanguards believe they know far more than their followers, but rather that they believe that their followers do not *understand* how much they know. This also suggests why ideological vanguards, who tend to be overconfident about their knowledge, tend to use more coercive measures.

Literature Review

Bueno de Mesquita (2010) studies the signaling role of vanguards in revolution. There is a vanguard and a continuum of citizens distinguished by their level of "anti-regime sentiments" that captures their *private* payoffs from a successful revolution. A revolution succeeds when the fraction of revolters exceeds a known threshold, but citizens do not know each other's private payoffs from a successful revolution. First, vanguards exert costly efforts to foment violence, which is publicly observed. Then, citizens decide whether or not to revolt. By assumption, the intensity of violence is a noisy public signal of the average level of anti-regime sentiments in the population—for any effort level of the vanguard, violence rises with the average level of anti-regime sentiments in the population. Thus, the vanguard's effort reduces the strategic risk of revolting, and hence can enhance coordination among citizens.[4]

Bueno de Mesquita (2013) focuses on rebel tactics: after observing the level of mobilization in the population, a rebel leader must decide among conventional war, irregular war, or withdrawal. Conventional war has a stronger complementary link

[4]However, the presence of vanguards need not lead to more revolution because the citizens' coordination game features one-sided limit dominance, so that if a citizen believes that others do not revolt, neither does he. Hence, if the vanguards believe that citizens never revolt, they exert a minimal effort.

with mobilization level than irregular war. Thus, irregular war occurs when the outside option is intermediate: when it is low, mobilization is high, and the leader chooses conventional war; when it is high, the leader withdraws from conflict all together. In a dynamic setting, the leader may continue conflict despite low mobilization levels, hoping that future exogenous shocks raise mobilization, i.e., conflict has an option value. Our analysis abstracts from the vanguard's choice of tactics, focusing on coordination and information aggregation issues and their consequences for state repose to dissent, the vanguard's use of coercion, and radicalism. However, loosely speaking, our analysis of a vanguard's use of coercion is related to rebel tactics—but its driving force is information aggregation rather than degrees of complementarity in fighting technology.

Majumdar and Mukand (2008) focus on a leader's information acquisition. A leader and a continuum of individuals must decide whether to participate in a costly action to change the status quo. The likelihood of changing the status quo rises with the measure of participants. First, individuals decide whether to "invest in activism" to reduce future costs of action. Next, the leader exerts costly efforts to learn about the payoffs of change, and reveals it to the population. Then, costs are realized and individuals decide whether to participate. They analyze the complimentary links between individuals' investments, the leader's information acquisition efforts, and participation levels. These complementarities can result in multiple equilibria, including one with no action, and hence give rise to the "threshold effects": when parameters such as the leader's ability to acquire information are below a threshold, in the unique equilibrium, no one acts so that the status quo is maintained. They also identify tradeoffs that arise when the leader's preference is incongruent with the population's: Such a leader has low credibility, but may exert more effort to learn the payoffs from change because, even when change is not beneficial to the individuals, the leader can still gain.

Much of the literature on leadership focuses on the role of "leader as communicator". In Hermalin (1998), the leader is a team member who has more information about the returns to costly efforts, and tries to credibly signal this information to other members to induce them to work harder to produce output that they will share. Komai et al. (2007) build on Hermalin (1998) to show that giving a member exclusive information can improve efficiency by reducing the incentives of others to shirk. Komai and Stegeman (2010) generalize these findings to settings with the possibility of coordination failure, so that a leader's signal can enhance coordination. Andreoni (2006) studies "leadership giving" in charitable fund-raising. A group of individuals distinguished by their income levels must decide how much to donate to a public good with an unknown common value that can be learned at a cost. The individual with highest income learns the value—low, medium, or high—and tries to credibly signal it by his donation choice in order to induce others to donate. Our focus is different: although a vanguard's action is informative about its private information, we do not focus on this signaling role. Instead, our focus is on information aggregation and coordination considerations that arise in equilibrium, and their interactions with the vanguard's radicalism and use of coercion, and the state's decisions to prevent revolution.

A more distant literature studies the tradeoffs between coordination and adaptation. In Dewan and Myatt (2008), the leaders are members of a group (e.g., party activists in a party conference) with noisy private signals about the unknown best action for the group, $\theta \in \mathbb{R}$. Each person sends a noisy public signal to other members who simultaneously choose actions. A member cares both about conformity with the group *and* taking the best action θ: his payoff from taking action $a \in \mathbb{R}$ is $-\pi(a-\theta)^2-(1-\pi)(a-\bar{a})^2$, where $\bar{a}$ is the average of members' actions and $\pi \in (0,1)$ is a weight.[5] Alonso et al. (2008) and Rantakari (2008) also study the tradeoffs between coordination and adaptation in the context of organizational design, in which, unlike Dewan and Myatt, members have conflicting interests. In their papers, a firm has two local managers and a headquarter manager. The best action for the local office $i \in \{1,2\}$ is θ_i, which is the private information of its local manager; but profits also depend on how coordinated local offices are: office i's profit from action a_i is $K_i-(a_i-\theta_i)^2-\delta(a_i-a_j)^2$, where $\delta \in [0,\infty)$ is a weight and K_i is a constant. Under centralization, each local manager sends a (cheap talk) message to the headquarters, which then decides which actions to be taken in local offices. With decentralization, each local manager sends a message to the other, and then each decides the action for her local office.

[5]See Calvert (1992), Dickson (2012), and Wilson and Rhodes (1997) for simpler models and experiments in which the vanguard sends public signals to enhance coordination.

Bolton et al. (2013) analyze the role of a leader's overconfidence and corporate culture in an organization. In their paper, the leader sends a public signal ("mission statement") revealing her original noisy private signal about the unknown state of the world $\theta \in \mathbb{R}$ to employees. Next, each employee $i \in [0, 1]$ receives his own noisy private signal about θ, and takes an action $a_i \in \mathbb{R}$ that the vanguard does not observe. After the employees' actions, the leader receives an additional noisy signal about θ, and decides on an action a_L ("organization's strategy"). An employee i's payoff from taking action a_i is $\Pi_i = -\int_j (a_j - \bar{a})^2 \, dj - (a_i - a_L)^2 - (a_L - \theta)^2$, where $\bar{a}$ is the average employees' actions; the leader's payoff is the sum of the employees' payoffs: $\Pi = \int_j \Pi_j \, dj$. To allow for "bottom-up information flow", they extend the game to allow the leader to observe a noisy signal of the *average* action $\bar{a}$ of employees, yielding multiple equilibria, and hence the role of corporate culture in which equilibrium is selected in reality.

The tradeoff between coordination and adaptation takes a different form in our model: Coordination concerns arise because acting alone is costly, and "adaptation" concerns arise because the best course of action (whether or not to revolt) depends on the unknown revolution payoffs. However, in that literature, there is no information aggregation in equilibrium—i.e., no learning about others' information via equilibrium actions (e.g., winner's curse (Wilson 1977), swing voter's curse or jury decisions (Feddersen and Pesendorfer 1996, 1998)). For example, in Dewan and Myatt, a party activist's strategy given his signal s_i in the party conference, denoted by $A(s_i)$, takes the form $A(s_i) = \pi E[\theta|s_i] + (1 - \pi)E[A(s_j)|s_i]$, implying that an activist only uses his own signal to estimate the common value θ. Further, unlike in the firms or party conferences settings studied in the literature, in our revolution, civil war or protest settings, contracting is implausible and communication is limited.

2 Model

A representative vanguard and a representative citizen (follower), sequentially decide whether or not to revolt. The vanguard moves first and the follower moves second. Figure 1 shows the sequence of moves and payoffs. The payoff θ from a successful revolution is uncertain, and players receive private signals s_1 and s_2 about θ. The other expected payoffs are common knowledge, with $h_i > l_i$. Thus, when only one player revolts, the revolution fails, and the sole revolter incurs a punishment cost.[6]

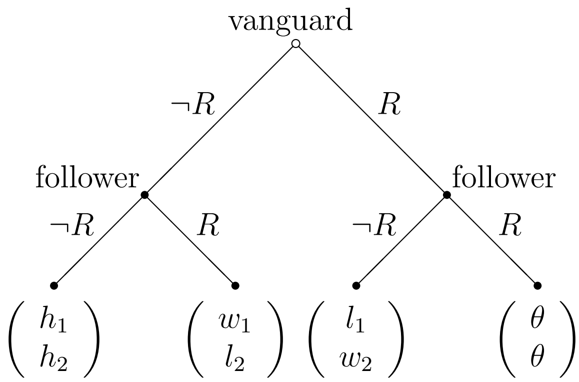

Figure 1: vanguard-follower Game. R indicates revolt and $\neg R$ indicates no revolt. There are miscoordination costs: $l_i < h_i$; and players have private information about the successful revolution payoff θ.

The signals and θ are jointly distributed with a strictly positive, continuously differentiable density $f(s_1, s_2, \theta)$ on $\mathbb{R}^3$. We assume that s_1, s_2 and θ are strictly affiliated, so that, for example, when the vanguard receives a higher signal s_1, the follower is more likely to receive higher signals s_2, and higher payoffs θ from successful revolution are more likely.[7] We impose minimal structure on the tail properties of $f(s_1, s_2, \theta)$: we require that expected revolution payoffs be very high following a very good signal, and very low following a very bad signal. Further, when an agent receives a very good or very bad signal, the other agent is very likely to receive a qualitatively similar very good or bad signal; and the likelihood that the follower's signal exceeds the vanguard's vanishes sufficiently fast as the vanguard's signal becomes very good:

[6]That revolt does not have an action cost is without loss of generality because only the relative (net) expected payoffs of revolting or not enter action choices. If revolt has an expected cost of c, the payoffs in Figure 1 capture it via normalization of w_i, h_i, and l_i. If the players' expected costs differ, i.e., $c_1 \neq c_2$, then the payoffs in Figure 3 capture them via normalization of payoffs.

[7]s_1, s_2 and θ are strictly affiliated if, for all z and $z' \in \mathbb{R}^3$, $f(\min\{z, z'\})f(\max\{z, z'\}) > f(z)f(z')$, where min and max are defined component-wise (Milgrom and Weber 1982, 1097-1100). See de Castro (2010) for a concise review of affiliation.

Assumption A1. For every k, for $i, j = 1, 2$, with $j \neq i$,

$$
\begin{aligned}
&(a) \lim_{s_i \to \infty} E[\theta | s_j < k, s_i] = \infty, \ \lim_{s_i \to -\infty} E[\theta | s_j > k, s_i] = -\infty \\
&(b) \lim_{s_i \to \infty} Pr(s_j > k | s_i) = 1, \ \lim_{s_i \to -\infty} Pr(s_j > k | s_i) = 0 \\
&(c) \lim_{k \to \infty} Pr(s_j > k | k) \ E[\theta | s_j > k, k] = 0 \\
&(d) \lim_{s_1 \to -\infty} Pr(s_2 > k | s_1) E[\theta | k, s_1] > -\infty.
\end{aligned}
$$

A1 obviously holds in an additive, normal noise signal setting where $s_i = \theta + \nu_i$, $i \in \{1, 2\}$, and θ and ν_is are independently distributed normal random variables.[8]

Strategies. A pure strategy for the vanguard is a function ρ_1 mapping her private signal s_1 into an action choice, $a_1 \in \{\neg R, R\}$, where $\neg R$ indicates no revolt and R indicates revolt. That is, $\rho_1 : S \to \{\neg R, R\}$. A pure strategy for the follower is a function ρ_2 mapping his private signal s_2 and the vanguard's action a_1 into an action choice, $a_2 \in \{\neg R, R\}$. That is, $\rho_2 : S \times \{\neg R, R\} \to \{\neg R, R\}$. The equilibrium concept is Perfect Bayesian Equilibrium.

3 Equilibrim

Strategic Forces. We start by identifying the forces that shape the incentives of the vanguard and follower to revolt. In the subgame in which the vanguard does not revolt, the positive punishment cost $\mu_2 \equiv h_2 - l_2 > 0$ means that the follower has a dominant strategy to do the same. Lemma 4 in the Appendix shows that when the vanguard revolts, the follower's best response takes a cutoff form in which he revolts whenever his signal s_2 exceeds a threshold k_2 that depends on the vanguard's strategy. Lemma 5 in the Appendix shows that if the follower's strategy takes a cutoff form, then so does the vanguard's strategy in equilibrium. That is, the vanguard revolts whenever $s_1 \geq k_1$ for some k_1. These lemmas imply that in any equilibrium in which the vanguard sometimes revolts, both agents adopt cutoff strategies. Lemma 1 identifies the strategic forces that shape a follower's incentive to revolt.

[8] In most revolution settings, the possibilities of pre-play communication are slim as regimes monitor such activities, limiting the channels of communication, creating a risk of premature revelation of intent to the regime. Hence, we do not allow for pre-play communication.

Lemma 1 *As the vanguard becomes more willing to revolt, the follower revolts less. That is, the follower's best response always features strategic substitutes:* $\frac{\partial k_2(k_1)}{\partial k_1} < 0$.

A vanguard who is more willing to revolt does so following worse signals about the payoff from revolution—her cutoff k_1 is lower. This lowers a follower's forecast of the payoff from successful revolution—$E[\theta|s_1 \geq k_1, s_2]$ falls—reducing his incentive to revolt.

The vanguard's strategic calculation is more complex, as it involves both coordination and information aggregation elements. Unlike the follower, the vanguard must decide whether to revolt without knowing whether she will be joined; and if the vanguard is the sole challenger to the regime, she expects to be punished. Thus, the vanguard faces a type of cost that her follower does not: miscoordination costs, $\mu_1 \equiv h_1 - l_1 > 0$, that she pays when she revolts, but her follower does not.[9] With a more compliant follower (someone who is more likely to follow the vanguard), the vanguard is less likely to incur the costs of miscoordination, and hence has more incentive to revolt. This provides a force for strategic complements in the vanguard's calculations.

However, as a follower grows more compliant it also means that he follows the vanguard after receiving worse signals. This reduces the vanguard's expectation of the payoff from successful revolution, i.e., $E[\theta|s_2 > k_2, s_1]$ falls, reducing her incentive to revolt. That is, excessive compliance by a follower deprives the vanguard of effectively aggregating the follower's information, information that can protect against a successful revolution that results in much worse outcomes than the status quo. This constitutes the force for strategic substitutes in the vanguard's calculations.

[9] While the follower faces miscoordination costs $\mu_2 = h_2 - l_2$ if he revolts alone, he never does so in equilibrium because he observes the vanguard's action before deciding whether to revolt. Thus, he has a dominant strategy not to revolt (to avoid μ_2) when the vanguard does not. Our results extend directly if there is a chance that revolution fails even when both players revolt: this amounts to a renormalization of payoffs. Also, our results extend qualitatively if there is a small probability p that a revolution succeeds even if only one party revolts. When p is positive and the vanguard does not revolt, then the follower will revolt whenever his signal about the payoff from successful revolution is so high that it offsets the low probability of success and the negative news conveyed by the vanguard's decision not to revolt. As long as $w_2 - h_2$ is not excessively large, the follower sets a higher cutoff for revolt when the vanguard does not act, reflecting that the vanguard's decision not to revolt conveys negative information about the expected payoffs from successful revolt, the low likelihood of success, and the high likelihood of being punished.

Lemma 2 shows that with a barely-compliant follower who rarely follows the vanguard, the force for strategic complements dominates: as the follower becomes more likely to follow, the vanguard revolts more. However, as the follower grows more compliant, the force for strategic substitutes rises relative to that for strategic complements. In fact, there is a unique threshold on the follower's level of compliance after which the force for strategic substitutes dominates: thereafter, as her follower grows more compliant, the vanguard revolts less.

Lemma 2 *Suppose* **A1** *holds. Then there is a critical level k^* of the follower's cutoff that determines whether a vanguard's best response features strategic complements or substitutes: if a follower is unlikely to revolt, so $k_2 > k^*$, the vanguard's best response features strategic complements; if, instead, $k_2 < k^*$, it features strategic substitutes.*

If the vanguard's best response featured global strategic complements, equilibrium would necessarily be unique—$k_1(k_2)$ would be strictly increasing, and $k_2(k_1)$ is strictly decreasing. However, because a vanguard's best response exhibits strategic substitutes when k_2 is low, multiple equilibria might exist. When w_2 is sufficiently large, the crossing can only occur on the strategic complement part of the vanguard's best response, so equilibrium is unique. To prove uniqueness more generally, we impose Assumption **A2**, which ensures that best responses cross at most once.

Assumption A2. For every x and y,

$$\frac{\partial E[\theta|x, s_2 \geq y]}{\partial x} \frac{\partial E[\theta|s_1 \geq x, y]}{\partial y} > \frac{\partial E[\theta|x, s_2 \geq y]}{\partial y} \frac{\partial E[\theta|s_1 \geq x, y]}{\partial x}.$$

Assumption **A2** states that the conditional expectation of θ is more sensitive to changes in a signal $s_i = x$ than to changes in the cutoff $x \leq s_i$. Lemma 3 shows this assumption holds for the standard signal structure used in the literature.

Lemma 3 *Assumption* ***A2*** *holds with an additive, normal noise signal structure, where $s_i = \theta + \nu_i$ and θ, ν_1 and ν_2 are independently distributed as $\theta \sim N(0, \sigma_0^2)$ and $\nu_i \sim N(0, \sigma_i^2)$.*

This leaves the possibility that, regardless of her signal, a vanguard never revolts. Then, a follower's beliefs cannot be determined via Bayes rule on an off-equilibrium path where the vanguard revolts. We impose a minimal plausibility condition on a follower's off-equilibrium beliefs: if the vanguard revolts, then $E[\theta|a_1 = R, s_2]$ is non-decreasing in the follower's signal s_2, reflecting the affiliation in their signals. With this condition, an equilibrium in which there is never revolution does not exist.

Proposition 1 *Suppose* **A1** *and* **A2** *hold. Then a unique equilibrium exists. In equilibrium, both agents adopt finite cutoff strategies, revolting if and only if their signals about θ are sufficiently high.*

Sacrifice and Coordination. This result contrasts with what happens when there is no vanguard—i.e., when citizens move simultaneously. Shadmehr and Bernhardt (2013) show that without a vanguard, there is always an equilibrium in which citizens never revolt, and this is the sole equilibrium if miscoordination costs are high. Absent a vanguard, coordination breaks down because if each citizen believes that the other does not revolt, then s/he does not revolt to avoid paying the miscoordination costs $h_i - l_i$. Proposition 1 shows that a vanguard always facilitates some coordination on revolution. By moving first, the vanguard mitigates miscoordination costs for the follower, eliminating the force for strategic complements in his considerations that underlies the multiple equilibria, including the one in which all coordination breaks down.

The revolutionary vanguard restores coordination by making a sacrifice: by being the first to challenge a regime, she takes on the risk of being the sole revolter, in which case she is punished, incurring miscoordination costs $\mu_1 = h_1 - l_1$. The level of μ_1 captures the magnitude of the vanguard's sacrifice. By publicly risking this sacrifice, the vanguard eliminates this risk for a follower by ensuring that if he revolts, he will not be alone. This coordination-enhancing effect of a vanguard resonates with the "tipping point" mechanism through which vanguards, by moving first, start a "snowballing process" that transforms "sparks to prairie fire" (Granovetter 1978; Kuran 1991; Lohmann 1994; Schelling 1969, 1971; see Bueno de Mesquita (2010) for a discussion). It contrasts with other coordination-enhancing mechanisms in which vanguards send public signals that reduce strategic risks precisely because they are publicly observed (Bolton et al. 2013; Dewan and Myatt 2007, 2008) and convey information about levels of anti-regime sentiments (Bueno de Mesquita 2010; Lohmann

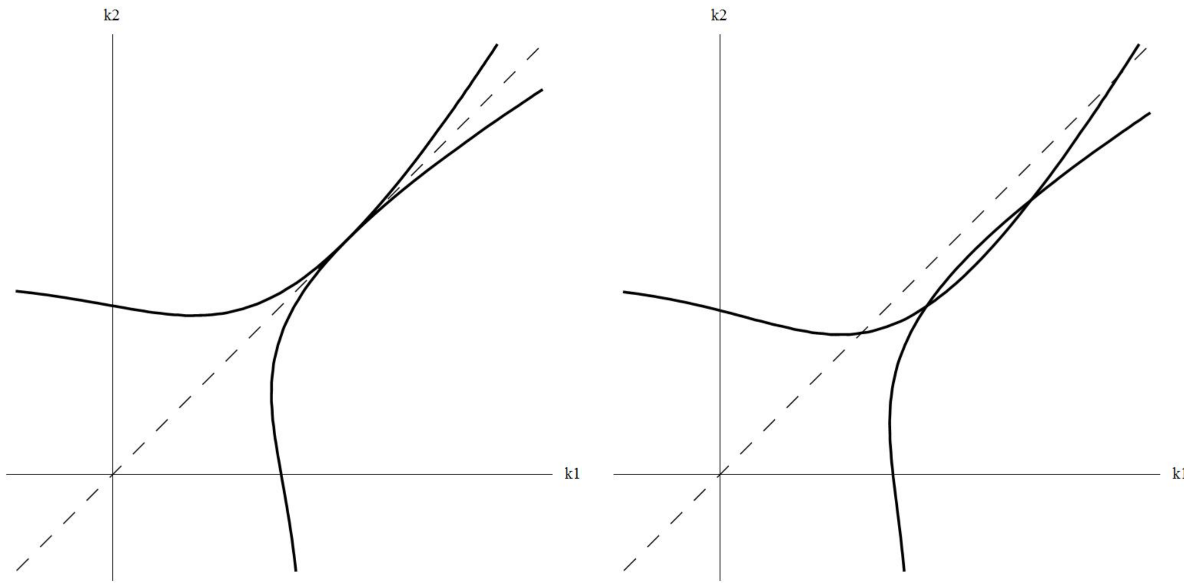

Figure 2: Best response curves when there is no vanguard (simultaneous move game). Left: symmetric miscoordination costs, $\mu_1 = \mu_2 = \mu^* \approx 0.16$. Right: asymmetric miscoordination costs, $\mu_1 = 0.23,\ \mu_2 = 0.09$.

1994) or the returns to effort (Hermalin 1998; see also Majumdar and Mukand 2008).

Such sacrifices play a critical role at facilitating revolt even when there is no vanguard to move first: When the agents move simultaneously, it facilitates coordination if miscoordination costs are divided asymmetrically, so that one citizen expects to bear a greater share of the sacrifice costs. To see why, consider the simultaneous move game analyzed in Shadmehr and Bernhardt (2013). For numerical investigation, we assume that citizens receive noisy, independently normally distributed signals about θ. The following result describes the equilibria of the game in cutoff strategies.

Result. *(Shadmehr and Bernhardt 2013) In the simultaneous move game, an equilibrium always exists in which citizens never revolt. There is a threshold μ^* for the punishment cost μ such that if $\mu > \mu^*$ then only this equilibrium exists. If $\mu = \mu^*$, there is an equilibrium in which citizens revolt with positive probability; and if $\mu < \mu^*$, there are two equilibria in which citizens revolt with positive probability.*

Figure 2 shows that when the punishment cost μ is close to μ^*, then asymmetric divisions of punishment costs, $(\mu_1, \mu_2) = (\mu - \epsilon, \mu + \epsilon)$ facilitate greater coordination.

Quite generally, asymmetric divisions of costs retrieve coordination on revolution when no coordination was originally possible (because $\mu > \mu^*$), and they raise coordination on revolution in cases where there was originally some coordination. That is, with a revolutionary vanguard that takes on a greater share of the risk of punishment, a situation can tip from one in which revolution does not occur to one in which it does.

The reason is that at μ^*, a slight ϵ reduction in μ_1 to $\mu_1 = \mu^* - \epsilon$ causes agent 1 to reduce his cutoff for revolting slightly, reducing the miscoordination risk for agent 2; and a slight ϵ increase in μ_2 to $\mu_2 = \mu^* + \epsilon$ causes agent 2 to raise his cutoff slightly, raising the miscoordination risk for agent 1. However, the asymmetry in miscoordination costs means that the high miscoordination cost agent 2 benefits more from the reduced miscoordination risk (giving rise to a relatively greater increased willingness to revolt), and the low miscoordination cost player 1 is hurt less by the increased miscoordination risk (so its reduction in willingness to revolt is less). As a result, greater coordination on revolution is sustained. More generally, more asymmetric (higher ϵ) divisions of miscoordination costs facilitate ever greater coordination on revolution.

4 Radicalism

Radicalism (conservatism) refers to the intensity of the preference to change (maintain) the status quo. To study the effects of the vanguard's level of radicalism, we modify the vanguard's payoffs from successful revolution, so that the vanguard receives $\theta + z$ when revolution succeeds—see Figure 3. Thus, z captures the relative extent of the vanguard's radicalism: a vanguard is more radical than her follower if $z > 0$, and she is more conservative if $z < 0$. A more radical vanguard revolts more often, which means that she revolts after receiving worse signals about θ, i.e., k_1 falls. This reduces her follower's incentive to revolt because she lowers the follower's estimate of θ, $E[\theta | s_1 \geq k_1, s_2]$. Formally,

Proposition 2 *Suppose that* **A1** *and* **A2** *hold. Then, the more radical is the revolutionary vanguard, the more likely she is to revolt, and the less likely is her follower:*

$$\frac{\partial \bar{k}_1(z)}{\partial z} < 0 < \frac{\partial \bar{k}_2(z)}{\partial z},$$

where $\overline{k}_1$ and $\overline{k}_2$ are the equilibrium cutoffs.

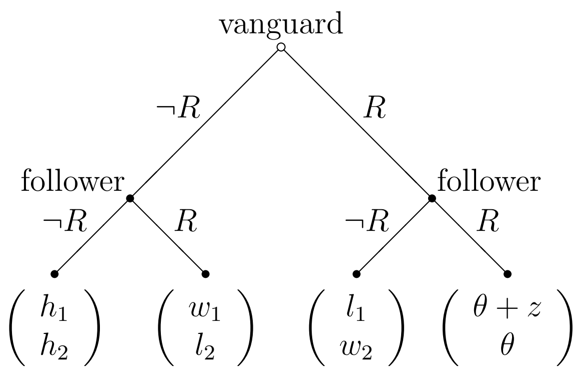

Figure 3: Radicalism. z captures a vanguard's radicalism: $z > 0$ ($z < 0$) indicates that the vanguard is more radical (conservative) than the follower.

Proposition 2 highlights important tradeoffs for a regime. One might think that greater rewards to a follower for defying a revolutionary vanguard, harsher punishments for failed revolt and more conservative vanguards are complementary tools for the state, and that increasing any of these measures always has value. Indeed, if a regime can punish failed revolters extremely harshly at a minimal expense, or radically raise the reward w_2 to a follower for defying his vanguard, or "deradicalize" a vanguard by decreasing z sufficiently then it can reduce the probability of successful revolt almost to zero.

However, when this is not possible, the state must be more careful. The state does not care about the probability of revolt, per se, but rather the probability that a revolt *succeeds*, and successful revolt requires both the vanguard and follower to act. As a result, to reduce the probability of successful revolt, rather than anoint a modestly conservative citizen as a puppet vanguard of the opposition (e.g., a union vanguard), the the regime may do better to *radicalize* the vanguard. A vanguard with a larger z is *more* eager to revolt, delegitimizing her in the eyes of a potential follower.

Increasing z has conflicting effects: (1) the direct, non-strategic effect is to increase the vanguard's incentive to revolt; (2) but the indirect, strategic effect is to *decrease* the follower's incentive to follow the vanguard. This strategic effect further feeds back into the revolutionary vanguard's strategic considerations, which can mitigate (amplify) the direct effect when the vanguard's best response features strategic complements (substitutes). When the strategic effect dominates, the likelihood of successful revolution falls as the vanguard become more radical—see Figure 4. Let P be the probability of successful revolution, i.e., $P = Pr(s_1 \geq k_1, s_2 \geq k_2)$. Then,

$$\frac{dP}{dz} = \left(\frac{\partial P}{\partial k_1} + \frac{\partial P}{\partial k_2}\frac{\partial k_2}{\partial k_1}\right) \frac{dk_1}{dz}.$$

At the interior maximum where the probability of successful revolution is highest,

$$-\frac{\partial P}{\partial k_1} \Big/ \frac{\partial P}{\partial k_2} = \frac{\partial k_2}{\partial k_1}. \tag{1}$$

The left-hand side is the marginal rate of substitution between the vanguard's will-

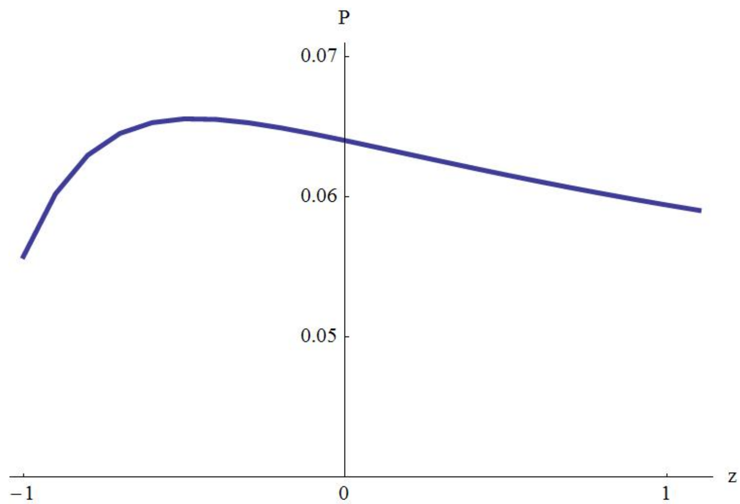

Figure 4: Probability of successful revolution, $Pr(s_1 \geq \overline{k}_1(z), s_2 \geq \overline{k}_2(z))$, as a function of the revolutionary vanguard's radicalism z. $s_i = \theta + \epsilon_i$, $\theta \sim N(0, 1)$, $\epsilon_i \sim N(0, 1)$, $h_1 = 0.1$, $l_1 = 0$, $w_2 = 1.2$.

ingness to revolt and the follower's willingness, while the right-hand side is the slope of the follower's best response curve. Equation (1) shows that as the revolutionary vanguard's cutoff k_1 is varied, the likelihood of successful revolution is highest when the marginal rate of substitution in the probability of successful revolution equals the slope of the follower's best response.

As Figure 4 shows, when a follower's predatory payoff w_2 is high, the probability of a successful revolution may be *highest* with a slightly *conservative* vanguard,

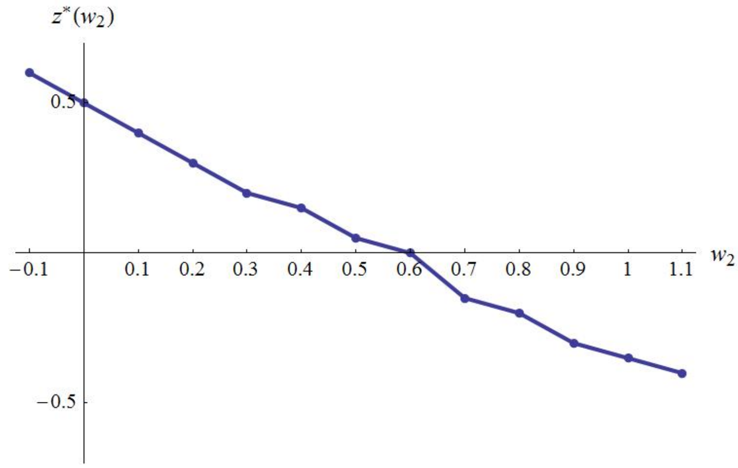

Figure 5: The revolutionary vanguard's level of radicalism $z^*(w_2)$ that maximizes the probability of successful revolution as a function of the follower's disloyalty rewards w_2. $s_i = \theta + \epsilon_i$, $\theta \sim N(0,1)$, $\epsilon_i \sim N(0,1)$, $h_1 = 0.1$, $l_1 = 0$.

whose payoff from successful revolt is *less* than her follower's. In turn, facing such a conservative vanguard, the state may want to *reduce* its punishment of a failed revolt, in order to make the vanguard more willing to revolt. Paradoxically, the vanguard's increased willingness to revolt makes the follower sufficiently less willing to follow that it reduces the likelihood that the regime is overthrown. In fact, Figure 5 numerically illustrates how with normally distributed uncertainty, when the state *raises* the follower's predatory payoff w_2, the level of the vanguard's radicalism that maximizes the likelihood of successful revolution *falls*. Thus, a regime must be wary about combining the twin tools of more generously rewarding a follower who turns on his vanguard, and of also punishing the vanguard of a failed revolt somewhat more harshly or anointing a more conservative puppet vanguard. Doing so can backfire and increase the probability of successful revolt.

5 Who is more willing to revolt?

When deciding whether to revolt, only the vanguard risks being punished in equilibrium. As a result, one may initially conjecture that unless a vanguard is more radical than her follower, i.e., unless $z > 0$, then she needs a more promising signal than her follower to revolt, i.e., $\overline{k}_1(z) > \overline{k}_2(z)$. However, further reflection reveals that the follower's payoff w_2 from *preventing* change by not supporting the vanguard must enter this calculation. When $w_2 < l_1$, the follower's payoff from preventing revolt is even worse than the vanguard's "punishment" when this happens. This might lead one to conjecture that when the quality of a vanguard's information is the same as its follower's, if $w_2 < l_1$, then a non-radical vanguard may be more willing to revolt than her follower, i.e., $w_2 < l_1$ could imply $\overline{k}_2(z = 0) > \overline{k}_1(z = 0)$. In fact, the *opposite* is true. Proposition 3 establishes that a necessary condition for a non-radical vanguard to revolt more than her follower is that $w_2 > h_1$, and hence $w_2 > l_1$. More generally, we establish that a non-radical vanguard revolts more than her follower if and only if the follower's predatory payoff w_2 is high, but not too high.

When the follower's predatory payoff w_2 is very low, the follower almost always supports the vanguard, but the vanguard revolts only when her signal is sufficiently high: $E[\theta|s_1] > h_1$. As w_2 increases, the follower tends to revolt less. Then, the strategic complements force acts to decrease the vanguard's incentive to revolt, while the strategic substitutes force acts to raise it. When a vanguard's status quo payoff is high, $h_1 \geq h_1^*$, the vanguard has a lot to lose from revolting alone. As a result, the strategic complements force is so strong that a vanguard never revolts more that her follower. When, instead, $h_1 < h_1^*$, the strategic complements force is less strong. As the follower's predatory payoff w_2 increases, he revolts less, and just as w_2 passes a threshold $\underline{w}(\mu_1)$, the vanguard's incentive to revolt exceeds her follower's. But as w_2 further increases, the follower revolts even less, and eventually, the strategic complements force dominates (Lemma 2), so that $\frac{\partial k_1(k_2)}{\partial k_2} > 0$. In addition, if the degree of strategic complementarities rises as the follower revolts less, it eventually reaches a level that the vanguard's incentive to revolt falls even *more* than her follower's with increases in w_2. Consequently, once w_2 passes a threshold $\overline{w}(\mu_1) > \underline{w}(\mu_1)$, the vanguard's incentive to revolt again becomes less than her follower's. Assumption **A3** guarantees that when the strategic complements force dominates, as the follower revolts even less, the degree of strategic complementarities in the vanguard's best responses grows. This assumption is satisfied by the standard additive noise, normal signal structure.

Assumption A3. The strategic complements segment of the vanguard's best response is convex, i.e., if $\frac{\partial k_1(k_2)}{\partial k_2} > 0$, then $\frac{\partial^2 k_1(k_2)}{\partial (k_2)^2} > 0$.

Proposition 3 *Suppose that* **A1**, **A2** *and* **A3** *hold.*[10] *Then*

- *There exist a threshold h_1^* on a vanguard's status quo payoff and thresholds $\overline{w}(h_1) > \underline{w}(h_1)$ on the follower's predatory payoff w_2 such that the vanguard sets a lower cutoff for revolt than her follower if and only if $h_1 < h_1^*$ and $\underline{w}(h_1) < w_2 < \overline{w}(h_1)$.*
- *$\underline{w}(h_1)$ is strictly increasing in h_1, while $\overline{w}(h_1)$ is strictly decreasing in h_1.*
- *If the signals of the vanguard and the follower have identical distributions, i.e., if $s_1 \sim s_2$, then $\underline{w}(h_1) > h_1$.*

When $w_2 > h_1$, the follower faces a higher opportunity cost of revolting. This makes the vanguard more likely to revolt than the follower as long as the likelihood of miscoordination, and its associated costs, which the vanguard alone incurs, are not too great. With symmetric signal structures, as $h_1 \to l_1$, we have $\overline{w}(h_1) \to \infty$ and $\underline{w}(h_1) \to h_1$: whether the vanguard is more eager to revolt than her follower boils down to a comparison of what they receive when they do not revolt.

6 What do vanguards and followers want in each other?

If a follower could select the vanguard, would he choose a radical vanguard that often initiates revolt or a conservative one that seldom does? Conversely, if a vanguard could punish her follower when he does not follow her lead, would she punish him harshly or mildly? In this section, we characterize a follower's preferred level of radicalism in a vanguard and the vanguard's choice of the coercion with which to threaten a follower.

[10] A weaker sufficient condition (implied by **A3**) for the results in Proposition 3 can be imposed on the vanguard's net expected payoff from revolting when she receives a signal that just equals the cutoff k_2 set by her follower. The vanguard is more willing to revolt than her follower if and only if at an equilibrium cutoff k_2, her net expected payoff from revolting when $s_1 = k_2$ is positive; and a sufficient condition for the results is that this net expected payoff be a single-peaked function of k_2.

6.1 The follower's preferred level of radicalism

The follower's preferred level of radicalism z induces the optimal choice by the vanguard from his perspective of when to revolt—i.e., it induces the vanguard to select the optimal cutoff from the follower's perspective.

Proposition 4 *Fixing the other parameters, there exists a $\bar{h} < w_2$ such that if $h_2 > \bar{h}$ then the follower's preferred level of radicalism in a vanguard is given by*

$$z^* = \frac{Pr(s_2 < k_2|k_1)\ (w_2 - l_1) - (h_2 - h_1)}{Pr(s_2 \geq k_2|k_1)}, \tag{2}$$

where k_1 and k_2 are endogenous equilibrium cutoffs. If, instead $h_2 < \bar{h}$ then the follower always wants the vanguard to revolt, i.e., $z^* = \infty$.

When w_2 exceeds h_2 sufficiently—when the follower's status quo payoffs are so much lower than his expected payoffs when the vanguard revolts—then because the value of information revealed about θ when the vanguard revolts more selectively is bounded, it follows that the follower prefers an extreme radical vanguard that is sure to revolt. The follower will then support the vanguard when $E[\theta|s_2] > w_2$, and turn on it when the inequality is reversed.

When, instead, his status quo payoffs are not so low, the follower wants a radical vanguard, i.e., $z^* > 0$, if and only if

$$Pr(s_2 < k_2|k_1)\ (w_2 - l_1) - (h_2 - h_1) > 0.$$

For example, if the follower and vanguard have the same status quo payoffs, i.e., $h_1 = h_2$, then the follower prefers a radical vanguard if and only if $w_2 > l_1$. In most settings, $w_2 >> l_1$, i.e., the follower's payoff when he does not support his vanguard in a revolution exceeds the vanguard's payoff from leading a failed revolt. But, it may well be that $h_2 > h_1$, i.e., the vanguard also dislikes the status quo by more than its follower.

To understand how the follower's preferred level of radicalism in a vanguard varies with their other characteristics, recognize that a choice of z amounts to choosing the equilibrium level of the vanguard's cutoff k_1. The optimal cutoff k_1 from the perspective of the follower only depends on his payoffs w_2 and h_2. However, the vanguard's payoffs h_1 and l_1 influence her willingness to revolt, and hence the level of radicalism that the follower seeks in a vanguard.

When the vanguard's status quo payoff, h_1, is higher, the vanguard is more reluctant to revolt. To reduce k_1 back to the follower's preferred level, the follower wants to increase z. Thus, z^* increases in h_1. So, too, when the vanguard's punishment for failed revolt falls, i.e., when l_1 is higher, the vanguard is more willing to revolt. To raise k_1 back to its optimal level, the follower wants to reduce z, i.e., z^* decreases in l_1. The follower also prefers a less radical vanguard when the follower's status quo payoff, h_2 is higher. Increasing the follower's status quo payoff h_2 does not affect the follower's equilibrium cutoff k_2 directly, and hence does not affect the vanguard's cutoff k_1 directly. However, with a higher status quo payoff, the follower wants the vanguard to revolt less. Hence, the follower wants a less radical vanguard, i.e., z^* decreases in h_2.

The effect of raising the follower's predatory payoff w_2 on the optimal level of the vanguard's radicalism is more complicated because the direct effect of raising w_2 is to cause the follower to revolt less, i.e., to increase k_2; and the impact of raising k_2 on k_1 depends on whether the vanguard's best response exhibits strategic complements (k_1 rises) or strategic substitutes (k_1 falls). For example, suppose $h_2 = h_1$ and $w_2 > l_1$, so that $z^* > 0$. Then, z^* is increasing in w_2 if the vanguard's best response features strategic substitutes. To see this, substitute $Pr(s_2 \geq k_2|k_1) = 1 - Pr(s_2 < k_2|k_1)$ into (2) and differentiate with respect to w_2 to get[11]

$$\frac{\partial}{\partial w_2}\left[\frac{(w_2 - l_1) - (h_2 - h_1)}{Pr(s_2 \geq k_2|k_1)} - (w_2 - l_1)\right]$$

$$= \frac{1}{Pr(s_2 \geq k_2|k_1)} - \frac{(w_2 - l_1) - (h_2 - h_1)}{[Pr(s_2 \geq k_2|k_1)]^2} \, \frac{dPr(s_2 \geq k_2|k_1)}{dw_2} - 1$$

$$= \frac{1 - Pr(s_2 \geq k_2|k_1) - (z^* + w_2 - l_1)\, \frac{dPr(s_2 \geq k_2|k_1)}{dw_2}}{Pr(s_2 \geq k_2|k_1)}, \tag{3}$$

[11]By the Implicit Function Theorem, for any parameter x, e.g., $x = w_2$, at an interior optimum z^* that maximizes $E[U_2]$, we have $\frac{dz^*}{dx} = -\frac{\partial}{\partial x}\frac{\partial E[U_2]}{\partial z^*} \Big/ \frac{\partial^2 E[U_2]}{\partial z^{*2}}$. Since $\frac{\partial^2 E[U_2]}{\partial z^{*2}} < 0$, $\text{sign}\left(\frac{dz^*}{dx}\right) = \text{sign}\left(\frac{\partial}{\partial x}\frac{\partial E[U_2]}{\partial z^*}\right)$. One can show that $\text{sign}\left(\frac{\partial}{\partial x}\frac{\partial E[U_2]}{\partial z^*}\right)$ is given by the sign of the partial derivative of (2) with respect to x.

where we have substituted for $z^* = \frac{(w_2-l_1)-(h_2-h_1)}{Pr(s_2 \geq k_2|k_1)} - (w_2 - l_1)$ in the last equality. When w_2 increases, k_2 increases, and if the vanguard's best response features strategic substitutes, k_1 falls, which implies $\frac{dPr(s_2 \geq k_2|k_1)}{dw_2} < 0$. Thus, (3) is positive.

6.2 Coercion

In settings such as civil wars or guerrilla movements, a vanguard can punish a follower who does not follow her lead. For example, consider a guerrilla organization (the vanguard) deciding whether to attack a government military post near a village. There is limited direct communication between the guerrillas and villagers to avoid identification and punishment by the state. The attack succeeds if the villagers (the follower) cooperate, e.g., do not reveal the likely hiding locations of guerrillas to government forces (Kalyvas 2006; Wood 2003). In this context, if the guerrillas do not attack, the status quo prevails as the villagers are not equipped to fight the government. If the guerrillas attack, but the villagers do not cooperate, the guerrillas incur costs, e.g., loosing some members to the government forces, but the organization can still punish the villagers, e.g., they can set crops on fire or kill or kidnap some villagers.[12] We now focus on the vanguard's use of punishment for non-compliant followers. When we say that a vanguard uses coercion, we mean that she imposes sanctions on those who do not follow her: by using coercion, a vanguard decreases her follower's payoff when the vanguard revolts, but the follower does not, reducing it from w_2 to $w_2 - c$. By threatening a very large c, a vanguard can coerce her follower to almost always follow her lead.

How much coercion should a vanguard employ? One might posit that, if possible, a vanguard should coerce her follower into almost always following her lead, in order to eliminate her risk of being punished by the regime. We now show that such high levels of coercion are *not* in the vanguard's interest. By punishing her follower severely whenever he does not follow, a vanguard induces her follower to revolt even when his information suggests that the outcome of successful revolution would be far worse than the status quo, i.e., when $E[\theta|s_2]$ is very low. This, in turn, hurts the vanguard

[12]More generally, as Ahlquist and Levi (2011) discuss in their review of the literature on leadership, in many settings, vanguards can use some coercion to make others follow their lead, and yet "no model so far encapsulates noninformational tools available to vanguards, such as coercion (p. 14)."

by preventing her from effectively using her follower's information. In other words, a vanguard should allow some level of "dissent" in order to make more effective use of a follower's information, thereby improving the vanguard's payoffs.

One might also expect that as the precision of a vanguard's information grows so that she becomes more confident in her information, she feels less need to rely on her follower's information, and hence finds harsher coercive measures optimal. In particular, one might think that if a vanguard's signal is very precise while her follower's is very imprecise, then the vanguard would like to coerce the follower so harshly that he almost always follows her. This reasoning is wrong: the follower *already* accounts for the precision of his vanguard's information in his strategic calculations. Thus, if he believes that the vanguard's information is far more precise than his, then he follows the vanguard as long as his own information suggests that the outcome will not be absolutely disastrous. In fact, the optimal level of coercion is unrelated to the quality of the follower's and vanguard's information:

Proposition 5 *Suppose* **A1** *and* **A2** *hold. If, prior to observing s_1, a vanguard can choose how much to punish her follower when he does not follow, then she chooses $c = \max\{w_2 - l_1, 0\}$.*

If $w_2 - l_1 < 0$, so that a follower gets hurt more than the vanguard when he does not follow her lead, then the vanguard would like to compensate her follower—if she could. This situation may arise in a civil war in which the government uses violence indiscriminately in a region with guerrilla activities. However, in such scenarios it is unlikely that guerrillas can protect the natives who do not cooperate with them. Thus, we focus on the case with $w_2 > l_1$ and hence $c > 0$. The follower's problem is simple—he revolts if $E[\theta | s_1 \geq k_1, s_2] \geq w_2$, and he does not revolt if the inequality is reversed. By having the follower internalize her payoff l_1 when he does not support it, the vanguard induces the follower to make the decision that is optimal from the vanguard's perspective based solely on the *follower's* information. As a result, sometimes the follower will not follow the vanguard, and the vanguard is punished, receiving l_1; however, when this happens the vanguard is protecting herself from leading a successful revolution that the follower's information suggests has an even worse expected payoff.[13]

The result that a vanguard wants to align her follower's preferences with hers contrasts sharply with what a follower wants in the vanguard. Most obviously, when the follower's status quo payoffs h_2 are sufficiently below his predatory payoffs w_2, the follower wants an extremely radical vanguard who always revolts. This allows the follower to avoid the low status quo payoffs and choose between the predatory and revolution payoffs, albeit at the expense of learning less about those revolution payoffs.

The magnitude of a vanguard's sacrifice is captured by the cost $\mu_1 = h_1 - l_1$ she pays when her follower does not support her revolt. Thus, a lower l_1 means that a vanguard sacrifices more. By reducing her follower's payoff for non-support to equal to her own—by setting $c = w_2 - l_1$—a vanguard ensures that if she has to make a greater sacrifice, then so does her follower.

Corollary 1 *Suppose that* **A1** *and* **A2** *hold. Then, the more a vanguard has to sacrifice, the more coercion she wants to use.*

One might naively conjecture that a follower must necessarily be hurt, at least in expectation, when his vanguard coercively threatens a significant punishment for defiance. However, this conjecture is off-base. It ignores the fact that the coercive measure raises the vanguard's confidence that her decision to revolt will be supported, making her more willing to act. As a result, coercion raises the likelihood that the follower receives the revolutionary payoff θ, rather than the status quo payoff h_2. To highlight most transparently that the follower can gain from coercion, suppose that $h_2 = h_1 = h$ so that with optimal coercion, the vanguard and follower's payoffs are the same. Then, when $E[\theta] >> h$ and w_2 is sufficiently large, not only the vanguard, but also the follower, benefit from the vanguard's coercion. When w_2 is very large,

[13]This result has a surface similarity to one in Landier et al. (2009). They show that some "dissent" in organizations can be optimal. They define "dissent" as divergent *preferences* between a decision maker, who chooses a project first, and an implementer, who exerts effort to implement the project chosen. A "dissenting" implementer fosters the use of objective information by the decision maker, which raises the likelihood that the project succeeds, and hence increases the implementer's effort. In our paper, the vanguard does not choose extreme punishment to foster information aggregation, while in Landier et al., "dissent" is optimal because it constrains selfish decisions by the decision maker. See also Che and Kartik (2009).

absent coercion, the vanguard almost never revolts, so the follower and vanguard almost always receive h, and the follower's expected payoff is only marginally above h (under **A1**). With coercion, the vanguard routinely revolts whenever her signal is high enough that she expects to gain over the status quo, and since the follower's payoff with coercion equals his vanguard's, the follower must also benefit.

The result that the vanguard wants to set coercion equal to $c = w_2 - l_1$ hinges on the assumption that a vanguard and her follower share common beliefs about the quality of each others' signals. If, instead, the vanguard has more faith in her information than does her follower, i.e., if she thinks that the follower believes that the vanguard's information is less precise than the vanguard believes her own information to be, then she wants to use more coercion to adjust for her follower's lesser incentive to revolt. Proposition 6 formalizes this finding. It uses Assumption **A4**, which is satisfied by the additive normal noise signal structure (see equation (9) in the Appendix) as well as most empirically-relevant distributions of noise.

Assumption A4. $E[\theta|s_1 \geq k_1, s_2]$ is strictly decreasing in the variance of $\theta|s_1$.

Proposition 6 *Suppose that* **A1**, **A2** *and* **A4** *hold. If the follower believes that the vanguard's information is less precise than the vanguard, herself, does, then the vanguard chooses harsher coercive measures, setting $c > w_2 - l_1$. That is, if the follower believes that $var(\theta|s_1)$ is higher than what the vanguard believes, then $c > w_2 - l_1$.*

For example, if the follower believes that the normally-distributed noise ν_1 in the vanguard's signal, $s_1 = \theta + \nu_1$, has variance $\tilde{\sigma}_\nu^2$, but the vanguard believes that the noise only has variance $\sigma_\nu^2 < \tilde{\sigma}_\nu^2$, then the vanguard chooses $c > w_2 - l_1$. In effect, when a follower believes that the vanguard's information is less precise, then from the vanguard's perspective, the follower underweights the positive information about the revolution payoffs conveyed by the vanguard's decision to revolt. This makes the follower less willing to revolt. To correct for this, the vanguard reduces the follower's payoff when he fails to support her.

7 Conclusion

Our paper considers the problem of a representative revolutionary vanguard who is unhappy with the status quo. The vanguard also knows that she does not know everything, and that her followers have valuable information that may bear on whether or not a revolution would be a good idea. The vanguard knows that if she initiates a revolution, but her follower is less sanguine about the revolution payoffs and hence fails to provide support, then the attempted revolution will fail, and the regime will punish her. But, if the vanguard does not act, the revolution has no chance, and a follower's information cannot influence the revolution outcome. What is the vanguard to do?

We characterize the strategic interactions between a vanguard and its follower. We show that under mild conditions a unique equilibrium exists. In equilibrium, the vanguard takes on the risk of being the only one to act: by observably moving, it eliminates that risk for a follower. Via its willingness to sacrifice and bear alone the risk of punishment, the vanguard precludes the possibility of failing to act even when each has very positive information about the benefits of revolution.

We establish how the vanguard would like to use coercion on her follower, not as a blunt tool that induces blind obedience, but rather as a delicate instrument that aligns the follower's incentives so that he sacrifices similarly to the vanguard when he does not follow her lead. An excessively-coerced follower will compliantly follow his vanguard even when he believes a "successful" revolution will turn out disastrously, so that the revolution outcome fails to reflect their collective information.

We show how a state that wants to minimizes the likelihood that it is overthrown must be careful in how it deploys the tools at its disposal when trying to maintain the status quo. In particular, when it cannot punish the vanguard of an unsuccessful revolt too harshly, if it raises the reward to a follower who turns on his vanguard, then to reduce the probability a successful revolt, it may also want to *reduce* the punishment for failed revolt. While a less reliable follower or harsher punishment both reduce the attraction to the vanguard of revolting, paradoxically, they may raise the probability of successful revolt, because when the vanguard acts it conveys better news to her follower about the payoffs from successful joint action.

We show that even though the vanguard alone risks punishment by the regime, she may still be more willing than a follower to act given the same signal about revolution payoffs. This is because the vanguard must also weigh the possible gains of acting in order to let her follower's information determine whether the revolution succeeds. In fact, the vanguard is more willing to act than her follower whenever the follower's payoff from not supporting the vanguard is high (so that the follower is not too willing to act), but not too high (so that the vanguard does not face an excessively high risk of miscoordination, and hence punishment).

Our analysis takes the leadership structure as given. We leave for future research the question: who chooses to become the vanguard? The endogenous emergence of leadership gives rise to a host of subtle considerations. Absent a formal leadership structure, any individual can choose to lead, but each individual may be less willing to risk leading and punishment, choosing to wait to see what others do. But then if no one acts, individuals update more negatively about the prospects of revolution, so that such free-riding can cause individuals to act too little.

8 Appendix

Lemma 4 *Suppose the vanguard sometimes revolts, i.e., that $\rho_1(s_1) = R$ for some s_1. Then the follower's best response to $\rho_1(\cdot)$ takes a cutoff form: There exists a finite cutoff $k_2(\rho_1)$ such that the follower revolts if and only if $s_2 \geq k_2(\rho_1)$.*

Proof of Lemma 4: The follower's best response following a_1 and s_2 is to take action R if and only if $E[\theta|\rho_1(s_1) = R, s_2] \geq w_2$. The limit properties in Part (a) of **A1** imply the existence of sufficiently good and bad signals, so that there exists some signal $s_2 = k_2$ such that $E[\theta|\rho_1(s_1) = R, k_2] = w_2$. Further, strict affiliation of signals implies that $E[\theta|\rho_1(s_1) = R, s_2] > w_2$ for all $s_2 > k_2$, and $E[\theta|\rho_1(s_1) = R, s_2] < w_2$ for all $s_2 < k_2$. $\square$

Lemma 5 *Suppose that $\rho_2(s_2, R) = R$ if and only if $s_2 \geq k_2$. Then, given* **A1**, *there exists a k_1 such that the vanguard's best response is to revolt if and only if $s_1 \geq k_1(k_2)$.*

Proof of Lemma 5: Given the follower's cutoff k_2, the vanguard's expected net payoff from revolt is $\Delta(s_1; k_2) \equiv Pr(s_2 \geq k_2|s_1)\ E[\theta|s_2 \geq k_2, s_1] + Pr(s_2 < k_2|s_1)l_1 - h_1$, which simplifies to

$$\Delta(s_1; k_2) = Pr(s_2 \geq k_2|s_1)\ (E[\theta|s_2 \geq k_2, s_1] - l_1) + l_1 - h_1. \tag{4}$$

Affiliation and the limit properties in **A1** imply the vanguard's best response to a follower's cutoff strategy is also a cutoff strategy with associated cutoff $k_1(k_2)$, where $\Delta(k_1(k_2); k_2) = 0$.

We have $Pr(s_2 \geq k_2|s_1) > 0$. Further, $Pr(s_2 \geq k_2|s_1)$ and $E[\theta|s_2 \geq k_2, s_1]$ rise with s_1 due to affiliation. Thus, from equation (4), if $\Delta(s_1 = x; k_2) = 0$, then $\Delta(s_1; k_2) > 0$ for all $s_1 > x$. From **A1**, $\lim_{s_1 \to -\infty} \Delta(s_1; k_2) < 0 < \lim_{s_1 \to +\infty} \Delta(s_1; k_2)$. Thus, for every k_2, there exists a unique $s_1 = k_1$ such that $\Delta(k_1; k_2) = 0$. Further, at $s_1 = k_2$,

$$\left.\frac{\partial \Delta(s_1; k_2)}{\partial s_1}\right|_{s_1 = k_1} > 0. \quad \square \tag{5}$$

Proof of Lemma 1: From Lemma 4, $E[\theta|s_1 \geq k_1, k_2] - w_2 = 0$. Thus,

$$\frac{\partial k_2(k_1)}{\partial k_1} = -\left(\frac{\partial E[\theta|s_1 \geq k_1, k_2]}{\partial k_2}\right)^{-1} \frac{\partial E[\theta|s_1 \geq k_1, k_2]}{\partial k_1}.$$

By affiliation, both terms are positive, and hence $\frac{\partial k_2(k_1)}{\partial k_1} < 0$. $\square$

Proof of Lemma 2:

$$\frac{\partial k_1(k_2)}{\partial k_2} = -\left(\frac{\partial \Delta(k_1; k_2)}{\partial k_1}\right)^{-1} \frac{\partial \Delta(k_1; k_2)}{\partial k_2}. \tag{6}$$

Rewrite equation (4) as

$$\Delta(k_1; k_2) = \int_{k^2}^{\infty} E[\theta|s_2, k_1]\ f(s_2|k_1)\ ds_2 + F(k_2|k_1)\ l_1 - h_1.$$

Let $\delta(k_2, k_1(k_2)) \equiv E[\theta|k_2, k_1(k_2)] - l_1$. Thus,

$$\frac{\partial \Delta(k_1; k_2)}{\partial k_2} = f(k_2|k_1)\ \ (-E[\theta|k_2, k_1] + l_1) \equiv -f(k_2|k_1)\ \delta(k_2, k_1), \tag{7}$$

and hence from equations (5), (6), and (7), $sign\left(\frac{\partial k_1}{\partial k_2}\right) = sign\left(\delta(k_2, k_1)\right)$. Next, we sign δ, establishing its monotonicity properties:

$$\begin{aligned}
\frac{d\delta(k_2, k_1(k_2))}{dk^2} &= \frac{dE[\theta|k_2, k_1(k_2)]}{dk^2} \\
&= \frac{\partial E[\theta|k_2, k_1(k_2)]}{\partial k_2} + \frac{\partial E[\theta|k_2, k_1(k_2)]}{\partial k_1}\ \frac{\partial k_1}{\partial k_2} \\
&= \frac{\partial E[\theta|k_2, k_1(k_2)]}{\partial k_2} - \frac{\partial E[\theta|k_2, k_1(k_2)]}{\partial k_1}\ \left(\frac{\partial \Delta(k_1; k_2)}{\partial k_1}\right)^{-1}\ \frac{\partial \Delta(k_1; k_2)}{\partial k_2} \\
&= \frac{\partial E[\theta|k_2, k_1(k_2)]}{\partial k_2} + \frac{\partial E[\theta|k_2, k_1(k_2)]}{\partial k_1}\ \frac{f(k_2|k_1)\ \delta(k_2, k_1)}{\frac{\partial \Delta(k_1;k_2)}{\partial k_1}}, \qquad (8)
\end{aligned}$$

where the third equality follows from equation (6) and the fourth from equation (7). Both $\frac{\partial E[\theta|k_2,k_1]}{\partial k_1}$ and $\frac{\partial E[\theta|k_2,k_1]}{\partial k_2}$ are positive because s_1, s_2, and θ are affiliated; and $\frac{\partial \Delta(k_1;k_2)}{\partial k_1} > 0$ from equation (5). Thus, $\frac{d\delta}{dk^2} > 0$ for all $\delta \geq 0$, which implies that $\delta(k_2, k_1(k_2))$ has a single-crossing property as a function of k_2. Next, we show δ changes sign from negative (strategic substitutes) to positive (strategic complements). From equation (4) and **A1**, $\lim_{k_2 \to -\infty} k_1(k_2) < \infty$ and $\lim_{k_2 \to \infty} k_1(k_2) > -\infty$. To see the latter, note that

$$\lim_{k_i \to -\infty} \lim_{k_j \to \infty} Pr(s_j > k_j|k_i)\ E[\theta|s_j > k_j, k_i] < \lim_{k \to \infty} Pr(s_j > k|k)\ E[\theta|s_j > k, k],$$

and hence part (c) of **A1** implies

$$\lim_{k_1 \to -\infty} \lim_{k_2 \to \infty} Pr(s_2 > k_2|k_1)\ E[\theta|s_2 > k_2, k_1] < 0.$$

Thus, from equation (4), $\lim_{s_1 \to -\infty} \lim_{k_2 \to \infty} \Delta(s_1; k_2) < 0$, and hence $\lim_{k_2 \to \infty} k_1(k_2) > -\infty$. Therefore, $\lim_{k_2 \to \pm\infty} \delta(k_2, k_1(k_2)) = \pm\infty$. $\square$

Proof of Lemma 3: For $i \in \{1, 2\}$, let $b_i = \sigma_0^2/(\sigma_0^2 + \sigma_i^2)$, $a_i = \sqrt{(1 + b_i)\sigma_i^2}$, and $\Sigma = \sigma_0^2\sigma_1^2 + \sigma_0^2\sigma_2^2 + \sigma_1^2\sigma_2^2$. Then,

$$E[\theta|k_i, s_j \geq k_j] = b_i k_i + \frac{\sigma_0^2\sigma_i^2 a_i}{\Sigma}\ \frac{\phi(x_i)}{1 - \Phi(x_i)}, \qquad (9)$$

where $x_i = (k_j - b_i k_i)/a_i$ and $\phi(x)$ and $\Phi(x)$ are pdf and cdf of standard normal distribution, respectively. Let $A(x) \equiv \frac{\partial}{\partial x}\ \frac{\phi(x)}{1-\Phi(x)}$. Moreover, $A(x) \in (0, 1)$ (Sampford 1953). Thus, Assumption **A2** holds if and only if

$$\left(1 - \frac{\sigma_0^2\sigma_1^2}{\Sigma}\ A(x_1)\right)\left(1 - \frac{\sigma_0^2\sigma_2^2}{\Sigma}\ A(x_2)\right)\ b_2 b_1 > \left(\frac{\sigma_0^2\sigma_2^2}{\Sigma}\ A(x_2)\right)\left(\frac{\sigma_0^2\sigma_1^2}{\Sigma}\ A(x_1)\right).$$

That is,

$$b_1 b_2 \left(1 - \frac{\sigma_0^2\sigma_1^2}{\Sigma}\ A(x_1) - \frac{\sigma_0^2\sigma_2^2}{\Sigma}\ A(x_2)\right) > (1 - b_1 b_2)\ \frac{\sigma_0^2\sigma_1^2}{\Sigma}\ \frac{\sigma_0^2\sigma_2^2}{\Sigma}\ A(x_1)A(x_2). \quad (10)$$

Next, observe that

$$b_1 b_2 = \frac{\sigma_0^4}{(\sigma_0^2+\sigma_1^2)(\sigma_0^2+\sigma_1^2)},\ \text{and hence } 1 - b_1 b_2 = \frac{\Sigma}{(\sigma_0^2+\sigma_1^2)(\sigma_0^2+\sigma_1^2)}. \quad (11)$$

Substituting (11) into (10) and rearrangement yields

$$\Sigma - \sigma_0^2\sigma_1^2\ A(x_1) - \sigma_0^2\sigma_2^2\ A(x_2) > \sigma_1^2\sigma_2^2\ A(x_1)A(x_2),$$

which is true because $A(x) \in (0,1)$, and hence

$$\Sigma - \sigma_0^2\sigma_1^2\ A(x_1) - \sigma_0^2\sigma_2^2\ A(x_2) > \Sigma - \sigma_0^2\sigma_1^2 - \sigma_0^2\sigma_2^2 = \sigma_1^2\sigma_2^2 > \sigma_1^2\sigma_2^2\ A(x_1)A(x_2). \quad \square$$

Proof of Proposition 1: First, we prove an equilibrium exists. The follower's best response $k_2(k_1)$ upon observing that the vanguard has revolted solves $E[\theta|s_1 \geq k_1, k_2(k_1)] = w_2$. Thus, from part (a) of **A1**, $\lim_{k_1\to+\infty} k_2(k_1) = -\infty$ and $\lim_{k_1\to-\infty} k_2(k_1)$ is finite. Moreover, $\Delta(k_1(k_2); k_2) = 0$, and hence from **A1**, $\lim_{k_2\to+\infty} k_1(k_2) > -\infty$. Then, the continuity of $k_1(k_2)$ and $k_2(k_1)$ implies that they cross, at least once.

Next, we prove uniqueness for the case where $l_1 = h_1$. Then, the vanguard best response satisfies $E[\theta|k_1, s_2 \geq k_2] = l_1$, and hence, *for the vanguard's best response*,

$$\frac{\partial k_1(k_2)}{\partial k_2} = -\left(\frac{\partial E[\theta|k_1, s_2 \geq k_2]}{\partial k_1}\right)^{-1}\ \frac{\partial E[\theta|k_1, s_2 \geq k_2]}{\partial k_2}. \quad (12)$$

Next, consider the follower. His best response satisfies $E[\theta|s_1 \geq k_1, k_2] = w_2$. Similar calculations *for the follower's best response* yields

$$\frac{\partial k_2(k_1)}{\partial k_1} = -\frac{\partial E[\theta|s_1 \geq k_1, k_2]}{\partial k_1}\ \left(\frac{\partial E[\theta|s_1 \geq k_1, k_2]}{\partial k_2}\right)^{-1}. \quad (13)$$

To prove uniqueness, it suffices to show that the the vanguard's beast response curve in (k_1, k_2)-space, for all relevant k_1s, always has a sharper negative slope that the follower's. That is, the inverse of (12) is more a larger negative number that (13), i.e.,

$$-\frac{\partial E[\theta|s_1 \geq k_1, k_2]}{\partial k_1}\ \left(\frac{\partial E[\theta|s_1 \geq k_1, k_2]}{\partial k_2}\right)^{-1} > -\frac{\partial E[\theta|k_1, s_2 \geq k_2]}{\partial k_1}\ \left(\frac{\partial E[\theta|k_1, s_2 \geq k_2]}{\partial k_2}\right)^{-1}.$$

Due to affiliation all terms are positive, and hence rearrangement yields

$$\frac{\partial E[\theta|s_1 \geq k_1, k_2]}{\partial k_1}\ \left(\frac{\partial E[\theta|k_1, s_2 \geq k_2]}{\partial k_2}\right) < \frac{\partial E[\theta|k_1, s_2 \geq k_2]}{\partial k_1}\ \left(\frac{\partial E[\theta|s_1 \geq k_1, k_2]}{\partial k_2}\right),$$

which hold by Assumption **A2**. If $h_1 > l_1$, then the slope of the strategic substitute segment of the vanguard's best response becomes even more negative. Thus, if a crossing happens on the strategic substitutes segment of the vanguard's best response, it is unique. Finally, clearly, crossing can happen only once on the strategic complements segment of the vanguard's best response. $\square$

Proof of Proposition 2: With the new vanguard's payoff $\theta + z$, the vanguard expected net payoff from taking action R, equation (4) becomes

$$\Delta(s_1; k_2, z) = Pr(s_2 \geq k_2|s_1)\ (E[\theta|s_2 \geq k_2, s_1] + z - l_1) + l_1 - h_1. \tag{14}$$

As in the proof of Lemma 5, one can show that the vanguard's best response is a finite cutoff strategy with associated cutoff k_1 such that $\Delta(k_1; k_2) = 0$, and $\frac{\partial \Delta(s_1;k_2,z)}{\partial s_1}|_{s_1=k_1} > 0$. Moreover, $\frac{\partial \Delta(k_1;k_2,z)}{\partial z} = Pr(s_2 \geq k_2|k_1) > 0$. Thus, $\frac{\partial k_1(z)}{\partial z} = -\left(\frac{\partial \Delta(k_1;k_2,z)}{\partial k_1}\right)^{-1} \frac{\partial \Delta(k_1;k_2,z)}{\partial z} < 0$ at any best response cutoff k_1, including the equilibrium cutoff $\overline{k_1}(z)$. Hence, since the follower's best response exhibits strategic substitutes (Lemma 1), $0 < \frac{\partial \overline{k_2}(z)}{\partial z}$. $\square$

Proof of Proposition 3: First, we prove a lemma.

Lemma 6 *There exists a h_1^* such that $k_1(k_2) > k_2$ if and only if $h_1^* < h_1$.*

Proof of Lemma 6: $k_1(k_2)$ is continuous in k_2, and $\lim_{k_2 \to -\infty} k_1(k_2) > -\infty$. Thus, $k_1(k_2) > k_2$ if and only if $\Delta(k_1(k_2) = k; k_2 = k) \neq 0$, where

$$\Delta(k_1(k_2) = k; k_2 = k) = Pr(s_2 \geq k|k)\ (E[\theta|s_2 \geq k, k] - l_1) + l_1 - h_1. \tag{15}$$

From **A1**, $\lim_{k \to \infty} \Delta(k; k) = l_1 - h_1 < 0$, and $\lim_{k \to -\infty} \Delta(k; k) = -\infty$. Moreover, $\Delta(k; k)$ is continuous, and hence $\Delta(k; k)$ has a global maximum in $\mathbb{R}$, call it k_{max}. Further, k_{max} is independent of h_1, but $\Delta(k; k)$ is uniformly decreasing in h_1. Thus, there exists an h_1^* such that $\Delta(k; k) < 0$ if and only if $h_1 > h_1^*$. $\square$

Next, we prove that, for $h_1 < h_1^*$, the vanguard's best response $k_1(k_2)$ crosses the 45 degree line at exactly two points. Lemma 6 implies that it must cross the 45 degree line at least twice. If $k_1(k^*) < k^*$, so that $k_1(k_2)$ first crosses the 45 degree line on its decreasing (strategic substitutes) segment, it means that the increasing segment of $k_1(k_2)$ starts from under 45 degree line. Then the convexity in Assumption **A3** implies that there is at most one crossing. If $k_1(k^*) > k^*$, so that $k_1(k_2)$ first crosses the 45 degree line on its increasing (strategic complements) segment, convexity also implies that $k_1(k_2)$ crosses the 45 degree line at most twice. Call the corresponding crossings $k_2 = \underline{k}$ and $k_2 = \bar{k}$, with $\underline{k} < \bar{k}$. Thus, $k_1(k_2)$ is above the 45 degree line between $k_2 = \underline{k}$ and $k_2 = \bar{k}$. Moreover, from equation (15), $\underline{k}$ rises in h_1 while $\overline{k}$ falls.

From the proof of Proposition 1, $k_2(k_1; w_2)$ is strictly decreasing in k_1, and uniformly increasing in w_2. Thus, there exists a unique $w_2 = \bar{w}$ such that $k_2(k_1 = \bar{k}; \bar{w}) = \bar{k}$, and a unique $w_2 = \underline{w} < \bar{w}$ such that $k_2(k_1 = \underline{k}; \underline{w}) = \underline{k}$. That $\underline{w}$ is increasing in h_1 and $\overline{w}$ is decreasing follows directly from the relationship between $\underline{k}$, $\overline{k}$ and h_1.

Finally, we prove that, under symmetry, $\underline{w} > h_1$. From the proof of Proposition 2, $Pr(s_2 \geq \underline{k} | s_1 = \underline{k})\ (E[\theta | s_2 \geq \underline{k}, s_1 = \underline{k}] - l_1) = h_1 - l_1$. Since $Pr(s_2 \geq \underline{k} | s_1 = \underline{k}) < 1$, it follows that $E[\theta | s_2 \geq \underline{k}, s_1 = \underline{k}] > h_1$. Moreover, from the proof of Proposition 1, $E[\theta | s_1 \geq \underline{k}, s_2 = \underline{k}] = \underline{w}$, and from symmetry, $E[\theta | s_2 \geq \underline{k}, s_1 = \underline{k}] = \underline{w}$. Thus, $\underline{w} > h_1$. $\square$

Proof of Proposition 4: Let $E[U_2|k_1, k_2]$ be the follower's ex ante expected utility given equilibrium cutoffs k_1 and k_2. Then

$$\begin{aligned}
\frac{dE[U_2|k_1,k_2]}{dz} &= \left(\frac{\partial E[U_2|k_1,k_2]}{\partial k_1} + \frac{\partial E[U_2|k_1,k_2]}{\partial k_2}\ \frac{\partial k_2}{\partial k_1} \right)\ \frac{dk_1}{dz} \\
&= \frac{\partial E[U_2|k_1,k_2]}{\partial k_1}\ \frac{dk_1}{dz}, \qquad (16)
\end{aligned}$$

where we use $\frac{\partial E[U_2|k_1,k_2]}{\partial k_2} = 0$ because k_2 is the follower's equilibrium cutoff. Moreover,

$$\begin{aligned}
E[U_2|k_1,k_2] &= Pr(s_1 < k_1)h_2 + Pr(s_1 \geq k_1, s_2 < k_2)w_2 \\
&\quad + Pr(s_1 \geq k_1, s_2 \geq k_2)\ E[\theta | s_1 \geq k_1, s_2 \geq k_2] \\
&= Pr(s_1 < k_1)h_2 + w_2 \int_{s_1=k_1}^{\infty} \int_{s_2=-\infty}^{k_2} f(s_1,s_2) ds_1 ds_2 \\
&\quad + \int_{s_1=k_1}^{\infty} \int_{s_2=k_2}^{\infty} E[\theta|s_1,s_2] f(s_1,s_2) ds_1 ds_2.
\end{aligned}$$

Hence, $\frac{\partial E[U_2|k_1,k_2]}{\partial k_1} =$

$$
\begin{aligned}
& f(k_1)\ h_2 - w_2\ f(k_1) \int_{-\infty}^{k_2} f(s_2|k_1) ds_2 - f(k_1)\ \int_{k_2}^{\infty} E[\theta|s_1 = k_1, s_2]\ f(s_2|s_1 = k_1)\ ds_2 \\
= \quad & f(k_1)\{h_2 - w_2 Pr(s_2 < k_2|s_1 = k_1) - Pr(s_2 \geq k_2|k_1) E[\theta|s_2 \geq k_2, k_1]\}. \qquad (17) \\
= \quad & f(k_1)\big\{h_2 - w_2 + Pr(s_2 \geq k_2|k_1)\big(E[\theta|s_1 \geq k_1, k_2] - E[\theta|k_1, s_2 \geq k_2]\big)\big\}. \qquad (18)
\end{aligned}
$$

First, substituting $\lim_{k_1 \to \infty} k_2(k_1) = -\infty$ into (17) reveals that $\frac{\partial E[U_2|k_1,k_2]}{\partial k_1} < 0$ for k_1 sufficiently large. Therefore, $\frac{dE[U_2|k_1,k_2]}{dz} > 0$ for sufficiently negative z, implying that $z^* > -\infty$.

Second, since $E[\theta|s_1 \geq k_1, k_2] = w_2$ and $\lim_{k_1 \to -\infty} k_2(k_1)$ is finite, we have $\lim_{k_1 \to -\infty} E[\theta|k_1, s_2 \geq k_2] = -\infty$ and $\lim_{k_1 \to -\infty} E[\theta|s_1 \geq k_1, k_2] = w_2$. Substituting these limits into (18) reveals that when $h_2 = w_2$, $\frac{\partial E[U_2|k_1,k_2]}{\partial k_1} > 0$ for sufficiently negative k_1. Therefore, when $h_2 = w_2$, $\frac{dE[U_2|k_1,k_2]}{dz} < 0$ for all sufficiently large z, i.e., $z^* < \infty$ when $h_2 = w_2$.

Third, if $Pr(s_2 \geq k_2|k_1) E[\theta|s_2 \geq k_2, k_1]$ is bounded from below, then fixing w_2, substituting sufficiently negative h_2 into from (17) reveals that $\frac{\partial E[U_2|k_1,k_2]}{\partial k_1}$ is always negative. Therefore, $\frac{dE[U_2|k_1,k_2]}{dz}$ is always positive, and hence $z^* = \infty$. Moreover, assumption A1 (d) implies $Pr(s_2 \geq k_2|k_1) E[\theta|s_2 \geq k_2, k_1]$ is bounded from below. To see this, note that (i) $\lim_{k_1 \to -\infty} Pr(s_2 \geq k_2|k_1) E[\theta|s_2 \geq k_2, k_1] \geq \lim_{k_1 \to -\infty} Pr(s_2 \geq k_2|k_1) E[\theta|k_2, k_1]$, and (ii) $k_2(k_1)$ is decreasing in k_1 with $\lim_{k_1 \to -\infty} k_2(k_1) = \overline{k}_2 \in \mathbb{R}$. Thus, for any $\epsilon > 0$, $\lim_{k_1 \to -\infty} Pr(s_2 \geq k_2|k_1) E[\theta|k_2, k_1] \geq \lim_{k_1 \to -\infty} Pr(s_2 \geq \overline{k}_2 - \epsilon|k_1) E[\theta|\overline{k}_2 - \epsilon, k_1] > -\infty$, where the last inequality follows from A1(d).

Finally, from (17), $\frac{\partial E[U_2|k_1,k_2]}{\partial k_1}$ rises in h_2 because h_2 does not affect k_1 or k_2. Hence, z^* falls with h_2. Thus, there exists $\bar{h} \in \mathbb{R}$ such that z^* is finite if and only if $h_2 > \bar{h}$.

Next, we derive z^* when it is finite. Because k_1 is the vanguard's equilibrium cutoff,

$$
Pr(s_2 \geq k_2|k_1) E[\theta|s_2 \geq k_2, k_1] = h_1 - l_1 Pr(s_2 < k_2|k_1) - Pr(s_2 \geq k_2|k_1) z.
$$

Substituting for $E[\theta|s_2 \geq k_2, k_1]$ using this equilibrium relationship yields

$$
\begin{aligned}
f(k_1) \frac{\partial E[U_2|k_1, k_2]}{\partial k_1} \quad &= \quad h_2 - w_2 Pr(s_2 < k_2|k_1) - h_1 + l_1 Pr(s_2 < k_2|k_1) + Pr(s_2 \geq k_2|k_1) z \\
&= \quad (h_2 - h_1) - (w_2 - l_1) Pr(s_2 < k_2|k_1) + Pr(s_2 \geq k_2|k_1) z. \qquad (19)
\end{aligned}
$$

Combining equations (16) and (19) yields

$$\frac{dE[U_2|k_1,k_2]}{dz} = \frac{dk_1}{dz}\ f(k_1)\ [(h_2-h_1) - Pr(s_2<k_2|k_1)\ (w_2-l_1) + Pr(s_2\geq k_2|k_1)\ z].$$

Solving this first-order condition yields the result in the proposition. $\square$

Proof of Proposition 5: Let $E[U_1|k_1,k_2]$ be the vanguard's ex ante expected utility given cutoffs k_1 and k_2:

$$\begin{aligned}
E[U_1|k_1,k_2] &= Pr(s_1\geq k_1, s_2\geq k_2)\ E[\theta|s_1\geq k_1, s_2\geq k_2] \\
&\quad +Pr(s_1\geq k_1, s_2<k_2)l_1 + Pr(s_1<k_1)h_1 \\
&= Pr(s_1\geq k_1)\int_{s_2=k_2}^{\infty} E[\theta|s_1\geq k_1, s_2]\ f(s_2|s_1\geq k_1)ds_2 \\
&\quad +Pr(s_1\geq k_1)Pr(s_2<k_2|s_1\geq k_1)l_1 + Pr(s_1<k_1)h_1. \qquad (20)
\end{aligned}$$

In equilibrium, k_1 and k_2 are best responses, and hence $\frac{\partial E[U_1|k_1,k_2]}{\partial k_1}=0$. Moreover, from Lemmas 4 and 5, only k_2 explicitly depends on w_2. Thus,

$$\begin{aligned}
\frac{\partial E[U_1|k_1(k_2(w_2)),k_2(w_2)]}{\partial w_2} &= \frac{\partial E[U_1]}{\partial k_2}\ \frac{\partial k_2}{\partial w_2} + \frac{\partial E[U_1]}{\partial k_1}\ \frac{\partial k_1}{\partial k_2}\ \frac{\partial k_2}{\partial w_2} \\
&= \frac{\partial E[U_1]}{\partial k_2}\ \frac{\partial k_2}{\partial w_2}.
\end{aligned}$$

It is easy to see that $\frac{\partial k_2}{\partial w_2}>0$, and hence the first order condition, $\frac{\partial E[U_1|k_1(k_2(w_2)),k_2(w_2)]}{\partial w_2}=0$, holds if and only if $\frac{\partial E[U_1]}{\partial k_2}=0$. From equation (20),

$$\begin{aligned}
\frac{\partial E[U_1]}{\partial k_2} &= Pr(s_1\geq k_1)\ (-E[\theta|s_1\geq k_1,k_2]\ f(k_2|s_1\geq k_1)\ ) + f(k_2|s_1\geq k_1)\ Pr(s_1\geq k_1)\ l_1 \\
&= Pr(s_1\geq k_1)\ f(k_2|s_1\geq k_1)\ (l_1 - E[\theta|s_1\geq k_1,k_2]).
\end{aligned}$$

Thus, $\frac{\partial E[U_1]}{\partial k_2}=0$ if and only if $E[\theta|s_1\geq k_1(k_2(w_2)),k_2(w_2)]=l_1$. Moreover, from Lemma 4, $E[\theta|s_1\geq k_1(k_2(w_2)),k_2(w_2)]=w_2$. Thus, the first order condition holds if and only if $w_2=l_1$. It is easy to see that this is a maximum. That is, the vanguard's expected payoff is maximized when $w_2=l_1$. Thus, it chooses c such that $w_2-c=l_1$, i.e., $c=w_2-l_1$. Clearly, when we restrict c to be positive, $c=\max\{w_2-l_1,0\}$. $\square$

Proof of Proposition 6: Let E_i be agent i's expectation. From **A4**, $E_2[\theta|s_1\geq k_1(k_2(w_2-c)),k_2(w_2-c)] < E_1[\theta|s_1\geq k_1(k_2(w_2-c)),k_2(w_2-c)]$. The equilibrium level of k_2 is determined by $E_2[\theta|s_1\geq k_1(k_2(w_2-c)),k_2(w_2-c)]=w_2-c$, and hence

$w_2 - c < E_1[\theta | s_1 \geq k_1(k_2(w_2 - c)), k_2(w_2 - c)]$ in equilibrium. From Proposition 5, the vanguard's optimal choice is to pick a c that yields a $k_2(w_2 - c)$ that satisfies $E_1[\theta | s_1 \geq k_1(k_2(w_2 - c)), k_2(w_2 - c)] = l_1$. This implies that $w_2 - c < l_1$, i.e., $w_2 - l_1 < c$. $\square$

9 References

Ahlquist, John S., and Margaret Levi. 2011. "Leadership: What It Means, What It Does, and What We Want to Know About It." *Annual Review of Political Science* 14: 1-24.

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek. 2008. "When Does Coordination Require Centralization?" *American Economic Review* 98: 145-79.

Andreoni, James. 2006. "Leadership Giving in Charitable Fund-Raising." *Journal of Public Economic Theory* 8: 1-22.

Bolton, Patrick, Markus K. Brunnermeier, and Laura Veldkamp. 2012. "Leadership, Coordination and Corporate Culture." *Review of Economic Studies* 80: 512-37.

Bueno de Mesquita, Ethan. 2010. "Regime Change and Revolutionary Entrepreneurs." *American Political Science Review* 104: 446-66.

Bueno de Mesquita, Ethan. 2013. "Rebel Tactics." *Journal of Political Economy* 121: 323-57.

Calvert, Randall L. 1992. "Leadership and Its Basis in Problems of Social Coordination." *International Political Science Review* 13: 7-24.

Che, Yeon-Koo, and Navin Kartik. 2009. "Opinions as Incentives." *Journal of Political Economy* 117: 815-60.

De Castro, Luciano I. 2010. "Affiliation, Equilibrium Existence and Revenue Ranking of Auctions." Mimeo.

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

Dewan, Turon, and David P. Myatt. 2007. "Leading the Party: Coordination, Direction, and Communication." *American Political Science Review* 101: 827-45.

Dewan, Turon, and David P. Myatt. 2008. "The Qualities of Leadership: Direction, Communication, and Obfuscation." *American Political Science Review* 102: 351-68.

Dickson, Eric S. 2012 "Leadership, Followership, and Beliefs About the World: An Experiment." *British Journal of Political Science, forthcoming.*

Feddersen, Timothy J., and Wolfgang Pesendorfer. 1996. "The Swing Voter's Curse." *American Economic Review* 86: 408-24.

Feddersen, Timothy J., and Wolfgang Pesendorfer. 1998. "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting." *American Political Science Review* 92: 23-35.

Granovetter, Mark. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83: 1420-43.

Hermalin, Benjamin E. 1998. "Toward an Economic Theory of Leadership: Leading by Example." *American Economic Review* 88: 1188-1206.

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

Komai, Mana, and Mark Stegeman. 2010. "Leadership Based on Asymmetric Information." *RAND Journal of Economics* 41: 35-63.

Komai, Mana, Mark Stegeman, and Benjamin E. Hermalin. 2007. "Leadership and Information." *American Economic Review* 97: 944-7.

Kuran, Timur. 1991. "Now Out of Never: The Element of Surprise in the East European Revolution of 1989." *World Politics*, 44(1): 7-48.

Landier, Augustin, David Sraer, and David Thesmar. 2009. "Optimal Dissent in Organizations." *Review of Economic Studies* 76: 761-94.

Lohmann, Susanne. 1994. "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany 1989-91." *World Politics* 47: 42-101.

Majumdar, Sumon, and Sharun W. Mukand. 2008. "The Vanguard as Catalyst: On Leadership and the Mechanics of Institutional Change." CESIFO Working Paper No. 2337.

Milgrom, Paul R., and Robert J. Weber. 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica* 50: 1089-1122.

Rantakari, Heikki. 2008. "Governing Adaptation." *Review of Economic Studies* 75: 1257-85.

Roemer, John. 1985. "Rationalizing Revolutionary Ideology." *Econometrica* 53: 85-108.

Sampford, M., R. 1953. "Some Inequalities on Mill's Ratio and Related Functions." *The Annals of Mathematical Statistics* 24: 130-2.

Schelling, Thomas. 1969. "Models of Segregation." *American Economic Review* 59: 488-93.

Schelling, Thomas. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1: 143-86.

Shadmehr, Mehdi, and Dan Bernhardt. 2013. "Information Aggregation in Coordination Games." Mimeo.

Wilson, Robert. 1977. "A Bidding Model of Perfect Competition." *Review of Economic Studies* 4: 511-18.

Wilson, Rick K., and Carl M. Rhodes. 1997. "Leadership and Credibility in N-Person Coordination Games." *Journal of Conflict Resolution* 41: 767-91.

Wood, Elisabeth J. 2003. *Insurgent Collective Action and Civil War in El Salvador.* New York: Cambridge University Press.