Inspiring Regime Change[1]

Stephen Morris[2] Mehdi Shadmehr[3]

April, 2019

[1]We are grateful for comments of Nemanja Antic, Dan Bernhardt, Raphael Boleslavsky, Georgy Egorov, Alex Jakobsen, Dimitri Migrow, Konstantin Sonin, seminar participants at the 2016 Columbia Conference on Political Economy, 2017 North American Summer Meeting of the Econometric Society, 2017 Wallis Institute Annual Conference on Political Economy, 2018 North American Winter Meeting of the Econometric Society, 2019 Midwest Political Science Association (MPSA) Meeting, University of Chicago, University of Michigan, University of Central Florida, University of Miami and Queen's University, and the discussions of Brenton Kenkel at 2019 MPSA, Mehmet Ekmekci at 2018 ASSA, Scott Tyson at the Wallis conference, and Alex Debs at the Columbia conference. We have received excellent research assistance from Elia Sartori. An earlier version of this paper circulated under the title "Reward and Punishment in a Regime Change Game."

[2]Department of Economics, Princeton University, Julis Romo Rabinowitz Building, Princeton, NJ 08544. Email: smorris@princeton.edu.

[3]Harris School of Public Policy, University of Chicago, and Department of Economics, University of Calgary, 1307 E 60th St, Chicago, IL 60637. E-mail: mshadmeh@gmail.com.

Abstract

We analyze how revolutionary leaders inspire participation in regime change. Drawing from sociological and historical literature, we model a leader's problem as designing reward schemes that assign psychological rewards to different anti-regime actions. Regime change occurs when the aggregate revolutionary effort from all citizens exceeds the uncertain regime's strength, about which citizens have private information. Citizens face a coordination problem in which each citizen has a private (endogenous) belief about the likelihood of regime change. Because optimistic citizens are easier to motivate, optimal inspiration entails optimal screening. Our approach decomposes this continuous action global game with endogenous rewards into two separate, tractable problems of coordination and screening. A key result is the emergence of a group akin to revolutionary vanguards, all of whom engage in the (endogenous) maximum level of revolutionary activity. In an extension, we show that heterogeneity among citizens (e.g., different levels of income or ideological convictions) reduces the likelihood of regime change.

1 Introduction

Late in December of 1776, the American Revolutionary Army was at the verge of disintegration as enlistments were about to expire and men were about to leave. That meant the end of the Revolution and the American Independence. When the officers' appeals to soldiers to stay did not work, Washington appeared in person to the troops and spoke to them "in the most affectionate manner," as a soldier later recalled. "My brave fellows," Washington said, "you have done all I asked you to do, and more than you could be reasonably expected; but your country is at stake, your wives, your houses, and all that you hold dear. You have worn yourselves out with fatigues and hardships, but we know not how to spare you. If you will consent to stay only one month longer, you will render that service to the cause of liberty and to your country which you probably never can do under any other circumstances" (McCullough 2005, p. 285). Nathanael Greene wrote later: "God Almighty inclined their hearts to listen to the proposal and they engaged anew." This is an episode of what we mean by inspiring regime regime: a leader makes a speech, convenes a meeting, writes a treatise or a pamphlet that appeals to, amplifies, or modifies the people's values, leading the people to contribute to the cause. We develop a model to analyze a leader's optimal inspiration strategy: how a leader with a given ability to inspire (e.g., charisma) should assign psychological rewards to different contribution levels to maximize the likelihood of regime change.

Our model involves two forms of interactions: a vertical interaction between the leader and citizens, and a horizontal interaction among citizens. The first, vertical form of interaction corresponds to a leader's decision how to assign psychological rewards to different contribution levels. Through speeches, meetings, and writings, a leader induces these contribution-reward mappings in the citizens' payoffs. For example, by elevating the values of honor, courage, or patriotism for contributions to the revolution, a leader increases a citizen's payoff from such contributions. But even very charismatic and skillful leaders can induce only so much contribution by invoking psychological mechanisms. Thus, we require that a leader is constrained to offer up to some maximum level of rewards.[1] The second, horizontal form of interactions corresponds to the coordination aspect of regime change. Citizens' payoffs from participating in the revolution depend on whether there is a regime change,[2] and, typically, a large number of citizens must contribute for a regime change to happen. These features imply that citizens care about others' contributions, making coordination a key strategic element of anti-regime movements. To maximize the likelihood of regime change, a leader must account for citizen coordination that follows his design of rewards for anti-regime contributions.

Our model is a coordination game model of regime change, where there is a continuum of ex-ante identical players (*citizens*) and a continuum of actions (*effort*, capturing the level

[1]This means, for example, that in the religious context, even the promise of unlimited comforts and pleasures in heaven in return for some contributions translates into some maximum psychological benefit from those contributions in the faithful's payoffs that determine their actions.

[2]For example, as we discuss below, a key form of motivation in revolution and protest contexts is "pleasure in agency" rewards, which are psychological benefits enjoyed only if the movement succeeds.

of participation/contribution). The revolution succeeds (*regime change*) if the aggregate effort, summing across citizens, exceeds a critical level, with the critical level depending on the state of the world. Any given effort level gives rise to a punishment (the *cost*). The leader must choose a reward (the *benefit*) for each effort level. The benefits are enjoyed only if there is regime change. There is an upper bound on the level of benefits that the leader can induce for any level of effort. We study this problem in a *global game* model (Carlsson and van Damme 1993) of regime change (Morris and Shin 1998), where citizens observe the state with private noise and perfect coordination is not possible. In sum, the leader designs a reward scheme, mapping contributions to rewards. Then, citizens observe their private signals and decide how much to contribute. Finally, the success or failure of the revolution is realized and payoffs are enjoyed.

Two forces drive our results. First, there is *strategic uncertainty.* Citizens observe heterogeneous signals, which through coordination considerations, give rise to endogenous heterogeneity in the population's beliefs about the likelihood of regime change. Some citizens will be endogenously more optimistic than others about the likelihood of regime change, and their degree of optimism is their private information. Second, there is *screening.* If the leader could condition rewards on a citizen's optimism (or some other payoff relevant heterogeneity), then the leader could fine-tune rewards at the individual level (e.g., a more optimistic citizen would get a lower reward for the same contribution than a more pessimistic citizen). Absent the ability to fine-tune at the individual level, the leader faces a more complicated tradeoff. For example, in an extreme case, where the leader offers the same (maximum) rewards for all contributions, all citizens will choose the minimum contribution, and the likelihood of regime change is minimized. Thus, the leader has to create incentives for higher contributions by giving higher rewards to higher contributions, while being cognizant of the strategic responses of citizens. If the leader increases the benefit for a given level of contribution, it induces higher contribution from those who would otherwise have chosen lower levels of contribution. But it also induces lower contribution from those who would otherwise have chosen higher levels of contribution, since they can get the reward at lower cost. The optimal reward scheme trades off these countervailing effects. In essence, the leader's problem is an exchange between a monopolist in possession of psychological rewards and a large number of costumers with private, heterogeneous needs for those rewards—heterogeneities that arise endogenously from their interactions. That is, when a leader decides how to map contributions to rewards, it is *as if* the leader is exchanging rewards for contributions in a population with endogenous and private valuation for these rewards. A key conceptual insight of this paper is to show that optimal inspiration entails optimal screening.

We show that optimal inspiration leads to three distinct groups of citizens: a set of professional revolutionaries engaging in the endogenous maximum level of effort and receiving the maximum level of rewards (bunching at the top); a set of part-time revolutionaries who participate at varying levels with corresponding variation in rewards; and a set of non-participants. The result reflects the fact that inducing more effort from the most optimistic citizens (who already exert more effort than others) is very costly in terms of the amount of effort that can be induced from less optimistic citizens; and inducing effort from the most pessimistic citizens generates second order gains for the leader at the expense of first order costs. The first group corresponds to the vanguard in Lenin's treatise, *What Is to be Done?*, Sons of Liberty in the American Revolution, and similar groups that emerge in many other revolutions. Thus, regardless of the limits of a leader's charisma, which determines the maximum psychological rewards that the leader can induce, and for a wide range of repression schemes, optimally inspiring regime change leads to the emergence of a revolutionary vanguard. The literature on revolutionary leadership divides leaders into "people oriented" leaders, who focus on inspiring people, and "task oriented" leaders, who focus on designing optimal strategies (Goldstone 2001; see Section 5). Our result that optimally inspiring regime change involves complex strategic considerations (a screening problem with endogenous heterogeneity) merges these seemingly distinct categories of leadership. Moreover, our characterization captures the real world feature that citizens make a varying degree of contribution in anti-government movements (e.g., graffitiing walls, joining peaceful demonstration, or taking up arms). Section 5 provides a detailed discussion of the nature participation, motivation and rewards as well as leadership and organization in anti-regime movements. We provide examples from the American Revolution and the early years of Islam to demonstrate how leaders have attempted to inspire contributions, and how their actions were consistent with the logic of screening.

Our theoretical framework can be extended in different directions to deliver further substantive predictions. As an example, we integrate exogenous heterogeneity among citizens by allowing them to have different propensities for rewards, for reasons such as different social status, religious convictions, or past interactions with the state. In particular, when the leader attempts to assign a benefit $B(e)$ to a contribution level e, it may generate a benefit of $\alpha_1 B(e)$ for one citizen and a benefit of $\alpha_2 B(e)$ for another, with $0 < \alpha_1 < \alpha_2$. To demonstrate, we specialize to linear costs and uniform distribution of exogenous heterogeneity in the population. We show that higher dispersion of such exogenous heterogeneities reduces the likelihood of regime change and the depth of the revolutionary coalition, defined as the fraction of participating citizens. The distribution of anti-regime activities in the citizenry may have important consequences for the post-revolution regime. For example, if regime change is mainly the result of high effort by a narrow subset of the population, the post-revolution regime is likely to reflect the preferences and ideological convictions of that narrow group. Movements that rely on intense contribution of a small subset of the population often have to rely on insurgencies (Bueno de Mesquita 2013), and successful movements based on insurgencies often lead to non-democratic outcomes (Wantchékon and García-Ponce 2014).

We now describe our results in more detail, before highlighting theoretical contributions and discussing the relation to the literature. We start by analyzing the regime change model with exogenous benefits, $B(e)$, and costs, $C(e)$, for a given level of contribution $e \geq 0$. A citizen who contributes e and believes that regime change happens with probability p has an expected payoff $pB(e) - C(e)$, or equivalently, $B(e) - C(e)/p$. Our choice of payoff structure is rooted in the literature in sociology and political science that identifies *pleasure in agency* psychological benefits as a key form of motivation for citizen contributions to anti-regime movements. In particular, Wood (2003) argues that participation is not based on citizens having the unreasonable beliefs that their chances of being pivotal compensate for the costs of participation. Nor does citizens' participation reflect expressive motives (cathartic benefits that participants receive regardless of the likelihood of success), reflecting the vast literature that followed Tilly (1978). Rather, Wood argues that participation is driven by pleasure in agency: a psychological benefit that depends on the likelihood of success, but does not depend on the ability to influence that likelihood. Pleasure in agency refers to psychological benefits that one gets from "successful assertion of intention" and from "being part of the making of history" (Wood 2003, p. 18-9; see Section 5). Benefits in our model correspond to these pleasure in agency payoffs, which are psychological but not purely expressive. As we will discuss, the consequence of the non-material nature of these rewards for our analysis is that they are non-rival, reflected only in the leader's constraint in assigning rewards to contributions.

The state, or strength of the regime, is the minimum total amount of citizens' efforts that gives rise to regime change. We assume there is incomplete information about the state, and focus on the case where the state is uniformly distributed and citizens observe noisy signals of the state—as is well known in the global game literature, the results extend to any smooth prior as long as citizens' signals are sufficiently accurate. We exploit a key statistical property that we call *uniform threshold belief.* For a given threshold state, a citizen's threshold belief is the probability that he assigns to the true state being below that threshold. At any given threshold state, there will be a realized distribution of threshold beliefs in the population. Guimaraes and Morris (2007) showed that this threshold belief distribution is uniform. An intuition for this surprising result is the following. At any given threshold state and for any given citizen, we can identify the citizen's rank in the threshold belief distribution, i.e., the proportion of citizens with higher signals. This rank is necessarily uniformly distributed: since a citizen does not know his rank, his belief about whether the true state is below the threshold state is uniform.

We consider monotone strategy profiles of the incomplete information game, where each citizen's effort is (weakly) decreasing in his signal. Any such strategy profile gives rise to a unique *regime change threshold,* such that there will be regime change if the true state is below that threshold. Using the uniform threshold belief property, we show that in a monotone equilibrium, the total effort at the regime change threshold is the integral of the optimal effort function between $p = 0$ and $p = 1$ (the *uniform threshold*). This implies that there is a unique monotone equilibrium where the regime changes only when the state is below this uniform threshold. That is, letting θ^* be the equilibrium regime change threshold, we show $\theta^* = \int_0^1 e^*(p)dp$, where $e^*(p)$ is the effort choice of a citizen with belief p about the likelihood of regime change. Critically, although different benefit functions may change the value of the uniform threshold, this characterization holds whatever the shape of the benefit function—a property that allows us to disentangle the citizens' coordination problem and the leader's screening problem of choosing pleasure in agency benefits.

Having established results for the case of exogenous benefit and cost functions, we then consider the problem of the leader choosing the benefit function, subject to the constraint that there is a maximum possible level of benefit. What benefit function would he choose, and what would be the implications for the level and distribution of effort choices in the population? Using the equilibrium characterization described above, we show that this problem reduces to finding the benefit function that maximizes the uniform threshold. This, in turn, reduces to a screening problem, where the leader chooses effort levels to maximize the threshold, subject to the constraint that effort levels are induced in an incentive compatible way using the benefit function. That is, the leader's problem becomes $\max_{B(\cdot)} \theta^*[B(\cdot)]$ subject to incentive compatibility, and the leader's constraint that $B(e) \in [0, M]$, for a given $M \geq 0$ that, for example, captures the leader's charisma. The results that we describe above come from characterizing the solution to this mechanism design problem. Thus, our model reveals the complex strategic considerations that arise when revolutionary leaders attempt to inspire revolutionary actions in a regime change setting with coordination and information frictions. Our approach disentangles this seemingly intractable problem into two tractable problems of coordination and screening.

Our paper builds on the literature on global games in general, and applications in political economy in particular (Bueno de Mesquita 2010; Edmond 2013; Chen and Suen 2016; Egorov and Sonin 2017; Shadmehr 2017; Tyson and Smith 2018). These papers assume a form of pleasure in agency payoffs—rewards from participation in a successful revolution. However, they do not study the nature and source of these rewards. In their binary settings, a leader would give the maximum reward to the high action, and no reward to the low action. A small literature studies the role of leaders in political settings (Dewan and Myatt 2007; Bueno de Mesquita 2010; Shadmehr and Bernhardt 2019; Lipnowski and Sadler 2019), focusing on signaling and coordinating effect of leaders. In contrast, we formalize and analyze a leader's problem of inspiring anti-regime actions, characterizing the optimal design of pleasure in agency rewards in a continuous action coordination setting.

The question of what motivates participation also arises in the context of elections: what induces citizens to vote? Two possible answers are again that voting is expressive, or that the probability of being pivotal compensates for voting costs. Feddersen and Sandroni (2006) argue that neither answer is satisfactory. The expressive approach is inconsistent with evidence of strategic voting, and the likelihood of pivotality is negligible in large elections. They develop a model in which citizens derive payoffs from "doing the right thing," defined as the behavior that maximizes a utilitarian welfare if followed by all members of a reference group (e.g., all supporters of a citizen's preferred candidate)—see also Coate and Conlin (2004). In the same vein, the pleasure in agency approach is a middle ground between expressive and "pivotality" approaches in the context of protest and revolution. As we will discuss in Section 5.2, our pleasure in agency approach is based on a large qualitative and quantitative literature in sociology and political science as well as suggestive historical evidence. Our focus on pleasure in agency rewards, however, does not mean that material rewards are entirely irrelevant, or even that psychological rewards are the primary motivation for a majority of participants in all conflicts. There is evidence suggesting that in some civil wars (e.g., in Siera Leon, Uganda, or Zimbabwe) material incentives played a key role to sustain rebel armies (Weinstein 2007; Humphreys and Weinstein 2008). However, this logic does not extend to all civil wars (e.g., in El Salvador (Wood 2003)), or to less militarized domestic conflicts. On the contrary, evidence suggests that psychological rewards are the primary motivation in many anti-government protests—e.g., in Turkey and

Ukraine (Aytaç and Stokes 2019), Morocco (Lawrence 2017), and Syria (Pearlman 2018). It is hard to imagine that participants in the Green Movement that followed the contested 2009 presidential election in Iran reasonably anticipated selective material rewards.

In the real world, some individuals who join a movement are motivated by psychological rewards, while others seek material rewards. Moreover, movements vary in terms of the distribution of these individuals and the intensity of their motivations. Our model should be interpreted as a simplified picture of anti-regime movements where pleasure in agency rewards are the primary motivation for anti-regime contributions. These motivations have received scant attention in the literature even though evidence suggests that they constitute the primary motivation in many protest and revolutionary movements. A future direction of research is to combine both material and non-material incentives and investigate their interactions.

Our main methodological contribution is to analyze a setting where coordination and screening are intertwined; a screening problem arises because of the endogenous heterogeneity that is introduced by strategic uncertainty. We show that this apparently complex problem can be cleanly reduced to solving a continuous action coordination game and a screening problem. Our analysis can be adopted to study other environments in which interactions between coordination and screening naturally arise. For example, consider a threshold public good problem (Corazzini, Cotton, and Valbonesi 2015) in which players decide how much to contribute to a project that succeeds whenever the aggregate contribution exceeds an uncertain threshold about which players have private information. The project manager or the fundraiser is then like the principal who seeks to maximize the likelihood of success by choosing recognition for donors of unknown types making heterogeneous contributions. Shen and Zou (2016) also address an interaction between strategic behavior and screening. They consider a binary action coordination game but allow players to receive a stimulus by self-selection. This contingent policy intervention allows coordination on a desirable outcome at minimal cost.

As a part of our analysis we solve a coordination game of regime change with continuous actions and exogenous payoffs, and characterize the essentially unique monotone equilibrium. Continuous action global games have been studied by Frankel, Morris and Pauzner (2003) in an abstract setting, and by Guimaraes and Morris (2007) in a model of currency attacks. Our work closely follows Guimaraes and Morris (2007) who identify the uniform threshold belief property insight. There is one important variation. Guimaraes and Morris (2007) study a supermodular payoff setting where there is a unique equilibrium that is also dominance solvable. Our problem does not have supermodular payoffs, but it gives rise to monotonic strategies in equilibrium, and we end up with their characterization, although we can establish only a weaker uniqueness result—among all monotone equilibria.[3]

As another ingredient to our analysis, we solve a screening problem in a setting without transfers where a principal chooses a reward scheme with no cost, constrained only by the

[3]Assuming that an optimal benefit scheme is increasing in effort is sufficient for supermodularity and hence uniqueness among all equilibria. However, when designing the benefits, benefit schemes are endogenous, and such an assumption is ex-ante restrictive.

requirement that rewards belong to a closed interval. Our analysis of the screening problem exploits classic arguments from the screening literature. Guesnerie and Laffont (1984) is a key early reference on a rich class of screening problems that embeds monopoly problems of choosing quality (Mussa and Rosen 1978) or quantity (Maskin and Riley 1984) and the government's regulation of a monopolist (Baron and Myerson 1982). In our problem, a principal gives "benefits" to an agent in exchange for "effort", but agent utility is not linear in effort and the principal's "budget" of rewards is not "smooth"— rewards up to a level are free to the principal, but higher rewards have infinite costs. The bunching-at-the-top feature of our screening solution also appears in Laffont and Robert's (1996) and Hartline's (2016, Ch. 8) analysis of optimal auction with financial constraints; because no buyer can bid above a known budget (e.g., due to an exogenous liquidity constraint), the seller cannot separate eager buyers. While non-standard, this screening problem is essentially that of Mookherjee and Png (1994) who study the problem of choosing the optimal likelihood of detection and optimal punishment schedule by an authority. Their key result is that the marginal punishment should remain lower than marginal harms of crime for low crime levels; in fact, if monitoring is sufficiently costly, a range of less harmful acts should be legalized. Although we use different techniques in our analysis, one can show that our screening problems can be transformed to each other. This implies that our total analysis with coordination also applies to a regime designing a repression scheme (expected costs of revolutionary efforts) to minimize the likelihood of regime change.

We describe and solve the exogenous benefits case in Section 2, describe and solve the optimal reward problem in Section 3, integrate exogenous heterogeneity in Section 4, and provide evidence for our model and results in Section 5.

2 Exogenous Rewards and Punishments

2.1 Model

There is a continuum of citizens, indexed by $i \in [0,1]$, who must decide how much effort $e \geq 0$ to contribute to regime change. Exerting effort e costs $C(e)$, independent of whether there is regime change, and gives a benefit $B(e)$ only if there is regime change. Thus, if a citizen thinks that the regime change occurs with probability p, his expected payoff from choosing effort e is

$$pB\left(e\right) - C\left(e\right).$$

Regime change occurs if and only if the total contribution $\int e_i di$ is greater than or equal to the regime's strength $\theta \in \mathbb{R}$. Thus, if citizens choose effort levels $\left(e_j\right)_{j\in[0,1]}$, the payoff to a citizen from effort level e_i is

$$B\left(e_i\right) I_{\int e_j dj \geq \theta} - C\left(e_i\right).$$

The regime's strength θ is uncertain, and citizens have an improper common prior that θ is distributed uniformly on $\mathbb{R}$. In addition, each citizen $i \in [0,1]$ receives a noisy private signal

$x_i = \theta + \nu_i$ about θ, where θ and the ν_is are independently distributed with $\nu_i \sim f(\cdot)$.[4]

2.2 A Key Statistical Property: Uniform Threshold Belief

Fix any state $\hat{\theta}$. A citizen's *threshold belief* about $\hat{\theta}$ is the probability that he assigns to the event that $\theta \leq \hat{\theta}$. The *threshold belief distribution* at $\hat{\theta}$ is the distribution of threshold beliefs about $\hat{\theta}$ in the population when the true state is $\hat{\theta}$. We will show that the threshold belief distribution is always uniform, regardless of the state $\hat{\theta}$ and the distribution of noise.

To state the uniform threshold belief property formally, write $H\left(.|\hat{\theta}\right)$for the cdf of the threshold belief distribution at $\hat{\theta}$. Thus, the proportion of citizens with threshold belief about $\hat{\theta}$ below p when the true state is $\hat{\theta}$ is $H\left(p|\hat{\theta}\right)$. The key to establishing the result is the observation that, because θ is distributed uniformly so that there is no prior information about θ, one can consider θ as a signal of x_i, writing $\theta = x_i - \nu_i$. So the threshold belief about $\hat{\theta}$ for a citizen observing x_i is simply the probability that $x_i - \nu_i \leq \hat{\theta}$, or $1 - F(x_i - \hat{\theta})$. Thus, a citizen observing x_i has a threshold belief less than or equal to p if $1 - F(x_i - \hat{\theta}) \leq p$, or $x_i \geq \hat{\theta} + F^{-1}(1-p)$. Now, if the true state is $\hat{\theta}$, the proportion of citizens observing a signal greater than or equal to $\hat{x}$ will be the proportion of citizens with $\hat{\theta} + \nu_i \geq \hat{x}$, whic is $1 - F\left(\hat{x} - \hat{\theta}\right)$. Combining these observations, we have that $H\left(p|\hat{\theta}\right) = 1 - F\left(\hat{x} - \hat{\theta}\right)$ where $\hat{x} = \hat{\theta} + F^{-1}(1-p)$. Substituting for $\hat{x}$, we have

$$H\left(p|\hat{\theta}\right) = 1 - F\left(\hat{x} - \hat{\theta}\right) = 1 - F\left(F^{-1}(1-p)\right) = 1 - (1-p) = p.$$

We have shown the uniform threshold belief property:

Lemma 1. ***(Guimaraes and Morris (2007))*** *The threshold beliefs distribution is always uniform on* $[0,1]$, *so that* $H\left(p|\hat{\theta}\right) = p$ for all $p \in [0,1]$ and $\hat{\theta} \in \mathbb{R}$.

To gain intuition for the result, suppose that noise itself was uniformly distributed on the interval $[0,1]$. If the true state was $\hat{\theta}$, a citizen observing x_i in the interval $\left[\hat{\theta}, \hat{\theta}+1\right]$ would have threshold belief $\hat{\theta} + 1 - x_i$, and we would have uniform threshold belief. But if the noise had some arbitrary distribution, we could do a change of variable, replacing the level of a citizen's signal with its percentile in the distribution. Because citizens do not know their own percentiles, the same argument then holds.

[4]The assumption of an improper common prior is standard in this literature. At the cost of further notation, one can replace this assumption with a prior that is uniformly distributed on a sufficiently large bounded interval. More importantly, results proved under the uniform prior assumption can also be reproduced with an arbitrary smooth prior when the noise is sufficiently small: see Carlsson and van Damme (1993), Morris and Shin (2003) and Frankel, Morris and Pauzner (2003). All our results thus hold with general priors in the limit as noise goes to zero.

2.3 Equilibrium

Our results will depend crucially on how citizens behave if they assign a fixed probability to the success of the revolution. It is convenient to first state our results depending only on the monotonicity and continuity of the solution to this problem, and then to identify conditions on the exogenous cost function $C(\cdot)$ and the exogenous (in this section) or endogenous (in the next section) benefit function $B(\cdot)$ that imply monotonicity and continuity.

Definition 1. The *optimal effort correspondence* is

$$e^*(p) = \arg\max_{e \geq 0} pB(e) - C(e);$$

we write

$$e_{min} = min\left(e^*(0)\right) \text{ and } e_{max} = max\left(e^*(1)\right).$$

We maintain the assumption that the maximum and minimum exist (and are finite).

Definition 2. The optimal effort correspondence is *weakly increasing* if $p_2 > p_1$ and $e_i \in e^*(p_i)$, $i \in \{1, 2\}$, imply $e_2 \geq e_1$.

That is, a citizen who is strictly more optimistic about the likelihood of success than another will put in at least as much effort as him. Note that if e^* is weakly increasing, it is almost everywhere single-valued. We first consider the complete information case where there is common knowledge of θ.

Proposition 1. *Suppose that the optimal effort correspondence is weakly increasing and that there is complete information. There is an equilibrium with regime change if $\theta \leq e_{max}$, e.g., with all players choosing effort level e_{max} ; and there is an equilibrium without regime change if $\theta > e_{min}$, e.g., with all players choosing effort level e_{min}. Thus, there are three cases: if $\theta \leq e_{min}$, there are only equilibria with regime change; if $\theta > e_{max}$, there are only equilibria without regime change; and if $e_{min} < \theta \leq e_{max}$, then there are equilibria with regime change and equilibria without regime change.*

Proof. If there is regime change in equilibrium, all citizens must choose $e \in e^*(1)$; this can give rise to regime change only if $e_{max} \geq \theta$. If there is no regime change in equilibrium, all citizens must choose $e \in e^*(0)$; this can give rise to no regime change only if $e_{min} < \theta$. $\square$

We now consider the incomplete information game where a citizen's only information is the signal that he receives, maintaining the assumption that the optimal effort correspondence is weakly increasing. Recall that because $e^*(p)$ is a weakly increasing correspondence, it is single valued almost everywhere. A strategy for a citizen i is a mapping $s_i : \mathbb{R} \to \mathbb{R}_+$, where $s_i(x_i)$ is the effort level of citizen i when he observes signal x_i. Each strategy profile $(s_i)_{i \in [0,1]}$ will give rise to aggregate behavior

$$\hat{s}(\theta) = \int_{i=0}^{1} \left(\int_{\nu_i=-\infty}^{\infty} s_i(\theta + \nu_i) f(\nu_i) d\nu_i \right) di.$$

If all citizens follow weakly decreasing strategies, then $\hat{s}$ is weakly decreasing, and there is a unique threshold θ^* such that $\hat{s}(\theta^*) = \theta^*$. Then, $\hat{s}(\theta) \geq \theta$ and there will be a regime change if and only if $\theta \leq \theta^*$. Thus, a citizen observing a signal x_i will assign probability $G(\theta^*|x_i)$ to the event that $\theta \leq \theta^*$ and thus to regime change. We conclude that each citizen must be following the strategy with

$$s^*(x_i) = e^*(G(\theta^*|x_i)).$$

Letting $p = G(\theta^*|x_i)$, and recalling that $H(p|\theta^*)$ is the cdf of $G(\theta^*|x_i)$ conditional on θ^*, this implies that the aggregate effort of citizens when the true state is θ^* is

$$\hat{s}(\theta^*) = \int_{p=0}^{1} e^*(p) \; dH(p|\theta^*).$$

By Lemma 1, we know that $H(p|\theta^*)$ is a uniform distribution. We also know that $\theta^* = \hat{s}(\theta^*)$. Thus,

$$\theta^* = \hat{s}(\theta^*) = \int_{p=0}^{1} e^*(p)\, dp.$$

Summarizing this analysis, we have:

Lemma 2. *If the optimal effort correspondence e^* is weakly increasing, then there is a unique monotonic equilibrium where all citizens follow the essentially unique common strategy*

$$s^*(x_i) = e^*(1 - F(x_i - \theta_{e^*})). \tag{1}$$

In this equilibrium, there is regime change whenever $\theta \leq \theta_{e^*}$, where the regime change threshold θ_{e^*} is given by

$$\theta_{e^*} = \int_{p=0}^{1} e^*(p)\, dp. \tag{2}$$

Guimaraes and Morris (2007) proved the stronger result that the corresponding equilibrium was unique among all equilibria, but under assumptions implying supermodular payoffs. The above proposition holds under weaker assumptions, requiring only that $e^*(p)$ is increasing in p. We have proved uniqueness within the class of monotonic equilibria, but leave open the question of whether non-monotonic equilibria exist.

We can provide a clean characterization of the regime change threshold if e^* is single-valued and continuous. Note that, under these assumptions, $e^*(0) = \{e_{min}\}$ and $e^*(1) = \{e_{max}\}$.

Lemma 3. *If the optimal effort correspondence e^* is single-valued, continuous and weakly increasing, and $e_{min} = 0$, then the regime threshold in the unique monotonic equilibrium is given by*

$$\theta_{e^*} = \int_{e=0}^{e_{max}} \left(1 - \frac{C'(e)}{B'(e)}\right) de. \tag{3}$$

We will use this characterization of the regime change threshold in our main result.

2.4 Optimal Effort

Having shown the existence and characterizations of a unique monotonic equilibrium for a given monotonic optimal effort correspondence, we now identify sufficient conditions on the exogenous cost function $C(\cdot)$ for the existence of such a monotonic optimal effort correspondence. We first present a weak sufficient condition for monotonicity, and then report standard sufficient conditions for continuity as well as monotonicity.

First suppose that punishments are increasing in effort, so that $C(e)$ is increasing.

Proposition 2. *If $C(e)$ is strictly increasing in e, then any selection from optimal effort correspondence is weakly increasing. Thus the unique monotone equilibrium has regime change threshold given by equation (2) of Lemma 2.*

Proposition 2 will be used to establish the existence and characterization of a monotone equilibrium for any reward scheme B.

We can also give completely standard conditions for continuity:

Proposition 3. *Suppose that costs and benefits are (1) twice continuously differentiable, (2) zero with zero effort ($C(0) = B(0) = 0$), (3) strictly increasing ($C'(e) > 0$ and $B'(e) > 0$ for all e) and (4) convex and strictly concave respectively ($C''(e) \geq 0$ and $B''(e) < 0$ for all e). Then e^* is a continuous and increasing function with $e^*(0) = 0$. Thus the unique monotone equilibrium has the regime change threshold given by equation (3) of Lemma 3.*

Clearly, one could give weaker conditions for continuity. The endogenous optimal reward scheme identified in the next section will not satisfy the restrictions on the exogenous reward scheme of Proposition 3. Nonetheless, we will show that the equilibrium effort function is continuous, and we will appeal to the characterization of the equilibrium regime threshold of equation (3).

We conclude this section by reporting a class of examples. Suppose that the exogenous reward scheme is $B(e) = \sqrt{e}$ and the punishment scheme is $C(e) = e^n$, $n \geq 1$. In this case, we have from Proposition 3 that $e^*(p) = (\frac{p}{2n})^{\frac{2}{2n-1}}$ and

$$\theta_{e^*} = \frac{2n-1}{2n+1} (2n)^{\frac{2}{1-2n}}.$$

Figure 1 illustrates.

This section provided an analysis of equilibrium with exogenous rewards and punishments. This characterization could be used to address many positive and policy questions. This paper will focus on one: optimal reward schemes.

3 Optimal Reward Schemes

We have analysed the equilibrium effort for fixed benefit and cost functions and identified the implied regime change threshold. We now fix an exogenous cost function C, maintaining the assumptions that C is twice continuously differentiable, strictly increasing and convex,

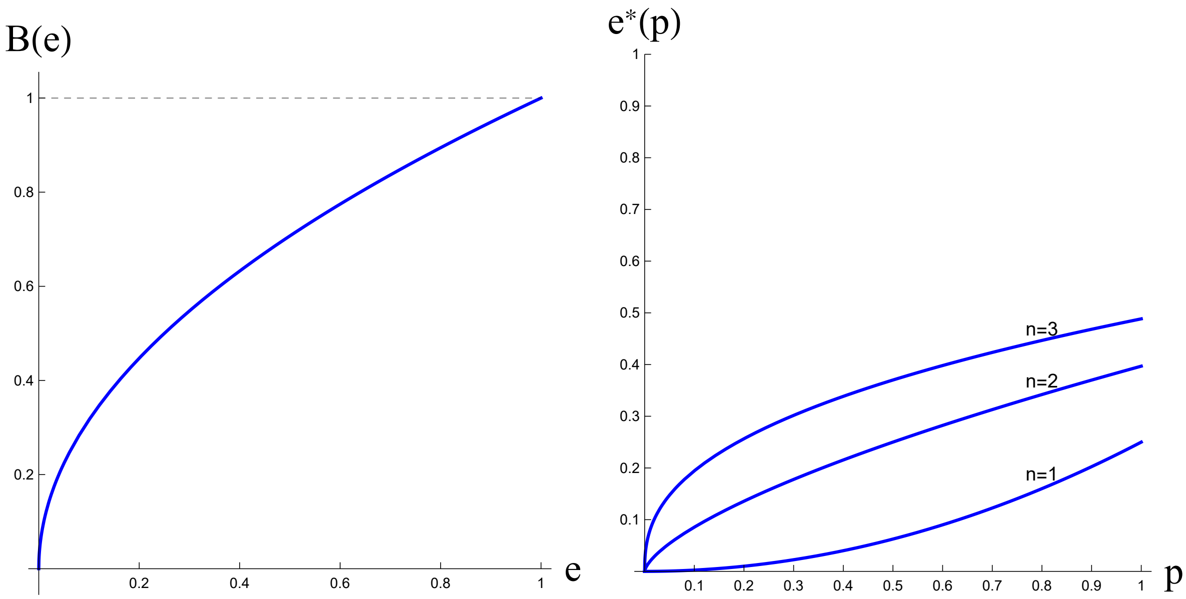

Figure 1: Optimal Effort for Exogenous Rewards. The exogenous reward scheme is $B(e) = \sqrt{e}$, and the exogenous punishment scheme is $C(e) = e^n,\ n \geq 1$. The right panel depicts $e^*(p)$ for $n = 1, 2, 3$.

with $C(0) = 0$. These are the properties of the cost function assumed in Proposition 3. We investigate the optimal design of the benefit function $B(e)$ by a revolutionary leader who aims to maximize the likelihood of regime change. Charismatic revolutionary leaders can inspire citizen participation by assigning psychological rewards to different levels of anti-regime activities. Through speeches, writings, and meetings, they activate, create, and manipulate intrinsic motivations by creating identities and innovative framing of events, and by "identification, idealization, and elevation of one or more values presumed basic to prospective constituents" (Snow et al. 1986, p. 469). Unlike material benefits, these psychological rewards are non-rival. For example, in a religious context, if a fighter tortured in the righteous struggle to bring down an unjust state or a wicked ruler is to receive one marble castle in heaven, God can build as many castles as there are fighters who deserve them. Still, even charismatic leaders can incite only so much intrinsic motivation in their potential followers. Thus, we require that $B(e) \leq M$ for some exogenous $M > 0$. The upper bound on rewards reflects limitations on the leader's skills and charisma, as well as other exogenous aspects of the environment, such as "cultural idioms" (Skocpol 1997) or "repertoire of common symbolics" (Dabashi 1993), that facilitate the leader's task of creating, activating, and manipulating intrinsic motivations for revolutionary activities.

Proposition 2 implies that $e^*(p)$ is weakly increasing (recall that this proposition did not impose any assumptions on the benefit function). By Proposition 2, we then know that—for any choice of $B(\cdot)$—there will be a unique equilibrium in which regime change occurs when $\theta \leq \theta_{e^*} = \int_{p=0}^{1} e^*(p)dp$. Thus the leader's optimization problem reduces to choosing reward scheme $B(e)$ to maximize the regime change threshold

$$\int_{p=0}^{1} e^*(p)dp \tag{4}$$

subject to the optimality of the effort function

$$e^*(p) \in \arg\max_{e\geq 0}\ p\ B(e) - C(e),$$

and the feasibility of the reward scheme

$$B(e) \in [0, M] \text{ for all } e. \tag{5}$$

A reward scheme B^* that maximizes this problem is an *optimal reward scheme* (there will be some indeterminacy in optimal reward schemes). An effort function e^* arising in this problem is an *optimal effort function* (there will be an essentially unique optimal effort function). The regime change threshold θ^* that results from the optimal effort function is the (essentially unique) *optimal regime change threshold.* To provide intuition and as a benchmark, we first present the analysis with linear costs, and then present our main result.

3.1 Linear Costs

Proposition 4. *Suppose costs are linear, with constant marginal cost c. An optimal reward scheme B^* is the step function*

$$B^*(e) = \begin{cases} 0 & ; e < M/2c \\ M & ; e \geq M/2c. \end{cases} \tag{6}$$

The optimal effort function e^* is the step function

$$e^*(p) = \begin{cases} 0 & ; p < 1/2 \\ M/2c & ; p \geq 1/2. \end{cases} \tag{7}$$

And the optimal regime change threshold is

$$\theta_{e^*} = M/4c.$$

Any reward scheme giving rise to the optimal effort function is optimal. Thus, it is enough, for example, that $B^*(0) = 0$, $B^*(M/2c) = M$ and $B^*(e) \leq 2ce$ for $e \leq M/2c$.

To prove the proposition, first assume the optimal reward scheme is a step function, where citizens get the maximum benefit M if they exert at least effort $\hat{e}$; otherwise they get no benefit. In this case, citizens either do not participate (i.e., choose effort 0) or participate with effort $\hat{e}$. A citizen with optimism (probability of regime change) p participates whenever $pM \geq c\hat{e}$, i.e., $p \geq c\hat{e}/M$. Thus, the total effort is $(1 - c\hat{e}/M)\,\hat{e}$. This is maximized by setting $\hat{e} = M/2c$. Thus we have a proof of the proposition if we could establish that the optimal reward schemes would be a step function.

To complete the proof, we show that we can restrict attention to step functions by showing that this problem reduces to the monopoly pricing problem. Consider a monopolist selling a single unit to buyers whose valuations are uniformly distributed on the interval $[0, M/c]$. The monopolist could sell using a posted price mechanism. At a posted price of $\hat{e}$, buyers with valuations above $\hat{e}$ would buy and pay $\hat{e}$, and buyers with valuations below $\hat{e}$ would not buy. However, the monopolist could also sell using a more complicated mechanism, offering a price schedule for probabilities of being allocated the object. By the revelation principle, we can restrict attention to direct mechanisms. If a type-p buyer has valuation pM/c, a direct mechanism is described by a payment $e^*(p)$ that buyer p will make to the seller and a probability $B\,(e^*(p))\,/M$ of receiving the object. The seller's revenue will now be $\int_{p=0}^{1} e^*(p)dp$ and incentive compatibility will require that $e^*(p) \in \arg\max_{e \geq 0}\ p\ B(e)/c - e$. The revenue is the maximand of our problem, and the incentive compatibility condition is the incentive compatibility condition of our problem. But Riley and Zackhauser (1983) established that the optimal mechanism in this problem is a posted price mechanism (see Börgers (2015, Ch. 2) for a modern textbook treatment), and simple calculations show that the optimal price is $M/2c$. This proves Proposition 4.

Our main result establishes that if the cost function is strictly convex, this step function is "smoothed": it remains optimal for a mass of the most optimistic citizens to choose a maximum effort level and receive the maximum reward. And it remains optimal (under some conditions) for a mass of the least optimistic citizens to not participate, i.e., choose 0 and receive no benefit. However, in between, there are strictly increasing benefits and effort.

We conclude our analysis of the linear case by discussing why bunching at the top starts exactly when $p = 1/2$. Recall that our strategic problem delivers a particular distribution over levels of optimism p: the uniform distribution. In the monopoly problem, this is equivalent to considering a linear demand curve. It is useful to consider what would have happened in the screening problem if p was not uniformly distributed, but distributed according to density f and corresponding cdf F. Under standard assumptions,[5] the problem reduces to solving for the critical effort level $\hat{e}$ to receive the (maximum) benefit. As before, the citizen will participate at this critical effort level if $p \geq c\hat{e}/M$. But now total effort would be $(1 - F\,(c\hat{e}/M))\,\hat{e}$. Letting e^* be the optimal $\hat{e}$, the first order condition implies

$$e^* = \frac{1 - F\,(ce^*/M)}{f\,(ce^*/M)} \frac{M}{c},$$

and the corresponding critical optimism level is

$$p^* = ce^*/M = \frac{1 - F\,(ce^*/M)}{f\,(ce^*/M)}.$$

[5] In particular, f should satisfy the standard regularity condition that $\frac{1-F(\cdot)}{f(\cdot)}$ is decreasing (decreasing marginal revenue in the monopoly case).

When p is distributed uniformly, the above equation implies that $ce^*/M = 1/2$. Thus, $p^* = 1/2$ arises as critical optimism because p is uniformly distributed.

3.2 Strictly Convex Costs and Main Result

Proposition 5. *Suppose costs are strictly convex.*

- For some "maximum effort" $\bar{e} > 0$, the optimal reward scheme is continuous, strictly increasing and strictly convex on the interval $[0, \bar{e}]$ with $B(0) = 0$ and $B(\bar{e}) = M$. Moreover, the marginal benefit of effort $B'(e)$ is strictly greater than the marginal cost of effort $C'(e)$ whenever the marginal benefit is non-zero

- The optimal effort function is continuous and weakly increasing; it is strictly increasing on an interval $\left[\underline{p}, 1/2\right]$, equal to 0 when $p \leq \underline{p}$, and equal to $\bar{e}$ when $p \geq 1/2$.

- The critical level of optimism $\underline{p}$ is strictly greater than 0 if and only if the marginal cost is strictly positive at zero effort ($C'(0) > 0$).

To illustrate the proposition, we graph the optimal reward scheme and optimal effort function for some examples. First, suppose $C(e) = e^n$, $n > 1$, and $M = 1$. Figure 2 illustrates the optimal reward and optimal effort schemes in this case for different values of n. The derivation is reported in the appendix. Note that, as n approaches 1, we approach the linear case described above (as noted earlier, there is an indeterminacy in the optimal reward scheme, and this limit is piece-wise linear, rather than the step function reported in Proposition 4). This class of examples has $C'(0) = 0$. Figure 3 illustrates an example where $C(e) = e^2 + 0.1e$, so that the marginal cost at 0 is strictly positive, $C'(0) = 0.1$.

Further details about the optimal reward scheme and the optimal effort function, as well as an algorithm for calculating them, are reported below in the proof of the proposition. In particular,[6]

$$\theta^* = \bar{e} - \frac{M}{4}\frac{1}{C'(\bar{e})}, \tag{8}$$

and $\bar{e}$ is the unique solution to

$$\sqrt{C'(\bar{e})} \int_0^{\bar{e}} \sqrt{C'(x)}\, dx = M/2. \tag{9}$$

Moreover, the strictly increasing segment of the optimal contribution schedule has a simple characterization

$$e^*(p) = C'^{-1}((2p)^2 C'(\bar{e})). \tag{10}$$

The basic features of the optimal reward scheme and optimal effort function are intuitive: both are weakly increasing and continuous, with optimal benefit strictly increasing

[6] Equation (9) follows from equations (12), (14) and (16). Equation (8) is calculated through integration by parts: θ^* is $\bar{e}$ minus the area above $e^*(p)$ and below $\bar{e}$. From equation (14), this area is λM. From (12) and (16), $\lambda = 1/4C'(\bar{e})$.

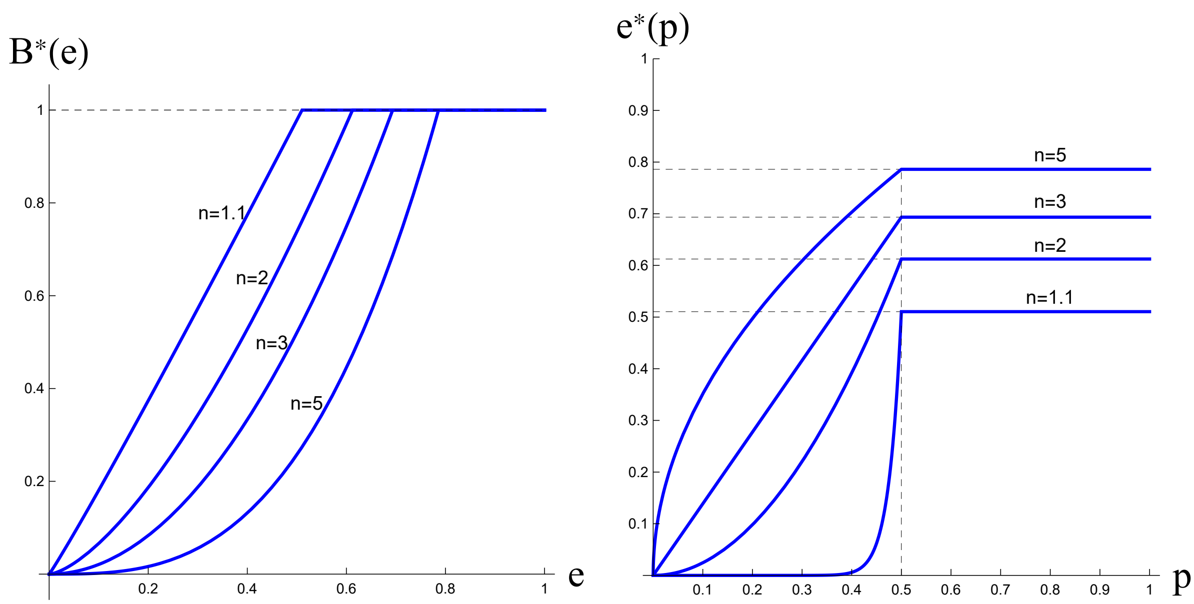

Figure 2: Optimal reward schedule, $B^*(e)$, and its induced effort schedule, $e^*(p)$, for $M = 1$ and cost function $C(e) = e^n$, $n > 1$. As the right panel illustrates, when the cost function approaches the linear $C(e) = e$, the effort schedule approaches a step function. Note that without screening, the leader would require $e^*(p) = p^{1/n}$ from a type-p citizen, and the structure of efforts would be very different, e.g., there would be no bunching.

and ranging from 0 to the maximum. Before presenting a proof, we provide an intuition for further properties. First, consider the shape of the optimal reward scheme when it is strictly increasing. In this case, marginal benefit exceeds the marginal cost. If the marginal cost did exceed the marginal benefit at some effort level, then any effort in the neighborhood of that level could not arise in equilibrium. But then it would be possible to replace the reward scheme with one that was constant in that neighborhood and increasing faster elsewhere, in a way that would increase the overall effort. More mechanically, the first order condition that $pB'(e^*(p)) = C'(e^*(p))$ implies that $B'(e^*(p)) > C'(e^*(p))$ when $p < 1$. To show strict convexity, observe that effort depends on the ratio of marginal costs to marginal benefits. Increasing marginal benefits will thus have more impact on effort when marginal cost is higher. But marginal cost is itself increasing in effort (by assumption) so marginal benefit will also be increasing. More formally, once we show that $e^*(p)$ is continuous and weakly increasing (and single-valued with $e^*(0) = 0$), then Lemma 3 characterizes θ_{e^*} that the leader seeks to maximize subject to the constraint that $B(e) \leq M$. Differentiating the integrand in θ_{e^*} with respect to $B'(e)$ yields $C'(e)/(B'(e))^2$. Because $C'(e)$ is increasing by assumption, the marginal gains of raising $B'(e)$ are higher for higher effort levels. In contrast, the marginal costs of raising $B'(e)$ to the leader are constant, because the leader has a fixed "budget of slopes" ($\int_{e=0}^{\bar{e}} B'(e)de \leq M$ for some $\bar{e}$). This implies that an optimal $B'(e)$ is increasing, and hence an optimal $B(e)$ is convex.

Now consider the "exclusion" regions, where effort is constant. First, consider exclusion

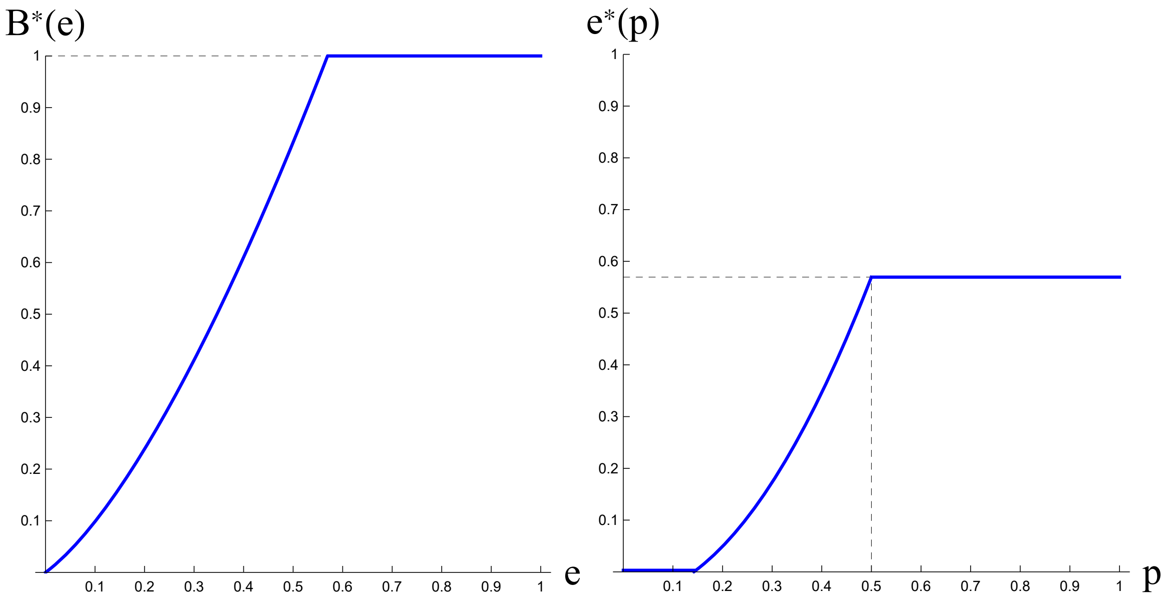

Figure 3: Optimal reward schedule, $B^*(e)$, and its induced effort schedule, $e^*(p)$, for $M = 1$ and cost function $C(e) = e^2 + 0.1e$. As the right panel illustrates, when $C'(0) > 0$, citizens with (endogenously) pessimistic beliefs about the likelihood of regime change do not exert any effort ($\underline{p} > 0$).

at the bottom. Recall that the optimal effort $e^*(p)$ is continuous and increasing, and consider the smallest level of optimism $\underline{p}$ after which $e^*(p)$ becomes strictly positive. Then, $\underline{p}\, B'(0) = C'(0)$, and hence $C'(0) > 0$ implies $\underline{p} > 0$. In fact, because the marginal costs of raising $B'(e)$ are constant, while its marginal gains are $C'(e)/(B'(e))^2$ as we discussed above, $C'(e)/B'(e)$ is proportional to $\sqrt{C'(e)}$,[7] and hence $\underline{p} > 0$ exactly when $C'(0) > 0$. Next, consider high optimism, and suppose that optimal effort was strictly increasing in a neighborhood around 1. The optimal reward scheme would then have to be strictly increasing in the neighborhood of $\bar{e}$. Now suppose that we considered benefit functions that were equal to M in an open neighborhood of $p = 1$. This would decrease effort in the open neighborhood, but would allow optimal effort to be increased at other levels of p. For very small neighborhoods, the latter effect would be of higher order than the former. More formally, the convexity of an optimal $B(e)$ (and that $B(e) \geq C(e)$ in the relevant range) implies that an optimal $B(e)$ hits its constraint M before $C(e)$ does. In turn, this implies that an optimal effort scheme is constant for high values of p.

Finally, we examine why effort reaches its maximum exactly at $p = 1/2$. We already gave an explanation of where $p = 1/2$ comes from in the linear case: it follows from the (endogenous) property that p is uniformly distributed on the interval $[0, 1]$. If p was drawn from a distribution $f(p)$, the corresponding threshold (in the linear cost case) would be when $\frac{1-F(p)}{f(p)} = p$, which corresponds to the point where marginal revenue is zero in the

[7] Formally, $C'(e)/(B'(e))^2 = \lambda \Rightarrow B'(e) = \sqrt{C'(e)/\lambda} \Rightarrow C'(e)/B'(e) = \sqrt{\lambda}\,\sqrt{C'(e)}$.

monopoly problem . The same trade-off between intensive and extensive margins arises in the non-linear case. One can show that with a general distribution f of p and under some regularity conditions, the screening problem of Proposition 5 yields the critical p^* that solves

$$\frac{1-F(p^*)}{f(p^*)} = -\frac{f\left(p^*\right)/p^*}{\left(f\left(p^*\right)/p^*\right)'}.$$

In the uniform case, this expression gives $1-p^* = -\frac{1/p^*}{-1/(p^*)^2} = p^*$, and hence $p^* = 1/2$.[8]

3.3 Proof

We now present the proof of Proposition 5, formalizing the earlier intuitions. We defer more technical steps to the appendix, where we formalize the leader's problem as a mechanism design problem, and prove that an optimal reward function $e^*(p)$ is continuous and increasing with $e^*(0) = 0$. Then, using Lemma 3, we show that the leader's problem is to choose $(\bar{e}, B')$ to maximize

$$\int_0^{\bar{e}} \left(1 - \frac{C'(e)}{B'(e)}\right) \; de$$

subject to the marginal benefit constraint that[9]

$$B\left(\bar{e}\right) = \int_{e=0}^{\bar{e}} B'(e)de = M.$$

This shows that one can think of the leader's problem as him deciding a maximum level of effort $\bar{e} \geq 0$ to induce, and then deciding how to allocate a fixed supply of marginal benefit to different effort levels $B' : [0, \bar{e}] \to \mathbb{R}$.

Thus, the leader's problem becomes a constrained point-wise optimization. The Lagrangian is

$$\begin{aligned} L &= \int_0^{\bar{e}} \left(1 - \frac{C'(e)}{B'(e)}\right) \; de + \lambda \left(M - \int_{e=0}^{\bar{e}} B'(e)de\right) \\ &= \int_0^{\bar{e}} \left(1 - \frac{C'(e)}{B'(e)} - \lambda \; B'(e)\right) \; de + \lambda \; M. \end{aligned}$$

Optimal $B'(e)$ simplifies to a point-wise maximization

$$\frac{\partial L}{\partial B'(e)} \left(1 - \frac{C'(e)}{B'(e)} - \lambda \; B'(e)\right) = \frac{C'(e)}{[B'(e)]^2} - \lambda = 0. \tag{11}$$

[8]In Section 4, we discuss an extension of the model where there is exogenous heterogeneity among citizens; the analysis of this problem formally reduces to allowing more general distributions on p.

[9]There are two other constraints: $\bar{e} \geq 0$ and $B'(e) \geq C'(e)$. However, as it will be clear from the solution, these two constraints are automatically satisfied.

This shows that at the optimum, the marginal gain of raising $B'(e)$, i.e., $C'(e)/[B'(e)]^2$, equals its marginal cost λ. Thus,

$$B'(x) = \frac{1}{\sqrt{\lambda}} \sqrt{C'(x)}, \tag{12}$$

and

$$B(e) = \int_0^e B'(x)dx = \frac{1}{\sqrt{\lambda}} \int_0^e \sqrt{C'(x)}\, dx. \tag{13}$$

Combining this with the constraint yields

$$B(\bar{e}) = \frac{1}{\sqrt{\lambda}} \int_0^{\bar{e}} \sqrt{C'(x)}\, dx = M. \tag{14}$$

Moreover, optimal $\bar{e}$ also satisfies the first order condition

$$\frac{\partial L}{\partial \bar{e}} = 1 - \frac{C'(\bar{e})}{B'(\bar{e})} - \lambda\, B'(\bar{e}) = 0 \;\Rightarrow\; 1 - \frac{C'(\bar{e})}{B'(\bar{e})} = \lambda\, B'(\bar{e}). \tag{15}$$

Substituting for λ from (11) into (15) yields

$$1 - \frac{C'(\bar{e})}{B'(\bar{e})} = \frac{C'(\bar{e})}{B'(\bar{e})} \;\Rightarrow\; \frac{C'(\bar{e})}{B'(\bar{e})} = \frac{1}{2}. \tag{16}$$

Thus, if $\bar{e}$ and λ are the solution to equations (14) and (16), we have solved for B.

From equation (12), we have $\frac{C'(e)}{B'(e)} = \sqrt{\lambda C'(e)} > 0$. This has two implications: (1) because $C(e)$ is strictly convex, $B(e)$ is strictly convex for $0 < e < \bar{e}$; and $\frac{C'(e)}{B'(e)}$ is strictly increasing. Thus, $\frac{C'(\bar{e})}{B'(\bar{e})} = \frac{1}{2}$ implies that $B'(e) \geq C'(e)$. (2) $e^*(p)$ is strictly increasing between $\underline{p} = \sqrt{\lambda}\sqrt{C'(0)}$ and $p = 1/2$; it is equal to 0 for $p \leq \underline{p}$ and equal to $\bar{e}$ for $p \geq 1/2$.

4 Exogenous Heterogeneity

So far, we have assumed that one citizen is distinguished from another only by his belief about the likelihood of regime change, which depends on his private information. Thus, citizens are easier or harder to motivate solely based on these private beliefs about the likelihood of regime change: optimistic citizens are easier to motivate, pessimistic ones are harder to motivate. However, in addition to their beliefs about the likelihood of success, differences in the citizens' past interactions with the state, their social and economic status, and their inherent psychological dispositions, such as religious convictions, can influence how they react to a leader's motivations. For example, richer citizens may be harder to motivate (low α below) (White 1989), or those who were treated worse by the state may be easier to motivate (high α below) (Wood 2003). In this section, we study the effect of increases in heterogeneity among anti-regime forces on the likelihood of regime change.[10]

[10] Adding ex ante heterogeneity is discussed in Section 6.1 of Guimaraes and Morris (2007).

We extend citizen payoffs, so that a citizen i is characterized by both his (endogenous) private belief p_i that the regime change occurs and his exogenous propensity to be influenced by the leader α_i. We assume that α_is are independent of θ and citizens' signals, and $\alpha_i \sim iid F_\alpha$. Thus, if a citizen i exerts effort e_i, his expected payoff is

$$p_i \alpha_i B(e_i) - C(e_i).$$

Because α and p are independent, the Uniform Threshold Belief property still holds and the distribution of beliefs is uniform. Thus, we can think of $p\alpha$ as the augmented type t, where $p \sim U[0,1]$, $\alpha \sim F_\alpha$, and $t \sim G$. Then, equation (2) becomes

$$\theta^* = \int e^*(t)\ dG(t). \tag{17}$$

If the leader knew the citizens' types, he could get a citizen with type $t = p\alpha$ to contribute $tM = e(t)$, and hence $\theta^* = M \cdot E[t]$. But because these types are private, the leader has a screening problem. In this extension, we focus on the case where the cost of effort is linear $C(e) = ce$, and $F_\alpha = U[\mu - \delta, \mu + \delta]$, with $\mu > 0$ and $\delta \in (0, \mu)$. As analyzed in the previous section, linear costs imply that the leader's optimal reward scheme is to give the maximum benefit M to those whose efforts exceed some optimally set threshold $\hat{e}$. In turn, those whose augmented type exceeds $c\hat{e}/M$ exert effort $\hat{e}$, and the rest do not participate. Thus, the leader's problem becomes

$$\max_{\hat{e}} (1 - G_\delta(c\hat{e}/M))\ \hat{e}, \tag{18}$$

where we have made explicit the dependence of this distribution on the degree of dispersion in exogenous heterogeneity, δ. The term $(1 - G_\delta(c\hat{e}/M))$ captures the extensive margin of participation at the equilibrium regime change threshold (this is implicit in $p \sim U[0,1]$, which is used in the construction of G_δ), and the term $\hat{e}$ captures the intensive margin.

Proposition 6. *Suppose costs are linear, and exogenous heterogeneity is distributed uniformly, $F_\alpha = U[\mu - \delta, \mu + \delta]$, with $\mu > 0$ and $\delta \in (0, \mu)$. The equilibrium regime change threshold $\theta^*(\delta)$ is strictly decreasing in the dispersion of exogenous heterogeneity: $d\theta^*(\delta)/d\delta < 0$. Moreover, the extensive margin of participation in revolution at the equilibrium regime change threshold $\theta^*(\delta)$ is decreasing in the dispersion of exogenous heterogeneity δ.*

To see the intuition, suppose $p = 1$, so that there is no coordination consideration, and the distribution of augmented types is $U[\mu - \delta, \mu + \delta]$. Now, increasing δ is a mean-preserving spread, which rotates the "demand curve" $(1 - G_\delta(t))$ counter clockwise around the median $t = \mu$. Thus, as Johnson and Myatt (2006) analyze in detail, when the leader's optimal strategy is to seek the participation of a majority of potential participants (i.e., when the marginal participating type is below μ), a mean-preserving spread lowers $\theta^*(\delta)$. Our results show that the coordination aspects of the game, which cause the distribution of augmented types to be more complex, do not change these properties.

The literature typically focuses on inequality in the whole population, showing that higher inequality in the whole population *increases* the likelihood of regime change. For example, in Acemoglu and Robinson (2001, 2006), the society is divided into two groups: the rich who want to maintain the status quo, and the poor who seek to change it. They find that as inequality between the rich and the poor increases, the likelihoods of instability and regime change increase. In contrast, our heterogeneity corresponds to inequality or other forms of heterogeneity among potential revolutionaries, e.g., higher dispersion of income or religious convictions within the poor. Our result shows that higher heterogeneity among potential revolutionaries *reduces* the likelihood of regime change. These findings support "divide and conquer" tactics as a useful counterrevolutionary strategy: the more one can instill heterogeneity among potential revolutionaries, the more one can maintain the status quo.

Proposition 6 also shows that, at the equilibrium regime change threshold, the fraction of citizens who contribute to the revolution is decreasing in the dispersion of exogenous heterogeneity—reflecting the combined effect of the change in the distribution of heterogeneity and the leader's adjustment. The size of the revolutionary coalition can have important consequences for the nature of post-revolution regime. As we discussed in the Introduction, when regime change happens mainly due to the efforts of a relatively small group who contributed to the cause far more than other polity members, the post-revolution regime is more likely to reflect their preferences and ideologies. In sum, our analysis suggests that higher dispersion of exogenous heterogeneity not only reduces the likelihood of regime change, but also reduces participation in the neighborhood of the equilibrium regime change threshold, with adverse consequences for the post-revolution regime. The goal of this section was to demonstrate that our model can be extended to allow for exogenous heterogeneity, and drive more empirical predictions. Thus, we focused on the special case of linear costs and uniform exogenous heterogeneity, which allow for a clean analytical solution. One must perform similar analyses for a broader class of examples and expand the model to include a post-revolution stage to arrive at more definitive and general conclusions. These directions are left for future research.

5 Discussion: Participation, Motivation, Leadership

5.1 Participation

What is the nature of participation in revolutions, civil wars, or other anti-regime collective actions? Charles Tilly defines *Contentious Performances* (Tilly 2008) to be political, collective claim-making actions that fall outside routine social interactions. Such contentious actions take various forms. Examples include machine-breaking, petitioning, arson (e.g., common in the Swing Rebellion), cattle maiming (e.g., in the Tithe War in Ireland), sit-ins in foreign embassies (e.g., in the 1905-07 Iranian Constitutional Revolution), noise-making through cacerolazo (banging on pots and pans, e.g., in 1989 Caracazo in Venezuela), shouting slogans from the roof-tops at night (e.g., in the 1979 Iranian Revolution), occupations of public space or government buildings, writing open letters, wearing particular colors (e.g., green wristbands in the 2009 Iranian green movement, orange ribbons in the Ukrainian Orange Revolution, or yellow ribbons in the Yellow Revolution in the Philippines), public meetings, boycotts (e.g., in the American Revolution), parades (e.g., women suffrage movement in the early 20th century), strikes, marches, demonstrations, freedom rides, street blockades, self-immolation, suicide bombings, assassinations, hijackings, and guerrilla wars.

Some contentious actions require more effort and entail more risks than others. For example, taking up arms against a regime takes far more effort and is far more dangerous than participating in a demonstration, which in turn is typically more costly than participating in a boycott campaign. That is, different levels of anti-regime contribution correspond to different contentious performances. Thus, a fundamental question for a revolutionary leader is: what kinds of contentious performances can be elicited and how? The answer depends on the nature of the people's motivation for joining the cause.

5.2 Motivation

What motivates citizens to participate in anti-regime collective actions, where individual effects are minimal, risks are high, and the fruits of success are public? This question is the essence of Tullock's (1971) "Paradox of Revolution," which specialized Olson's (1965) *Logic of Collective Action* to revolution settings. One possible answer is that participants somehow receive selective material rewards. In many civil wars, "lootable resources" such as oil and diamond have been linked to the intensity of conflict (Ross 2006; Blattman and Miguel 2010), and selective material incentives seem to have played a role in the decision of some individuals to join rebel forces (Popkin 1979; Humphreys and Weinstein 2008). However, even in those settings, the literature distinguishes between opportunists, who participate in anticipation of immediate material rewards, and more committed activists (Weinstein 2007). Moreover, in the context of civil war, receiving material rewards in exchange for joining the rebel may not indicate the primacy of material incentives; individuals often have to give up their economic activities that generate income for their families, and the material rewards may only partially compensate for the foregone income. Besides internal validity concerns, one should also be cautious of the external validity (the scope and generalizability) of these studies even in the context of civil wars. While material rewards may have been the key motivation in Siera Leon civil war (Humphreys and Weinstein 2008), as Wood's (2003) study indicates, psychological rewards were the key motivation in El Salvadoran civil war. As Blattman and Miguel (2010) argue in their review of the civil war literature, even in such risky environments, "non-material incentives are thought to be common [even] within armed groups. Several studies argue that a leader's charisma, group ideology, or a citizen's satisfaction in pursuit of justice (or vengeance) can also help solve the problem of collective action in rebellion" (p. 15). More importantly, civil wars with rebel armies are a particular form of conflict that are more prone to mercenary-like recruitment and more sensitive to the availability of "lootable resources"; arms and logistic equipments are expensive, and more so in black markets where rebels often have to operate. Less militarized conflicts are less likely to be susceptible to pay-for-rebel incentives.

In the broader context of protests and revolutions, researchers have persistently pointed to the critical role of non-material incentives among the participants. Survey data suggests that psychological incentives were the primary motivation for participants in Gezi Park protests in Turkey and Euromaidan in Ukraine (Aytaç and Stokes 2019), protests in Syria (Pearlman 2018) and Morocco in the Arab Spring, and in the resistance movements of the Eastern Europe under Nazi and Soviet domination (Petersen 2001). For example, Pearlman (2018) highlights the role of "moral identity" (e.g., self-respect and "joy of agency") in Syrian protests. In sum, even if material motivations are the primary form of motivation in some civil wars, that logic does not extends to less violent anti-government movements. Hundreds of thousands of Iranians who participated in the Green Movement that followed the contested 2009 presidential election could not have reasonably expected to receive any significant selective material rewards for their participation.

What exactly is the nature of these non-material incentives? In this paper, we focus on the notion of pleasure in agency. Wood (2003) develops the notion of pleasure in agency to capture individuals' motives for participating in contentious collective actions. Pleasure in agency is "the positive effect associated with self-determination, autonomy, self-esteem, efficacy, and pride that come from the successful assertion of intention" (p. 235). It is "a frequency-based motivation: it depends on the likelihood of success, which in turn increases with the number participating (Schelling 1978; Hardin 1982). Yet the pleasure in agency is undiminished by the fact that one's own contribution to the likelihood of victory is vanishingly small" (p. 235-6). For example, in her interviews with insurgents during and after the El Salvadoran civil war, she found that insurgents "repeatedly asserted their pride in their wartime activities and consistently claimed *authorship* of the changes that they identified as their work" (p. 231). "They took pride, indeed pleasure, in the successful assertion of their interests and identity...motivated in part by the value they put on being part of the making of history" (p. 18-9). Pleasure in agency rewards contrast with the Gurrian psychological theories of revolution, delineated in *Why Men Rebel?* (Gurr 1970), which claim that citizens receive expressive payoffs from participation in revolutionary movements. Rather, pleasure in agency falls into the dominant, Tillyan theories of social movements, delineated in *From Mobilization to Revolution* (Tilly 1978), in which individuals take into account the costs, benefits, and likelihood of success when deciding whether to protest. In the same vein, there is no pleasure in agency without success, and hence Wood's agents estimate the likelihood of success before deciding whether to participate. In essence, pleasure in agency rewards are *selective* psychological rewards from participating in a successful movement. But if pleasure in agency rewards are a key motivating factor for participation, where do these pleasure in agency rewards come from?

5.3 Leadership

Although many scholars agree that "an account of participation in insurgency requires a consideration of the moral and emotional dimensions of participation" (Wood 2003, p. 225), they typically consider these intrinsic motivations as the automatic result of past experiences or the interactions with the regime in periods of contention. However, an extensive leadership literature that cuts across the disciplines of management, sociology and political science argue that some people have abilities, skills, or the charisma to inspire costly actions (Burns 1978, 2003; Bass 1985; Snow et al. 1986; Goldstone 2001). These leaders have been called transformational, transformative, or people-oriented leaders. Combining the insights of these literatures, we argue that some leaders can create, activate, or manipulate pleasure in agency rewards by tapping into values, feelings and emotions, and "cultural idioms" (Skocpol 1997) of a population. One can summarize the question that we ask in this paper as: what happens when these *transformative leaders* manipulate *pleasure in agency* rewards to influence the citizens' choice of *contentious performances* to maximize the likelihood of *regime change*?

In his review of the literature on revolutions, Goldstone (2001) identifies two distinct types of revolutionary leaders: people-oriented and task-oriented. "People-oriented leaders are those who inspire people, give them a sense of identity and power, and provide a vision of a new and just order" (p. 157). These are leaders who can create, amplify, or transform people's feelings and identities, e.g., by innovative framing of events and experiences or by their personality traits, such as their charisma (Snow et al. 1986). In the terminology of organizational behavior, they are "transformational leaders" who have the ability to create inspirational motivations through a variety of psychological mechanisms (Burns 1978, 2003; Bass 1985). In contrast to people-oriented leaders, in Goldstone (2001), "task-oriented leaders are those who can plot a strategy suitable to resources and circumstances" (p. 157), including the repressive capacity of the state and available revolutionary skills, to effectively transform anti-regime sentiments into concrete actions. Combining these separate branches of the literature, we show how a"people-oriented" leader can optimally create and manipulate "pleasure in agency" feelings among potential revolutionaries to maximize the likelihood of regime change. That even the design of optimal "pleasure in agency schemes" requires "a strategy suitable to resources and circumstances" merges the seemingly separate categories of people-oriented and task-oriented leadership.

5.4 The American and the Islamic Revolutions

To demonstrate our theoretical framework in concrete settings, we provide examples from the American Revolution as well as the development of Islam in its early years. We show that leaders appeal to people's values (e.g., honor, justice, faith, bravery, and morality) and insist that their endeavor has good chances of success. This dual emphasis is consistent with the nature of pleasure in agency rewards, which are obtained only if the movement succeeds. In Washington's words to the Congress, "The honor of making a brave defense does not seem to be a sufficient stimulus, when the success is very doubtful and the falling into the Enemy's hands probable" (Middlekauff 2005, p. 342). Indeed, Washington's victory in the battle of Trenton after a series of defeats improved recruiting for the revolution by improving the people's beliefs that the revolution could succeed; as the loyalist Nicholas Cresswell of Virginia recorded in his diary in 1777, after the battle of Trenton: "The minds of the people are much altered. A few days ago they had given up the cause for lost. Their late successes have turned the scale and now they are all liberty mad again" (Rhodehamel

2001, p. 264).

During the American Revolution, both religious and secular leaders called upon the colonists to rise against the British. Paine's first essay of *The American Crisis* (Paine 1995), in December of 1776, was so effective that Washington had it read for the troops; "These are the times that try men's souls. The summer soldier and the sunshine patriot will, in this crisis, shrink from the service of their country; but he that stands by it now, deserves the love and thanks of man and woman." He continued to establish the likelihood of success: "I have as little superstition in me as any man living, but my secret opinion has ever been, and still is, that God Almighty will not give up a people to military destruction, or leave them unsupportedly to perish, who have so earnestly and so repeatedly sought to avoid the calamities of war.... Neither have I so much of the infidel in me, as to suppose that He has relinquished the government of the world, and given us up to the care of devils." Returning to inspiration, he added: "The heart that feels not now is dead; the blood of his children will curse his cowardice, who shrinks back at a time when a little might have saved the whole, and made them happy.... 'Tis the business of little minds to shrink; but he whose heart is firm, and whose conscience approves his conduct, will pursue his principles unto death."

Washington, too, attempted to inspire people to contribute to the cause. In November of 1777, he wrote to the militia: "I therefore call upon you, by all that you hold dear to rise up as one man and rid your country of its unjust invaders." Following some examples of successes, he ends by asking them to "Reflect upon these things, and I am convinced that every man who can bear a musket will take it up and without respect to time or place give his services in the field for a few weeks, perhaps only a few days" (Washington 1890, vol. VI, p. 213-4). A few months later, Washington wrote to the inhabitants of the Middle Colonies, Maryland, and Virginia, asking "to put up and feed immediately as many of their stock cattle as they can spare.... the proprietors may assure themselves, that they will render a most essential service to the illustrious cause of their country, and contribute in a great degree to shorten this bloody contest. But should there be any so insensible to the common interest, as not to exert themselves upon these generous principles, the private interest of those, whose situation makes them liable to become immediate subjects to the enemy's incursions, should prompt them at least to a measure, which is calculated to save their property from plunder, their families from insult, and their own persons from abuse, hopeless confinement, or perhaps a violent death" (Washington 1890, vol. VI, p. 382-3). By invoking the people's sense of honor and patriotism, Washington was assigning different rewards to different levels of participation: those who would join the army for a few weeks were more honorable and better patriots than those who would join for a few days, who in turn may be better than those who would contribute only their cattle. Those who would refuse even that could still contribute simply by hiding their stock from the enemy.

When inspirational attempts succeed, their impact can be dramatic. For example, on December 31, 1776, enlistments were about to expire, and the army was about to disintegrate. The seriousness of the matter is reflected in Washington's confidential letter to his brother earlier that month: "In a word, my dear Sir, if every nerve is not restrained to recruit the new army with all possible expedition, I think the game is pretty near up.... You can form no idea of the perplexity of my situation. No man, I believe, ever had a greater choice of difficulties, and less means to extricate himself from them" (Washington 1890, vol. V, p. 112).[11] The men were about to leave when Washington appeared in person to one brigade: "My brave fellows, you have done all I asked you to do, and more than you could be reasonably expected; but your country is at stake.... If you will consent to stay only one month longer, you will render that service to the cause of liberty and to your country which you probably never can do under any other circumstances" (McCullough 2005, p. 285). It worked, and the men stayed. As Middlekauff (2005) argues, "Courage, honor, gallantry in service of liberty, all those words calculated to bring a blush of embarrassment to jaded twentieth-century men, defined manhood for the eighteenth century" (p. 515).

Paine's writings perhaps represent the best of the secular inspirational writings of the Revolutionary era. Abraham Keteltas's 1777 sermon, "God pleads his cause" (Sandoz 1998, p. 579-605), is among the best religious inspirational speeches of the Revolution. Drawing heavily from the old and the new testaments, he argued "the cause of this American continent...is the cause of God.... It is the cause, for which heroes have fought, patriots bled, prophets, apostles, martyrs, confessors, and righteous men have died: Nay, it is a cause, for which the Son of God came down from his celestial throne, and expired on a cross—it is a cause, for the sake of which, your pious ancestors forsook all the delights and enjoyments of England, that land of wealth and plenty, and came to this once howling wilderness, destitute of houses, cultivated fields, the comforts and conveniencies of life." He ended the sermon by saying that the revolution will succeed, against all odds, because "God will effectually plead it [his cause]; he will plead it by his almighty word, his all conquering spirit, and his over ruling providence.... America will be a glorious land of freedom, knowledge, and religion, an asylum for distressed, oppressed, and persecuted virtue," all of which he interprets in detail based on scriptures. Based on these foundations, Keteltas called for contributions: "Let this exhilerating thought, fire your souls, and give new ardor and encouragement to your hopes—you contend not only for your own happiness, for your dear relations; for the happiness of the present inhabitants of America; but you contend for the happiness of millions yet unborn. Exert therefore, your utmost efforts, strain every nerve, do all you can to promote this cause.... Be ready to fight for it, and maintain it to the last drop of your blood. *Herein was the love of God manifested, that he laid down his life for us, and we ought to lay down our lives for the brethren* [emphasis in the original]."

Perhaps a prime example of religious inspiration for revolutionary movements can be gleaned from the development of Islam in its early years, when, based on numbers and resources, the prospects of the movement were bleak even inside the Arabian peninsula. Muhammad needed believers to contribute to the cause with their money as well as their bodies and weapons. Many Quranic verses are to get people to make such contributions. Quran repeatedly insists that success is near, and that God will help in many hidden ways, including sending warriors from the skies. Thus, like Washington, Paine, and Keteltas,

[11] In McCullough's (2005) words: "From the last week of August to the last week of December, the year 1776 had been as dark a time as those devoted to the American cause had ever known—indeed, as dark a time as any in the history of the country" (p. 291).

Muhammad had to emphasize that the movement will succeed.

We showed that the problem is inspiring revolutionary action can be perceived as an exchange between a seller, trading with a buyer with unknown valuation. Indeed, Quran uses the language of trade to describe the the nature of rewards designed to motivate contributions: "Allah hath purchased of the believers their persons and their goods; for theirs (in return) is the garden (of Paradise)" (Repentance 9:111). Moreover, rewards increase even with marginal contributions: "Nor could they spend anything (for the cause)—small or great—nor cut across a valley [to go to war], but the deed is inscribed to their credit: that Allah may requite their deed with the best (possible reward)" (Repentance 9:121). Contributions that had higher expected costs warranted higher rewards: "Not equal among you are those who spent (freely) and fought, before the Victory, (with those who did so later). Those are higher in rank than those who spent (freely) and fought afterwards" (Iron 57:27). Giving drinks to pilgrims, maintaining the Sacred Mosque in Mecca, staying behind during the war to attend to home affairs, and fighting battles have all been assigned different degrees of rewards:

- Not equal are those believers who sit (at home) and receive no hurt, and those who strive and fight in the cause of Allah with their goods and their persons. Allah hath granted a grade higher to those who strive and fight with their goods and persons than to those who sit (at home). Unto all (in Faith) Hath Allah promised good: But those who strive and fight Hath He distinguished above those who sit (at home) by a special reward. (Women 4:95)

- Do ye make the giving of drink to pilgrims, or the maintenance of the Sacred Mosque, equal to (the pious service of) those who believe in Allah and the Last Day, and strive with might and main in the cause of Allah? They are not comparable in the sight of Allah. (Repentance 9: 19-20)

In sum, the leaders' dual emphasis on the citizens' values and rewards and the chances of success, both in the American and in the Islamic Revolutions, fits the nature of pleasure in agency rewards. It is also worth emphasizing that the mechanism for obtaining these rewards is in turn consistent with the logic of screening. If leaders knew the citizens' beliefs about the likelihood of success, they would condition rewards on those beliefs. Not knowing the beliefs, however, the leaders assign different rewards to different participation levels. But this mechanism maps the leaders' optimal motivation schemes into screening problems, with the accompanying tradeoffs. For example, raising the value of joining the cause as militiamen "only a few days" increases the incentives of those who would have otherwise limited their contribution to giving up a chicken to feed the army; but it also reduces the incentive of those who would have otherwise stayed "for a few weeks."

Of course, leaders may have a variety of other constraints, unmodeled here, which prevent them from designing optimal pleasure in agency schemes. But when they do, our Proposition 5 implies that, regardless of the the leader's ability M or the exact shape of the costs (as long as they remain increasing and convex), the leader's optimal design of (pleasure in agency) rewards leads to the emergence of a group of citizens who engage in the maximum level of anti-regime activities. In the American Revolution, this group corresponds to the Sons of Liberty (Maier 1972), and during the Revolutionary War, perhaps a subset of regulars who remained with the Continental Army during the toughest times. Middlekauff's (2005) description of them, based on a large body of primary and secondary sources, is illuminating: "By winter 1779-1780 the Continentals were beginning to believe that they had no one save themselves to lean on.... Dissatisfaction in these months slowly turned into a feeling of martyrdom. They felt themselves to be the martyrs to the 'glorious cause.' They would fulfill the ideals of the Revolution and see things through to independence because the civilian population would not.... Their country might ignore them in camp, might allow their bellies to shrivel and their backs to freeze, might allow them to wear rags, but in battle they would not be ignored. And in battles they would support one another in the knowledge that their own moral and professional resources remained secure" (p. 513-4).

More broadly, the emergence of this cadre is consistent with the notion of professional revolutionaries in Lenin's treatise, *What Is to be Done?*. Although this group is distinct from other citizens contributing to the cause, there is little difference in the magnitude of duties and the level of revolutionary activity among the members of the group. Contrasting the organization of workers, who engage in various degrees of contentious activity, with the organization of "revolutionary social-democratic party," Lenin insisted that "the organization of the revolutionaries must consist first and foremost of people who make revolutionary activity their profession.... In view of this common characteristic of the members of such an organization, all distinctions as between workers and intellectuals, not to speak of distinctions of trade and profession, in both categories, must be effaced" (p. 71). Of course, the members of this group are not identical; some have stronger beliefs in the likelihood of regime change than others, for example, as a result of their different interpretations of Marxist theory. However, despite these differences in their beliefs, they constitute a distinct group that exert maximum effort in service of the revolution—in this sense, they are professional revolutionaries.

Appendix: Examples and Proofs

Example

Suppose $C(e) = e^n$, $n > 1$, and $M = 1$. What is the optimal $B(e)$? From (13),

$$B(e) = \frac{1}{\sqrt{\lambda}} \int_0^e \sqrt{C'(x)}\; dx = \frac{1}{\sqrt{\lambda}} \int_0^e \sqrt{n}\; x^{\frac{n-1}{2}}\; dx = \frac{1}{\sqrt{\lambda}}\, \frac{2\sqrt{n}}{n+1}\; e^{\frac{n+1}{2}}. \tag{19}$$

From (14), $B(\bar{e}) = 1$, which implies

$$\frac{1}{\sqrt{\lambda}}\, \frac{2\sqrt{n}}{n+1}\; \bar{e}^{\frac{n+1}{2}} = 1. \tag{20}$$

From (16), $\frac{C'(\bar{e})}{B'(\bar{e})} = \frac{1}{2}$, which implies

$$\frac{n\bar{e}^{n-1}}{\sqrt{\frac{n}{\lambda}}\; \bar{e}^{\frac{n-1}{2}}} = \sqrt{n\lambda}\; \bar{e}^{\frac{n-1}{2}} = \frac{1}{2}. \tag{21}$$

From (20), $\sqrt{\lambda} = \frac{2\sqrt{n}}{n+1}\; \bar{e}^{\frac{n+1}{2}}$. Substituting this $\sqrt{\lambda}$ into equation (21) yields $\frac{2n}{n+1}\; \bar{e}^n = \frac{1}{2}$, and hence

$$\bar{e} = \left(\frac{n+1}{4n}\right)^{\frac{1}{n}} \quad \text{and} \quad \sqrt{\lambda} = \frac{2\sqrt{n}}{n+1}\; \left(\frac{n+1}{4n}\right)^{\frac{n+1}{2n}}.$$

Substituting these into equation (19) yields the optimal $B(e) = \left(\frac{4n}{n+1}\right)^{\frac{n+1}{2n}}\; e^{\frac{n+1}{2}}$ for $e \leq \bar{e}$. Moreover, $B(e) \leq M = 1$ for $e > \bar{e}$. For the purposes of this example, we choose $B(e) = M = 1$ for $e > \bar{e}$. Thus,

$$B^*(e) = \begin{cases} \left(\frac{4n}{n+1}\right)^{\frac{n+1}{2n}}\; e^{\frac{n+1}{2}} & ; e \leq \bar{e} = \left(\frac{n+1}{4n}\right)^{\frac{1}{n}} \\ \\ M = 1 & ; e > \bar{e}. \end{cases}$$

As expected from our earlier discussion, $B^*(e)$ is *convex* for $e \in [0, \bar{e}]$. Moreover,

$$B^{*'}(e) = \begin{cases} \frac{n+1}{2}\; \left(\frac{4n}{n+1}\right)^{\frac{n+1}{2n}}\; e^{\frac{n-1}{2}} & ; e < \bar{e} = \left(\frac{n+1}{4n}\right)^{\frac{1}{n}} \\ \\ 0 & ; e > \bar{e}. \end{cases}$$

How does $e^*(p)$ look like? Recall that $e^*(p) = \arg\max_{e \geq 0}\; p\; B^*(e) - e^n$, and hence

$$e^*(p) = \begin{cases} 4^{\frac{1}{n(n-1)}}\; \left(\frac{n+1}{n}\right)^{\frac{1}{n}}\; p^{\frac{2}{n-1}} & ; p \in [0, \frac{1}{2}) \\ \\ \bar{e} & ; p \in [\frac{1}{2}, 1]. \end{cases}$$

Figure 2 illustrated this solution.

Proofs

Proof of Lemma 3: By assumption, $e^*(p)$ is a weakly increasing function. Whenever $e^*(p)$ is strictly increasing, we have $pB'(e^*(p)) = C'(e^*(p))$. Let $[p_1, p_2]$ be an interval on which e^* is strictly increasing and continuous. In this case,

$$\begin{aligned}\int_{p=p_1}^{p_2} e^*(p)\,dp &= [p\,e^*(p)]_{p=p_1}^{p_2} - \int_{p=p_1}^{p_2} p\,[e^*]'(p)\,dp \\ &= p_2\,e^*(p_2) - p_1\,e^*(p_1) - \int_{e=e^*(p_1)}^{e^*(p_2)} \frac{C'(e)}{B'(e)}de,\end{aligned}$$

where the last inequality uses the change of variables $e = e^*(p)$, and the derivative of $B'(e)$ at $e^*(p_2)$ is the left derivative and at $e^*(p_1)$ is the right derivative.

Now let $[p_2, p_3]$ be an interval on which e^* is constant. In this case,

$$\begin{aligned}\int_{p=p_2}^{p_3} e^*(p)\,dp &= (p_3 - p_2)\,e^*(p_3) \\ &= p_3 e^*(p_3) - p_2 e^*(p_2).\end{aligned}$$

We conclude that

$$\int_{p=p_1}^{p_3} e^*(p)\,dp = p_3 e^*(p_3) - p_1 e^*(p_1) - \int_{e=e^*(p_1)}^{e^*(p_2)} \frac{C'(e)}{B'(e)}de.$$

Now, consider a partition of $[0,1]$, and suppose that $e^*(p)$ is strictly increasing on the intervals $[p_1, p_2]$,$[p_3, p_4]$,...,$[p_{2n-1}, p_{2n}]$, where $p_1 < p_2 < ... < p_{2n}$. Then,

$$\begin{aligned}\int_{p=0}^{1} e^*(p)\,dp &= e_{max} - \sum_{m=1}^{n}\left(\int_{e=e^*(p_{2m-1})}^{e^*(p_{2m})} \frac{C'(e)}{B'(e)}de\right) \\ &= e_{max} - \int_{e=0}^{e_{max}} \frac{C'(e)}{B'(e)}de \\ &= \int_{e=0}^{e_{max}} \left(1 - \frac{C'(e)}{B'(e)}\right)de.\end{aligned}$$

$\square$

Proof of Proposition 2: Let $p_2 > p_1 > 0$, $e_i \in e^*(p_i) = \arg\max_{e \geq 0} p_i B(e) - C(e)$, $i \in \{1, 2\}$. We establish that $e_2 \geq e_1$ by way of contradiction. Suppose not, so that $e_2 < e_1$. First, from the optimality of e_1 and e_2, we have:

$$p_2\ B(e_2) - C(e_2) \geq p_2\ B(e_1) - C(e_1) \quad \Leftrightarrow \quad p_2\ [B(e_2) - B(e_1)] \geq C(e_2) - C(e_1) \tag{22}$$

$$p_1\ B(e_1) - C(e_1) \geq p_1\ B(e_2) - C(e_2) \quad \Leftrightarrow \quad C(e_2) - C(e_1) \geq p_1\ [B(e_2) - B(e_1)] \tag{23}$$

Because $C(e)$ is strictly increasing, $e_2 < e_1$ implies $C(e_2) < C(e_1)$. Thus, from (23):

$$0 > C(e_2) - C(e_1) \geq p_1\ [B(e_2) - B(e_1)] \quad \Rightarrow \quad B(e_2) - B(e_1) < 0,$$

and hence:

$$p_1\ [B(e_2) - B(e_1)] > p_2\ [B(e_2) - B(e_1)]. \tag{24}$$

However, combining (22) and (23), we have:

$$p_1\ [B(e_2) - B(e_1)] \leq C(e_2) - C(e_1) \leq p_2\ [B(e_2) - B(e_1)]. \tag{25}$$

Hence, (24) and (25) contradict each other, and hence our assumption that $e_2 < e_1$ must be false. The proposition is then implied by Lemma 2. □

Proof of Proposition 3: By standard arguments, e^* is continuous, weakly increasing, and single-valued. Minimum effort $e_{min} = 0$, since 0 is the unique maximizer of $-C(e)$. The proposition is then implied by Lemma 3. □

Proof of Proposition 5: As a first step in analyzing this problem, we apply the revelation principle to transform the problem into one where the leader chooses effort levels and benefit levels $\{(e(p), B(p))\}$ depending on the threshold probability p that a citizen assigns to regime change. This mechanism design approach amounts to treating the citizens' endogenous beliefs *as if* they were exogenous and distributed uniformly on $[0, 1]$. However, determining $\{(e(p), B(p))\}$ generates a set of "recommended" revolutionary efforts $E \equiv \{e(p)\ s.t.\ p \in [0, 1]\}$, and the corresponding rewards $B(e)$ for those $e \in E$. It remains to characterize $B(e)$ for $e \notin E$. The only requirement for such $B(e)$ is that players do not choose it. For example, one could set $B(e) = 0$ for all $e \notin E$. Of course, this choice is not unique. We choose $B(e)$ for $e \notin E$ such that $B(e)$ is constant for $e \notin E$, and $B(e)$ is continuous for all $e \geq 0$. Therefore, we can write the leader's problem as:

$$\begin{aligned} &\max_{\{(e(p),B(p))\}} \int_{p=0}^{1} e(p)dp \\ s.t. \quad & pB(p) - C(e(p)) \geq 0, && \forall\ p \in [0, 1] \\ & pB(p) - C(e(p)) \geq p\ B(p') - C(e(p')), && \forall\ p, p' \in [0, 1] \\ & B(p) \in [0, M], && \forall\ p \in [0, 1]. \end{aligned}$$

Observe that under this program, the designer can assign any cost to any citizen through their choice of e. To simplify notation, write $h(p) = C(e(p))$.

We can use this new representation of the problem to use standard arguments from screening models. We first establish that $B(p)$ is weakly increasing, and hence $B(p)$ is piecewise continuously differentiable. The incentive compatibility constraints imply $pB(p) - h(p) \geq p\ B(p') - h(p')$ and $p'B(p') - h(p') \geq p'\ B(p) - h(p)$. Adding these inequalities implies: $(p - p')[B(p) - B(p')] \geq 0$. Hence, $B(p)$ is weakly increasing, and hence $B(p)$ is piecewise continuously differentiable. Thus, a necessary first order condition is $pB'(p) - h'(p) = 0$ almost everywhere, with the corresponding second order condition $pB''(p) - h''(p) \leq 0$. Differentiating the FOC w.r.t. p yields $B'(p) + pB''(p) - h''(p) = 0$. Thus, the SOC simplifies to $B'(p) \geq 0$. Moreover, because $p\ B(p) - h(p)$ is increasing in p, the condition $pB(p) - C(e(p)) \geq 0,\ \forall\ p \in [0, 1]$, simplifies to $C(e(0)) = h(0) = 0$. Further, because $B'(p) \geq 0$, the constraint $B(p) \in [0, M],\ \forall\ p \in [0, 1]$, can be replaced by $B(0) \geq 0$ and $B(1) \leq M$. Combining these results, the leader's problem becomes:

$$\max_{\{(e(p), B(p))\}} \int_{p=0}^{1} e(p)\ dp \tag{26}$$
$$pB'(p) - h'(p) = 0, \quad h(0) = 0 \tag{27}$$
$$B'(p) \geq 0 \tag{28}$$
$$B(0) \geq 0, \quad B(1) \leq M. \tag{29}$$

We can re-write this problem, letting $\Pi(\cdot) = C^{-1}(\cdot)$, so that $\Pi(h(p)) = e(p)$. Then, the leader's problem (26)-(29) becomes:

$$\max_{\{(B(p), h(p))\}} \int_{p=0}^{1} \Pi(h(p))\ dp \tag{30}$$
$$h'(p) = pB'(p), \quad h(0) = 0 \tag{31}$$
$$B'(p) \geq 0,\ B(0) \geq 0,\ \text{and } B(1) \leq M. \tag{32}$$

Because incentives are created by the slope of benefits, it is clear that $B(0) = 0$ and $B(1) = M$. To proceed, we use the optimal control techniques by defining two state variables. Taking a similar approach to Kamien and Schwartz (2012, p. 244-6), let (h, B) be the state and B' be the control, so that $h' = pB'$. The Hamiltonian and Lagrangian are:

$$\begin{aligned} H &= \Pi(h) + \lambda_h\ pB' + \lambda_B\ B'. \\ L &= H + \mu B'. \end{aligned}$$

Then, by the maximum principle,

$$\frac{\partial L}{\partial B'} = \lambda_B + \lambda_h\ p + \mu = 0,\ \mu \geq 0,\ \mu B' = 0. \tag{33}$$
$$\lambda_B'(p) = -\frac{\partial L}{\partial B} = 0. \tag{34}$$
$$\lambda_h'(p) = -\frac{\partial L}{\partial h} = -\Pi'(h(p)), \quad \lambda_h(1) = 0. \tag{35}$$
$$B' \quad \text{must maximize } H \text{ given optimal } (B, h) \text{ and } (\lambda_B, \lambda_h)\ . \tag{36}$$

Moreover, we recognize that $B(1) = M$ is captured by having a free $\lambda_B(1)$ as opposed to $\lambda_h(1) = 0$ that stems from a free $h(1)$. Now, because $C(e)$ is strictly convex, its inverse $\Pi(\cdot)$ is strictly concave. This allows us to apply Arrow's (1966) approach together with a theorem from Seierstad and Sydsæter (1993) to show that $h(p)$ and $B(p)$ have no jumps.

Lemma 4. *Optimal $h(p)$ and $B(p)$ have no jump.*

Proof of Lemma 4: We show that an optimal $h(p)$ has no jump, which then implies that an optimal $B(p)$ has no jump. First, we prove that there is no interior jumps at any $p \in (0,1)$. Our proof is based on Arrow (1966, p. 11-3). Suppose $h(p)$ has a jump at $\tau \in (0,1)$. Let $h(\tau^+) \equiv \lim_{p\to\tau^+} h(\tau)$ and $h(\tau^-) \equiv \lim_{p\to\tau^-} h(\tau)$. Because $h(p)$ is increasing, $h(\tau^-) < h(\tau^+)$. From (34) and (35), recall that $\lambda'_B(p)+p\lambda'_h(p) = -p\Pi'(h(p))$ in $(\tau-\epsilon,\tau)\cup(\tau,\tau+\epsilon)$ for some $\epsilon > 0$. Because $\Pi(x)$ is strictly concave, $\Pi'(h(\tau^+)) < \Pi'(h(\tau^-))$, and hence:

$$\lambda'_B(\tau^-) + \tau\ \lambda'_h(\tau^-) < \lambda'_B(\tau^+) + \tau\ \lambda'_h(\tau^+). \tag{37}$$

Moreover, from (33), $\lambda_B(p)+p\lambda_h(p) \le 0$ and $\lambda_B(\tau)+\tau\lambda_h(\tau) = 0$.[12] Further, from condition (74) of Theorem 7 in Seierstad and Sydsæter (1993, p. 197), $\lambda_h(p)$ and $\lambda_B(p)$ are continuous at τ. Thus, in a right neighborhood of τ, we have $\lambda'_B(p)+p\lambda'_h(p)+\lambda_h(p) \le 0$, and in a left neighborhood of τ, we have $\lambda'_B(p) + p\lambda'_h(p) + \lambda_h(p) \ge 0$. But this contradicts (37).

Next, we prove that there is no jump at $p = 0$ or $p = 1$. From (34), $\lambda_B(p) = \textit{constant} \equiv \bar{\lambda}$ on $p \in (0,1)$. Moreover, $\lambda_B(p)$ is continuous (Seierstad and Sydsæter 1993, p. 197). Thus, $\lambda_B(p) = \bar{\lambda}$, for $p \in [0,1]$. From conditions (74) and (75) in Theorem 7 of Seierstad and Sydsæter (1993, p. 197), $\lambda_B(\tau) + \tau\lambda_h(\tau) = 0$. Combining this with $\lambda_B(p) = \bar{\lambda}$ and (35), $\bar{\lambda} + \tau\int_{x=\tau}^{1} \Pi'(h(x))dx = 0$. In particular, if $\tau = 0$ or $\tau = 1$, then $\bar{\lambda} = 0$. But from (33), $\bar{\lambda} + \lambda_h\ p \le 0$ and $\lambda_h(p) = \int_{x=p}^{1} \Pi'(h(x))dx$ is positive for some $p \in (0,1)$. Thus, $\bar{\lambda} < 0$. A contradiction. □

From (34) and (35),

$$\lambda_B(p) = \textit{constant} = \bar{\lambda} \quad \text{and} \quad \lambda_h(p) = -\int_{x=p}^{1} \lambda'_h(x)dx = \int_{x=p}^{1} \Pi'(h(x))dx. \tag{38}$$

Moreover, from (33), $\bar{\lambda} + p\ \lambda_h(p) \le 0$. Because at $p \in (0,1)$, $p\ \lambda_h(p) = p\int_{x=p}^{1} \Pi'(h(x))dx > 0$, we must have $\bar{\lambda} < 0$.

Lemma 5. *There is no interior bunching. That is, there is no $0 < p_1 < p_2 < 1$ such that $h'(p) = 0$ for $p \in (p_1, p_2)$, with $h'(p) > 0$ in a left neighborhood of p_1 and in a right neighborhood of p_2 .*

[12] More formally, $\lambda_B(\tau) + \tau\lambda_h(\tau) = 0$ obtains from conditions (74) and (75) in Theorem 7 of Seierstad and Sydsæter (1993, p. 197).

Proof of Lemma 5: Suppose not. Let $h(p) = \bar{h}$ for $p \in [p_1, p_2]$. Because $h(p)$ and hence $B(p)$ must be strictly increasing in a left neighborhood of p_1 and in a right neighborhood of p_2, $\mu = 0$ in those neighborhoods. Moreover, $\lambda_B(p)$ and $\lambda_h(p)$ are continuous. Therefore, (33) implies that $\lambda_B(p_1)+p_1\lambda_h(p_1) = 0 = \lambda_B(p_2)+p_2\lambda_h(p_2)$. Because $\lambda_B(p) = \bar{\lambda}$ from (38), we have,

$$p_1\lambda_h(p_1) = p_2\lambda_h(p_2). \tag{39}$$

Further, from (38),

$$\lambda_h(p_1) = \int_{p_1}^{p_2} \Pi'(h(x))dx + \int_{p_2}^{1} \Pi'(h(x))dx = (p_2 - p_1)\ \Pi'(\bar{h}) + \lambda_h(p_2). \tag{40}$$

Combining equations (39) and (40) yields:

$$p_1\ \Pi'(\bar{h}) = \lambda_h(p_2). \tag{41}$$

Next, consider $p_3 \in (p_1, p_2)$. From (33), $\lambda_B(p_3) + p_3\lambda_h(p_3) \leq 0 = \lambda_B(p_2) + p_2\lambda_h(p_2)$. Hence,

$$p_3\lambda_h(p_3) \leq p_2\lambda_h(p_2). \tag{42}$$

Mirroring the calculations of equation (40), we have:

$$\lambda_h(p_3) = (p_2 - p_3)\ \Pi'(\bar{h}) + \lambda_h(p_2). \tag{43}$$

Combining equations (42) and (43) yields:

$$p_3\ \Pi'(\bar{h}) \leq \lambda_h(p_2). \tag{44}$$

From equations (41) and (44), $p_3\ \Pi'(\bar{h}) \leq p_1\ \Pi'(\bar{h})$. However, because $p_3 > p_1$ and $\Pi'(\bar{h}) > 0$, we must have $p_3\ \Pi'(\bar{h}) > p_1\ \Pi'(\bar{h})$, a contradiction. □

Next, we solve the problem ignoring the constraint $B'(p) \geq 0$, and subsequently check whether and when this constraint binds. Without $B'(p) \geq 0$, $\mu = 0$, and hence from (33) and (38),

$$p\int_{x=p}^{1} \Pi'(h(x))dx + \bar{\lambda} = 0, \text{ and hence } \frac{d}{dp}\left\{p\int_{x=p}^{1} \Pi'(h(x))dx + \bar{\lambda}\right\} = 0 \tag{45}$$

Differentiating (45) with respect to p yields $\Pi'(h(p)) = -\frac{\bar{\lambda}}{p^2}$, and hence

$$h(p) = \Pi'^{-1}\left(-\frac{\bar{\lambda}}{p^2}\right). \tag{46}$$

Moreover, Because $\Pi(h)$ is strictly concave, $h(p)$ is strictly increasing, and hence $B(p)$ is strictly increasing as far as the constraints $B(0) = 0$ and $B(1) = M$ are satisfied. Therefore, an optimal $B(p)$ takes the following form:

$$B(p) = \begin{cases} 0 & ; p \in [0, p_1] \\ \text{strictly increasing function} & ; p \in [p_1, p_2] \\ M & ; p \in [p_2, 1], \end{cases}$$

for $0 \leq p_1 < p_2 \leq 1$. This, in turn, implies a similar form for $h(p)$, and hence for optimal effort schedule:

$$e^*(p) = \begin{cases} 0 & ; p \in [0, \underline{p}] \\ \text{strictly increasing and satisfies the first order condition} & ; p \in [\underline{p}, \overline{p}] \\ \bar{e} & ; p \in [\overline{p}, 1], \end{cases} \tag{47}$$

for some $0 \leq \underline{p} < \overline{p} \leq 1$ and $\bar{e} > 0$.

Equation (47) together with Lemma 3 allows us to write the leader's objective function as:

$$\theta^* \;\; = \;\; \int_{e=0}^{\bar{e}} \left(1 - \frac{C'(e)}{B'(e)}\right) \, de,$$

where we recognize that $e^*(p)$ satisfying the first order condition implies $B'(e) \geq C'(e)$, and that $\underline{p} = \frac{C'(0)}{B'(0)}$ and $\overline{p} = \frac{C'(\bar{e})}{B'(\bar{e})}$. Thus, we can formulate the leader's problem as:

$$\begin{aligned} &\max_{B',\, \bar{e}} \;\; \int_0^{\bar{e}} \left(1 - \frac{C'(e)}{B'(e)}\right) \, de \\ &s.t. \; B(\bar{e}) = M, \; B(0) = 0, \; B'(e) \geq C'(e), \; \bar{e} \geq 0, \end{aligned} \tag{48}$$

Writing $B(\bar{e})$ as $B(\bar{e}) = \int_{e=0}^{\bar{e}} B'(e) de$, (48) can be written as:

$$\begin{aligned} &\max_{B',\, \bar{e}} \;\; \int_0^{\bar{e}} \left(1 - \frac{C'(e)}{B'(e)}\right) \, de \\ &s.t. \; \int_{e=0}^{\bar{e}} B'(e) de = M, \; B(0) = 0, \; B'(e) \geq C'(e), \; \bar{e} \geq 0. \end{aligned}$$

This is the maximization problem that we analyze in Section 3.3 of the text. The rest of the proof is in the text. □

Proof of Proposition 6: First, we prove a lemma.

Lemma 6. *Suppose X and Y are two independent random variables, where $X \sim U[0, 1]$ and $Y \sim U[\mu - \delta, \mu + \delta]$, with $0 < \delta \leq \mu$. Then, $X \cdot Y \sim G(\cdot)$, where*

$$G(t) = \begin{cases} \max\{0, t \; log(\frac{\mu+\delta}{\mu-\delta})/2\delta\} & ; t \leq \mu - \delta \\ \left(t \; log\left(\frac{\mu+\delta}{t}\right) + t - (\mu - \delta)\right)/2\delta & ; \mu - \delta \leq t \leq \mu + \delta \\ 1 & ; \mu + \delta \leq t. \end{cases}$$

Moreover, $G(t)$ is differentiable except at 0.

Poof of Lemma 6: $G(t)$ is

$$Pr(X \cdot Y \leq t) = \int_{\mu-\delta}^{\mu+\delta} Pr(X \leq t/Y)\ \frac{dY}{2\delta} = \begin{cases} 0 & ; t \leq 0 \\ \int_{\mu-\delta}^{\mu+\delta} \frac{t}{Y}\ \frac{dY}{2\delta} & ; 0 \leq t \leq \mu-\delta \\ \int_{\mu-\delta}^{t} \frac{dY}{2\delta} + \int_{t}^{\mu+\delta} \frac{t}{Y}\ \frac{dY}{2\delta} & ; \mu-\delta \leq t \leq \mu+\delta. \\ 1 & ; t \geq \mu+\delta. \end{cases}$$

The rest follows from calculating the integrals, differentiating $G(t)$, and checking its differentiability at $t = \mu \pm \delta$. $\square$

From (18) and Lemma 6, the leader maximizes

$$(1-G(t))\ t = \begin{cases} \left(1 - \frac{t}{2\delta}\ log(\frac{\mu+\delta}{\mu-\delta})\right)\ t & ; t \in [0, \mu-\delta] \\ \left(1 - \frac{t}{2\delta}\ log\left(\frac{\mu+\delta}{t}\right) - \frac{t-(\mu-\delta)}{2\delta}\right)\ t & ; t \in [\mu-\delta, \mu+\delta]. \end{cases} \tag{49}$$

Differentiating with respect to t yields

$$\frac{d[(1-G(t))\ t]}{dt} = \begin{cases} 1 - \frac{t}{\delta}\ log(\frac{\mu+\delta}{\mu-\delta}) & ; t \in (0, \mu-\delta] \\ 1 - \frac{t}{\delta}\ log\left(\frac{\mu+\delta}{t}\right) - \frac{t-(\mu-\delta)}{2\delta} & ; t \in [\mu-\delta, \mu+\delta). \end{cases}$$

When $t \approx 0$, the derivative is clearly positive, and when $t \approx \mu+\delta$, using Taylor series expansion, one can show that the derivative is negative. Next, observe that

$$\left.\frac{d[(1-G(t))\ t]}{dt}\right|_{t=\mu-\delta} = 1 - \frac{\mu-\delta}{\delta}\ log(\frac{\mu+\delta}{\mu-\delta}).$$

One can show that $\frac{\mu-\delta}{\delta}\ log(\frac{\mu+\delta}{\mu-\delta})$ is strictly decreasing in δ, and using L'Hospital's rule, one can show that it approaches 2 when $\delta \to 0$, and it approaches 0 when $\delta \to \mu$. Thus, there is a unique $\delta \in (0, \mu)$ that solves $1 = \frac{\mu-\delta}{\delta}\ log(\frac{\mu+\delta}{\mu-\delta})$. Let $\hat{\delta}$ be that solution.

$$1 = \frac{\mu-\hat{\delta}}{\hat{\delta}}\ log(\frac{\mu+\hat{\delta}}{\mu-\hat{\delta}}). \tag{50}$$

Then,

$$\left.\frac{d[(1-G(t))\ t]}{dt}\right|_{t=\mu-\delta} > 0 \ \Leftrightarrow\ \delta > \hat{\delta}. \tag{51}$$

Further, observe that

$$\frac{d^2[(1-G(t))\ t]}{dt^2} = -\frac{1}{\delta}\ log(\frac{\mu+\delta}{\mu-\delta}) < 0, \quad t \in (0, \mu-\delta). \tag{52}$$

Thus, if $\delta < \hat{\delta}$, then there is a local maximum $t^* \in (0, \mu - \delta)$, satisfying the first order condition:

$$1 - \frac{t^*}{\delta}\ log(\frac{\mu+\delta}{\mu-\delta}) = 0 \ \Leftrightarrow\ t^* = \delta / log(\frac{\mu+\delta}{\mu-\delta}). \tag{53}$$

This yields the result if it is also the global maximum, which we will show below. Moreover, it can be shown that $t^*(\delta) = \delta/log(\frac{\mu+\delta}{\mu-\delta})$ is strictly decreasing in $\delta \in (0, \mu)$.

If $\delta > \hat{\delta}$, then t^* must be in $(\mu - \delta, \mu + \delta)$ and satisfy the first order condition:

$$1 - \frac{t^*}{\delta}\ log\left(\frac{\mu+\delta}{t^*}\right) - \frac{t^* - (\mu-\delta)}{2\delta} = 0. \tag{54}$$

First, one can confirm that $t^* = \frac{\mu-\hat{\delta}}{\mu+\hat{\delta}}\ (\mu + \delta)$ satisfies this equation, where we recall the definition of $\hat{\delta}$ from equation (50). To see this solution is unique, rewrite equation (54) as:

$$\frac{\mu+\delta}{2t^*} - \frac{1}{2} = log\left(\frac{\mu+\delta}{t^*}\right) \ \Leftrightarrow\ z - 1 = 2\ log(z),\ \text{where } z \equiv \frac{\mu+\delta}{t^*}.$$

Because $t^* \in (\mu - \delta, \mu + \delta)$ implies $z \in (1, \infty)$, the solution to $z - 1 = 2log(z)$ is unique. This solution corresponds to a maximum because we showed that the derivative is negative when $t \approx \mu + \delta$, and it is positive when $t \approx \mu - \delta$ and $\delta > \hat{\delta}$ (see (51)).

It remains to show that, if $\delta < \hat{\delta}$, then the local maximum in (53) is the global maximum. Because we showed in (52) that the objective function is concave for $t \in (0, \mu - \delta)$, it suffices to show that the global maximum is not in $(\mu - \delta, \mu + \delta)$. Suppose it is. This implies that there are more than one extremum in $(\mu - \delta, \mu + \delta)$. But we showed that the first order condition (54) has a unique solution, which is contradition. Thus, we have shown

$$t^*(\delta) = \begin{cases} \frac{\delta}{log\left(\frac{\mu+\delta}{\mu-\delta}\right)} & ;\delta \leq \hat{\delta} \\ & \end{cases}$$

Substituting this into (49) yields $\theta^*(\delta)$.

$$\theta^*(\delta) = [1 - G(t^*(\delta), \delta)]\ t^*(\delta) = \begin{cases} \frac{1}{2}\ t^*(\delta) & ;\delta \leq \hat{\delta} \\ \\ \frac{(t^*)^2}{2\delta}\ log\left(\frac{\mu+\delta}{t^*}\right) = \frac{\mu-\hat{\delta}}{\mu+\hat{\delta}}\ log\left(\frac{\mu+\hat{\delta}}{\mu-\hat{\delta}}\right)\ \frac{(\mu+\delta)^2}{2\delta} & ;\delta \geq \hat{\delta}. \end{cases}$$

For $\delta \geq \hat{\delta}$, $t^*(\delta)$ is strictly increasing, and it is easy to confirm that $\theta^*(\delta)$ is strictly decreasing in $\delta \in (0, \mu)$. Finally, we show that $t^*(\delta)$ is strictly decreasing in $\delta \in (0, \hat{\delta})$. Differentiating $\frac{\delta}{log\left(\frac{\mu+\delta}{\mu-\delta}\right)}$ with respect to δ, and substituting for $y = \frac{\mu+\delta}{\mu-\delta}$ yields

$$\frac{d}{d\delta}\left(\delta \Big/ log\left(\frac{\mu+\delta}{\mu-\delta}\right)\right) < 0,\ \text{for } \delta \in (0, \mu) \ \Leftrightarrow\ 2log(y) - y + \frac{1}{y} < 0,\ \text{for } y > 1.$$

The latter inequality is true because its derivative is $-(1 - 1/y)^2 < 0$. $\square$

6 References

Acemoglu, Daron, and James A. Robinson. 2001. "A Theory of Political Transition." *American Economic Review* 91: 938-63.

Acemoglu, Daron, and James A. Robinson. 2006. *Economic Origins of Democracy and Dictatorship.* New York: Cambridge University Press.

Arrow, Kenneth. 1966. "Optimal Capital Policy with Irreversible Investment." Institute for Mathematical Studies in Social Science, Stanford University.

Aytaç, Erdem, and Susan Stokes. 2019. *Why Bother?: Rethinking Participation in Elections and Protests.* New York, NY: Cambridge University Press.

Baron, David, and Roger Myerson. 1982. "Regulating a Monopolist with Unknown Costs." *Econometrica* 50 (4): 911-30.

Bass, Bernard. 1985. *Leadership and Performance Beyond Expectations.* NY: Free Press.

Blattman, Christopher, and Edward Miguel. 2010. "Civil War." *Journal of Economic Literature* 48: 3-57.

Börgers, Tilman. 2015. *An Introduction to the Theory of Mechanism Design.* New York: Oxford University Press.

Bueno de Mesquita, Ethan. 2010. "Regime Change and Revolutionary Entrepreneurs." *American Political Science Review* 104 (3): 446-66.

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

Burns, James MacGregor. 1978. *Leadership.* New York: Harper & Row Publishers.

Burns, James MacGregor. 2003. *Transforming Leadership: A New Pursuit of Happiness.* New York: Grove Press.

Chen, Heng, and Wing Suen. 2016. "Falling Dominoes: A Theory of Rare Events and Crisis Contagion." *American Economic Journal: Microeconomics* 8: 1-29.

Coate, Stephen, and Michael Conlin. 2004. "A Group Rule–Utilitarian Approach to Voter Turnout: Theory and Evidence." *American Economic Review* 94: 1476-1504.

Corazzini, Luca, Christopher Cotton, and Paola Valbonesi. 2015. "Donor Coordination in Project Funding: Evidence from a Threshold Public Goods Experiment." *Journal of Public Economics* 128: 16-29.

Dabashi, Hamid. 1993. *Theology of Discontent: The Ideological Foundation of the Islamic Revolution in Iran.* New York: New York University Press.

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

Edmond, Chris. 2013. "Information Manipulation, Coordination and Regime Change." *Review of Economic Studies* 80: 1422-58.

Egorov, Georgy, and Konstantin Sonin. 2017. "Elections in Non-Democracies." Mimeo.

Feddersen, Timothy, and Alvaro Sandroni. 2006. "A Theory of Participation in Elections." *American Economic Review* 96: 1271-82.

Frankel, David, Ady Pauzner and Stephen Morris. 2003. "Equilibrium Selection in Global Games with Strategic Complementarities." *Journal of Economic Theory* 108: 1-44.

Goldstone, Jack A. 2001. "Toward a Forth Generation of Revolutionary Theory." *Annual Review of Political Science* 4: 139-87.
Guesnerie, Roger, and Jean-Jacques Laffont. 1984. "A Complete Solution to a Class of Principal-Agent Problems with an Application in the Control of a Self-Managed Firm." *Journal of Public Economics* 25: 329-69.
Guimaraes, Bernardo, and Stephen Morris. 2007. "Risk and Wealth in a Model of Self-fulfilling Currency Attack." *Journal of Monetary Economics* 54: 2205-30.
Gurr, Ted. 1970. *Why Men Rebel?* Princeton, NJ: Princeton University Press
Hardin, Russell. 1982. *Collective Action*. Baltimore, MD: Johns Hopkins University Press.
Hartline, Jason D. 2016. *Mechanism Design and Approximation.* Online Manuscript.
Humphreys, Macartan, and Jeremy Weinstein. 2008. "Who Fights? The Determinants of Participation in Civil War." *American Journal of Political Science* 52(2): 436-55.
Johnson, Justin, and David Myatt. 2006. "On the Simple Economics of Advertising, Marketing, and Product Design." *American Economic Review* 96: 756-84.
Kamien, Morton, and Nancy Schwartz. 2012. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management (2nd Edition).* Dover.
Laffont, Jean-Jacques, and Jacques Robert. 1996. "Optimal Auction with Financially Constrained Buyers." *Economics Letters* 52: 181-86.
Lawrence, Adria. 2017. "Repression and Activism among the Arab Spring's First Movers: Evidence from Morocco's February 20th Movement." *British Journal of Political Science* 47(3): 699-718.
Lipnowski, Elliot, and Evan Sadler. 2019. "Peer-Confirming Equilibrium." *Econometrica* 87: 567-91.
Maier, Pauline. 1972. *From Resistance to Revolution: Colonial Radicals and the Development of American Opposition to Britain, 1765-1776.* New York: Vintage Books.
Maskin, Eric, and John Riley. 1984. "Monopoly with Incomplete Information." *RAND Journal of Economics* 15 (2): 171-96.
McCullough, David. 2005. *1776.* New York: Simon & Schuster.
Middlekauff, Robert. 2005. *The Glorious Cause: The American Revolution, 1763-1789.* New York, NY: Oxford University Press.
Mookherjee, Dilip, and I. P. L. Png. 1994. "Marginal Deterrence in Enforcement of Law." *Journal of Political Economy* 102 (5): 1039-66.
Morris, Stephen and Hyun Song Shin. 1998. "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks." *American Economic Review* 88 (3): 587-597.
Morris, Stephen, and Hyun Song Shin. 2003. "Global Games: Theory and Application." In *Advances in Economics and Econometrics, Theory and Applications, Eighth World Congress, Volume I,* edited by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky. New York: Cambridge University Press.
Mussa, Michael, and Sherwin Rosen. 1978. "Monopoly and Product Quality." *Journal of Economic Theory* 18: 301-17.
Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups.* Cambridge, MA: Harvard University Press.

Paine, Thomas. 1995. *Thomas Paine: Collected Writings.* New York, NY: Library of America.
Pearlman, Wendy. 2018. "Moral Identity and Protest Cascades in Syria." *British Journal of Political Science* 48: 877-901.
Petersen, Roger D. 2001. *Resistance and Rebellion: Lessons from Eastern Europe.* New York: Cambridge University Press.
Popkin, Samuel. 1979. *The Rational Peasant: The Political Economy of Rural Society in Vietnam.* Berkeley, CA: University of California Press.
Rhodehamel, John. 2001. *The American Revolution: Writings from the War of Independence 1775-1783.* New York, NY: The Library of America.
Riley, John, and Richard Zackhauser. 1983. "Optimal Selling Strategies: When to Haggle, When to Hold Firm." *Quarterly Journal of Economic* 98 (2): 267-289.
Ross, Michael. 2006. "A Closer Look at Oil, Diamonds, and Civil War." *Annual Review of Political Science* 9: 265-300.
Sandoz, Ellis. 1998. *Political Sermons of the American Founding Era: 1730-1805. 2nd Ed. Vol 1.* Indianapolis, IN: Liberty Fund.
Schelling, Thomas C. 1978. *Micromotives and Macrobehavior*. New York: Norton.
Seierstad, Atle, and Knut Sydsæter. 1993. *Optimal Control Theory with Economic Applications.* Amsterdam, The Netherlands: Elsevier.
Shadmehr, Mehdi. 2017. "Investment in the Shadow of Conflict: General Equilibrium Foundations of Regime Change." Mimeo.
Shadmehr, Mehdi, and Dan Bernhardt. 2019. "Vanguards in Revolution." *Games and Economic Behavior* 115: 146-66.
Shen, Lin, and Junyuan Zou. 2016. "Intervention with Voluntary Participation in Global Games." Mimeo.
Skocpol, Theda. 1997. "Cultural Idioms and Political Ideologies in the Revolutionary Reconstruction of State Power: A Rejoinder to Sewell." In *Social Revolutions in the Modern World*. Ed. by Theda Skocpol. New York: Cambridge University Press.
Snow, David, Burke Rochford, Jr., Steven Worden, and Robert Benford. 1986. "Belief Alignment Processes, Micromobilization, and Movement Participation." *American Sociological Review* 51: 464-81.
Tilly, Charles. 1978. *From Mobilization to Revolution.* Reading, MA: Addison-Wesley.
Tilly, Charles. 2008. *Contentious Performances.* New York: Cambridge University Press.
Tullock, Gordon. 1971. "The Paradox of Revolution." *Public Choice* 11: 89-99.
Tyson, Scott, and Alastair Smith. 2018. "Dual-Layered Coordination and Political Instability." *Journal of Politics* 80: 44-58.
Wantchékon, Léonard, and Omar García-Ponce. 2014. "Critical Junctures: Independence Movements and Democracy in Africa." Mimeo.
Washington, George. 1890. *The Writings of George Washington, (1777-1778). Collected and edited by Worthington Chauncey Ford.* New York: G. P. Putnam's Sons. Available at: http://oll.libertyfund.org/titles/2404
Weinstein, Jeremy. 2007. *Inside Rebellion: The Politics of Insurgent Violence.* New York, NY: Cambridge University Press.

White, Robert. 1989. "From Peaceful Protest to Guerrilla War: Micromobilization of the Provisional Irish Republican Army." *American Journal of Sociology* 94: 1277-1302.
Wood, Elisabeth J. 2003. *Insurgent Collective Action and Civil War in El Salvador.* New York, NY: Cambridge University Press.