Protest Puzzles: Tullock's Paradox, Hong Kong Experiment, and the Strength of Weak States*

Mehdi Shadmehr†

Abstract

Tullock's (1971) Paradox of Revolution uses an Olsonian logic to conclude that revolutions should not happen in large societies. Cantoni et al.'s (2019) Hong Kong Experiment shows that, in sharp contrast to the literature that models protest as a coordination problem, actions can be strategic substitutes. We develop a model to address these standing puzzles, and investigate its empirical implications. We show that when the movement's goal is modest, free-riding concerns dominate the citizens' interactions, making their actions strategic substitutes. By contrast, when the movement's goal is to topple the regime, coordination concerns dominate, and actions become strategic complements. Moreover, with natural other-regarding preferences, some citizens participate in costly revolt even in large societies. A new empirical implication of the model is that as a regime grows stronger in the sense that a larger fraction of citizens is needed to overthrow it, the likelihood of regime change may rise.

Keywords: Tullock's Paradox, Hong Kong Experiment, Protest, Strategic Complements, Strategic Substitutes, Pivotality

JEL Classification: D74, D8, H4.

*I am grateful for comments of Dan Bernhardt, Raphael Boleslavsky, Chris Cotton, Sergei Guriev, Stephen Morris, David Myatt, Jesse Shapiro, Konstantin Sonin, David Yang, Noam Yutchman, seminar participants and the discussion of Richard Van Weelden at 2019 ASSA. I have received excellent research assistance from Sina Ferdowsi.

†Department of Economics, University of Calgary. E-mail: `mshadmeh@gmail.com`

There are two standing puzzles in the literature on protests: Tullock's (1971) Paradox of Revolution, which is old and theoretical, and Cantoni et al.'s (2019) Hong Kong Experiment, which is new and empirical. Tullock's Paradox posits that (1) a successful revolution is just a public good, and (2) societies are large and one person's effect on the success of the revolution is negligible, while participating in a revolution is costly, so that revolutions should not occur. Nonetheless, revolutions happen. Cantoni et al.'s experiment shows that, in the context of the Hong Kong Democracy Movement, potential protesters who were presented with the information that others were more likely to protest, became less likely to protest—actions are strategic substitutes. This result suggests that the strategic interactions among potential protesters is a free-riding problem. However, based on anecdotal evidence, almost all current models of protest frame the strategic interactions among potential protesters as a coordination problem, so that when a citizen believes that others are more likely to protest, he becomes more inclined to protest—actions are strategic complements. We propose a model to address these puzzles, and investigate its empirical implications.

We adopt a collective action, regime change model. N citizens simultaneously decide whether to protest. Citizens have private, correlated costs of protest, and the protest succeeds whenever the fraction of protesters exceeds a threshold, which captures the regime's strength. There are no selective benefits, so regime change is a public good. We are interested in settings with large N. In contrast to the literature, players use the logic of pivotality. A citizen revolts whenever he believes that the likelihood that he is pivotal in determining the outcome is sufficiently large relative to his expected costs of protest. A key insight of our approach is to describe the strategic environment of protest as the mixture of the two extreme cases of pure coordination (and hence always featuring strategic complements) and pure free-riding (and hence always featuring strategic substitutes). As we will discuss below, this allows us to simultaneously address Tullock's Paradox and Cantoni et al.'s Hong Kong Experiment, and to generate novel empirical predictions.

The paper has three main results. The first result is about Tullock's Paradox. We show that if each citizen values a fraction of each other citizen's payoffs, then with a finite or countably infinite number of citizens, there is a unique monotone equilibrium in which some citizens with positive costs revolt. That is, there is a threshold $c^* > 0$ on private costs such that all citizens with costs below c^* revolt. As society grows larger, the influence of a single individual declines, but his benefit grows because he also cares about the payoffs of others. Thus, when the number of citizens is very large, the overall effect may still favor participation even when it is costly. Critically, we show that the likelihood that a citizen is pivotal in equilibrium falls at the rate of $1/N$. This implies that if a citizen's payoff from successful protest takes the natural form of $b_0 + b_1 N$, some citizens with positive costs protest even when N is very large. The form of $b_0 + b_1 N$ is natural because it represents that each citizen values a fraction of each other citizen's payoffs from a successful protest. Different forms of other-regarding preferences such as $b_0 + b_1 N^6$ or $b_0 + b_1 N^{1/2}$ are both difficult to interpret and have the unrealistic prediction that, in large societies, either no one revolts or everyone does.

Caring for others has always been an integral part of protests and social movements. Leaders who appeal to the people's sense of justice and ask them to help make the world a better place tap into these other-regarding preferences. Abraham Keteltas's 1777 sermon, "God Pleads His Cause," in the context of the American Revolution is an example (Sandoz 1998, p. 579-605):

> America will be a glorious land of freedom, knowledge, and religion, an asylum for distressed, oppressed, and persecuted virtue. Let this exhilarating 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.

The second result is about Cantoni et al.'s (2019) Hong Kong Experiment. We show that free-riding incentives dominate and actions become strategic substitutes if and only if the fraction of citizens needed for success is below a threshold. Thus, when a regime is weak or when protest goals are modest (e.g., to voice dissatisfaction with corruption or to keep the movement alive rather than toppling the regime) as in Cantoni et al.'s (2019) Hong Kong Experiment, a citizen is less likely to protest if he believes that others are more likely to protest. In contrast, when a regime is strong and the goal is to topple the regime (e.g., the protests preceding the 1979 Iranian Revolution), a citizen is more likely to protest if he believes that others are more likely to protest. The logic is that when success requires a relatively low fraction of citizens to protest, when others become more likely to protest, a citizen believes that he is less likely to be pivotal because there will likely be more than enough protesters. That is, for "easy" goals, the free-riding element of strategic interactions dominates.

To convey the basic logic, suppose the protest succeeds whenever qN out of N players protest, with $1 < qN < N$. If a citizen believes that each other citizen protests with a probability p (which will be endogenous), then he knows the probability that he is pivotal is proportional to $p^{qN}(1-p)^{(1-q)N}$. This probability is unimodal in p with a maximum at $p = q$; when a citizen believes that others are more likely to protest (i.e., as p increases), his estimate of being pivotal, and hence his incentives to protest, first rises and then falls. That is, both coordination and free-riding considerations co-exist. This logic implies that when a regime is weak or when a movement's goals are modest, so that the fraction of protesters needed for success (q) is small, free-riding is salient and actions are strategic substitutes. By contrast, when a regime is strong and a movement's goals are grand, coordination is salient and actions are strategic complements.

More broadly, this result sheds light on the conflicting empirical findings in the experimental literature that study the effect of beliefs about others' contributions on an individual's contribution to public goods. For example, in Hager et al.'s (2019a) study of right and left wing rallies in Germany, interactions among left activists exhibit strategic complements, while interactions among right activists exhibit strategic substitutes. In Frey and Meier (2004) and Shang and Croson (2009), actions are strategic complements, while in Hager et al. (2019b) actions are strategic substitutes. In this literature, strategic complements and substitutes are considered as opposite implications of distinct theories, e.g., free-riding versus desire for conformity. In contrast, in our model of pivotal revolutionaries, actions exhibit strategic complements or substitutes depending on environment parameters in a natural and intuitive way.

The third result is a new empirical implication of the model. Specializing to the Normal distribution, we characterize when making a regime stronger (raising q) has the paradoxical effect of increasing its risk of collapse. The logic is that when a higher fraction of citizens are needed for regime change, a citizen may believe that he is more likely to be pivotal, raising his incentive to revolt; and this strategic effect can swamp the direct effect of having a stronger regime. This is the potential strength of weak states: the very fact that weak states can be easily overturned may so exacerbate the free-riding problem among citizens that it makes those states more stable.

To establish our results, we use two key statistical properties, one from global games (Morris and Shin 2003), and one from Bayes' memoirs that is used in the literature on large elections (Good and Mayer 1975; Chamberlain and Rothschild 1981; Hummel 2012; Myatt 2015, 2017). An analytical challenge is that, due to the logic of pivotality, net expected payoffs are non-monotone. Therefore, the best response to a cutoff strategy need not be a cutoff strategy, precluding the existence of cutoff equilibria. With correlated private costs, when a citizen's costs are low, he expects many others, too, to have low costs and revolt, reducing his expectation of pivotality and with it, his incentives to act. When cutoff equilibria do not exist, one has to search for more complex equilibria, e.g., equilibria in which a citizen revolts when his costs are neither too high, nor too low, but rather are in a bounded interval (Chen and Suen 2017; Shadmehr and Bernhardt 2017). We show that the best response to a cutoff strategy is also a cutoff strategy if and only if the noise in signals is *not* too small.

In the Online Appendix, we study the simple and unrealistic case of independent private costs. We also show that in revolution games with uncertainty about post revolution payoffs and additive normal noise signal structure (Bueno de Mesquita 2010; Shadmehr and Bernhardt 2011), the logic of pivotality precludes monotone equilibria in large societies.

Literature. This paper is related to the literature on protests and revolutions (Bueno de Mesquita 2010, 2013; Shadmehr and Bernhardt 2011; Boix and Svolik 2013; Edmond 2013; Casper and Tyson 2014; Guriev and Treisman 2015; Chen and Suen 2017; Egorov and Sonin 2017; Tyson and Smith 2018). Given that selective material benefits are hard to justify in revolution settings, models of revolution and protest explicitly or implicitly use psychological incentives to circumvent Tullock's Paradox. These rewards take two forms. Some papers posit selective "warm glow" benefits from participating in a successful revolution (Bueno de Mesquita 2010). Others posit expressive benefits by presuming that citizens derive psychological benefits from participating in a revolution, regardless of whether or not it succeeds (Egorov and Sonin 2017). Regardless of the validity of either of these theories,[1] neither the expressive nor pleasure-in-agency approaches resolve Tullock's Paradox; rather, they claim that it was not a paradox in the first place. The pleasure-in-agency approach presumes that there are selective (psychological) benefits, while the expressive benefits approach presumes that there are no net costs in participating due to the psychological benefits of expressing one's emotions. Rather than circumvent Tullock's Paradox, this paper accepts its assumptions, but acknowledges that citizens also care about their fellow citizens, showing that with natural other-regarding preferences, even in large societies, citizens participate in costly protest.

This paper is also related to the literature on voter turnout and costly voting (Palfrey and Rosenthal 1985; Börgers 2004; Myatt 2015). However, in that literature, voting for all alternatives is costly, and costs are uncorrelated. A closer literature studies strategic and protest voting in large elections with aggregate uncertainty (Razin 2003; Myatt 2007; Dewan and Myatt 2007). Key shared features are the uncertainty about the aggregate turnout or candidate votes, and the application of asymptotic pivotality results from Good and Mayer (1975) and Chamberlain and Rothschild (1981). The closest is Myatt (2017), in which citizens can protest against their favored party by voting for a protest candidate. Citizens want enough protest votes to change the party's behavior, but not too much to lose the election. Thus, there is an endogenous cost of protest vote because it may cause the party to lose. In both

[1]Each of these approaches has a genealogy in social sciences that offers evidence for its validity. As Morris and Shadmehr (2018) discuss, Wood's (2003) notion of "pleasure in agency" captures the selective psychological benefits that a citizen receives from participating in a movement that succeeds. Pleasure in agency fits within the 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 contrast, the expressive payoffs approach is an implication of Gurrian psychological theories of revolution, delineated in *Why Men Rebel?* (Gurr 1971), which claims that citizens use revolution to pursue a cathartic release of their grievances.

papers, monotone equilibria do not exist if private signals are too informative. In contrast to this paper, what matters in Myatt's (2017) model is the *ratio* of the likelihood of pivotality in one outcome rather than another. This together with the payoff structure and Normal distributions cause payoffs to be linear in signals and actions to be always strategic substitutes.

1 Benchmark: Standard Models of Revolution

To demonstrate the approach of the literature, we adapt the standard global game model in Figure 1. A continuum of citizens, indexed by $i \in [0,1]$, must simultaneously decide whether or not to revolt. The payoff of a citizen who does not revolt is normalized to 0. A citizen who participates in a successful revolution receives an exogenous "warm glow" payoff $b > 0$. A citizen who revolts incurs a cost or receives expressive benefits, c_i, where $c_i = \theta + \sigma\ \epsilon_i$, and θ and ϵ_is are independent. Citizens share an improper prior that θ is distributed uniformly on $\mathbb{R}$, and $\epsilon_i \sim F$ with full support on $\mathbb{R}$. The regime collapses whenever the fraction of revolters, n, exceeds a threshold $q \in (0,1)$. This game is a special case of Morris and Shin (2003), where

		outcome	
		$n > q$	$n \leq q$
citizen i	*revolt*	$b - c_i$	$-c_i$
	no revolt	0	0

Figure 1: Regime Change Game with Selective Benefits $b > 0$.

the net expected payoff from revolting versus not revolting grows in the fraction of players who revolt. Actions are always strategic complements: a citizen is more likely to revolt if he believes that others are more likely to revolt. As Morris and Shin (2003) show, without loss of generality, we can focus on symmetric cutoff strategies in which a citizen revolts whenever his cost is below a threshold c^*. Given a regime's strength θ, this strategy implies that the fraction of revolters is $Pr(c_i < c^*|\theta)$. Thus, a regime collapses whenever $\theta < \theta^*$, where $Pr(c_i < c^*|\theta^*) = q$. Now, given that a citizen's belief must be consistent with strategies, a citizen i believes that the regime collapses with probability $Pr(\theta < \theta^*|c_i)$, and hence revolts whenever his expected payoff exceeds his cost of revolt: $Pr(\theta < \theta^*|c_i) \cdot b > c_i$. Thus, equilibrium is characterized by a pair of thresholds, (c^*, θ^*), that satisfy the consistency of beliefs with strategies and the indifference condition of the marginal citizen with signal $c_i = c^*$:

$$Pr(c_i < c^*|\theta^*) = q \quad \text{and} \quad Pr(\theta < \theta^*|c_i = c^*) \cdot b = c^*.$$

One can simplify the analysis by exploiting a key statistical property. When the prior is uniform or the noise goes to zero, we have (Morris and Shin 2003): $Pr(c_i < \hat{c}|\theta = \hat{\theta}) = Pr(\theta > \hat{\theta}|c_i = \hat{c})$ for all $\hat{\theta}$ and $\hat{c}$.[2] This lets us write the above equations as:

$$F\left(\frac{c^* - \theta^*}{\sigma}\right) = q \quad \text{and} \quad \left(1 - F\left(\frac{c^* - \theta^*}{\sigma}\right)\right) \cdot b = c^*.$$

Thus,

$$c^* = b\ (1-q) \quad \text{and} \quad \theta^* = c^* - \sigma\ F^{-1}(q) = b\ (1-q) - \sigma\ F^{-1}(q). \tag{1}$$

Proposition 1 formally states these standard results as well as the intuitive comparative statics with respect to the regime's strength q.

Proposition 1 *Actions are always strategic complements. There is a unique monotone equilibrium characterized by (c^*, θ^*) given in (1). In equilibrium, a citizen revolts whenever his signal is below c^*, and the regime collapses whenever $\theta < \theta^*$. In particular, without warm glow payoffs, only citizens with expressive benefits revolt: $c^*(b = 0) = 0$. Moreover, when the regime becomes stronger in the sense that more citizens must revolt for the regime to collapse, then in equilibrium, both less citizens revolt and the regime is more likely to survive:*

$$\frac{\partial \theta^*}{\partial q} < \frac{\partial c^*}{\partial q} < 0.$$

[2]To see this, note that when θ is uniformly distributed, so that there is no prior information about it, there is no difference between the signal c_i and the fundamental θ. Thus, we can think of θ as a signal of c_i: $\theta = c_i - \sigma\ \epsilon_i$. Then, $Pr(\theta > \hat{\theta}|c_i = \hat{c}) = Pr(c_i - \sigma\ \epsilon_i > \hat{\theta}|c_i = \hat{c}) = Pr(\frac{\hat{c}-\hat{\theta}}{\sigma} > \epsilon_i) = Pr(c_i < \hat{c}|\theta = \hat{\theta})$. It can be shown that the same result is obtained with well-behaved distributions in the limit when the noise goes to zero.

2 A Model of Pivotal Revolutionaries

Although the standard model is simple and elegant, it neither addresses Tullock's (1971) Paradox of Revolution nor Cantoni et al.'s (2019) Hong Kong Experiment. Actions are always strategic complements, and without warm glow payoffs, only citizens with expressive benefits protest. We propose an alternative model of protest to address these puzzles, while maintaining as much of the standard model structure as possible.

Consider $N+1 \in \mathbb{N}$ citizens, and assume that the revolution succeeds whenever the number of revolters exceeds qN for some $q \in (0,1)$. To ease exposition, we assume that $qN \in \mathbb{N}$. For example, $N=3$ and $q=2/3$ implies that $qN=2$, and hence at least three out of the total of four must revolt for the revolution to succeed. We will focus on large N. To preserve similarity with the game analyzed earlier, we analyze the game represented in Figure 2. A citizen who revolts incurs a cost or receives expressive benefit, c_i, where $c_i = \theta + \sigma\ \epsilon_i$, and θ and ϵ_is are independent. Citizens share an improper prior that θ is distributed uniformly on $\mathbb{R}$, and $\epsilon_i \sim F$, where F is twice continuously differentiable with full support on $\mathbb{R}$.

citizen i	outcome: $n > qN$	outcome: $n \leq qN$
revolt	$u(N) - c_i$	$-c_i$
no revolt	$u(N)$	0

Figure 2: Regime Change Game with Public Benefits $u(N) > 0$, with $u'(N) \geq 0$, where we make the potential dependence of benefits on the total size of the society explicit.

We focus on symmetric equilibria in cutoff strategies, so that a citizen revolts if and only if his signal is below a threshold. If other citizens take a cutoff strategy and revolt whenever their signals are below c^*, then a citizen i with signal c_i revolts if and only if:

$$B(c_i) = u(N)\int_{\theta=-\infty}^{\infty} \binom{N}{qN} \left(F\left(\frac{c^*-\theta}{\sigma}\right)\right)^{qN} \left(1-F\left(\frac{c^*-\theta}{\sigma}\right)\right)^{(1-q)N} pdf(\theta|c_i)\; d\theta > c_i. \tag{2}$$

The left hand side, $B(c_i)$, is citizen i's payoff $u(N)$ times the probability that he assigns to being pivotal. That is, the probability that exactly qN other citizens revolt. A citizen j revolts whenever $c_j < c^*$, but i does not observe c_j. If he knew θ, then he would believe that j revolts with probability $Pr(c_j < c^*|\theta) = F(\frac{c^*-\theta}{\sigma})$. Moreover, conditional on θ, each citizen's decision is independent of others, so that we can use binomial distribution. Thus, given θ, citizen i would believe that he is pivotal with probability $\binom{N}{qN} \left(F\left(\frac{c^*-\theta}{\sigma}\right)\right)^{qN} \left(1-F\left(\frac{c^*-\theta}{\sigma}\right)\right)^{(1-q)N}$. But he does not know θ, and he has to estimate this probability given his signal c_i.

Critically, the best response to a monotone strategy need not be monotone. As a citizen's signal increases, his beliefs that he is pivotal first rise, and then fall. But *if* the best response to a monotone strategy is monotone, then all cutoff equilibria with associated cutoff c^* are characterized by the indifference condition of the marginal citizen whose signal equals the cutoff. This indifference condition looks complex at first, and it seems that there can be multiple equilibria. Remarkably, one can exploit two statistical properties to simplify the indifference condition in two steps, showing that the equilibrium is unique and that is takes a simple form:

$$\begin{aligned} c^* &= u(N)\int_{\theta=-\infty}^{\infty} \binom{N}{qN} \left(F\left(\frac{c^*-\theta}{\sigma}\right)\right)^{qN} \left(1-F\left(\frac{c^*-\theta}{\sigma}\right)\right)^{(1-q)N} pdf(\theta|c_i=c^*)\; d\theta \\ &= u(N)\int_{z=0}^{1} \binom{N}{qN} z^{qN}\,(1-z)^{(1-q)N}\; dz \\ &= \frac{u(N)}{1+N}. \end{aligned} \tag{3}$$

These calculations make use of two key statistical properties from two literatures:

1. The first step (second equality) exploits a statistical property that is often used in global games (Morris and Shin 2003; Morris and Shadmehr 2018). We used the same property in our analysis of standard global game models of revolution, leading to equation (1). When either

the prior is very diffuse or the noise is very small, for any pair of thresholds $\hat{\theta}$ and $\hat{c}$ we have:

$$Pr(\theta < \hat{\theta}|c_i = \hat{c}) = 1 - Pr(c_i < \hat{c}|\hat{\theta}) = 1 - F\left(\frac{\hat{c} - \hat{\theta}}{\sigma}\right). \tag{4}$$

Now consider the marginal citizen with signal $c_i = c^*$. He does not know θ, and has a belief about the probability that another citizen revolts. Applying (4) reveals that this belief is uniform:

$$\begin{aligned} Pr\left(F\left(\frac{c^* - \theta}{\sigma}\right) < A \middle| c_i = c^*\right) &= Pr(c^* - \sigma\ F^{-1}(A) < \theta | c_i = c^*) \\ &= F\left(\frac{c^* - c^* + \sigma\ F^{-1}(A)}{\sigma}\right) \quad \text{(from (4))} \\ &= A. \end{aligned}$$

That is, from the perspective of the marginal citizen whose signal is exactly the cutoff, the probability that another citizen revolts is uniformly distributed on $[0, 1]$. That is, a change of variable from θ to $z = F\left(\frac{c^* - \hat{\theta}}{\sigma}\right) = 1 - Pr(\theta < \hat{\theta}|c_i = c^*)$, allows us to write $dz/d\theta = -pdf(\theta = \hat{\theta}|c_i = c^*)$ for any $\hat{\theta}$.

2. The second step (last inequality) in (3) is due to Bayes in *The Doctrine of Chances* (Gillies 1987). To see that it is true, we use the argument of Chamberlain and Rothschild (1981). Consider $N + 1$ random variables $\{X_0, X_1, \cdots, X_N\}$ with $X_i \sim iid\ U[0, 1]$, and let x_i denote a realization of X_i. Now, consider a random draw for each and rank them in the usual order. First, observe that because these random variables are identical, the probability that the realization x_0 is the $qN + 1$st smallest is $\frac{1}{1+N}$: Each of the $N + 1$ random variables are equally likely to be the $qN + 1$st smallest one. Next, observe that if we knew $x_0 = z$, then the probability that x_0 was the $qN + 1$st smallest one would be $\binom{N}{qN}\ z^{qN}\ (1 - z)^{(1-q)N}$: qN draws must be lower than z (each happening with probability z) and the remaining $(1 - q)N$ must be above z (each happening with probability $1 - z$). Of course, we do not know that $x_0 = z$. We have a uniform prior that X_0 is uniformly distributed between 0 and 1. Thus, to calculate the overall probability that x_0 is the $qN + 1$st smallest, we must integrate over those probabilities. But this is exactly the integral above (3). Combining these two observations, we conclude that the integral must be $\frac{1}{1+N}$.

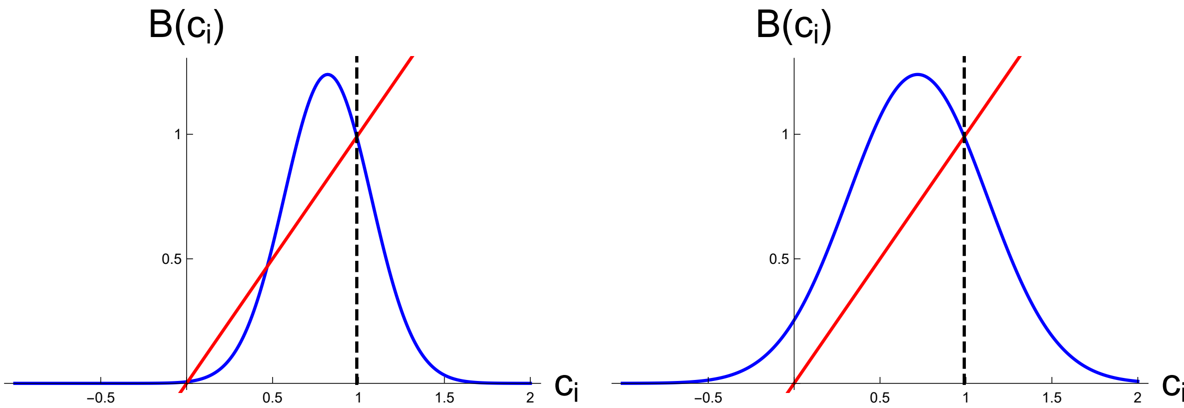

Figure 3: The curve is the left hand side of (2), $B(c_i; c^*, \sigma)$, and the straight line is the 45 degree line. In the left panel, the noise is small ($\sigma = 0.25$) and the best response to a monotone strategy with cutoff $c^* = 1$ is non-monotone. In the right panel, the noise is larger ($\sigma = 0.4$), and the best response to the monotone strategy with cutoff $c^* = 1$ is a cutoff strategy with cutoff c^*. Parameters: $c_i \sim N(0, \sigma)$, $u(N) = N$, $q = 0.75$, and $N = 100$.

The analysis above *assumed* that the best response to a cutoff (monotone) strategy is a cutoff strategy. This enabled us to fully characterize the equilibrium with the indifference condition (3). But the logic of pivotality together with the correlated nature of costs cause non-monotonicities that can preclude monotone equilibria. Suppose all citizens except citizen i take a cutoff strategy with cutoff c^*. Then, citizen i's net expected payoff from revolting versus not revolting is generally non-monotone. As i's signal c_i increases, two conflicting economic forces arise: (i) the direct, non-strategic effect reduces i's incentives to revolt; but (ii) because costs are correlated, citizen i believes that others will also reduce their participation, and this can raise the likelihood that i is pivotal, raising his incentives to revolt. Figure 3 demonstrates. Despite these non-monotonicities, we show that when the noise in private signals is not too small, so that the second effect is relatively weak, the best response to cutoff strategies of others is a cutoff strategy—see the proof of Proposition 2 in the Appendix.

Finally, we turn to the equilibrium likelihood of regime change implied by the citizens' equilibrium strategies. The likelihood of regime change is the probability that at least $qN+1$

citizens revolt. Think of each citizen's decision as a binary random variable $X_i \in \{0, 1\}$, where $X_i = 1$ corresponds to revolt. Thus, the regime collapses if and only if

$$n(c^*, \theta) > qN \quad \Leftrightarrow \quad \frac{\sum_{i=1}^{N} X_i}{N} > q.$$

Conditional on θ, in equilibrium, the likelihood that a citizen revolts ($X_i = 1$) is $F\left(\frac{c^*-\theta}{\sigma}\right)$, and these random variables are independent. Thus, by the Law of Large Numbers, as N grows unboundedly, $\frac{\sum_{i=1}^{N} X_i}{N}$ goes to $E[X_i]$, which is $F\left(\frac{c^*-\theta}{\sigma}\right)$. Therefore, revolution succeeds if and only if

$$\theta < \theta^*, \text{ where } \theta^* = \lim_{N \to \infty} c^*(N) - \sigma F^{-1}(q).$$

Proposition 2 *Suppose the noise in private signals is not too small. Then there is a unique equilibrium in symmetric finite-cutoff strategies, in which a citizen revolts if and only if his private signal is below a threshold:*

$$c^* = \frac{u(N)}{1+N}.$$

In particular, c^* does not depend on the regime's strength q. Moreover, in the limit as $N \to \infty$, the regime collapses if and only if:

$$\theta < \theta^* = \lim_{N \to \infty} \frac{u(N)}{1+N} - \sigma F^{-1}(q).$$

Corollary 1 ***(Tullock's Paradox)*** *If $u(N) = b_0 + b_1 N$, with $b_1 > 0$, then in the limit as $N \to \infty$, $c^* = b_1 > 0$. That is, some citizens with positive costs of revolt participate in the revolution.*

We conclude that Tullock's Paradox of Revolution is resolved in the sense that, with natural of other-regarding preferences, even some of those citizens without expressive or pleasure in agency payoffs will revolt. The nature of these other-regarding preferences is simple: each citizen values each other citizen's payoff as a fraction of his own payoff.

We now turn to Cantoni et al.'s (2019) Hong Kong Experiment, which shows that citizens' actions can be strategic substitutes in protest settings. To analyze whether actions are strategic complements or substitutes in equilibrium, let $B(c_i; c^*)$ be a citizen i's net expected benefit from revolting versus not revolting when other citizens choose a cutoff c^*—i.e., the left hand side of equation (2). When c^* is an equilibrium cutoff, we have $B(c_i; c^*) - c_i = 0$ at $c_i = c^*$, because a citizen is indifferent between revolting and not revolting at the equilibrium cutoff. To address Cantoni et al.'s (2019) Experiment, we want to know whether citizen i's incentives to revolt increase or decrease if all other citizens marginally raise their cutoff from c^* in equilibrium. Thus, we need the sign of $\frac{\partial B(c_i;c^*)}{\partial c^*}\Big|_{c_i=c^*}$. In equation (7) of the Appendix we show that

$$\lim_{N\to\infty} \frac{\partial B(c_i; c^*, \sigma)}{\partial c_i} = \frac{1}{\sigma} \frac{f'\left(\frac{c_i - c^*}{\sigma} + F^{-1}(q)\right)}{f(F^{-1}(q))} \lim_{N\to\infty} \frac{u(N)}{1+N},$$

which, in turn, implies

$$\lim_{N\to\infty} \frac{\partial B(c_i; c^*)}{\partial c^*}\Bigg|_{c_i=c^*} = -\lim_{N\to\infty} \frac{\partial B(c_i; c^*)}{\partial c_i}\Bigg|_{c_i=c^*} = -\frac{1}{\sigma} \frac{f'(F^{-1}(q))}{f(F^{-1}(q))} \lim_{N\to\infty} \frac{u(N)}{1+N}.$$

Thus, we have:

Proposition 3 ***(Hong Kong Experiment)*** *Suppose $f(\cdot)$ is strictly unimodal, with $q_m = F(mode)$. At equilibrium, actions are strategic substitutes if $q < q_m$ and strategic complements if $q > q_m$.*

Proposition 3 shows that when the necessary fraction of protesters for a successful protest is below a threshold (q_m), actions are strategic substitutes in equilibrium. That is, when success is "easy," free-riding dominates coordination considerations, and when a citizens believes that others are more likely to protest, he has less incentives to protest. The ease of the success depends both on the regime's strength and the goal of the movement. q is lower when the regime is relatively weak, or when the movement's goals are modest, e.g., keeping the movement alive rather than bringing about major changes.

Cantoni et al.'s (2019) Hong Kong Experiment analyzed the beliefs and behavior of a sample of students from the Hong Kong University of Science and Technology (HKUST) around the July 1, 2016, protests. These protests were part of the annual July 1 protests, which have been organized yearly since the British "handover" of Hong Kong to China in the late 1990s, and grew in popularity in the early 2000s. The key goals were, "first, to denounce the perceived corruption of Beijing-backed Chief Executive C. Y. Leung; second, to mobilize support for democratic—especially the newly established localist—political parties in the run-up to the 2016 LegCo elections" (Cantoni et al. 2019, p. 1030), and to keep the movement alive and set the stage for some future time when major democratization goals (e.g., democratic election of the chief executive) could be achieved. Moderate protest goals correspond to lower qs. For such settings, Proposition 3 suggests that the strategic interactions between potential protesters resemble classic free-riding problems—e.g., contributing to building a public bridge, or providing public education. In contrast, when the regime is strong and goals are grand, so that q is high (e.g., during the months preceding the 1979 Iranian Revolution), actions become strategic complements, and we fall into the realm of standard protest models.

3 The Strength of Weak States

To investigate what happens when the prior is not uniform, we specialize to normal noise signal settings, where the distributions of the prior and the noise are both Normal: $\theta \sim N(0, \sigma_0)$ and $\epsilon_i \sim iidN(0, \sigma)$, with θ and ϵ_i being independent. Let $\phi(\cdot)$ be the pdf and $\Phi(\cdot)$ be the cdf of the standard normal distribution. Recall that

$$\theta | c_i \sim \frac{1}{\sqrt{\beta\sigma^2}} \; \phi\left(\frac{\theta - \beta c_i}{\sqrt{\beta\sigma^2}}\right) \quad \text{with} \quad \beta \equiv \frac{\sigma_0^2}{\sigma^2 + \sigma_0^2}.$$

Using the same change of variables as in (3), $z = F\left(\frac{c^* - \theta}{\sigma}\right)$, the left hand side of (2), $B(c_i)$, can be written as:

$$u(N) \int_{z=0}^{1} \binom{N}{qN} \; (z)^{qN} \; (1-z)^{(1-q)N} \; \left\{ \frac{1}{\sqrt{\beta}} \; \phi\left(\frac{c^* - \beta c_i}{\sqrt{\beta\sigma^2}} - \frac{\Phi^{-1}(z)}{\sqrt{\beta}} \right) \Big/ \phi(\Phi^{-1}(z)) \right\} dz. \qquad (5)$$

Because the prior is not uniform, the marginal citizen's belief about the probability (conditional on θ) that another citizen will revolt is not uniform, and we cannot obtain a simple closed form solution as in (3), nor can we conclude that the equilibrium is unique in general. Although this equation looks complicated, it significantly simplifies in the limit as $N \to \infty$, using a result from Good and Mayer (1975) and Chamberlain and Rothschild (1981) that, as

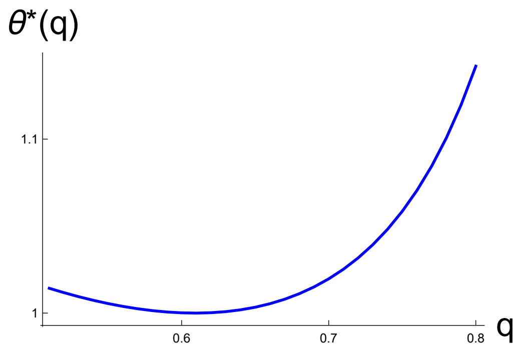

Figure 4: Parameters: $\sigma = 0.6$, $\sigma_0 = 1$, $u(N) = N$, and $N = 1000$.

$N \to \infty$, the term $z^{qN}(1-z)^{(1-q)N}$ becomes very sharp at $z = q$, putting almost all weight in a vanishingly small neighborhood of q. This means that only the value at $z = q$ of any continuous term that multiplies $z^{qN}(1-z)^{(1-q)N}$ will matter:

$$\lim_{N\to\infty} B(c_i; c^*, \sigma, \sigma_0) = \lim_{N\to\infty} \frac{u(N)}{1+N} \times \frac{1}{\sqrt{\beta}}\, \phi\left(\frac{c^* - \beta c_i}{\sqrt{\beta\sigma^2}} - \frac{\Phi^{-1}(q)}{\sqrt{\beta}}\right) \Big/ \phi(\Phi^{-1}(q)).$$

Thus, if the best response to a monotone strategy is monotone (e.g., when the noise in private signals is sufficiently large), then the equilibria are characterized by the solutions to the indifference condition,

$$\lim_{N\to\infty} \frac{u(N)}{1+N} \times \frac{1}{\sqrt{\beta}}\, \phi\left(\frac{(1-\beta)c^*}{\sqrt{\beta\sigma^2}} - \frac{\Phi^{-1}(q)}{\sqrt{\beta}}\right) \Big/ \phi(\Phi^{-1}(q)) = c^*.$$

Focusing on the case with $u(N) = b_0 + b_1 N$, there are generically either one or three equilibria. Critically, unlike the case with a uniform prior, now the equilibrium cutoff c^* depends on the regime's strength q. Moreover, in contrast with existing models, when a regime becomes stronger in the sense that a larger fraction of citizens is needed to overthrow the regime, more citizens may revolt: $c^*(q)$ can be increasing in q. Proposition 4 characterizes when this strategic effect dominates the direct effect of raising q, thereby making stronger regimes more unstable.

Proposition 4 *In any stable equilibrium, increasing the regime's strength q increases the likelihood of regime change if and only if that likelihood is sufficiently high:*

$$\frac{\partial \theta^*(q)}{\partial q} > 0 \quad \textit{if and only if} \quad \theta^* > \bar{\theta},$$

where $\bar{\theta}$ *depends on* q, σ *and* σ_0.

The intuition reflects the logic of pivotality. When a regime becomes stronger, the marginal citizen with signal $c_i = c^*$ believes that the likelihood that he is pivotal rises, increasing his incentives to revolt. This strategic effect can swamp the direct effect so that the overall likelihood of revolution *increases* with the regime's strength: $\frac{\partial c^*}{\partial q} > \frac{\partial \theta^*}{\partial q} > 0$. Because $\theta^* = c^* - \sigma \Phi^{-1}(q)$, the condition $\theta^* > \bar{\theta}$ means that for the strategic effect to dominate, c^* must be sufficiently large. But free-riding considerations are severe exactly when c^* is large, so that each citizen believes that others are likely to revolt. Then, raising q mitigates these free-riding considerations by increasing the marginal citizen's belief that he will be pivotal.

Figure 4 illustrates an example of this phenomenon in a case where the equilibrium is unique. In weaker states, a citizen believes that he is less likely to be pivotal for regime change, and therefore he has less incentives to revolt. This strategic effect can dominate, making weaker states more stable. We call this phenomenon the strength of weak states.

4 Conclusion

This paper develops a model of pivotal revolutionaries that resolves the two standing puzzles in the protest literature. It also provides a novel prediction for the relationship between the strength of regimes and their stability—the strength of weak states. Although the paper focuses on the context of protests and revolutions, our model and results routinely apply to the literature on threshold public good games (Palfrey and Rosenthal 1984; Andreoni 1998; Corazzini et al. 2015), by replacing the word "revolter" with "contributor." For example, the strength of weak states then translates to the difficulty of easy projects. Therefore, our paper brings the insights from global games methodology as well as the asymptotic properties of the logic of pivotality to the study of private provision of public goods. Delineating the implications of our analysis for the theoretical and experimental literature on threshold public goods is left for future research.

Several additional directions for future research stand out. We analyzed a setting in which players had binary actions and differed only in their costs. As Morris and Shadmehr (2018) demonstrate, regime change global games can accommodate both continuous actions and other forms of player heterogeneity. (1) What happens if actions are continuous? (2) What happens if we introduce uncorrelated heterogeneity, e.g., by positing that players also differ in their utility from a successful regime change (or provision of public good)? What is the effect of more heterogeneity? (3) We analyzed the game in the absence of common knowledge about the uncertain, shared cost component. What happens if this assumption is relaxed? (4) What happens in alternative settings in which uncertainty is about the threshold of success? As we show in the Online Appendix, analyzing such settings is complex because monotone equilibria may not exist.

Appendix: Proofs

Proof of Proposition 2: It remains to show the monotonicity of best responses. Recall that we must show the left hand side of (2), as a function of c_i, crosses the 45 degree line at a unique point and from above. It suffices to show that the slope of the left hand side is less than 1. Using the same change of variables as in (3), $z = F\left(\frac{c^*-\theta}{\sigma}\right)$, the left hand side of (2) can be written as:

$$B(c_i; c^*, \sigma) \equiv u(N) \int_{z=0}^{1} \binom{N}{qN} z^{qN}(1-z)^{(1-q)N} \frac{f\left(\frac{c_i-c^*}{\sigma} + F^{-1}(z)\right)}{f(F^{-1}(z))} \, dz. \qquad (6)$$

For the marginal citizen with $c_i = c^*$, the distribution of z is uniform and the distribution term simplifies to 1. For others, this distribution is not uniform in general, complicating the analysis. Differentiating with respect to c_i yields:

$$\frac{\partial B(c_i; c^*, \sigma)}{\partial c_i} = u(N) \int_{z=0}^{1} \binom{N}{qN} z^{qN}(1-z)^{(1-q)N} \frac{1}{\sigma} \frac{f'\left(\frac{c_i-c^*}{\sigma} + F^{-1}(z)\right)}{f(F^{-1}(z))} \, dz.$$

Using the result from Good and Mayer (1975) and Chamberlain and Rothschild (1981) that let us simplify (5), we can provide a relatively simple characterization when N is large:

$$\begin{aligned}
\lim_{N\to\infty} \frac{\partial B(c_i; c^*, \sigma)}{\partial c_i} &= \lim_{N\to\infty} u(N) \int_{z=0}^{1} \binom{N}{qN} z^{qN}(1-z)^{(1-q)N} \frac{1}{\sigma} \frac{f'\left(\frac{c_i-c^*}{\sigma} + F^{-1}(z)\right)}{f(F^{-1}(z))} \, dz. \\
&= \frac{1}{\sigma} \frac{f'\left(\frac{c_i-c^*}{\sigma} + F^{-1}(q)\right)}{f(F^{-1}(q))} \lim_{N\to\infty} u(N) \int_{z=0}^{1} \binom{N}{qN} z^{qN}(1-z)^{(1-q)N} \, dz. \\
&= \frac{1}{\sigma} \frac{f'\left(\frac{c_i-c^*}{\sigma} + F^{-1}(q)\right)}{f(F^{-1}(q))} \lim_{N\to\infty} \frac{u(N)}{1+N}. \qquad (7)
\end{aligned}$$

Thus, for sufficiently large σ, the slope of $B(c_i; c^*, \sigma)$ is always less than one, and hence the best response to a finite-cutoff strategy is a finite-cutoff strategy. □

Proof of Proposition 4: Let $z = \Phi^{-1}(q)$, and $b_1 = \lim_{N\to\infty} \frac{u(N)}{1+N}$. Rewrite the indifference condition as:

$$\frac{b_1}{\sqrt{\beta}} \, \phi\left(\frac{(1-\beta)c^*}{\sigma\sqrt{\beta}} - \frac{z}{\sqrt{\beta}}\right) - \phi(z) \, c^* = 0. \qquad (8)$$

To ease exposition, define $\gamma = \frac{(1-\beta)c^*}{\sigma\sqrt{\beta}} - \frac{z}{\sqrt{\beta}}$, and recall that that for the Normal distribution, we have $\phi'(x) = -x\phi(x)$. Thus,

$$\begin{aligned}\frac{\partial c^*(z)}{\partial z} &= -\frac{\frac{b_1}{\sqrt{\beta}}(-\gamma)\phi(\gamma)\left(-\frac{1}{\sqrt{\beta}}\right) - (-z)\phi(z)c^*}{\frac{b_1}{\sqrt{\beta}}(-\gamma)\phi(\gamma)\left(\frac{1-\beta}{\sigma\sqrt{\beta}}\right) - \phi(z)} = \frac{\frac{b_1}{\sqrt{\beta}}\gamma\phi(\gamma)\frac{1}{\sqrt{\beta}} + z\phi(z)c^*}{\frac{b_1}{\sqrt{\beta}}\gamma\phi(\gamma)\frac{1-\beta}{\sigma\sqrt{\beta}} + \phi(z)} \\ &= \frac{\frac{b_1}{\sqrt{\beta}}\gamma\phi(\gamma)\frac{1}{\sqrt{\beta}} + z\frac{b_1}{\sqrt{\beta}}\phi(\gamma)}{\frac{b_1}{\sqrt{\beta}}\gamma\phi(\gamma)\frac{1-\beta}{\sigma\sqrt{\beta}} + \phi(z)} \quad \text{(substituting for } c^* \text{ from equation (8))}, \qquad (9)\end{aligned}$$

where we recognize that in a stable equilibrium the denominator is positive.

Now, recall that $\theta^*(q) = c^*(q) - \sigma\Phi^{-1}(q) = c^*(q) - \sigma z$. Thus,

$$\frac{\partial \theta^*(q)}{\partial q} = \left(\frac{\partial c^*(z)}{\partial z} - \sigma\right)\frac{\partial z}{\partial q} > 0 \;\Leftrightarrow\; \frac{\partial c^*(z)}{\partial z} > \sigma,$$

which, using equation (9) yields:

$$\frac{b_1}{\sqrt{\beta}}\gamma\phi(\gamma)\frac{1}{\sqrt{\beta}} + z\frac{b_1}{\sqrt{\beta}}\phi(\gamma) > \frac{b_1}{\sqrt{\beta}}\gamma\phi(\gamma)\frac{1-\beta}{\sqrt{\beta}} + \sigma\phi(z).$$

I.e.,

$$b_1\gamma\phi(\gamma) + z\frac{b_1}{\sqrt{\beta}}\phi(\gamma) > \sigma\phi(z) \quad\Leftrightarrow\quad \frac{b_1}{\sqrt{\beta}}\frac{\phi(\gamma)}{\phi(z)}\,(\sqrt{\beta}\,\gamma + z) > \sigma.$$

Substituting from (8) yields:

$$c^*(\sqrt{\beta}\,\gamma + z) > \sigma.$$

Substituting from the definition of γ yields:

$$c^*\left(\frac{(1-\beta)c^*}{\sigma} - z + z\right) > \sigma \quad\Leftrightarrow\quad (c^*)^2 > \frac{\sigma^2}{1-\beta}.$$

Thus, recalling the definition of β, we have:

$$\frac{\partial \theta^*(q)}{\partial q} > 0 \;\Leftrightarrow\; c^* > \sqrt{\sigma^2 + \sigma_0^2} \;\Leftrightarrow\; \theta^* > \sqrt{\sigma^2 + \sigma_0^2} - \sigma\;\Phi^{-1}(q).$$

Letting $\bar{\theta} = \sqrt{\sigma^2 + \sigma_0^2} - \sigma\;\Phi^{-1}(q)$ completes the proof. $\square$

References

Andreoni, James. 1998. "Toward a Theory of Charitable Fund-Raising." *Journal of Political Economy* 106: 1186-1213.

Börgers, Tilman. 2004. "Costly Voting." *American Economic Review* 94: 57-66.

Boix, Carles, and Milan Svolik. 2013. "The Foundations of Limited Authoritarian Government: Institutions and Power-Sharing in Dictatorships." *Journal of Politics* 75: 300-16.

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

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

Cantoni, Davide, David Yang, Noam Yuchtman, and Jane Zhang. 2019. "Protests as Strategic Games: Experimental Evidence from Hong Kong's Democracy Movement." *Quarterly Journal of Economics* 134: 1021-77.

Casper, Brett, and Scott Tyson. 2014. "Popular Protest and Elite Coordination in a Coup d'etat." *Journal of Politics* 76: 548-64.

Chamberlain, Gary, and Michael Rothschild. 1981. "A Note on the Probability of Casting a Decisive Vote." *Journal of Economic Theory* 25: 152-62.

Chen, Heng, and Wing Suen. 2017. "Aspiring for Change: A Theory of Middle Class Activism." *The Economic Journal* 127: 1318-47.

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

Dewan, Torun, and David 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. "Incumbency Advantage in Non-democracies."

Available at http://www.nber.org/papers/w20519.

Frey, Bruno, and Stephan Meier. 2004. "Social Comparisons and Pro-Social Behavior: Testing 'Conditional Cooperation' in a Field Experiment." *American Economic Review* 94: 1717-22.

Gillies, Donald. 1987. "Was Bayes a Bayesian?" *Historia Mathematica* 14: 325-46.

Good, Irving, and Lawrence Mayer. 1975. "Estimating the Efficacy of a Vote." *Behavioral Science* 20: 25-33.

Guriev, Sergei, and Daniel Treisman. 2015. "How Modern Dictators Survive: An Informational Theory of the New Authoritarianism." Available at http://www.nber.org/papers/w21136.

Gurr, Ted. 1971. *Why Men Rebel?* Princeton, NJ: Princeton University Press

Hager, Anselm, Hensel, Lukas, Johannes Hermle, and Christopher Roth. 2019a. "Strategic Interdependence in Political Movements and Countermovements." Mimeo.

Hager, Anselm, Hensel, Lukas, Johannes Hermle, and Christopher Roth. 2019b. "Political Activists as Free-Riders: Evidence from a Natural Field Experiment." Mimeo.

Hummel, Patrick. 2012. "Sequential Voting in Large Elections with Multiple Candidates." *Journal of Public Economics* 96: 341-8.

Morris, Stephen, and Mehdi Shadmehr. 2018. "Inspiring Regime Change." Available at https://ssrn.com/abstract=2987073.

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 Dewatripont, Hansen, and Turnovsky. Cambridge University Press.

Myatt, David. 2007. "On the Theory of Strategic Voting." *Review of Economic Studies* 74: 255-81.

Myatt, David. 2015. "A Theory of Voter Turnout." Available at http://dpmyatt.org/uploads/turnout-2015.pdf.

Myatt, David. 2017. "A Theory of Protest Voting." *Economic Journal* 127: 1527-67.

Olson, Mancur. 1965. T*he Logic of Collective Action: Public Goods and the Theory of*

Groups. Cambridge, MA: Harvard University Press.

Palfrey, Thomas, and Howard Rosenthal. 1984. "Participation and the Provision of Discrete Public Goods: A Strategic Analysis." *Journal of Public Economics* 24: 171-93..

Palfrey, Thomas, and Howard Rosenthal. 1985. "Voter Participation and Strategic Uncertainty." *American Political Science Review* 79: 62-78.

Razin, Ronny. 2003. "Signaling and Election Motivations in a Voting Model." *Econometrica* 71: 1083-119.

Sandoz, Ellis. 1998. *Political Sermons of the American Founding Era: 1730-1805. 2nd Ed. Vol 1.* Indianapolis, IN: Liberty Fund.

Shadmehr, Mehdi, and Dan Bernhardt. 2011. "Collective Action with Uncertain Payoffs." *American Political Science Review* 105: 829-51.

Shadmehr, M., and Dan Bernhardt. 2017. "Monotone and Bounded Interval Equilibria in a Coordination Game with Information Aggregation." *Mathematical Social Sciences 89: 61-9.*

Shang, Jen, and Rachel Croson. 2009. "A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods." *Economic Journal* 119: 1422-39.

Tilly, Charles. 1978. *From Mobilization to Revolution.* Addison-Wesley.

Tullock, Gordon. 1971. "The Paradox of Revolution." *Public Choice* 11: 89-99.

Tyson, Scott, and Alastair Smith. 2018. "Dual-Layered Coordination and Political Instability: Repression, Cooptation, and the Role of Information." *Journal of Politics* 18: 44-58.

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

Online Appendix

Uncorrelated Signals

The difficulties in ensuring the existence of monotone equilibria in our setting raise the question of why we do not consider a setting with uncorrelated private signals, where we know that the best response to a monotone strategy is monotone. One answer is that such a setting is unnatural because costs must reflect some common factor, in which case a citizen's cost realization contains some information about the costs of others. Moreover, as we will see, this setting offers a less natural resolution of the Tullock's Paradox.

Our setting is the same as before except that now, $c_i \sim iid\ F$, where $F(\cdot)$ has full support on $\mathbb{R}$. The best response to a monotone strategy is clearly monotone: Higher c_i only reduces i's incentive to revolt without changing his beliefs about others' behavior. The equilibria are characterized by the indifference condition:

$$u(N)\binom{N}{qN}\ F(c^*)^{qN}\ (1-F(c^*))^{(1-q)N} = c^*. \tag{10}$$

It is beneficial to do a change of variables $z^* = F(c^*)$, so that (10) becomes:

$$u(N)\binom{N}{qN}\ [z^*]^{qN}\ [1-z^*]^{(1-q)N} = F^{-1}(z^*),\ \text{with}\ \ z^* \in [0,1]. \tag{11}$$

A key simple observation is that as N increases, the maximum of $[z^*]^{qN}\ [1-z^*]^{(1-q)N}$ becomes very sharp, even though the whole expression approaches zero. In fact, using the Stirling approximation, one can identify the rate of convergence as $N \to \infty$:

$$\binom{N}{qN}\ z^{qN}\ (1-z)^{(1-q)N} \approx \frac{1}{\sqrt{\pi N}}\ \frac{1}{\sqrt{2q(1-q)}}\ \left(\frac{z}{q}\right)^{qN}\ \left(\frac{1-z}{1-q}\right)^{(1-q)N}. \tag{12}$$

Because Stirling approximation is close even when N is small, (12) provides a good approximation even for small N. The maximum of the estimated probability of pivotality (the left hand side of the indifference condition (11)), which happens at $z = q$, approaches:

$$\lim_{N\to\infty} \max\left\{\binom{N}{qN}\ z^{qN}\ (1-z)^{(1-q)N}\right\} = \lim_{N\to\infty} \frac{1}{\sqrt{N}}\ \frac{1}{\sqrt{2\pi q(1-q)}}. \tag{13}$$

If $u(N)$ does not depend on N or grows with N at a rate smaller than $N^{1/2}$, then in the limit, there is a unique equilibrium with $\lim_{N\to\infty} c^*(N) = 0$. Moreover, if $u(N)$ grows with N at the rate N, then in the limit, there are three equilibria: There is an equilibrium, in which $\lim_{N\to\infty} c^*(N) = 0$, and there are two equilibria in which $\lim_{N\to\infty} F(c^*(N)) = q$, one from below and one from above. These observations follow from the inspection of equations (12) and (13). Our simulations with Normal distribution suggest that when $\lim_{N\to\infty} c^*(N) = q^+$, the likelihood of success approaches 1, and when $\lim_{N\to\infty} c^*(N) = q^-$, the likelihood of success approaches 0. Next, suppose $u(N)$ increases at the rate $N^{1/2}$. From (12),

$$u(N) = b_0 + b_1\sqrt{N} \quad \Rightarrow \quad \lim_{N\to\infty} u(N) \binom{N}{qN} [z^*]^{qN} \, [1 - z^*]^{(1-q)N} = \begin{cases} \frac{b_1}{\sqrt{2\pi q(1-q)}} & ; z^* = q \\ 0 & ; z^* \neq q \end{cases}$$

Figure 5 illustrates the left and right hand side of equation (10), which characterizes the equilibrium, for a few cases of N, when $q = 0.75$, and F is the Normal distribution. Clearly, as long as N is moderately large, there is always an equilibrium with $c^* \approx 0$. In addition, because $u(N) \binom{N}{qN} [z^*]^{qN} \, [1-z^*]^{(1-q)N}$ remains single-peaked, when σ is not too large, there are multiple equilibria, in which $F(c^*(N))$ approaches q as N grows. When σ is larger, for large N, any equilibrium with $c^* > 0$ disappears, and we are left with $c^* \approx 0$. Dzuida et al. (2020) provide a characterization of how the likelihood of success varies with q for a given N and u.[3]

[3] Dziuda, Wioletta, Arda Gitmez, and Mehdi Shadmehr. 2020. "The Difficulty of Easy Projects." *Forthcoming, American Economic Review: Insights.*

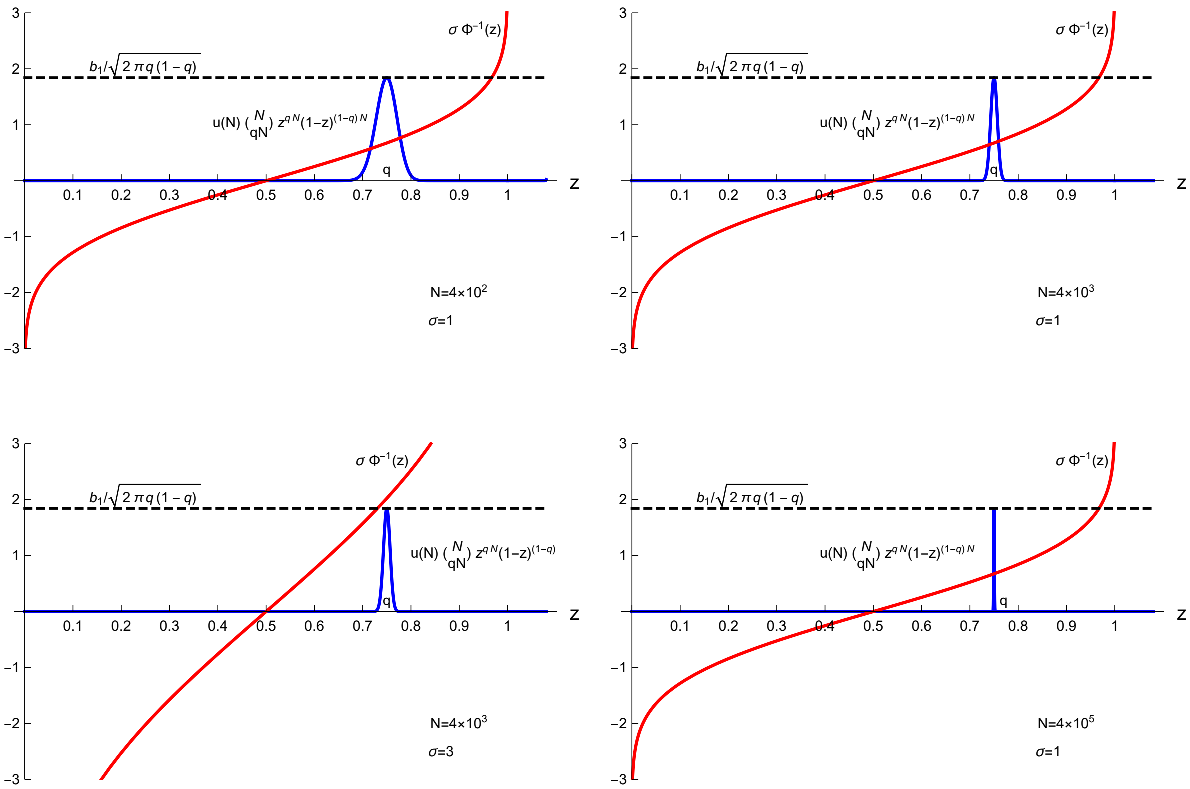

Figure 5: The unidomal curve is the left hand side of the indifference equation (10), and the increasing cure is its right hand side. The dashed line is (13). Parameters: $c_i \sim N(0, \sigma)$, $q = 0.75$, $b_1 = 2$, $b_0 = 0$, N and σ are shown on the graph.

Alternative Models of Revolution

Another class of games used in the literature on revolutions contains uncertainty about the revolution payoff that is received when there is a regime change (Bueno de Mesquita 2010; Shadmehr and Bernhardt 2011).

Common Value Payoffs. Consider the game in Figure 6 with a continuum of players, indexed by $i \in [0,1]$. The revolution succeeds whenever the fraction of revolters exceeds a threshold $q \in (0,1)$. The status quo payoff is 0. If the revolution succeeds, everyone gets θ, and those who participated in a successful revolution, get an additional $\alpha\theta$, with $\alpha \in (0,1)$. As before, a citizen i receives private signals $x_i = \theta + \sigma\ \epsilon_i$, where θ and ϵ_is are independent. Citizens share an improper prior that θ is distributed uniformly on $\mathbb{R}$, and $\epsilon_i \sim F$ with full support on $\mathbb{R}$. There is always an equilibrium in which no one revolts. We focus on finite-cutoff strategies, where i revolts if and only if $x_i > x^*$. Then, the regime collapses if and only if $\theta > \theta^*$, where

$$Pr(x_i > x^*|\theta^*) = 1 - F\left(\frac{x^* - \theta^*}{\sigma}\right) = q, \text{ so that } x^* = \theta^* + \sigma\ F^{-1}(1-q). \qquad (14)$$

		outcome	
		$n > q$	$n \leq q$
citizen i	*revolt*	$(1+\alpha)\theta - c$	$-c$
	no revolt	θ	0

Figure 6: A common value version of the revolution model of Bueno de Mesquita (2010).

The indifference condition is:

$$\begin{aligned}
\frac{c}{\alpha} &= Pr(\theta > \theta^* | x_i = x^*) E[\theta | x_i = x^*, \theta > \theta^*] \\
&= \int_{\theta^*}^{\infty} \theta \; pdf(\theta | x^*) \; d\theta \\
&= \int_{\theta^*}^{\infty} \theta \; \frac{1}{\sigma} \; f\left(\frac{x^* - \theta}{\sigma}\right) \; d\theta \quad \text{(because the prior is uniform)} \\
&= \int_{-\infty}^{z^* \equiv z(\theta = \theta^*)} (x^* - \sigma \; z) \; f(z) \; dz, \quad z = \frac{x^* - \theta}{\sigma} \\
&= \int_{-\infty}^{F^{-1}(1-q)} (x^* - \sigma \; z) \; f(z) \; dz \quad \text{(from equation (14))} \\
&= x^* \; F(F^{-1}(1-q)) - \sigma \; F(F^{-1}(1-q)) \; E[\epsilon_i | \epsilon_i < F^{-1}(1-q)] \\
&= (1-q) \; (x^* - \sigma \; E[\epsilon_i | \epsilon_i < F^{-1}(1-q)]).
\end{aligned}$$

Thus,

$$\begin{aligned}
x^* &= \frac{c}{\alpha} \frac{1}{(1-q)} + \sigma \; E[\epsilon_i | \epsilon_i < F^{-1}(1-q)]. \\
\theta^* &= \frac{c}{\alpha} \frac{1}{(1-q)} + \sigma \; \{E[\epsilon_i | \epsilon_i < F^{-1}(1-q)] - F^{-1}(1-q)\}.
\end{aligned} \tag{15}$$

The term $\sigma \; E[\epsilon_i | \epsilon_i < F^{-1}(1-q)]$ is decreasing in q, indicating a force that increases the citizens' incentives to revolt when the regime is stronger. This force stems from learning-in-equilibrium incentives generated by common value payoffs: When the regime becomes stronger so that citizens become more hesitant to revolt, the information content of their actions is a better news of θ, and hence the expected revolution payoff conditional on regime change is higher. However, when F is logconcave (An 1998, p. 357), the curly bracket in θ^* is increasing in q.[4] Thus, as the regime becomes stronger (q increases), θ^* increases. The analysis is far simpler in a private value setting, where a citizen's payoff is his signal x_i rather

[4] An, Mark Yuying. 1998. "Logconcavity versus Logconvexity: A Complete Characterization." *Journal of Economic Theory* 80: 350-69.

than the uncertain fundamental θ. Then, the indifference condition is:

$$\begin{aligned} c &= Pr(\theta > \theta^*|x_i = x^*)\ \alpha\ x^* \\ &= [1 - Pr(x_i > x^*|\theta^*)]\ \alpha\ x^* \\ &= (1-q)\ \alpha\ x^*. \qquad\qquad \text{(from equation (14))} \end{aligned}$$

Thus,

$$x^* = \frac{c}{\alpha}\frac{1}{1-q} \ \text{ and } \ \theta^* = x^* - \sigma F^{-1}(1-q) = \frac{c}{\alpha}\frac{1}{1-q} - \sigma F^{-1}(1-q) \tag{16}$$

Proposition 1 *The equilibria in finite-cutoff strategies are characterized by (x^*, θ^*), so that a citizen revolts whenever his signal is above x^* and the regime collapses whenever $\theta > \theta^*$. When the prior is uniform or the noise in private signals approaches zero, the equilibrium is unique and is given by (15) for the common value and by (16) for the private value setting. In both settings, as the regime becomes stronger (q increases), the revolution is less likely.*[5]

Pivotality. Now, consider the setting with $N+1$ players, which features the logic of pivotality. We show that for large N, with strictly unimodal distributions like Normal, the best response to a finite-cutoff strategy is not a finite-cutoff strategy. Because revolting is costly, i only revolts if he is pivotal, i.e., only if the number of revolters is qN. Thus, i's net expected payoff from revolting versus not revolting is:

$$Pr(piv|x_i)\ E[u(\theta, N)|x_i, piv] - c = \int_{\theta=-\infty}^{\infty} Pr(piv|\theta)\ u(\theta, N)\ pdf(\theta|x_i)\ d\theta - c,$$

where piv denotes the event of i being pivotal. We focus on symmetric monotone strategies, so that a citizen revolts if and only of his signal exceeds a threshold: $x_i > x^*$. If the best response to a monotone strategy was also a monotone strategy, then the equilibrium would be characterized by the indifference condition of the marginal player whose signal is the exact cutoff:

$$Pr(piv|x_i = x^*) \cdot E[u(\theta, N)|piv, x_i = x^*] = \int_{\theta=-\infty}^{\infty} Pr(piv|\theta)\ u(\theta, N)\ pdf(\theta|x_i = x^*)\ d\theta = c.$$

[5]In the common value setting, when the prior is uniform, but the noise is not vanishingly small, we also require that F be logconcave as a sufficient condition.

$$Pr(piv|\theta) = \binom{N}{qN}\left[1-F\left(\frac{x^*-\theta}{\sigma}\right)\right]^{qN}\left[F\left(\frac{x^*-\theta}{\sigma}\right)\right]^{(1-q)N}.$$

Thus, focusing on $u(\theta, N) = (b_0 + b_1 N)\theta$ to match the standard games of the literature, the indifference condition that characterizes the equilibrium cutoffs is:

$$\begin{aligned}
c &= \int_{\theta=-\infty}^{\infty} u(\theta, N)\binom{N}{qN}\left[1-F\left(\frac{x^*-\theta}{\sigma}\right)\right]^{qN}\left[F\left(\frac{x^*-\theta}{\sigma}\right)\right]^{(1-q)N}\frac{1}{\sigma}f\left(\frac{x^*-\theta}{\sigma}\right)d\theta \\
&= \int_{z=0}^{1} u(x^*-\sigma\ F^{-1}(1-z), N)\binom{N}{qN} z^{qN}(1-z)^{(1-q)N}\,dz \\
&= (b_0+b_1N)\int_{z=0}^{1}(x^*-\sigma\ F^{-1}(1-z))\binom{N}{qN} z^{qN}(1-z)^{(1-q)N}\,dz \\
&\overbrace{=}^{large\ N} b_1\ (x^*-\sigma\ F^{-1}(1-q)) \quad \text{(from Chamberlain and Rothschild (1981))}.
\end{aligned}$$

Thus,

$$x^* = \frac{c}{b_1} + \sigma\ F^{-1}(1-q).$$

Ignoring the directs costs of revolting, the net expected payoffs from revolting versus not revolting for a citizen i with signal x_i is:

$$\begin{aligned}
& \int_{\theta=-\infty}^{\infty} u(\theta, N)\binom{N}{qN}\left[1-F\left(\frac{x^*-\theta}{\sigma}\right)\right]^{qN}\left[F\left(\frac{x^*-\theta}{\sigma}\right)\right]^{(1-q)N}\frac{1}{\sigma}f\left(\frac{x_i-\theta}{\sigma}\right)d\theta \\
= & \int_{z=0}^{1} u(x^*-\sigma\ F^{-1}(1-z), N)\binom{N}{qN} z^{qN}(1-z)^{(1-q)N}\frac{f\left(\frac{x_i-x^*}{\sigma}+F^{-1}(1-z)\right)}{f(F^{-1}(1-z))}\,dz \\
= & (b_0+b_1N)\int_{z=0}^{1}(x^*-\sigma\ F^{-1}(1-z))\binom{N}{qN} z^{qN}(1-z)^{(1-q)N}\frac{f\left(\frac{x_i-x^*}{\sigma}+F^{-1}(1-z)\right)}{f(F^{-1}(1-z))}\,dz \\
\overbrace{=}^{large\ N} & b_1\ (x^*-\sigma\ F^{-1}(1-q))\ \frac{f\left(\frac{x_i-x^*}{\sigma}+F^{-1}(1-q)\right)}{f(F^{-1}(1-q))}. \\
\overbrace{=}^{in\ equilibrium} & c\ \frac{f\left(\frac{x_i-c/b_1}{\sigma}\right)}{f(F^{-1}(1-q))}.
\end{aligned}$$

In sum, we have established that if the best response to a cutoff strategy is indeed a cutoff strategy, then there is a unique equilibrium with x^* given above. Now, given this x^* that characterizes the strategies of other citizens, the net expected payoff from revolting versus not revolting for a citizen i with signal x_i is:

$$c \times \left(\frac{f\left(\frac{x_i - c/\alpha}{\sigma}\right)}{f(F^{-1}(1-q))} - 1 \right),$$

implying that i revolts if and only if

$$f\left(\frac{x_i - c/\alpha}{\sigma}\right) > f(F^{-1}(1-q)).$$

When f is strictly unimodal (e.g., Normal distribution), this expression does *not* have a single-crossing property: Either there is no crossing and i never revolts, or it has two crossings and i's best response is non-monotone.

Private Value Payoffs. Now, consider a private value payoff structure, so that a citizen with signal x_i receives $u(x_i, N)$. Then, mirroring the calculations for the common value case, we have:

$$x^* = \frac{c}{b_1} \quad \text{and} \quad \theta^* = x^* - \sigma\ F^{-1}(1-q) = \frac{c}{b_1} - \sigma\ F^{-1}(1-q), \tag{17}$$

where we recognize that, similar to our setting in the text, the fraction of citizens who participate in a revolution does not change with the regime's strength. Again, mirroring the calculations for the common value case, we have:

$$
\begin{aligned}
B(x_i;x^*) &= \int_{\theta=-\infty}^{\infty} u(x_i,N)\binom{N}{qN}\left[1-F\left(\frac{x^*-\theta}{\sigma}\right)\right]^{qN}\left[F\left(\frac{x^*-\theta}{\sigma}\right)\right]^{(1-q)N}\ \frac{1}{\sigma}f\left(\frac{x_i-\theta}{\sigma}\right)d\theta \\
&= \int_{z=0}^{1} u(x_i,N)\ \binom{N}{qN}\ z^{qN}\ (1-z)^{(1-q)N}\ \frac{f\left(\frac{x_i-x^*}{\sigma}+F^{-1}(1-z)\right)}{f(F^{-1}(1-z))}\ dz \\
&= (b_0+b_1N)\ \int_{z=0}^{1} x_i\ \binom{N}{qN}\ z^{qN}\ (1-z)^{(1-q)N}\ \frac{f\left(\frac{x_i-x^*}{\sigma}+F^{-1}(1-z)\right)}{f(F^{-1}(1-z))}\ dz \\
&\overbrace{=}^{large\ N}\ b_1\ x_i\ \frac{f\left(\frac{x_i-x^*}{\sigma}+F^{-1}(1-q)\right)}{f(F^{-1}(1-q))}. \\
&= b_1\ x_i\ \frac{f\left(\frac{x_i-c/b_1}{\sigma}+F^{-1}(1-q)\right)}{f(F^{-1}(1-q))}\quad \text{(in equilibrium, from (17))}.
\end{aligned}
$$

Recall that given other citizens' cutoff strategy with associated cutoff x^*, citizen i with signal x_i revolts if and only if $B(x_i;x^*) > c$. Next, observe that $B(0;x^*) = \lim_{x_i\to\infty} B(x_i;x^*) = 0$. Thus, the best response to a finite-cutoff strategy is not a finite-cutoff strategy.