Optimal Size of Rebellions: Trade-off between Large Group and Maintaining Secrecy *

Congyi Zhou†

September 25, 2015

Abstract

This paper studies a regime change model in which, when a rebel leader mobilizes supporters, he faces a trade-off between increase the size of the rebel group and the risk of information leaks. I find that when the authority implement collective punishment to repress the rebellion, in which both rebel participants and those who knew about but did not report the rebellion are punished, it may result in a smaller rebel group size compared to targeted punishment, in which only the rebel participants are held accountable for their actions. Meanwhile, choosing collective punishment comes at a price, by forcing many to side with the insurgency, which may decrease the authority's survival chances. My findings also indicate that targeted punishment is more useful to prevent a revolution by ordinary citizens, while collective punishment should be adopted to prevent a coup staged by the politicians. Furthermore, when both authorities and rebel leaders compete for support by threatening retribution against the non-supporters, both parties tend to prefer the use of harsh methods to force civilians to choose sides.

Key words: Regime change, Coordination game, Rebellion, Collective Punishment, Targeted Punishment, Coalition size, Survival probability

1 Introduction

Coordination games have been widely used to study collective decisions under individuals' uncertainty about other players' private information. This model type is established based on an

*I am grateful to Scott Ashworth, Ethan Bueno de Mesquita, for the extremely helpful advice they gave me. I would also like to thank Stephen Ansolabehere, Chris Berry, Dan Black, Alessandra Casella, Stephen Chaudoin, Oeindrila Dube, Anthony Fowler, Scott Gehlbach, Olle Folke, William Howell, Siyuan Liu, Zachary Peskowitz, Pablo Montagnes, Salvatore Nunnari, Giacomo Ponzetto, Robert Powell, Kris Ramsey, Mehdi Shadmehr, Milan Svolik, Dustin Tingley, Richard Van Weelden, Razvan Vlaicu, Stephane Wolton, David Yanagizawa-Drott for helpful comments and workshop and conference participants at University of Chicago Political Economy Lunch, Chicago Economics Theory working group and MPSA annual conference. Previously, this article was circulated under name "Regime Change and the Optimal Size of Illegal Coalitions"

†Harris School, the University of Chicago, Email: czhou@uchicago.edu

underlying assumption that the more players take the same action, the easier it is for that action to succeed.[1] The regime change model is one of its applications. The model focuses on what conditions or mechanisms can the insurgency motivate more participants (Bueno de Mesquita (2010)).[2] However, a rebel group always faces a dilemma during the process of achieving a successful insurgency. On the one hand, rebel leaders want to motivate as many supporters as possible to participate in rebellions. On the other hand, maintaining confidentiality prior to taking action is crucial to success. As the number of people involved in an action increases, the probability of information leak also increases. A famous example in China is the failure of the Wu Hsu Reform in 1898, as the revolutionary party members were betrayed by Yuan Shikai, a potential supporter of a coup.

As such, I present a model to study the trade-off between increasing the number of supporters and the growing risk of information leaks faced by rebel leaders. This article advances the literature on the coordination game by allowing rebel leaders to choose their coalition size endogenously. This leeway makes it different from existing models in the sense that the players make decisions by inferring other players' actions according to the signals that each individual receives. In the current model, the individuals need to consider not only the signals they receive but also how the coalition size can affect other players' actions. The model provides a way to study the optimal coalition size that the rebel leaders should form under different environments.

The basic model can be used to answer another natural question in the regime change model, namely, how to prevent an insurgency in advance from the government's perspective. In this article, I focus on comparing two widely adopted ex ante punishment rules: *collective punishment* and *targeted punishment*. Collective punishment is favored because it takes advantage of people's fears to instill mistrust in the target population and encourages people to monitor each other's behaviour, which helps the authorities gather more information to prevent potential threats. Nevertheless, this harsh punishment has been gradually abandoned and replaced by targeted punishment, under which only those who participate in illegal actions are punished. Such abandonment is due to the idea that contemporary society upholds the idea of human rights. However, the effectiveness of these two punishments varies (in Section 2, I discuss historic and political evidence). In this article, I calculate the optimal rebel coalition size under these two punishment rules and identify the conditions under which punishment rule is more likely to prevent insurgencies.

In the model, a rebel leader wants to mobilize civilians to join a rebellion. Each player has a private type that represents individual anti-government sentiments. Given that the players dislike the government, they suspect that other players do too, although they are uncertain of their peers'

[1]In a 2×2 game, two players need to take the same action to achieve a goal (Carlsson and Van Damme (1993)). In a game with an infinite number of players, a threshold exists such that only a certain number of players need to take one action to achieve the goal. This condition is true in such models as the pricing debt model (Morris and Shin (2004)) and the currency crisis model (Morris and Shin (1998))

[2]Other examples: Boix and Svolik (2013); Edmond (2011); Egorov et al. (2009); Persson and Tabellini (2009); Tyson and Smith (2013), and Angeletos et al. (2007) and Little (2014)addresses a dynamic global game in regime change.

exact views. Once the leader decides to stage a rebellion, he can endogenously choose the coalition size by sharing his private type with coalition members.[3] Each member has three choices: i) to participate in the rebellion, ii) to maintain neutrality (i.e., be a free rider), or iii) to alert the government. The rebellion succeeds if and only if the number of participants is sufficiently large and the number of people who alert the government is sufficiently small. This model assumes two incentives that drive the civilians to desire a regime change. The first one is the regime change incentive, which measures how much an individual values the regime change through the personal anti-government types. The second, the pecuniary incentive, represents a direct financial benefit.

Two interesting results are obtained. First, under collective punishment, the optimal coalition size is small and fixed, and it is independent from the rebel leader's type. Intuitively, the rebel leader wants to maximize the coalition size to guarantee the participation rate within the limits of potential betrayal. Moreover, no one wants to remain neutral under collective punishment, because free riders share the same risk as participants, but obtain fewer rewards if the rebellion succeeds. Owing to the lack of free riders, the optimal rebel group size should be fixed to satisfy the successful conditions of a rebellion. Conversely, a free riding problem may exist under targeted punishment because some players wish to remain neutral to avoid taking any risk regardless of the outcome. Given the free riding problem, a weak leader will find it hard to motivate others. Consequently, he needs to recruit more people to obtain a more sufficient number of participants than that under collective punishment. However, when the leader is strong enough to convince others that the rebellion has a high chance of success, the free riding problem can be eliminated and the optimal coalition size can remain as small as that under collective punishment.

Second, from the government's perspective, choosing collective punishment comes at a cost because forcing neutral citizens to pick a side might negatively affect the regime's survival probability. When the incentives are considerable, free riding motivates citizens to remain neutral to avoid risk. By choosing targeted punishment and thereby tolerating the existence of free riders, the authorities can reduce the number of people joining the rebellion. However, the collective punishment rule forces free riders to choose sides, potentially pushing more neutrals to side against authority because of the considerable incentives that the success of the rebellion offers. On the contrary, when the incentives to be a free rider are low, pushing free riders to choose sides by collective punishment may bring more people on the government's side, increasing the survival probability of the regime.

This framework provides valuable insights into the endogenous institution choice problem. When an authority faces a rebellion from civilians who are primarily driven by anti-government sentiment rather than any desire for pecuniary reward, targeted punishment may be effective in quelling the rebellion. In this case, being a free rider might provide a citizen with a regime change

[3] In the collective action literature, a common theme is sending public signals to motivate people to join the actions, such as Angeletos et al. (2006); Bueno de Mesquita (2010); Egorov et al. (2009); Chwe (2000, 2013)). Baliga and Sjostrom (2012), Bueno de Mesquita (2010) and Ginkel and Smith (1999), allow the signal can be manipulated which is different with this article.

incentive if the rebellion succeeds, and could enable the citizen to avoid risk if it fails. Therefore, using targeted punishment prompts citizens to become free riders and lowers the rebellion's success probability.[4] However, when a ruler faces a threat from the people who are driven primarily by direct pecuniary reward or by their ambitions for power, collective punishment is more effective than that of targeted punishment.[5] Intuitively, using targeted punishment to encourage free riders does not work well in this scenario because free riders can only obtain regime change incentives, which hold less importance in this case.[6]

Another extension of the model is the strategic competition between rebel leaders and the government. Assuming the leader can also threaten to punish the free riders when the rebellion succeeds, this harsh threat can eliminate free riders even under targeted punishment. In particular, when the authority prefers targeted punishment to encourage free riders, the rebel leader will use the harsh threat to push free riders into joining the rebellion. Therefore, the government must switch to collective punishment to defeat the rebellion. This explanation demonstrates why an equivalent retaliation strategy between governments and rebel groups is frequently observed.[7]

Finally, I consider a special case in which a rebel leader threatens to punish traitors severely once the rebellion succeeds. In this case, collective punishment also results in a free rider problem because this severe threat to punish traitors increases citizens incentives to be free riders to avoid punishment from the rebel leader. This effect offsets the effect of eliminating free riders caused by the use of collective punishment. Therefore, for the leader, threatening traitors with severe punishment is more effective in reducing betrayal than recruiting participants.

The article proceeds as follows. Section 2 discusses the historical and political evidence. Section 3 indicates the model setup and solve the equilibria under collective punishment. Section 4 solves the equilibria under targeted punishment. Section 5 compares the regime's survival probabilities under two punishment rules. Section 6 is the comparative static analysis and the endogenous institution choice problem. Section 7 considers the strategic competition between rebel leaders and authorities. Section 8 studies the case that rebel leaders prefer using severe threat to punish traitors to avoid betrayal, and Section 9 is the conclusion.

---

[4]For example, in the Jasmine Revolution in Tunisia and Lotus Revolution in Egypt, young people were motivated by anger over corruption and dictatorship. The direct monetary reward is relatively small for individual participants.

[5]This situation is observed more often in coups as exemplified by Sophia Alekseyevna who led the rebellion of the Streltsy to seize the highest power of Russia in 1682 (Hughes (1985)).

[6]Some works under the regime-change topic emphasize the risk for coups to fail, and treat the military as a unitary actor (Galetovic and Sanhueza (2000); Svolik (2009); Svolik (2013); Besley and Robinson (2010); Acemoglu et al. (2010)).

[7]One example is the civil war caused by the rebellion of the Liberation Tigers of Tamil Eelam (LTTE) in Sri Lanka. During the war, both LTTE and Sri Lankan military committed war crimes, including attacks on civilians and civilian buildings. Equivalent retaliation between the governments and rebel groups derived in this article is not the unique reason for war crimes in many similar situations. However, equivalent retaliation provides a possible explanation for civilians being harmed because of cruel competition when both sides want to obtain support from civilians.

2 Historical and Political Evidence

{sec: eviden

Little empirical evidence is available on the direct measurement of the size of the rebel coalition. In contrast, the effectiveness of collective punishment has gained much attention in the academe. The adoption of this harsh punishment rule is based on the logic that "collective sanctions mobilize groups to monitor and control the conduct of their members" (Levinson (2003)), which is exactly the problem encountered by every rebel group leader in maintaining secrecy.

Academic views on the effectiveness of collective punishment are mixed. Advocates of collective sanctioning strategies claim that collective sanctions not only leverage but also build group solidarity (Levinson (2003)). Arguably, punishments are likely to be most effective when the right actors or interest groups are affected; in many cases however, the optimal targets may be bystanders rather than the perpetrators of the objectionable deeds (Kaempfer and Lowenberg (1988); Major and McGann (2005)). Other scholars take an opposing view and argue that collective punishment is likely to be counterproductive in many settings. For example, in the contexts of counterinsurgency and counterterrorism, tactics involving indiscriminate coercion have backfired because they induced moral outrage (DeNardo (2014)). Generally, collective punishments in a wide variety of contexts may increase solidarity of the targeted group (Galtung (1967); Khawaja (1993)).

Recent studies on the Baojia system in the history of mainland China and Taiwan provide more details that can help link the punishments from the government and the problem of maintaining secrecy among rebel groups. After Japan took Taiwan from China in 1895, the Japanese encountered intense anti-Japanese activities from the Taiwanese. Governor-General Kodama Gentaro (1898 – 1906) began to implement the Baojia system to suppress the uprisings in 1898.[8] The primary purpose of Baojia was to prevent local residents from sheltering criminals and concealing crimes by applying the rule of collective responsibility. The Baojia system successfully suppressed the uprisings of the Taiwanese. By 1902, the police and Baojia network were perfected and support from the people for the guerrillas was completely cut off (Chen (1975)).

The deterrent potential of the system further demonstrated itself in the 1910s. When the revolutionary party overthrew the Qing government in 1911, a number of Taiwanese Liberals sought Chinese revolutionary assistance in overthrowing the Japanese regime. Six major incidents occurred between 1912 and 1915 as a result of the renewed anti-Japanese movement. However, locals reported all of these incidents to the Japanese authority through the Baojia system, and thus, the Japanese were able to prevent four of the incidents from transpiring (Chen (1975)). The largest of these anti-Japanese uprisings was the Tapani incident in 1915. During the early stage of the uprising, the leader, Yu Qingfang, attempted to motivate many participants through local religious activities without any effort to control information leaks. Consequently, several local residents from the Baojia system alerted the Japanese police on the potential uprising. The uprising failed; nevertheless, more than 3200 villagers were massacred(Xue (2005), Katz (2005)).

[8] the Hoko Law and the Regulations Governing the Execution of the Hoko Law were promulgated in August 1989.

Another example indicates the failure of collective punishment. During the latter years of the Republic of China era, Chiang Kai-shek pressed for the reintroduction of the Baojia system to suppress the uprising in Jianxi led by the communist party in 1934. After temporary success, the Baojia system was extended to the entire country, particularly in rural areas. However, this system became one of the instruments that propelled young peasants to join the communist party army after the World War II, which ultimately ended the regime of Chiang Kai-shek (Li and Ran (2005), Xue (2005)).

The difference between the Baojia system in Taiwan and China provides consistent evidence for our theoretical model. In Taiwan, the Japanese authorities used the Baojia system not just to suppress insurgency but also to improve local administration and economic development, such as building infrastructure, retaining sanitation, and preparing for natural disasters. As the economy and social security improved, the ordinary Taiwanese lost the incentive to join the uprising to change the regime. Therefore, when the exogenous shock from China came to Taiwan in the 1910s, collective punishment drew more residents who were against the uprisings to the side of the Japanese authorities. By contrast, the Baojia system implemented by Chiang Kai-shek only brutally suppressed the people. The system contributed nothing to improving the living conditions in rural China, and the years of civil wars only worsened the economic situation of the country. Therefore, given the choice between the government and the rebels, collective punishment forced a number of people to join the communist party's revolution.

3 Collective Punishment

{sec:set up}

3.1 Model Setup

{subsec:basic

Players: There are two types of player: a rebel leader who plans to instigate a rebellion to overthrow the government; and a unit mass of population member, called followers, indexed by $i \in [0,1]$.

Timeline: Stage 1: the leader receives his private type $x_0 = \theta + \epsilon_0$, and each follower receives his/her type $x_i = \theta + \epsilon_i$.

x represents personal anti-government sentiments, which also measures how much an individual values regime change. θ is the common underlying state variable drawn by nature from a normal distribution $N(m_\theta, \sigma_\theta^2)$. This value represents the true situation of the country such as the status of economic development or human rights. The larger θ is, the worse the situation. ϵ is an individual idiosyncratic variable drawn from a normal distribution $N(0, \sigma_\epsilon^2)$. The players only observe x, not θ or ϵ, but the distributions are known by all players.

Stage 2: The leader chooses the coalition size $L \in [0,1]$ ($L = 0$ represents that no rebellion is staged). Then L members are randomly picked from $[0,1]$ to form rebel coalition.

	Join	Remain neutral	Turn in the leader
Succeed	$Tx_i + t$	vx_i	$sx_i - q + I$
Fail	$-b$	$-b$	I

Table 1: The follower's payoff under collective punishment {table:payoff

Stage 3: Each coalition member i learns the leader's type x_0 and coalition size L. Subsequently, i decides from three options: join the rebellion, remain neutral or turn in the leader to the government.

Stage 4: The rebellion occurs and payoffs are realized.

A rebellion can succeed if and only if two conditions are satisfied:

Participation condition: $L_c > C$;

Maintaining secrecy condition: $L_s < E$.

L_c and L_s denote the number of players who join the rebellion and turn in the leader respectively. C and E are constants, with $1 > C > E > 0$ and $E + C \leq 1$. The participation condition indicates that a successful rebellion requires sufficient supporters (greater than C) to join the action. The maintaining secrecy condition implies that there cannot be too many the players (less than E) who turn in the leader.

Payoffs: When a rebellion succeeds, the leader receives a lump sum payment R, which represents the value of the entire country. When a rebellion fails, the payoff to the leader is $-b$ as punishment from the government, with $b > 0$ (i.e., imprisonment, exile, or execution).

The follower's payoffs are summarized in Table 1. When a rebellion succeeds, the payoff to a participant is $Tx_i + t$. Tx_i is the regime change reward representing how many he/she values the regime change, where T is the multiplier. t is the lump sum pecuniary reward for rebellion participants. If a rebellion fails, the payoff to a participant is $-b$.

The payoff to a neutral follower is vx_i when a rebellion succeeds. This payoff indicates that a free rider can still benefit from the regime change but not with a direct pecuniary reward.[9] When a rebellion fails, the payoff to a neutral follower is $-b$

The player who turns in the leader receives the lump sum reward I from the government whether a rebellion succeeds or not, given that betrayal always occurs before the rebel group takes action. If a rebellion fails, a traitor keeps the reward I. If a rebellion succeeds, the payoff to a traitor is $sx_i - q + I$, where sx_i is the payoff from the regime change and $-q$ as the punishment from the rebel leader, with $q > I$.

I assume the multipliers have $T > v > s > 0$, which means that for a given x_i, the regime change benefits from directly involving the rebellion, indirectly involving, and siding the government follow

[9]Followers who are not in the rebel coalition can also receive vx_i if a rebellion succeeds. However, they are ignored because they cannot take any action during the game.

a decrease order. Meanwhile I also assume that $\frac{t}{T-v} \geq \frac{q-I}{v-s}$.[10] This assumption means that the leader who attracts supporters relies primarily on their desire for reward not on their fear of punishment. The case $\frac{t}{T-v} < \frac{q-I}{v-s}$ is discussed in Section 8.

3.2 Discussion

{subsec:discu

Before proceeding with the analysis, it is necessary to discuss some assumptions in the model. When the leader forms the coalition, I assume he randomly chooses L member from $[0, 1]$ without targeting specific individuals. In reality, although revolutionary vanguards may start to mobilize those they trust the most, they still cannot ascertain people's credibility. When people face choices involving high risk actions, such as revolutions or coups, betrayal can happen among even a close-knit group. This assumption in the model can be interpreted as meaning that the population in $[0, 1]$ is the leader's most trusted group, such as the residents of their home town or their tribe, however, the leader cannot further distinguish the credibility of individual coalition members, which is related to their types but cannot be observe by the leader directly.

The model assumes the leader shares his true type with the coalition members without any cheap talk or signal manipulation (Baliga and Sjostrom (2012),Bueno de Mesquita (2010)). This is a technical assumption, because this article focuses on the scale of the rebel group, rather than the communication strategy. The truth-telling assumption represents the key concept that the larger the leader's type, the easier it is to mobilize a large group of followers. In reality, even if a rebel leader provides misleading information to motivate civilians, his communication skills, personality, and other factors can affect the efficacy of communications. Therefore, when the players receive a noisy signal from the leader, as long as that signal is distributed as a mean preserving spread from the leader's type, it does not change the main purpose of this paper, but increases the computational complexity.

One condition for a successful rebellion is $L_s < E$, with $E > 0$. E can be viewed as the monitoring threshold chosen by the government exogenously. This assumption can be shown in the statement that the government only begins to suppress a rebellion if sufficient evidence is collected. It reflects the facts that rumors about revolutions or coups always exist in any regime. Authorities cannot afford to take action with every single rumor, instead they set up a threshold of quantity of evidence, beyond which they begin serious investigations and suppression activities. The details of choosing a value for E are discussed in Section 6.

This model is different with global game, because it includes the individual type into the utility function, which violates the two-sided limit dominance property(Carlsson and Van Damme (1993),

[10]This assumption indicates that when the rebellion succeeds, the ratio of the monetary increment to the regime change reward increment is higher when a follower switches from a free rider to a participant($\frac{t}{T-v}$) than the ratio when he/she switches from turning in the leader to a free rider ($\frac{q-I}{v-s}$).

Morris and Shin (2003)).[11] The model in this paper develops the idea that no matter how anti-government a citizen, it is never a dominant strategy to rebel. This is because, if a citizen believes no one else will attack the regime, he is certain the regime will not fall, so he does not want to rebel (Bueno de Mesquita (2010), Baliga and Sjostrom (2012)).[12]

3.3 Beliefs and Equilibrium Concept:

{subsec:belie

Applying Bayes' rule for the case of normal prior and normal signals, the leader, after observing his own signal, acquires a posterior belief about θ that is distributed normally with mean $m_0 = \lambda x_0 + (1-\lambda)m_\theta$ and variance $\sigma^2 = \lambda\sigma_\epsilon^2$, where $\lambda = \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\epsilon^2}$. Intuitively, the more anti-government a leader is, the more anti-government he believes others are likely to be. Similarly, i, after observing his/her own type and the leader's type x_0, acquires a posterior belief about θ that is distributed normally with mean $\bar{m}_i = \psi x_0 + (1-\psi)\lambda x_i + (1-\psi)(1-\lambda)m_\theta$ and variance $\bar{\sigma}^2 = \psi\lambda\sigma_\epsilon^2$, where $\psi = \frac{\lambda}{1+\lambda}$.

A pure strategy for the leader is a mapping $L(x_0) : \mathbb{R} \to [0,1]$; from the personal type into a decision of the size of rebellion coalition L, in which $L = 0$ represents the case that the leader does not stage a rebellion. A pure strategy for follower $i \in G$ is a mapping $s(x_i, x_0, L) : \mathbb{R} \times \mathbb{R} \times (0,1] \to \{-1, 0, 1\}$; from the personal type, the leader's type and the coalition size into a decision of turning in the leader (-1), remaining neutral (0) or joining a rebellion (1). The solution concept is a pure strategy Perfect Bayesian Equilibrium (PBE), and I focus on those pure strategies PBE of the full game in which the followers follow cutoff rules.

The first result in the following proposition indicates that when a follower believes no one else will join a rebellion, the best strategy for his/her is to turn in a leader always. During this scenario, a leader will not stage a rebellion.

{PA:trivialeq

Proposition 1. *There is always an equilibrium of the game that all followers in the coalition turn in the leader, and the leader does not stage a rebellion.*

Next, I focus on analyzing the decisions of followers with finite cut-off rules through backward induction, then I revert to the decision of the leader.

[11] The limit dominance implies that participating the rebellion is a dominant strategy for the sufficiently high type players and turning in the leader is a dominant strategy for the sufficiently low type players

[12] Another example: the player with $x_i < \frac{-q+I-t}{T-s}$ always want to side the government no matter the regime change probability. This example reflect some phenomenon in the reality that even if a government will definitely be overthrown, it will still gain some support from officials' relatives, or some religious factions, or the residents from the dictators home town.

3.4 Followers' decisions:

{subsec:PA_fo

In the situation in which a type x_0 leader stages a rebellion with a coalition size L, the expected payoff for a type x_i follower to join the rebellion is expressed as follows:

$$u_i^c(x_i, x_0, L) = P(L_c > C \cap L_s < E | x_i, x_0, L, s_{-i})(Tx_i + t + b) - b,$$

where $P(L_c > C \cap L_s < E | x_0, x_i, L, s_{-i})$ is a follower's assessment of the probability of regime change given his/her own type x_i, the leader's type x_0, the coalition size L, and the strategies of the other followers s_{-i}. Similarly, the expected payoff for being neutral is $u_i^n(x_i, x_0, L) = P(vx_i + b) - b$, while that for turning the leader in is $u_i^b(x_i, x_0, L) = P(sx_i - q) + I$.

These expected payoffs indicate that when $x_i > -\frac{t}{T-v}$, then $u_i^c > u_i^n$. This scenario implies that acting as a free rider is not a wise choice for the followers with $x_i > -\frac{t}{T-v}$ as they must share the same risk as the participants, despite obtaining a smaller reward. When $x_i < -\frac{q-I}{v-s}$, then $u_i^b > u_i^n$. This scenario implies that followers who like the current regime should side with the government rather than wait for a regime change. Therefore, remaining neutral is a weakly dominated strategy for followers according to the assumption that $\frac{t}{T-v} > \frac{q-I}{v-s}$, and I assume that followers either join the rebellion or turn the leader in.

There may exist cutoff equilibria with positive participants, in which followers switch strategies according to a finite threshold $k(x_0, L)$. If such a threshold exists, the equilibrium strategy for followers should be:

$$s(x_i, x_0, L) = \begin{cases} 1 & \text{if } x_i \geq k(x_0, L), \\ -1 & \text{else.} \end{cases}$$

Next, I perform the following steps to discuss the existence of finite $k(x_0, L)$: 1. It is first necessary to derive the equilibrium conditions to solve the finite $k(x_0, L)$ for a given x_0 and L. 2. Then the conditions related to x_0 and L that are necessary for the existence of finite $k(x_0, L)$ must be computed.

Step 1. First, the subject belief of the regime change probability need to be calculated for follower i. From the perspective of i, the total participants are $L_c = L(1 - \Phi(\frac{k-\theta}{\sigma_\epsilon}))$, where Φ is the standard normal CDF. Therefore, the participation condition $L_c > C$ should be equivalent to $\theta > k - \sigma_\epsilon \Phi^{-1}(1 - \frac{C}{L})$.[13] Similarly, the maintaining secrecy condition $L_s < E$ is equivalent to $\theta > k - \sigma_\epsilon \Phi^{-1}(\frac{E}{L})$. Given that the conditions for a successful rebellion are $L_c > C$ and $L_s < E$,

[13] From the perspective of i, when all other players follow the same cutoff rule, then the number of followers with $x_i > k \Rightarrow \epsilon_i > k - \theta$ should join the rebellion. Since ϵ follows $N(0, \delta_\epsilon)$ and the coalition size is L. This grouping is expressed as $L_c = L(1 - \Phi(\frac{k-\theta}{\sigma_\epsilon}))$. Since the participation strictly increases in θ (i.e., the worse the underlying situation of the country is, the more participants are recruited).The right-hand side of the inequality represents the minimum θ, in which the participation condition is maintained.

θ must be greater than both $k - \sigma_\epsilon \Phi^{-1}(\frac{E}{L})$ and $k - \sigma_\epsilon \Phi^{-1}(1 - \frac{C}{L})$ to ensure the success of the rebellion. The minimum level of θ that is required for regime change can be referred to as

$$\bar{\theta} \equiv \max\{k - \sigma_\epsilon \Phi^{-1}(1 - \frac{C}{L}), k - \sigma_\epsilon \Phi^{-1}(\frac{E}{L})\}. \tag{1}$$

Recall that i believes θ is normally distributed as $N(\bar{m}_i, \bar{\sigma}^2)$ derived in Section 3.3. Therefore i's subjective belief of the regime change probability is

$$\begin{aligned} P(L_c > C \cap L_s < E) &= P(\theta > \bar{\theta}(x_0, L)) \\ &= 1 - \Phi(\frac{\bar{\theta}(x_0, L) - (1-\psi)\lambda x_i - \psi x_0 - (1-\psi)(1-\lambda) m_\theta}{\bar{\sigma}}). \end{aligned} \tag{2}$$ {e8}

It is worth to point out that when $L \leq E + C$, $\bar{\theta} = k - \sigma_\epsilon \Phi^{-1}(1 - \frac{C}{L})$ with $k - \sigma_\epsilon \Phi^{-1}(1 - \frac{C}{L}) \geq k - \sigma_\epsilon \Phi^{-1}(\frac{E}{L})$. Intuitively, followers care more about whether the number of participants is adequate, rather than maintaining secrecy during decision-making when the coalition size is small. In other words, the participation condition dominates the maintaining secrecy condition in this scenario. On the contrary, when $L > E + C$, $\bar{\theta} = k - \sigma_\epsilon \Phi^{-1}(\frac{E}{L})$. It means that the maintenance of secrecy is more important than recruiting a sufficient number of participants to mobilize followers when the coalition size is large (i.e., the maintaining secrecy condition dominates the participant condition).

Follower i's subjective belief must be consistent on the equilibrium path, i.e. k, must be applied by all followers. Given the monotonicity of the payoff functions in x_i, the follower whose type is equal to k would be indifferent to both joining the rebellion and turning in the leader. Therefore, with a given x_0 and L, the equilibrium condition for a finite k should be:

$$u_i^c(x_i, k, x_0, L)|_{x_i = k} = u_i^b(x_i, k, x_0, L)|_{x_i = k}. \tag{3}$$

By plugging in equation (2), this equilibrium condition can be rewritten as

$$(1 - \Phi(\alpha k - M(x_0, L) - \beta))((T - s)k + b + q + t) - b - I = 0, \tag{4}$$ {PA:eq cond}

where $\alpha = \frac{1-(1-\psi)\lambda}{\bar{\sigma}}$, $\beta = \frac{(1-\lambda)(1-\psi)}{\bar{\sigma}} m_\theta$; and $M = \frac{\sigma_\epsilon}{\bar{\sigma}} \Phi^{-1}(1 - \frac{C}{L}) + \frac{\psi}{\bar{\sigma}} x_0$, when $L \leq E + C$; and $M = \frac{\sigma_\epsilon}{\bar{\sigma}} \Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}} x_0$, when $L > E + C$.

Step 2. $M(x_0, L)$ in equation (4), exclusively includes the information shared by the leader, i.e. x_0 and L. It can be referred as the *"influence"* exerted by the leader on followers. Then, the relationship between x_0, L and the existence of the finite k can be derived by analysing *influence* M. Here it is necessary to introduce a new notation to denote the left hand side of (4): $\hat{u}(k, M(x_0, L)) \equiv (1 - \Phi(\alpha k - M(x_0, L) - \beta))((T - s)k + b + q + t) - b - I$.

The next lemma describes the shape of $\hat{u}(k, M)$, which is illustrated in Figure 1.

{u_hat proper

Lemma 1. *For all parameter values:*

1. *For a given M, $\hat{u}(k, M)$ is single peaked of k;* $\lim_{k \to +\infty} \hat{u}(k, M) = -(I + b)$*;*

and $\lim_{k\to-\infty} \hat{u}(k,M) = -\infty$.

2. For a given k*,* $\hat{u}(k,M)$ *increases in* M*.*

(All proofs are in the appendix)

The shape of $\hat{u}$ signifies that there are generically two finite cutoff rules (the curve of $\hat{u}(k,M)$ in Figure 1). Given $M(x_0,L)$, I label the low $k^l(x_0,L)$ and the high $k^h(x_0,L)$ (for low and high). Since $\hat{u}(k,M)$ increases in M, it implies a knife-edge case that a minimal $M^{\min}$ exists such that finite k can be solved from equation (4) for any $M \geq M^{\min}$. In this knife-edge scenario, only one finite threshold exists, that is, $k^l = k^h$ (the curve of $\hat{u}(k,M^{min})$ in Figure 1). For any $M' < M^{min}$, no finite k can be solved from the equilibrium condition (the curve of $\hat{u}(k,M')$ in Figure 1).

From Lemma 1, it is easy to find that $M^{\min}$ represents the minimal influence level that the leader must possess to motivate positive participants. By the definition of $M(x_0,L)$, it is an increasing function of x_0, hence there exists a $x_0^{\min}$ which represents the least type of leader who can exert the influences on the followers above $M^{\min}$. In other words, when the leader's type is sufficiently weak ($x_0 < x_0^{\min}$), no positive participants can be convinced to join the rebellion regardless of L. When $x_0 > x_0^{\min}$, for any small coalition size ($L < E + C$), the influence $M = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$ increases with L because the participation condition dominates the maintaining secrecy condition. Consequently, as long as the leader chooses L such that $M^{\min} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$ for a given x_0, positive amount of followers will join the rebellion. When the coalition is large ($L \geq E + C$), the maintaining secrecy condition dominates the participation condition. Moreover, $M = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0$ decreases with L. For a given x_0, the selected L must satisfy $M^{\min} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0$ to recruit positive participants. The above discussion can be displaced in Figure 2 and the results are summarized in the next lemma.

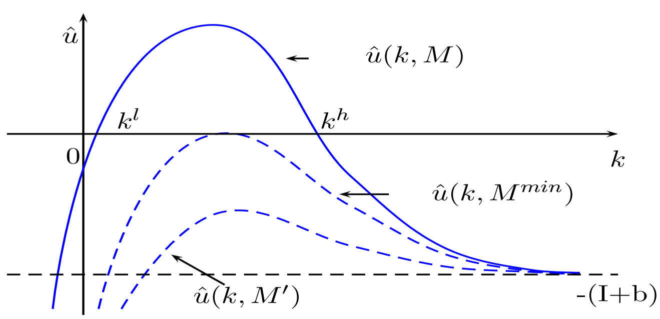

Figure 1: $M' < M^{\min} < M$

{fig:u}

{PA:le:minina

Proposition 2. *There exist a* $x_0^{\min}$ *and a constant* $M^{\min}$ *such that:*

1. For any $x_0 < x_0^{\min}$*, there is no finite* k *which is consistent with the cutoff equilibrium.*

2. For any $x_0 \geq x_0^{\min}$*, finite* $k(x_0,L)$*, which is consistent with equilibrium, exists if and only if either (i)* $L \leq E + C$ *and* $M^{\min} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$*; or (ii)* $L > E + C$ *and* $M^{\min} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0$*.*

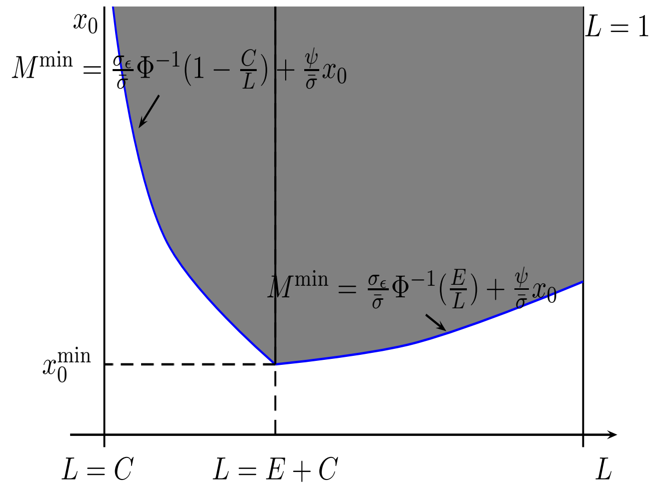

Figure 2: Any point (L, x_0) in the shadow region represents that the finite cutoff $k(x_0, L)$ can be solved from equation (4). When $L \leq E + C$, i.e. the participation condition dominates the maintaining secrecy, curve $M^{\min} = \frac{\sigma_\epsilon}{\sigma}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\sigma}x_0$ represents the boundary composed of x_0 and L such that condition (4) holds. When $L > E+C$, i.e. the maintaining secrecy condition dominates the participation condition, curve $M^{\min} = \frac{\sigma_\epsilon}{\sigma}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\sigma}x_0$ represents the boundary composed of x_0 and L such that condition (4) holds. {fig:PA}

Since k solved from the equilibrium condition (4) are not unique, I follow the equilibrium selection criterion applied by Bueno de Mesquita (2010) to assume that no follower adheres to threshold k^h, in which the followers' strategies are unstable from the perspective of a dynamic learning procedure.[14] Moreover, the strategy with k^h is also counter-intuitive: fixing all other parameters, when the reward from turning in the leader, I, increases, rebellion participation increases, i.e., k^h decreases.

Assumption 1. (Equilibrium Selection) *Followers do not adopt the strategy using* $k^h(x_0, L)$. {ass2}

Based on Assumption 1, I employ $k(x_0, L)$ instead of $k^l(x_0, L)$ by dropping superscript 'l' hereafter to prevent confusion. $k(x_0, L)$ is a function of x_0 and L. x_0 exerts only one effect on k. When x_0 increases, the followers believe that anti-government sentiments are high. As a result, participation is high. In other words, a high-type leader can persuade many followers to join the rebellion. L induces two varied effects on k depending on its scale. When L is small, the participation condition dominates the maintaining secrecy condition; an increase in L causes followers to believe that participation increases as a result. Consequently, the cutoff threshold k decreases. When L is large, the maintaining secrecy condition dominates the participation condition. Therefore, k increases when L increases beyond $E + C$. As a result, participation is low. These facts are summarized as follows:

{PA: L prop}

[14]Details are cited in Bueno de Mesquita (2010) Assumption 2

Corollary 1. *1.* $\frac{\partial k(x_0,L)}{\partial L} < 0$ *when* $L < E + C$, $\frac{\partial k(x_0,L)}{\partial L} > 0$ *when* $L \geq E + C$.
2. $\frac{\partial k(x_0,L)}{\partial x_0} < 0$.

3.5 Leader's decision:

{PA: subsec l

For a given x_0 and under Assumption 1, if the leader selects the coalition size $L > 0$, his expected payoff is

$$u_0(x_0, L) = (1 - \Phi(\frac{\bar{\theta}(x_0, L) - \lambda x_0 - (1-\lambda)m_\theta}{\sigma}))(R + b) - b, \quad (5)$$

{eq:PA_leader

where $1-\Phi(\frac{\bar{\theta}(x_0,L)-\lambda x_0-(1-\lambda)m_\theta}{\sigma})$ represents the leader's subjective belief of the probability of regime change.

When the leader decides to plot a rebellion, his optimal choice of L is to minimize the cutoff threshold $k(x_0, L)$, through which he can recruit more participants. More precisely, when $L < E+C$, the participation condition dominates the maintaining secrecy condition. In this case, enlarging L can attract more participants by decreasing k (Lemma 1), therefore $u_0(x_0, L)$ increases with L. When $L > E + C$, the maintaining secrecy condition dominates the participation condition. As a result, reducing L is more effective in recruiting participants, and $u_0(x_0, L)$ decreases with L. Therefore, the best choice for a leader is $E + C$, which helps the leader achieve the maximal influence over followers by balancing the participation and maintaining secrecy conditions. This result is summarized as follow.

{PA: best L}

Lemma 2. *For given* x_0, $u_0(x_0, L)$ *increases with* L *when* $L < E+C$ *and decreases when* $L > E+C$. *Therefore, the optimal coalition size is* $L^* = E + C$ *at which* $u_0(x_0, L)$ *achieves the maximal value.*

Lemma 2 suggests that L^* is independent of x_0 under collective punishment. This is because if L^* changes with x_0, the effect of the change in L^* offsets the effect from the change of x_0 concerning either an increasing number of participants or a decreasing number of traitors. This result provides us an interesting finding that the rebel group size should be always kept small under collective punishment no matter how strong the non-government motivation is. Intuitively, the optimal choice for the leader is to maximize the coalition size under the tolerance of potential betrayal. Meanwhile, since collective punishment does not leave rooms for free riders. Any follower who is approached by the rebel leader has to choose a side. From the leader's perceptive, any coalition size smaller $E + C$ can undermine the opportunity of recruiting sufficient supporters, on the other hand, any coalition size greater than $E+C$ increases the risk of betray beyond too much. Therefore, the leader should choose the coalition size slight larger than the necessary scale C by adding additional E followers which is the threshold he can afford for the betrayal.

	Join	Remain neutral	Turn in the leader
Succeed	$Tx_i + t$	vx_i	$sx_i - q + I$
Fail	$-b$	0	I

Table 2: The follower payoffs under targeted punishment {table:payoff

Since the leader's expected utility function for staging a rebellion is a monotone increasing function of his type x_0, i.e. it is easy for a leader with a high type to persuade followers to join a rebellion. Therefore, a unique threshold x_0^* exists for a rebel leader to decide whether to stage one.

{PA: start re

Proposition 3. *There exists a unique x_0^* such that the leader does not start a rebellion when $x_0 < x_0^*$, and starts a rebellion with the coalition size $E + C$ when $x_0 \geq x_0^*$.*

4 Targeted Punishment

{sec:pp}

Borrowing the method developed in the previous section, the discussion can be extended to analyse targeted punishment. In this case, the government only punishes the participants if the rebellion fails, and the player who remains neutral (free rider) receives zero payment (Table 2). The rest of the model setup is the same as collective punishment in Section 3.

4.1 Followers' decision

{subsec:PP_fo

Since the free rider will not be punished when the rebellion fails, remaining neutral is not always a weakly dominated strategy under the targeted punishment rule. Two potential cutoff strategies may be considered on the equilibrium path. First, a one cutoff threshold $k(x_0, L)$ may be generated such that a follower joins a rebellion if his/her type is greater than k, and betrays the leader otherwise (referred to as the *one-cutoff strategy*).

Second, a two cutoff thresholds pair (k^c, k^s) with $k^c > k^s$ may be generated. In this case, a follower joins the rebellion if his/her type is greater than k^c; he/she betrays the leader if his/her type is less than k^s; and he/she remains neutral if his/her type is between k^c and k^s (referred to as the *two-cutoff strategy*).

Next, I focus on determining under what conditions of x_0 and L, finite thresholds k, and/or (k^c, k^s) exist which are consistent with the cutoff equilibrium.[15] Similar as the discussion under collective punishment, whether the leader can motivate positive participants is determined by the influence he can exert on the followers, which is a function of x_0 and L. For a given x_0, when a small L is chosen, the participation condition dominates the maintaining secrecy condition in the

[15]Finding the equilibrium conditions for both the one-cutoff strategy and the two-cutoff strategy follows a similar procedure developed in Section 3. The technical details to solve the equilibrium conditions can be found in Appendice.

followers' decision-making, then the influence $M = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0$. A lower bound $M_C^{\min}$ can be drawn for the leader's influence such that when $M_C^{\min} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0$, positive followers will join the rebellion. Similarly, when a large L is chosen, the maintaining secrecy condition dominates the participation condition, then $M = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, meanwhile, a lower bound $M_E^{\min}$ can be found such that when $M_E^{\min} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, the leader can recruit positive participants.

When the leader's influence is low but enough to motivate positive participants, some followers will choose to remain neutral to avoid the risk of rebellion failure. In this scenario, two cutoff thresholds, k^s and k^c can be found which are consistent with the equilibrium. Furthermore, the number of free riders shrinks when the leader's influence increases.[16] It is because that the marginal benefit of switching from a free rider to a participant is large than that of switching from a traitor to a free rider, when the followers believe that regime will change with a high probability. Consequently, free riders are eliminated (one-cutoff strategy) when the leader's influence is sufficiently large. Natural, a threshold M^{med} for the leader's influence should exist, beyond which only one cutoff threshold k can be found i.e. no free rider. These results are summarized in the next lemma and illustrated in Figure 3

{PP:le:eq exi

Proposition 4. *1. There exists a constant M^{med} such that, for any given x_0 and L, one cutoff threshold $k(x_0, L)$ exist, which is consistent with the equilibrium, if and only if $M^{med} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$ and $M^{med} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0$.*

2. There exist constant $M_C^{\min}$ and $M_E^{\min}$ with $M_C^{\min}, M_E^{\min} \leq M^{med}$ such that, for any given x_0 and L, two cutoff thresholds $(k^s(x_0, L), k^c(x_0, L))$ exist, which is consistent with the equilibrium, if and only if (1) $M_C^{\min} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$ and $M_E^{\min} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, and (2) either $M^{med} > \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, or $M^{med} > \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0$.

Since Lemma 4 indicates the existence of the lower bounds of the leader's influence to motivate positive participants ($M_C^{\min}$ and $M_C^{\min}$), then it also implies that when the leader's type is sufficiently low, no participant can be recruited regardless the coalition size. Furthermore, since M^{med} is the lower bound of the leader's influence to eliminate the free riders, it implies that if the leader's type is not sufficiently high, he cannot get rid of the free riding problem.

{cor:min and

Corollary 2. *There exist $x_0^{\min}$ and x_0^{med} such that when $x_0 < x_0^{\min}$, no positive participant can be motivated by the leader regardless L; when $x_0^{\min} < x_0 < x_0^{med}$, no one-cutoff strategy exists for the followers.*

Similar as the case under collective punishment, when the finite cutoff rules exist on the equilibrium path, they are not unique. I follow the same equilibrium selection criterion (Bueno de Mesquita (2010)) to assume that when finite cutoff rules exist, followers do not adopt any strategy associated with the large thresholds.[17]

[16] More precisely, for a given coalition size L, the number of free riders shrinks when x_0 increase. For a given x_0, the number of free riders shrinks when the influence increases by choosing different L

[17] More precisely, in the one-cutoff strategy, given a x_0 and L, when the finite threshold exists, k solved from the

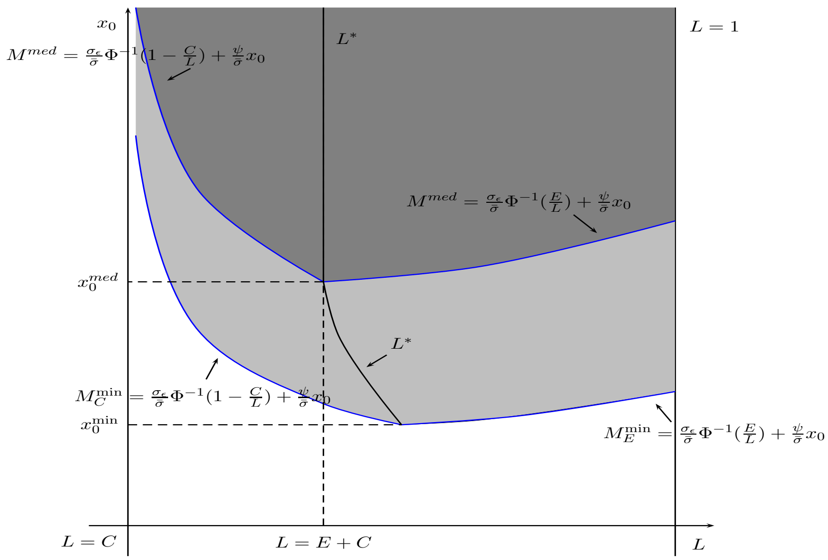

Figure 3: 1. Any point (L, x_0) in the dark shadow region represents the finite cutoff $k(x_0, L)$ existing. 2. Any point (L, x_0) in the light shadow region represents the finite cutoff pair (k^c, k^s) existing. 3. L^* represents the balance points for the participation and maintaining secrecy conditions. 4. $M^{\min_C}$ and $M^{\min_E}$ represent the influence thresholds separating the two-cutoff strategy and no finite cutoff case when the participation condition is the dominant condition and when the maintaining secrecy condition is the dominant condition respectively. 5. M^{med} represents the influence threshold separating the one-cutoff strategy and the two-cutoff strategy. {fig:PP eq}

4.2 Leader's decision: {PP: subsec l

For a given x_0, the optimum coalition size chosen by a leader balances the participation and maintaining secrecy conditions if he wants to plot a rebellion. Corollary 2 shows that when $x_0 > x_0^{med}$, the leader can achieve an influence level greater than M^{med} to eliminate free riders. Therefore, the balance point for the participation and maintaining secrecy conditions can be achieved with $L^*(x_0) = E + C$, which is exactly the same as that under collective punishment. When $x_0^{\min} < x_0 < x_0^{med}$, the maximum influence the leader can exert cannot reach M^{med} any longer. Because the leader knows some coalition members will be free riders in this situation, he has to recruit more followers to satisfy the participation condition than the case when no free rider exist. Furthermore, as x_0 decreases, it becomes harder for the leader to persuade people join the rebellion, which leads to the increase of free riders, as a consequence, the leader has to enlarge the coalition sizes further. {PP: leader's

Proposition 5. *Under targeted punishment, the leader's optimal coalition size is $L^*(x_0) = E + C$ if $x_0 \geq x_0^{med}$; and $L^*(x_0) > E + C$ and decreases in x_0, if $x_0^{\min} \leq x_0 < x_0^{med}$.*

---

equilibrium conditions is entirely the same as it is solved from (4) in Section 3. So there exist two thresholds, k^l and k^h with $k^l \leq k^h$. Similarly, in the two-cutoff strategy, when finite threshold pair (k^c, k^s) exists, there are two pairs of solutions that satisfy the equilibrium conditions, (k^{cl}, k^{sl}) and (k^{ch}, k^{sm}) with $k^{cl} \leq k^{cm}$ and $k^{sl} \leq k^{sm}$. This assumption means followers do not adopt any strategy using threshold k^h or pair (k^{ch}, k^{sh}).

Proposition 5 implies that different punishment rule does not affect a high type leader's choice of the coalition size. Large coalition size occurs only when the leader is weak. Since a high type leader also implies a high chance for the regime change, it indicates that a small scale rebel group maybe a more vital threat to the government. Finally, it is easy to show that a starting point x_0^* exists for the leader to stage a rebellion under the targeted punishment rule.

{PP: leader d

Proposition 6. *A unique x_0^* exists such that the leader would not start a rebellion when $x_0 < x_0^*$, and start a rebellion with the coalition size $L^*(x_0)$ when $x_0 \geq x_0^*$.*

5 Survival Probability Analysis

{sec:survival

After calculating equilibria under two punishment rules, the question of which rule is more likely to prevent a rebellion should be addressed.

Recall that $\bar{\theta}(x_0, L)$ is the minimum level of the underlying situation variable θ required for regime change. On the equilibrium path, for a given θ, once a rebellion is staged by a type x_0 leader with the optimal coalition size $L^*(x_0)$, the rebellion succeeds if and only if $\theta > \bar{\theta}(x_0, L^*(x_0))$. Since $\bar{\theta}$ is clearly a monotonic decreasing function of x_0, a $\hat{x}_0$ exists such that $\bar{\theta}(\hat{x}_0, L^*(\hat{x}_0)) = \theta$. Intuitively, $\hat{x}_0(\theta)$ is the minimum leader's type to implement a successful rebellion. $\hat{x}_0(\theta)$ is a decreasing function of θ because when the underlying situation of the government worsens, the leader can use a smaller x_0 to motivate followers. Therefore, a rebellion is staged and succeeds if and only if the leader's type x_0 is greater than both x_0^* and $\hat{x}_0(\theta)$. Now the regime's survival probability can be written as:

$$P_{sur}(\theta) = 1 - P(x_0 > \max\{x_0^*, \hat{x}_0(\theta)\})$$

To avoid confusion when we discuss different punishment rules, hereafter we employ CP and TP as superscribes to refer the collective and targeted punishment respectively. Under targeted punishment, when the rebellion starting point x_0^{*TP} is greater than x_0^{med}, followers only use the one-cutoff strategy on the equilibrium path, which is exactly the same as the equilibrium strategy under the collective punishment rule. Consequently, two punishment rules have the same effect in preventing a rebellion.[18]

I then focus on the case when $x_0^{*TP} < x_0^{med}$. The next two propositions indicate that when the underlying situation of a government is good, both the collective and targeted punishment rules have the same effect in preventing rebellions, whereas when the underlying situation worsens, they might result in disparate survival probabilities for a government. In the first case, when the regime change incentive T and v are sufficiently large with relatively small difference, targeted punishment has a higher survival probability than that of collective punishment. Intuitively, although a large

[18] Intuitively, $x_0^{*TP} > x_0^{med}$ means that the leader's starting point for a rebellion is high, in other words, leading a rebellion is unattractive to the leader. Then, both punishments provide the regime with the same survival probability.

T makes joining a rebellion attractive, a large v implies that being a free rider is an equally good choice, without taking risk. However, given that no free rider problem exists under collective punishment, this harsh punishment forces those followers who could be free riders under targeted punishment to choose a side. The collective punishment rule could urge more followers to the side opposing the government when both T and v are large and the underlying situation θ is poor. This result is summarized as:

{Sur:prop: P

Proposition 7. *When $T - v = g$ where g is a constant, and v is greater than a constant $\hat{v}$, then there exists a $\hat{\theta}$ such that when $\theta \leq \hat{\theta}$, both collective and targeted punishment rules have the same survival probability, i.e., $P_{sur}^{CP}(\theta) = P_{sur}^{TP}(\theta)$; when $\theta > \hat{\theta}$, targeted punishment has a higher survival probability, i.e., $P_{sur}^{CP}(\theta) < P_{sur}^{TP}(\theta)$.*

The next proposition indicates that if $v - s$ is small enough, and $v - s$ and $q - I$ are proportional, collective punishment is more effective in preventing a rebellion. Intuitively, when remaining neutral is unattractive to followers in comparison to turning in a leader, and the threat from a leader (q) is not significantly large than the government reward (I), collective punishment is a better choice for a government.

{Sur:prop: P

Proposition 8. *When $q - I = (v - s)h$ where h is a constant, and $v - s$ is less than a constant v', then there exists $\hat{\theta}'$ such that when $\theta < \hat{\theta}'$, both collective and targeted punishment rules have the same survival probability, i.e., $P_{sur}^{CP}(\theta) = P_{sur}^{TP}(\theta)$; and when $\theta > \hat{\theta}'$, collective punishment has a higher survival probability, i.e., $P_{sur}^{CP}(\theta) > P_{sur}^{TP}(\theta)$.*

Proposition 7 and 8 indicate that targeted punishment allows institutions some flexibility by allowing free riders when the underlying situation worsens. Therefore, targeted punishment provides two paths for the prevention of a rebellion. The first is a direct method for increasing traitors, and the second is an indirect method for allowing free riders, which use the people's hesitation to join a risky action. However, collective punishment employs only the direct method to prevent a rebellion by attracting traitors. The flexibility offered by targeted punishment is not functional when the incentive for being a free rider is small. However, targeted punishment is more likely to prevent rebellion than collective punishment when joining a rebellion and being a free rider are both attractive.

6 Comparative Static and Endogenous Institution Choice

{sec: compar

I begin with a discussion of t, the pecuniary reward for rebellion participants. Joining a rebellion becomes increasingly attractive as t increases. As a consequence, more participants and fewer traitors are expected under both the collective and targeted punishments. However, the increase in t has disparate effects on the participants and free riders under targeted punishment. The former can benefit both from the increases in the direct monetary reward and the rebellion's probability of success, while the latter can only benefit from the increase in the rebellion's probability of success. Therefore, as t increases under targeted punishment, the role of the free rider problem decreases, and thus the difference between two punishment rules reduces.

{CS: Prop: t

Proposition 9. *For a given θ, when t is sufficiently large, there is no free rider under targeted punishment; and both the collective and targeted punishment rules provide the same survival probability for the government, i.e., $P_{sur}^{CT}(\theta) = P_{sur}^{TP}(\theta)$.*

Proposition 9 implies that the leader can attract more supporters by sharing his return. This proposition accords with historic facts, particularly regarding small coups during which the leader promises a huge reward to recruit key supporters. However, recruiting by increasing reward t also creates other problems that are beyond the scope of this paper, but are interesting to mention for future research.[19] First, promising huge rewards to supporters reduces the payoff for a leader, which decreases the incentive for a leader to stage a rebellion. Meanwhile, a huge reward also presents an issue of commitment; the leader might renege on his promise once he assumes the highest power. Second, a constraint for increasing t might exist. For example, the Chinese Communist Party promised land reform to attract farmers to join a revolution. However, natural local constraints and a large population drew an upper bound for rewards to village families.

Another interesting parameter is the monitoring threshold, E. When E decreases, it implies that a government enhances the monitoring of the targeted population, which makes it harder for a rebel group to achieve success.[20] The discussion can be extended to the question about the optimal choice of E from the government's perspective. Assume the government is subject to a cost function, $B(N, E)$, when monitoring a targeted population, where N is the exogenous population size and E is the monitoring threshold that the government sets up. It is also reasonable to assume that B increases with the size of the target population (N), and decreases with the monitoring threshold (E). When the primary threat to a government is an insurgency begun by a large population, for whom the regime change incentive is more important than the direct monetary reward (large T and v, small t). According to Proposition 7, selecting targeted punishment provides a higher survival probability for the government than selecting collective punishment because the former allows free riders. Considering that monitoring the entire population is costly (large N), adopting targeted punishment can also allow the government to select a high E to save on costs. This scenario can be used to explain revolutions instigated by ordinary civilians. In general, when an ordinary citizen decides to join a risky revolution that aims to overthrow a dictator, the regime change incentive is more important than the direct monetary reward. An example of such a case is the Jasmine Revolution in Tunisia, during which young men were motivated to join protests because of anger

[19] I thank Scott Gehlbach and Mehdi Shadmehr for raising this issue.

[20] It is easy to find k and x_0^{*CP} under collective punishment and k, k^c, k^s, and x_0^{*CP} under targeted punishment all increase when E decreases. Briefly, selecting a small E is effective in preventing a rebellion.

that spread through social networks. The dictatorship of Ben Ali was overthrown in a month.[21] Under this circumstance, a collective punishment rule does not help maintain the regimes stability, owing to a lack of tolerance for protests.[22]

On the contrary, when a government faces a potential rebellion plotted by a small population primitively driven by ambitions for power or direct monetary reward (large t, small v, and s), such as palace coups, collective punishment might be a good choice for the government, according to Propositions 8. The government might also afford to select a small E to prevent the rebellion. This scenario can explain why dictators prefer using secret police to monitor officials to prevent coups. In addition, punishment can be implemented frequently without transparently enforcing the rule of law.

7 Cruel Competition

{sec:competit

In the previous sections, I focus on the case in which the leader punishes only those who betray him once a rebellion succeeds. In reality, a rebel group leader can employ rewards and threats to recruit coalition members, so I extend the discussion to a case in which the leader can also threaten to punish free riders when a rebellion succeeds.

When the leader punishes followers who remain neutral once a rebellion succeeds, the payoff for a free rider is $vx_i - q$. I call this scenario the *harsh threat* scenario to distinguish it from the previous case, referred to as the *easy threat* scenario hereafter, in which the leader punishes only traitors. When the government chooses collective punishment, being a free rider is a dominated strategy, and the harsh threat from the leader does not change the results presented in Section 3.

When a leader switches from an easy threat to a harsh threat under targeted punishment, the original free riders with high types join a rebel group because being a free rider becomes less attractive than joining a rebellion when they think the probability of a rebellion's success is high. However, free riders with low types betray the leader because such actors feel the probability of success is low, and thus would rather turn in the leader to gain some reward from the government than wait for the punishment from a leader as free riders. These changes imply that the size of free riders decreases under a harsh threat scenario. On the equilibrium path, free riders do not exist when the punishment q is large enough. The following lemma summarizes this result.

{CC: le: fre

Lemma 3. *Under targeted punishment and a leader chooses a harsh threat, if q is large enough, then free rider does not exist.*

[21] More details about the Arab Spring can be found in Anderson (2011). The effect and role of social media during the Arab Spring Revolution can be found in Howard et al. (2011).

[22] Another authority with a low tolerance for protests is China, which is subject to a large cost. Although no evidence shows that the color revolution can happen in China (Swartz (2011)), the cost of stability maintenance (Weiwen), such as managing social protest or Internet sanction, continuously increases from CNY 549 billion in 2010 to CNY 769 billion in 2013. This cost increases faster than China's military expenditure (Benney (2013))

Now, I assume that the condition for Lemma 3 holds. Consider a 2 by 2 game played by a government and a rebel leader. The government can choose either the collective or targeted punishments, and the leader can choose either the easy or harsh threat. I assume that: 1) the leader always wants to start a rebellion;[23] and 2) the government's payoff is $-b$ when the regime is changed. For a given θ, the payoff table is listed below. The first payoff in each cell is the expected payoff for the leader, and the second payoff is for the government. $P^1_{sur}(\theta)$ and $P^2_{sur}(\theta)$ are the survival probabilities for the government when remaining neutral is a dominated strategy and when it is not, respectively.[24]

Leader\King	*Collective*	*Targeted*
Easy	$R - P^1_{sur}(R+b), \quad P^1_{sur}(R-b)+b$	$R - P^2_{sur}(R+b), \quad P^2_{sur}(R-b)+b$
Harsh	$R - P^1_{sur}(R+b), \quad P^1_{sur}(R-b)+b$	$R - P^1_{sur}(R+b), \quad P^1_{sur}(R-b)+b$

For a given θ, the government prefers *Collective* to *Targeted* when $P^1_{sur}(\theta) < P^2_{sur}(\theta)$. In this case, the leader prefers *Harsh* to *Easy* , because when the government does not want to use *collective* to force potential free riders to choose sides, a leader must urge free riders to choose sides because most would join the side of the leader when they cannot remain neutral. (*Harsh*, *Collective*) and (*Harsh*, *Targeted*) are two pure strategy equilibria. If $P^1_{sur}(\theta) > P^2_{sur}(\theta)$, then (*Easy*, *Collective*) and (*Harsh*, *Collective*) are two pure strategy equilibria because the leader is indifferent between these two choices when the government prefers *Collective*.

From the preceding discussion, it is simple to find that the use of a harsh way to oppose the other side is always an option for one side on the equilibrium path. Moreover, for any θ, (*Harsh*, *Targeted*) is always an equilibrium. This result provides a possible explanation for why we often observe an equivalent retaliation strategy between governments and rebel groups in reality. The Liberation Tigers of Tamil Eelam in Sri Lanka and the Revolutionary Armed Forces of Colombia are two cases that fit this explanation.

8 Threat More Than Reward

{sec:threat m

In this section, I focus on the condition $\frac{t}{T-v} < \frac{q-I}{v-s}$ in comparison to the assumption stated in Section 3. This inequality implies that a leader prefers threat to reward when recruiting supporters to join a rebellion. When $\frac{t}{T-v} < \frac{q-I}{v-s}$, remaining neutral may not be weakly dominated strategy under collective punishment. Similarly to Section 4, there may also exist one-cutoff strategy or the two-cutoff strategy for the followers.

In the case that only traitors are punished by the leader when the rebellion succeeds, the leader's threat is more useful in reducing traitors than increasing participants. When the leader's

[23]When the leader always starts a rebellion, I assume when his type is less than $x_0^{\min}$ under both punishment rules, he can choose any coalition size, because the rebellion will fail for sure. Also the regime's survival probability becomes $P_{sur}(\theta) = 1 - P(x_0 > \hat{x}_0(\theta))$. Proposition 7 and Proposition 8 still hold.

[24]Actually, $P^1_{sur}(\theta) = P^{CP}_{sur}(\theta)$ and $P^2_{sur}(\theta) = P^{TP}_{sur}(\theta)$.

type increases, the followers believe that the rebellion has a high chance to win. Consequently, more followers would like to switch from being traitors to free riders to avoid the treat from the leader than those who switch from free riders to participants. As a result, the number of free riders increases with the leader's type, meanwhile, the leader will increase the coalition size to not only increases the participants but also deduce the rate of traitors. However when the coalition size increases, some traitors always exist, therefore increasing a coalition's size is limited by an upper bound to avoid breaking the maintaining secrecy condition. These results are summarized as follow: {TR:balance p

Proposition 10. *1. There exist a constant $x_0^{\min}$ such that When $x_0 < x_{0TR}^{\min}$, no finite cutoff equilibrium exist.*
2. There exist a constant $x_{0_{TR}}^{med}$ with $x_0^{med} \geq x_{0_{TR}}^{\min}$, such that When $x_{0_{TR}}^{\min} \leq x_0 < x_{0_{TR}}^{med}$, only one-cutoff strategy for the followers exists on the equilibrium path, and the optimal coalition size for the leader is $L_{TR}^* = E + C$.
3. When $x_0 \geq x_{0_{TR}}^{med}$, only two-cutoff strategy for the followers exists on the equilibrium path; and the optimal coalition size for the follower is $L_{TR}^*(x_0) > E + C$ which increases with x_0
4. $\lim_{x_0 \to +\infty} L_{TR}^*(x_0) \leq 1$

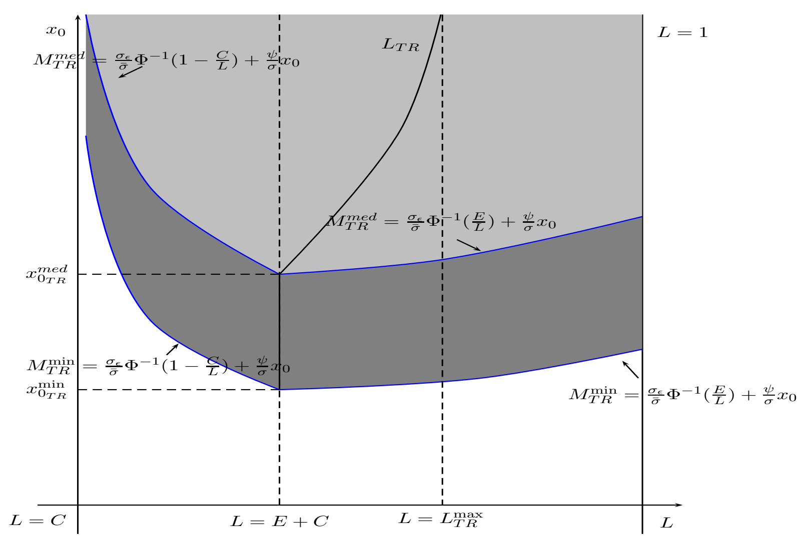

Figure 4: 1. Any point (L, x_0) in the dark shadow region represents the finite cutoff $k(x_0, L)$ existing. 2. Any point (L, x_0) in the light shadow region represents finite cutoff (k^c, k^s) existing. 3. L_{TR}^* is the optimal coalition size which represents the balance points for the participation and maintaining secrecy conditions. 4. $M_{TR}^{\min}$ represents the influence threshold separating the one-cutoff strategy and no finite cutoff case. 5. M_{TR}^{med} represents the influence threshold separating the one-cutoff strategy and the two-cutoff strategy. {fig:last eq}

Under targeted punishment, the results about the equilibrium strategies and optimal coalition sizes are very similar to Proposition 10. Only one fact is worth pointing out: the optimal coalition size under targeted punishment is greater than that under collective punishment. It is because when the rebellion fails, free riders will not be punished by the government; therefore, the marginal benefit of being a free rider is larger under targeted punishment than under collective punishment. Consequently, the leader needs to recruit more followers to satisfy the participation condition in this case.

9 Conclusion

{sec:conclusi

In this article, I present a coordination model in which a rebel leader forms a rebel group to overthrow an incumbent. This model is different with the existing literature in three aspects: First, I model a trade-off, faced by a rebel leader, between trying to mobilize more supporters and preventing information leaks. In such a scenario, rebellions are not necessarily likely to succeed because of the increased number of involved players. Second, this model, to my knowledge, is the first coordination game that allows the player to endogenously select the coalition size. Third, this model tests two widely adopted punishment rules to determine which one is more useful in preventing a rebellion.

I identify equilibria under two punishment rules and find that the rebel group size is smaller under collective punishment than that under targeted punishment because the latter is subject to a free-rider problem. From the perspective of institutional design, targeted punishment is more effective in preventing a rebellion when social anti-authority sentiment is high and being a free rider is attractive because an authority can become flexible and make people hesitate joining a risky rebellion. Thus, participation is lowered.

The model suggests several valuable implications. When the primary threat to authority comes from citizens who are driven by the anti-government sentiment, targeted punishment should be used to prevent rebellions. However, when the primary threat comes from a rebellion staged by people driven by monetary rewards, collective punishment should be adopted. This model also addresses the case when a leader can select whether to punish free riders if the rebellion succeeds; in this instance, both the authority and the rebel leader may use harsh methods to push free riders to choose sides.

This model sheds light on future research in regime change and coordination models. First, following the majority corroboration game, this study assumes that a single individuals decision will not affect the equilibrium result because of the continuity of the players' set. However, a discrete set of players should also be considered, such as a small scale coup, in which each individuals decision may be pivotal in the game. Second, when the coalition size is endogenously determined, trust among rebel group members is another important issue, which can lead the regime change model to a multidimensional case. Finally, another possible extension is to endogenize the reward sharing rule among rebel group members by considering the commitment problem.

Appendix:

Proof of Proposition 1. :

If a follower plays a strategy that turn in the leader for sure regardless the type x_i, x_0 and size L, i.e. $s(x_i, x_0, L) = -1$, then no rebellion can succeed. Therefore each follower has no incentive to deviate given other followers' strategies. Given the followers' strategies, It is obvious the best response for the leader is to choose $L = 0$. $\square$

Proof of Lemma 1. :

1. Let $f = \alpha k - M - \beta$. $\hat{u}(k, M)$ is increasing in k if and only if $\frac{1-\Phi(f)}{\phi(f)}$ is greater than the finite positive constant $\alpha(k + \frac{b+q+t}{T-s})$. Since f is increasing in k and $\frac{1-\Phi(f)}{\phi(f)}$ is decreasing monotonically in k by the monotone hazard rate property of the normal density function. It implies that $\frac{1-\Phi(f)}{\phi(f)(k+\frac{b+q+t}{T-s})}$ is decreasing monotonically in k for $k > -\frac{b+q+t}{T-s}$. Thus we need to show that $\frac{1-\Phi(f)}{\phi(f)(k+\frac{b+q+t}{T-s})}$ passes through α. First we have $\lim\limits_{k \to -\frac{b+q+t}{T-s}} \frac{1-\Phi(f)}{\phi(f)(k+\frac{b+q+t}{T-s})} = +\infty$, it is because $\frac{1-\Phi(f)}{\phi(f)}$ is finite when $k = -\frac{b+q+t}{T-s}$.

$$
\begin{aligned}
& \lim_{k \to +\infty} \frac{1-\Phi(f)}{\phi(f)(k + \frac{b+q+t}{T-s})} \\
= & \lim_{k \to +\infty} \frac{-\phi(f) f_k}{\phi(f) - \phi(f) f f_k (k + \frac{b+q+t}{T-s})} \\
= & \lim_{k \to +\infty} \frac{f_k}{f f_k (k + \frac{b+q+t}{T-s}) - 1} \\
= & \quad 0
\end{aligned}
$$

The first equality is due to l'Hopital's rule and the fact that $\phi'(x) = -x\phi(x)$, the second equality is algebra, and the last equality uses the fact that $f_k = \alpha$ and f is increasing in k. Therefore, $\hat{u}$ is single peaked,

When k approaches $+\infty$, $1 - \Phi(f)$ goes to 0, therefore, $\lim\limits_{k \to +\infty} \hat{u}(k.M) = -(I + b)$. Similarly k goes to $-\infty$, $1 - \Phi(f)$ approaches 1, Therefore $\lim\limits_{k \to -\infty} \hat{u}(k.M) = -\infty$.

2. It is straightforward when take derivative with respect to M. $\square$

Proof of Proposition 2. :

Since $\hat{u}$ is single peaked and increases in M, therefore there exists a $M^{\min}$, such that the peak of $\hat{u}$ tangent to zero, i.e $\max\limits_{k} (1 - \Phi(\alpha k - M^{\min} - \beta))((T - s)k + q + t + I) - b - I = 0$.

1. If $L \le E + C$, $M = \frac{\sigma_\epsilon}{\bar{\sigma}} \Phi^{-1}(1 - \frac{C}{L}) + \frac{\psi}{\bar{\sigma}} x_0$ and increases with L. Therefore, when $x_0 < \frac{\bar{\sigma}}{\psi}(M^{\min} - \frac{\sigma_\epsilon}{\bar{\sigma}} \Phi^{-1}(1 - \frac{C}{E+C}))$, no matter what L is chosen, M cannot exceed $M^{\min}$. If $L > E + C$, $M = \frac{\sigma_\epsilon}{\bar{\sigma}} \Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}} x_0$ and decreases with L. Consequently when $x_0 < \frac{\bar{\sigma}}{\psi}(M^{\min} - \frac{\sigma_\epsilon}{\bar{\sigma}} \Phi^{-1}(\frac{E}{E+C}))$, M cannot exceed $M^{\min}$ either. In summary, when $x_0 < x_0^{\min} \equiv \frac{\bar{\sigma}}{\psi}(M^{\min} - \frac{\sigma_\epsilon}{\bar{\sigma}} \Phi^{-1}(\frac{E}{E+C}))$, no finite k can be solved from equation (4).

2. Sufficiency: When $L \leq E+C$, $M = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, then when $M^{\min} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, we have $\max\limits_k \hat{u} \geq 0$. i.e. there exist finite k satisfying equation (4). Similarly, when $L > E+C$, $M = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, then when we have $M^{\min} \leq \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0$. Therefore there exist finite k satisfying (4).

Necessity: When the finite k satisfying (4) exists, then we must have $\max\limits_k \hat{u} \geq 0$. When $L \leq E+C$, $M = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, then we must have $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0 \geq M^{\min}$. Similarly, when $L > E+C$, we must have $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0 \geq M^{\min}$. □

Proof of Corollary 1. :

1. By implicit function theory,

$$\frac{\partial k}{\partial L} = -\frac{\frac{\partial \hat{u}}{\partial x_0}}{\frac{\partial \hat{u}}{\partial k}} = \begin{cases} -\frac{\tilde{\phi}\tau(\Phi_c^{-1})'\frac{C}{L^2}}{-\tilde{\phi}\alpha((T-s)k+t+2b+I)+(1-\tilde{\Phi})(T-s)} < 0, & \text{if } L \leq E+C, \\ \frac{\tilde{\phi}\tau(\Phi_E^{-1})'\frac{E}{L^2}}{-\tilde{\phi}\alpha((T-s)k+t+2b+I)+(1-\tilde{\Phi})(T-s)} > 0, & \text{if } L > E+C, \end{cases}$$

where $\Phi_C^{-1} = \Phi^{-1}(1-\frac{C}{L})$, $\Phi_E^{-1} = \Phi^{-1}(\frac{E}{L})$, $\tilde{\phi} = \phi(\frac{\bar{\theta}(x_0,L)-(1-\psi)\lambda k-\psi x_0-(1-\psi)(1-\lambda)m_\theta}{\bar{\sigma}})$, $\tilde{\Phi} = \Phi(\frac{\bar{\theta}(x_0,L)-(1-\psi)\lambda k-\psi x_0-(1-\psi)(1-\lambda)m_\theta}{\bar{\sigma}})$ and $\tau = \frac{\psi}{\bar{\sigma}}$.

2.

$$\frac{\partial k}{\partial x_0} = -\frac{\frac{\partial \hat{u}}{\partial x_0}}{\frac{\partial \hat{u}}{\partial k}} = \frac{\tilde{\phi}(1-\psi)((T-s)k+t+2b+I)}{-\tilde{\phi}\alpha((T-s)k+t+2b+I)+(1-\tilde{\Phi})(T-s)} < 0.$$

□

Proof of Lemma 2.

$$\frac{\partial u_0}{\partial L} = \begin{cases} -\frac{(R+b)}{\sigma}(\frac{\partial k(x_0,L)}{\partial L} - \sigma_\epsilon(\Phi_c^{-1})'\frac{C}{L^2})\phi(\frac{\bar{\theta}(x_0,L)-\lambda x_0-(1-\lambda)m_\theta}{\sigma}) > 0 & \text{if } L \leq E+C \\ -\frac{(R+b)}{\sigma}(\frac{\partial k(x_0,L)}{\partial L} + \sigma_\epsilon(\Phi_E^{-1})'\frac{E}{L^2})\phi(\frac{\bar{\theta}(x_0,L)-\lambda x_0-(1-\lambda)m_\theta}{\sigma}) < 0 & \text{if } L > E+C \end{cases}$$

By, Corollary 1, u_0 is increasing with L when $L \leq E+C$, and decreasing with L when $L > E+C$. Therefore $L = E+C$ can let $u_0(x_0, L)$ achieve the maximal value. □

Proof of Proposition 3. :

From Lemma 1, u_0 increases with x_0 because $\frac{\partial u_0(x_0,E+C)}{\partial x_0} = -\phi(\frac{\partial k(x_0,E+C)}{\partial x_0} - \lambda)\frac{1}{\sigma} > 0$. When $x_0 \to -\infty$, we have $u_0(x_0, E+C) \to -\infty$; and when $x_0 \to +\infty$, $u_0(x_0, E+C) \to +\infty$. Therefore there exists unique x_0^*, such that $u_0(x_0^*, E+C) = 0$. □

Define the following notations:

$$\begin{aligned} u'_c(k, M) &\equiv (1-\Phi(\alpha k - M - \beta))(Tk+t+b) - b, \\ u'_n(k, M) &\equiv (1-\Phi(\alpha k - M - \beta))vk, \\ u'_b(k, M) &\equiv (1-\Phi(\alpha k - M - \beta))(sk-q) + I, \\ \Delta u'_{cn}(k, M) &\equiv u'_c(k) - u'_n(k) = (1-\Phi(\alpha k - M - \beta))((T-v)k+t+b) - b, \\ \Delta u'_{nb}(k, M) &\equiv u'_n(k) - u'_b(k) = (1-\Phi(\alpha k - M - \beta))((v-s)k+q) - I. \end{aligned}$$

I describe the intuition of proof of Proposition 4 first. When $M \to +\infty$, the smaller root of $\Delta u'_{cn}(k,M)$ converges to $-\frac{t}{T-v}$ which is less than the smaller root of $\Delta u'_{nb}(k,M)$ converging to $-\frac{q-I}{v-s}$. It implies that when M is large, k^c solved from $\Delta u'_{cn}(k,M)$ is less than k^s solved from $\Delta u'_{nb}(k,M)$, therefore two-cutoff strategy does not exists, only one-cutoff strategy exists. When M decreases until $u'_c(k,M) = u'_n(k,M) = u'_b(k,M)$, then two-cutoff strategy begin to replace one-cutoff strategy.

Proof of Proposition 4. :
First, we prove some properties for the notations defined above. Similar as Lemma 1, u'_{cn} is also single peaked, and increases with M. Then there exists a $M_C^{\min}$ such that $\Delta u'_{cn}(k, M_C^{\min}) = 0$ has one finite root. For any $M > M_C^{\min}$, $\Delta u'_{cn}(k,M) = 0$ has two roots. For any $M < M_C^{\min}$, $\Delta u'_{cn}(k,M) = 0$ has no finite root.

We use k^l and k^h to denote the small and large root respectively, for $\Delta u'_{cn}(k) = 0$, when they exist. Since we only care about small root k^l, I only focus on k^l, k^h has the similar properties. By the Implicate Function Theory, it is easy to should k^l is a decreasing function of M.

Next, we can find that $u'_c(k^l(M),M)$ decreases with M. It is because $\frac{\partial u'_n(k^l,M)}{\partial M} = \phi(vk^l+b)[(-\alpha+\frac{1-\Phi}{\phi k^l})\frac{\partial k^l}{\partial M}+1]$. Since $u'_c(k^l(M),M) = u'_n(k^l(M),M)$, take derivative w.r.t M on both side, we have $\frac{\partial k^l}{\partial M} = 1/(\alpha - \frac{1-\Phi}{\phi}\frac{1}{k^l+\frac{t+b}{T-v}})$. Plug this form into $\frac{\partial u'_n(k^l,M)}{\partial M}$, we have $\frac{\partial u'_n(k^l,M)}{\partial M} = -\phi v k^l(\frac{\alpha-\frac{1-\Phi}{\phi k^l}}{\alpha-\frac{1-\Phi}{\phi}\frac{1}{k^l+\frac{t}{T-v}}} - 1)$. Similar as the proof in Lemma 1, we have $\alpha k^l < \frac{1-\Phi}{\phi}$, when $-\frac{t}{T-v} < k < 0$, we have $\alpha - \frac{1-\Phi}{\phi k^l} > 0$ and $\alpha - \frac{1-\Phi}{\phi}\frac{1}{k^l+\frac{t}{T-v}} < 0$, so $\frac{\partial u'_n(k^l(M),M)}{\partial M} < 0$. When $k > 0$, we have $\alpha - \frac{1-\Phi}{\phi k^l} < 0$, $\alpha - \frac{1-\Phi}{\phi}\frac{1}{k^l+\frac{t}{T-v}} < 0$ and $\frac{\alpha-\frac{1-\Phi}{\phi k^l}}{\alpha-\frac{1-\Phi}{\phi}\frac{1}{k^l+\frac{t}{T-v}}} - 1 > 0$. Therefore $\frac{\partial u'_n(k^l(M),M)}{\partial M} < 0$.

Similarly there exists a $M_E^{\min}$ such that $\Delta u'_{nb}(k, M_E^{\min}) = 0$ has one finite root. For any $M > M_E^{\min}$, $\Delta u'_{nb}(k,M) = 0$ has two roots. For any $M < M_E^{\min}$, $\Delta u'_{nb}(k,M) = 0$ has no finite root. We use k'^l to denote the small root for $\Delta u'_{nb}(k) = 0$, if it exists. It is easy to show that k'^l is a decreasing function of M and $u'_c(k'^l(M),M)$ is decreasing with M.

Let M^{med} be the M such that $u'_c(k,M)$, $u'_n(k,M)$ and $u'_b(k,M)$ intersect at the same point, i.e. $k^l(M^{med}) = k'^l(M^{med})$. We have $u'_c(k^l(M^{med}),M^{med}) = u'_n(k^l(M^{med}),M^{med}) = u'_b(k^*(M^{med}),M^{med})$. Since $T > v > s$ we can wisely choose α and β which are functions of σ_ϵ, σ_{ϵ_0}, σ_θ and m_θ to guarantee the existence of M^{med}, and it is easy to see that $M^{med} > M_E^{\min}$ and $M^{med} > M_C^{\min}$.

Since when M approaches $+\infty$, we have $k^l \to -\frac{t}{T-v}$ and $k'^l \to -\frac{q-I}{v-s}$. By $\frac{t}{T-v} > \frac{q-I}{v-s}$, we have $k^l < k'^l$. Then we can find that $\frac{\partial u'_n(k^l)-u'_b(k^l)}{\partial M} = -\phi((v-s)k^l+b+t)(\frac{\alpha-\frac{1-\Phi}{\phi}\frac{1}{k^l+\frac{b+I}{v-s}}}{\alpha-\frac{1-\Phi}{\phi}\frac{1}{k^l+\frac{b+t}{T-v}}} - 1) < 0$, and $u'_b(k^l) > u'_b(k'^l)$ so $k^l(M) - k'^l(M)$ decreases with M. Since $u'_c(k,M^{med}) = u'_n(k,M^{med}) = u'_b(k,M^{med})$, then we have

$$M > M^{med} \Rightarrow k^l < k'^l, \text{and} \quad M < M^{med} \Rightarrow k^l > k'^l \tag{6}$$ {appd:propert

.

Let $(x_0^{\min}, L^m)$ be the pair satisfying $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L^m})+\frac{\psi}{\sigma}x_0^{\min} = M_C^{\min}$ and $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L^m})+\frac{\psi}{\sigma}x_0^{\min} = M_E^{\min}$. In other words, $(x_0^{\min}, L^m)$ is the intersection of $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0 = M_C^{\min}$ and $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L})+\frac{\psi}{\bar{\sigma}}x_0 = M_E^{\min}$.

We next proof the first part of the proposition.

1. Sufficiency: For any (x_0, L) such that $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0 > M^{med}$ and $L \leq E+C$, then $M = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0$. Since $M > M^{med}$, then equation (4) does have finite solutions, we only consider the smaller solution which is denoted as k. Also $(1-\Phi(\alpha k-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})-\frac{\psi}{\bar{\sigma}}x_0-\beta))vk < (1-\Phi(\alpha k-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})-\frac{\psi}{\bar{\sigma}}x_0-\beta))(Tk+b+t)-b$, which is because $M > M^{med}$. Similar conclusions are hold when $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L})+\frac{\psi}{\bar{\sigma}}x_0 > M^{med}$ and $L > E+C$.

Next, we will show that there does not exist two thresholds pair (k^s, k^c) which is consistent with the cutoff equilibrium in this case. If there exists a pair (k^s, k^c) which is consistent with the equilibrium for given (x_0, L), it must satisfies:

$$(1-\Phi(\alpha k^c - M(x_0,L)-\beta))((T-v)k^c+b+t)-b=0, \tag{7}$$
{PP:two-cut e

$$(1-\Phi(\rho k^c - \eta k^s - M(x_0,L)-\beta))((v-s)k^s+q)-I=0, \tag{8}$$
{PP:two-cut e

$$k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) \geq k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L}), \tag{9}$$
{PP:two-cut e

$$k^c \geq k^s, \tag{10}$$
{PP:two-cut e

where $\rho = \frac{1}{\bar{\sigma}}$, $\eta = \frac{(1-\psi)\lambda}{\bar{\sigma}}$ and $M = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0$;[25] or

$$(1-\Phi(\rho k^s - \eta k^c - M(x_0,L)-\beta))((T-v)k^c+b+t)-b=0, \tag{11}$$
{PP:two-cut e

$$(1-\Phi(\alpha k^s - M(x_0,L)-\beta))((v-s)k^s+q)-I=0, \tag{12}$$
{PP:two-cut e

$$k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) < k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L}), \tag{13}$$
{PP:two-cut e

$$k^c \geq k^s, \tag{14}$$
{PP:two-cut e

where $M = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L})+\frac{\psi}{\bar{\sigma}}x_0$.[26]

[25] These system of equations are based on the assumption that $\bar{\theta} = k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L})$, i.e. the maintaining secrecy condition dominates the participation condition, the equilibrium conditions should be: Equation (7) indicates a type k^c follower should be indifferent between joining the rebellion and remaining neutral. Equation (8) indicates that a type k^s follower must be indifferent between remaining neutral and turning in the leader. Inequality (9) indicates that k^c and k^s solved from (7) and (8) must be consistent with the assumption that the participation condition dominates the maintaining secrecy condition. Inequality (10) is needed to indicate that k^c should be the threshold that separates the acts of joining a rebellion and of remaining neutral, whereas k^s must be the threshold that isolates the actions of remaining neutral and of turning the leader in.

[26] These system of equations are based on the assumption that $\bar{\theta} = k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$, i.e. the maintaining secrecy condition dominates the participation condition, the equilibrium conditions should be:

Without loss of generality, we assume (k^s, k^c) satisfies (7)-(10) with $M = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$. (7) indicates that $u_c'(k^c, M) = u_n'(k^c, M)$. Now we use k^c replace k^s in (8), then we have

$$(1 - \Phi(\alpha k^c - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1 - \frac{C}{L}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta))((v-s)k^c + q) - I > 0,$$

it is because the left hand side of (8) is a increasing function of k^s and $k^c > k^s$. Then we have $u_n'(k^c, M) > u_b'(k^c, M)$ which is a contradiction when $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) - \frac{\psi}{\bar{\sigma}}x_0 > M^{med}$. Therefore, there is no (k^c, k^s) with $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) \geq k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$ satisfying (7)-(10). Similarly it is easy to show that there is no (k^s, k^c) satisfying (11)-(14).

Necessity: When one cutoff threshold $k(x_0, L)$ exist, then equation (4) must hold. Similar as the proof in Proposition 2, we have $M^{med} < \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$ when $L \leq E + C$ and $M^{med} < \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0$ when $L > E + C$.

Next, we focus on the second part of the proposition.

2. Sufficiency: For any (x_0, L) with $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0 > M_C^{\min}$ and $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0 > M_E^{\min}$, let $k^c(x_0, L)$ and $k^s(x_0, L)$ be the solutions for equation (7) and (12) respectively.

For this fixed x_0, if $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) > k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$, then when we choose a larger L, the solution $k^s(x_0, L)$ will increase and $k^c(x_0, L)$ will decrease, therefore the left hand side of this inequality will decrease and the right hand side of this inequality will increase. We continuously increase L until $L = L'$ such that $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L'}) = k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L'})$. That L' exists because $\Phi^{-1}(1-\frac{C}{L})$ approaches $+\infty$ and $\Phi^{-1}(\frac{E}{L})$ approaches $-\infty$, and both $k^c(x_0, L)$ and $k^s(x_0, L)$ exists with finite value when L is sufficiently large.

Similarly, when $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) < k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$, we can decrease L to reach the equality, and let L'' be the solution to hold the equality.

We must have $L'(x_0) = L''(x_0)$, it is because any $k^c(x_0, L)$ satisfying (7) is monotonously decreasing with L, and $k^s(x_0, L)$ satisfying (12) is monotonously increasing with L, therefore $(k^c(x_0, L), k^s(x_0, L))$ satisfying (7), (12) and $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) = k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$ is unique for a given x_0.

For a given x_0 and L, first, we focus on the case when $M_C^{min} < \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0 < M^{med}$ and $L < L'(x_0)$.

{claim2}

Claim 1. *For given (x_0, L) When $M_C^{\min} < \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0 < M^{med}$ and the pair $(k^c(x_0, L), k^s(x_0, L))$ satisfies (7)-(9), it must satisfies (10).*

Proof of Claim 1: When $M_C^{\min} < \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0 < M^{med}$. Let k^c be the solution solved from (7). For a given k^c solved from (7), let k^s be the solution solved from (8), assume (10) is not satisfied.

When (k^c, k^s) satisfies $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) \geq k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$, we have

$$
\begin{aligned}
& (1-\Phi(\frac{k^c-(1-\psi)\lambda k^c}{\bar{\sigma}}-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})-\frac{\psi}{\bar{\sigma}}x_0-\beta))vk^c \\
= \;& u_n'(k^c) \\
> \;& u_b'(k^c) \\
= \;& (1-\Phi(\frac{k^c-(1-\psi)\lambda k^c}{\bar{\sigma}}-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})-\frac{\psi}{\bar{\sigma}}x_0-\beta))(sk^c-q)+I. \qquad (15)
\end{aligned}
$$

The first equality comes from that k^c is solved from (7) and the definition of u_n'. The second inequality comes from that $M=\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0<M^{med}$ and (6). The last equality is the definition of u_b'. Since

$$
(1-\Phi(\frac{k^c-(1-\psi)\lambda k^s}{\bar{\sigma}}-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})-\frac{\psi}{\bar{\sigma}}x_0-\beta))((v-s)k^s+q).
$$

is an increasing function of k^s, and we have

$$
(1-\Phi(\frac{k^c-(1-\psi)\lambda k^c}{\bar{\sigma}}-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})-\frac{\psi}{\bar{\sigma}}x_0-\beta))((v-s)k^c+q)>I.
$$

Therefore we have k^s solved from (8) is less than k^c solved from (7), which is a contradiction. □,

Define x_0^{med} satisfies $M^{med}=\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0^{med}=\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L})+\frac{\psi}{\bar{\sigma}}x_0^{med}$.

{claim3}

Claim 2. *For any (x_0,L), when $M_C^{\min}<\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0<M^{med}$, $x_0>x_0^{med}$, and $k^c(x_0,L)$, $k^s(x_0,L)$ satisfying (7)-(8), we have $k^c-\sigma_\epsilon\Phi^{-1}(1-\frac{C}{L})>k^s-\sigma_\epsilon\Phi^{-1}(\frac{E}{L})$.*

Proof of Claim 2: Assume we have $k^c-\sigma_\epsilon\Phi^{-1}(1-\frac{C}{L})\le k^s-\sigma_\epsilon\Phi^{-1}(\frac{E}{L})$. Since $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0<M^{med}$, then we can find L^a with $L^a>L$ such that $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L^a})+\frac{\psi}{\bar{\sigma}}x_0=M^{med}$. Then we have $k(x_0,L^a)=k^c(x_0,L^a)=k^s(x_0,L^a)$ satisfying

$$
(1-\Phi(\alpha k-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L^a})-\frac{\psi}{\bar{\sigma}}x_0-\beta))((T-s)k+q+b+t)-I-b=0 \qquad (16)
$$

$$
(1-\Phi(\alpha k^c-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L^a})-\frac{\psi}{\bar{\sigma}}x_0-\beta))((T-v)k^c+b+t)-b=0 \qquad (17)
$$

and

$$
(1-\Phi(\alpha k^s-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L^a})-\frac{\psi}{\bar{\sigma}}x_0-\beta))((v-s)k^s+q)-I=0 \qquad (18)
$$

with $k^c(x_0,L^a)-\sigma_\epsilon\Phi^{-1}(1-\frac{C}{L^a})>k^s(x_0,L^a)-\sigma_\epsilon\Phi^{-1}(\frac{E}{L^a})$.

By the continuity of the solution, there must exist a L^b such that $k^s(x_0,L^b)$, $k^c(x_0,L^b)$ satisfy (7) and (8) with $k^c(x_0,L^b)-\sigma_\epsilon\Phi^{-1}(1-\frac{C}{L^b})=k^s(x_0,L^b)-\sigma_\epsilon\Phi^{-1}(\frac{E}{L^b})$ and $L\le L^b<L^a$. We also have that $L^b<E+C$, therefore we must have that $k^c(x_0,L^b)<k^s(x_0,L^b)$ which is a contradiction with Claim 1. Therefore the claim is proved. □

Now we can claim that for any given x_0 and L with $x_0>x_0^{med}$ and $M_C^{\min}<\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0<M^{med}$, then two cutoff thresholds $(k^s(x_0,L),k^c(x_0,L))$ exist. It is because, when

$M_C^{\min} < \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, finite k^c and k^s can be solved from (7) and (8), and these solutions are consistent with the condition (9) and (10) due to Claim 1 and Claim 2.

Next, define $x_0^{\min}$ satisfies $M_C^{\min} = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0^{\min} = M_E^{\min} = \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0^{\min}$. For any (x_0, L), when $x_0^{\min} < x_0 < x_0^{med}$, we have the following claim:

{claim4}

Claim 3. *For any (x_0, L), when $M_C^{\min} < \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0 < M^{med}$, $x_0^{\min} < x_0 < x_0^{med}$, and $k^c(x_0, L)$, $k^s(x_0, L)$ satisfying (7) and (8), then we have $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) > k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$.*

Proof of Claim 3: For the given x_0 with $x_0^{\min} < x_0 < x_0^{med}$, we know there exists a $L'(x_0)$ such that $(k^s(x_0, L'(x_0)), k^c(x_0, L'(x_0)))$ satisfy (7) and (12) with $k^c(x_0, L'(x_0)) - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L'(x_0)}) = k^s(x_0, L'(x_0)) - \sigma_\epsilon\Phi^{-1}(\frac{E}{L'(x_0)})$.

$k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) = k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$ is not possible, because $L < L'(x_0)$ and $L'(x_0)$ is the unique L satisfying satisfy (7) and (12) with $k^c(x_0, L'(x_0)) - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L'(x_0)}) = k^s(x_0, L'(x_0)) - \sigma_\epsilon\Phi^{-1}(\frac{E}{L'(x_0)})$.

Now assume $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) < k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$. By the continuity of the solution, we can choose a L^c with $L' - L^c < \epsilon^c$ such that the solution $k^c(x_0, L^c)$ and $k^s(x_0, L^c)$ solved from (7) and (8) has the following property that $k^c(x_0, L^c) - k^c(x_0, L'(x_0)) > k^s(x_0, L^c) - k^s(x_0, L'(x_0))$. It means that we choose L^c smaller than but close enough to $L'(x_0)$ such that the increase of k^c is larger than the increase of k^s. Then we have $k^c(x_0, L^c) - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L^c}) > k^s(x_0, L^c) - \sigma_\epsilon\Phi^{-1}(\frac{E}{L^c})$.

Next, we need to show that this L^c can be found.

$$\frac{\partial k^s(x_0, L)}{\partial L} = \frac{\frac{1-\hat{\Phi}}{\hat{\phi}}\frac{1}{k^c + \frac{t+b}{T-v}} + \frac{(1-\psi)\lambda}{\bar{\sigma}}}{\frac{1-\tilde{\Phi}}{\tilde{\phi}}\frac{1}{k^s + \frac{q}{v-s}} + \frac{(1-\psi)\lambda}{\bar{\sigma}}}\frac{\partial k^c(x_0, L)}{\partial L}$$

where $\hat{\Phi} = \Phi(\alpha k^c - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L^a}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta)$, $\hat{\phi}$ is $\hat{\Phi}$'s density function and $\tilde{\Phi} = \Phi(\frac{k^s - (1-\psi)\lambda k^c}{\bar{\sigma}} - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta)$ and $\tilde{\Phi}$'s density function.

At $x_0, L'(x_0)$, $\hat{\Phi} = \tilde{\Phi}$ and $\hat{\phi} = \tilde{\phi}$. We also know $k^c(x_0, L'(x_0)) + \frac{t+b}{T-v} > k^s(x_0, L'(x_0)) + \frac{q}{v-s} > 0$. Therefore $|\frac{\partial k^s(x_0,L)}{\partial L}| < |\frac{\partial k^c(x_0,L)}{\partial L}|$ at $(x_0, L'(x_0))$, furthermore, it means there exists an ϵ^c such that any L satisfies $0 < L'(x_0) - L < \epsilon^c$ can be our L^c.

After that, since $k^c(x_0, L^c) - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L^c}) > k^s(x_0, L^c) - \sigma_\epsilon\Phi^{-1}(\frac{E}{L^c})$ and $k^c(x_0, L) - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) < k^s(x_0, L^c) - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$, by the continuity, we must have a L^d such that $k^s(x_0, L^d)$ and $k^c(x_0, L^d)$ satisfy (7) and (8) with $k^c(x_0, L^d) - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L^d}) = k^s(x_0, L^d) - \sigma_\epsilon\Phi^{-1}(\frac{E}{L^d})$. It is a contradiction with the uniqueness of $L'(x_0)$.

Therefore we have $k^c(x_0, L) - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) > k^s(x_0, L) - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$. □

So far, we prove that when $M_C^{\min} < \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0 < M^{med}$ and $L \leq L'(x_0)$, there exist $(k^s(x_0, L), k^c(x_0, L))$, which is consistent with the cutoff equilibrium.

The case that when $M_E^{min} < \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0 < M^{med}$ and $L > L'(x_0)$ is similar with the discussion above.

Necessity: When two cutoff thresholds $(k^s(x_0, L), k^c(x_0, L))$ exist, then x_0 and L must satisfy either $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0 < M^{med}$ or $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0 < M^{med}$, otherwise, there only exist one cutoff threshold case.

Also, for any x_0 and L with $M_C^{\min} < \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$ but $M_E^{\min} > \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, then k_c and k_s can only be solved from (7) and (8), however, these solutions cannot satisfy (9). Similarly, for any x_0 and L with $M_E^{\min} < \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$ but $M_C^{\min} > \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, then k_c and k_s can only be solved from (11) and (12), however, these solutions cannot satisfy (13). Finally, when $M_C^{\min} > \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$ but $M_E^{\min} > \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0$, then no finite k^c and k^s can be solved.

□

Proof of Corollary 2. It is straightforward from the proof of Proposition 4. □

Define some notations:

$$\Phi_1 = \Phi(\alpha k^c - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta),$$

and ϕ_1 is Φ_1's density function.

$$\Phi_2 = \Phi(\alpha k^s - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta),$$

and ϕ_2 is Φ_2's density function.

$$\Phi_3 = \Phi(\alpha k - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta),$$

and ϕ_3 is Φ_3's density function.

On the equilibrium path, k^c is solved from

$$(1-\Phi(\alpha k^c - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta))((T-v)k^c + t + b) - b = 0$$

By Implicit Function Theory,

$$\begin{aligned}\frac{\partial k^c}{\partial x_0} &= -\frac{1}{\frac{1-\Phi_1}{\phi_1}\frac{1}{k^c+\frac{t+b}{T-v}} - \alpha}(\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi_c'^{-1}\frac{C}{L^2}\frac{\partial L}{\partial x_0} + \frac{\psi}{\bar{\sigma}}) \\ &\equiv -\frac{1}{B_1}(\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi_c'^{-1}\frac{C}{L^2}\frac{\partial L}{\partial x_0} + \frac{\psi}{\bar{\sigma}})\end{aligned}$$

where $B_1 \equiv \frac{1-\Phi_1}{\phi_1}\frac{1}{k^c+\frac{t+b}{T-v}} - \alpha$

k^s is solved from

$$(1-\Phi(\alpha k^s - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta))((v-s)k^s + q) = I$$

By Implicit Function Theory,

$$\begin{aligned}\frac{\partial k^s}{\partial x_0} &= -\frac{1}{\frac{1-\Phi_2}{\phi_2}\frac{1}{k^s+\frac{q}{v-s}} - \alpha}(-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi_E'^{-1}\frac{E}{L^2}\frac{\partial L}{\partial x_0} + \frac{\psi}{\bar{\sigma}}) \\ &\equiv -\frac{1}{B_2}(-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi_E'^{-1}\frac{E}{L^2}\frac{\partial L}{\partial x_0} + \frac{\psi}{\bar{\sigma}})\end{aligned}$$

where $B_2 \equiv \frac{1-\Phi_2}{\phi_2}\frac{1}{k^s+\frac{q}{v-s}} - \alpha$
k is solved from

$$(1-\Phi(\alpha k - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta))((T-s)k+t+b+q) = b+I$$

By Implicit Function Theory,

$$\begin{aligned}
\frac{\partial k}{\partial x_0} &= -\frac{1}{\frac{1-\Phi_3}{\phi_3}\frac{1}{k+\frac{t+b+q}{T-s}} - \alpha}(\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi_c'^{-1}\frac{C}{L^2}\frac{\partial L}{\partial x_0} + \frac{\psi}{\bar{\sigma}}) \\
&\equiv -\frac{1}{B_3}(\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi_c'^{-1}\frac{C}{L^2}\frac{\partial L}{\partial x_0} + \frac{\psi}{\bar{\sigma}})
\end{aligned}$$

where $B_3 \equiv \frac{1-\Phi_3}{\phi_3}\frac{1}{k+\frac{t+b+q}{T-s}} - \alpha$.

Proof of Proposition 5. For any given $x_0 > x_0^{med}$, if there exist L such that $(k^s(L, x_0), k^c(L, x_0))$, by Proposition 4 we must have either $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) > k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$ or $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) < k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$.

When $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) > k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$, we have $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0 < M^{med}$

we know $bar\theta(x_0, L) = k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L})$, therefore the leader's utility u_0 is an increasing function of L. When L increases such that $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L}) + \frac{\psi}{\bar{\sigma}}x_0 = M^{med}$, then by the continuity of the solutions we have $k^s = k^c = k$. After that, when L continuously increasing, there only exists one cutoff threshold $k(x_0, L)$ which is consistent with the cutoff equilibrium and $\bar{\theta}(x_0, L) = k - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L})$, then u_0 is still an increasing function of L until $L = E + C$.

Similarly, when $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) < k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$, we have $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0 < M^{med}$. When $\bar{\theta}(x_0, L) = k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$, the leader's utility is an decreasing function of L. When L decreases such that $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L}) + \frac{\psi}{\bar{\sigma}}x_0 = M^{med}$, then by the continuity of the solutions we have $k^s = k^c = k$. After that, when L continuously decreasing, there only exists one cutoff threshold $k(L, x_0)$ which is consistent with the cutoff equilibrium and $\bar{\theta}(x_0, L) = k - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$, then the leader's utility is still an decreasing function of L until $L = E + C$.

Therefore the best choice of L is $E + C$.

For a given x_0 with $x_0^{\min} < x_0 < x_0^{med}$, we know there is no L such that one-cutoff strategy exists. Therefore, the optimal coalition size under two-cutoff case is L such that $k^c(x_0, L) - \sigma_\epsilon(\Phi^{-1}(1-\frac{C}{L})) = k^s(x_0, L(x_0)) - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$, through which $\bar{\theta}(L, x_0)$ can reach the minimal point and the leader's utility achieves the maximal point.

On the equilibrium path, since $k^c - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L}) = k^s - \sigma_\epsilon\Phi^{-1}(\frac{E}{L})$, we rewrite the equation as $k^c - k^s = \sigma_\epsilon(\Phi^{-1}(1-\frac{C}{L}) - \Phi^{-1}(\frac{E}{L}))$ which is greater than 0. It implies $L > E + C$ when $x_0 > x_0^{med}$. Since when $x_0 = x_0^{med}$, the best L is $E + C$ on the equilibrium path, and $L'(x_0)$ is continuous and differential with respect to x_0 when $x_0 < x_0^{med}$, then we know $\frac{\partial L'(x_0)}{\partial x_0} < 0$ in a small interval $(x_0^{med} - \epsilon, x_0^{med}]$.

Assume $\frac{\partial L(x_0)}{\partial x_0} < 0$ is not always true, then there exists x_0' such that $\frac{\partial L'(x_0)}{\partial x_0} = 0$. Furthermore $\frac{\partial k^c - k^s}{\partial x_0} = \frac{\partial}{\partial x_0}(\sigma_\epsilon(\Phi^{-1}(1-\frac{C}{L}) - \Phi^{-1}(\frac{E}{L}))) = 0$ at x_0'. Also we have $\frac{\partial k^c}{\partial x_0} = -\frac{1}{B_1}\beta$ and $\frac{\partial k^s}{\partial x_0} = -\frac{1}{B_2}\beta$.

However, since $k^c > k^s$, we have $\frac{1-\Phi_1}{\phi_1} < \frac{1-\Phi_2}{\phi_2}$, then $B_1 < B_2$. So at x_0', $\frac{\partial k^c - k^s}{\partial x_0} \neq 0$. It is a contradiction. □

Proof of Proposition 6. Similar as Proposition 3, it is because leader's utility function increases in x_0. □

When $x_0 < x_0^{med}$ in TP, the leader would select $L^*(x_0)$ to stage a rebellion and the threshold for the followers would be $k^s(L^*(x_0), x_0)$ and $k^c(L^*(x_0), x_0)$ on the equilibrium path. The thresholds satisfy the following equations:

$$(1-\Phi(\alpha k^c(L^*, x_0) - \tau\Phi^{-1}(1-\frac{C}{L^*}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta))((T-v)k^c(L^*, x_0) + b + t) = b, \tag{19}$$ {PP: eq cond1

$$(1-\Phi(\alpha k^s(L^*, x_0) - \tau\Phi^{-1}(\frac{E}{L^*}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta))((v-s)k^s(L^*, x_0) + q) = I, \tag{20}$$ {PP: eq cond1

$$k^c(L^*, x_0) - \sigma_\epsilon\Phi^{-1}(1-\frac{C}{L^*}) = k^s(L^*, x_0) - \sigma_\epsilon\Phi^{-1}(\frac{E}{L^*}). \tag{21}$$ {PP: eq cond1

When $x_0 = x_0^{med}$, the leader chooses $L^* = E + C$ and we have $k^c(L^*, x_0) = k^s(L^*, x_0)$. With the same $x_0 = x_0^{med}$ in CP, the leader would also choose $L^* = E + C$ to stage a rebellion and the cutoff threshold for the followers satisfies the following equation.

$$(1-\Phi(\alpha k(L^*, x_0) - \tau\Phi^{-1}(1-\frac{C}{L^*}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta))[(T-s)k(L^*, x_0) + t + b + q] = I + b, \tag{22}$$

we have $k^c(E+C, x_0^{med}) = k^s(E+C, x_0^{med}) = k(E+C, x_0^{med})$.

Proof of Proposition 7. On the equilibrium path, when $T - v = a$, for any given ϵ_0, there exist a T_1 and a v_1 such that when $v > v_1$, then $-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi(\frac{E}{L})'^{-1}\frac{E}{(E+C)^2}\frac{\partial L}{\partial x_0} > \epsilon_0 > 0$. It is because for given T and v, $\frac{\partial L}{\partial x_0} > 0$ at $x_0 < x_0^{med}$; and when v increases k^s decreases, and $k^c - k^s$ increases, therefore $\frac{\partial L}{\partial x_0}$ is an increasing function of v on the equilibrium path when $x_0 < x_0^{med}$. Furthermore when $x_0 = x_0^{med}$, $\frac{\partial L}{\partial x_0 \partial v} > 0$.

When $v > v_1$, we have

$$\begin{aligned}
& \frac{1}{B_2}(-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi_E'^{-1}\frac{E}{L^2}\frac{\partial L}{\partial x_0} + \frac{\psi}{\bar{\sigma}}) - \frac{1}{B_3}\frac{\psi}{\bar{\sigma}}; \\
> \; & \frac{1}{B_2}(\epsilon_0 + \frac{\psi}{\bar{\sigma}}) - \frac{1}{B_3}\frac{\psi}{\bar{\sigma}}; \\
= \; & \frac{1}{B_2}\epsilon_0 + (\frac{1}{B_2} - \frac{1}{B_3})\frac{\psi}{\bar{\sigma}};
\end{aligned} \tag{23}$$ {Sur:eq: PP>

There exists a v_2 such that when $v > v_2$ we have $0 < (\frac{1}{B_3} - \frac{1}{B_2}) < \epsilon$ and $\frac{1}{B_2}\epsilon_0 + (\frac{1}{B_2} - \frac{1}{B_3})\frac{\psi}{\bar{\sigma}} > \frac{1}{B_2}\epsilon_0 - \epsilon\frac{\psi}{\bar{\sigma}} > 0$

Therefore when $v > \bar{v} \equiv \max\{v_1, v_2\}$, we have that $\frac{\partial k^s}{\partial x_0} < \frac{\partial k}{\partial x_0} < 0$ at x_0^{med}. By the continuity, we have $k^s > k$ at least in a small interval $(x_0^{med} - \epsilon', x_0^{med}]$.

Start from any $x_0 < x_0^{med}$ with $k^s(x_0, L^*(x_0)) > k(x_0, E + C)$, we have $k^s(x_0, L^*(x_0)) - \sigma\Phi^{-1}(\frac{E}{L(x_0)}) > k(x_0, E + C) - \sigma\Phi^{-1}(\frac{E}{E+C})$ and $\alpha k^s(x_0, L^*(x_0)) - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L^*(x_0)}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta > \alpha k - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{E+C}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta$. Therefore $\frac{1-\Phi_2}{\phi_2} < \frac{1-\Phi_3}{\phi_3}$ and $B_2 < B_3$ (under the condition $v > v_2$ and $T - v = a > 0$). As a results $\frac{\partial k^s}{\partial x_0} < \frac{\partial k}{\partial x_0}$. It means once $x_0 < x_0^{med}$, $k^s > k$ always. Furthermore $\bar{\theta}^{TP} = k^s(x_0, L^*(x_0)) - \sigma\Phi^{-1}(\frac{E}{L^*(x_0)}) > \bar{\theta}^{CP} = k(x_0, E+C) - \sigma\Phi^{-1}(\frac{E}{E+C})$ Therefore the leader has a higher starting point in TP i.e. $x_0^{*CP} < x_0^{*TP}$.

Since when $x_0 > x_0^{med}$, equilibrium results are the same in CP and TP, so $\hat{x}^{CP}(\theta)$ and $\hat{x}^{TP}(\theta)$ are identical when they are above x_0^{med}. Let $\hat{\theta}$ be the point such that $\hat{x}_0^{TP}(\bar{\theta}) = x_0^{med}$. When $\theta > \hat{\theta}$, since for any given x_0 with $x_0^{*CP} < x_0^{*TP} < x_0^{med}$, $k^s(x_0, L^*(x_0)) - \sigma\Phi^{-1}(\frac{E}{L(x_0)}) > k(x_0, E + C) - \sigma\Phi^{-1}(\frac{E}{E+C})$, it means $\hat{x}_0^{CP}(\theta) < \hat{x}_0^{TP}(\theta)$.

Therefore, TP has higher survival probability in this case when $\theta > \hat{\theta}$. □

Proof of Proposition 8. The proof is similar as the proof of Proposition 7, we just give a sketch as follow. Let $t + I$ and $v - s$ smaller enough to guarantee $\frac{1}{B_1} - \frac{1}{B_3}$ smaller enough, then since $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi'^{-1}_C \frac{C}{L^2}\frac{\partial L}{\partial x_0} + \frac{\psi}{\bar{\sigma}} < \frac{\psi}{\bar{\sigma}}$. Then we have $\frac{\partial k^c}{\partial x_0} < \frac{\partial k}{\partial x_0}$ at x_0^{med}. $\frac{q-I}{v-s} = h$ is to guarantee $\frac{\partial L}{\partial x_0}$ has a uniform low bound which is greater than 0, when $q - I$ and $v - s$ are smaller enough. By the continuity, we have $k^c < k$ at least in a small interval $(x_0^{med} - \epsilon'', x_0^{med}]$. Start from any $x_0 < x_0^{med}$ with $k^c(x_0, L^*(x_0)) < k(x_0, E+C)$, we have $k^c(x_0, L^*(x_0)) - \sigma\Phi^{-1}(1 - \frac{C}{L(x_0)}) < k(x_0, E+C) - \sigma\Phi^{-1}(1 - \frac{C}{E+C})$ and $\alpha k^c(x_0, L^*(x_0)) - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1 - \frac{C}{L^*(x_0)}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta < \alpha k - \frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1 - \frac{C}{E+C}) - \frac{\psi}{\bar{\sigma}}x_0 - \beta$. Therefore $\frac{1-\Phi_1}{\phi_1} > \frac{1-\Phi_3}{\phi_3}$ and $B_1 > B_3$. As a results $\frac{\partial k^c}{\partial x_0} < \frac{\partial k}{\partial x_0}$. It means once $x_0 < x_0^{med}$, $k^c < k$ always. Furthermore $\hat{\theta}^{TP} = k^c(x_0, L^*(x_0)) - \sigma\Phi^{-1}(1 - \frac{C}{L^*(x_0)}) < \bar{\theta}^{CP} = k(x_0, E + C) - \sigma\Phi^{-1}(1 - \frac{C}{E+C})$. Consequently, the leader has a higher starting point in CP i.e. $x_0^{*CP} > x_0^{*TP}$.

Since when $x_0 > x_0^{med}$, equilibrium results are the same in CP and TP, so $\hat{x}^{CP}(\theta)$ and $\hat{x}^{TP}(\theta)$ are identical when they are above x_0^{med}. Let $\hat{\theta}'$ be the point such that $\hat{x}_0^{TP}(\bar{\theta}') = x_0^{med}$. When $\theta > \hat{\theta}'$, since for any given x_0 with $x_0^{*TP} < x_0^{*CP} < x_0^{med}$, $k^c(x_0, L^*(x_0)) - \sigma\Phi^{-1}(1 - \frac{C}{L(x_0)}) < k(x_0, E + C) - \sigma\Phi^{-1}(1 - \frac{C}{E+C})$, it means $\hat{x}_0^{CP}(\theta) > \hat{x}_0^{TP}(\theta)$.

Therefore, CP has higher survival probability in this case when $\theta > \hat{\theta}'$. □

Proof of Proposition 9. In TP, by Proposition 4, M^{med} is the M such that $u'_c(k, M)$, $u'_b(k, M)$, and $u'_b(k, M)$ intersect at the same point. Since $u'_c(k, M)$ is an increasing function of t, therefore M^{med} is an decreasing function of t. It means x_0^{med} decreases with t. When t is larger enough, then there will be a intersection of $u'_c(k, M)$, $u'_n(k, M)$ and $u'_c(k, M)$. It means in the large t case, when M decrease, $u'_n(k, M)$ and $u'_b(k, M)$ will separate before the intersection of $u'_c(k, M)$, $u'_n(k, M)$ bypass the intersection of $u'_n(k, M)$, $u'_b(k, M)$. Therefore there is only one threshold which is consistent with the cutoff equilibrium. □

Proof of Lemma 3. In this case, the type x_i follower's expected payoffs for remaining neutral under TP is

$$(1-\Phi(\frac{\bar{\theta}^{TP}(x_0,L)-(1-\psi)\lambda x_i-\psi x_0-(1-\psi)(1-\lambda)m_\theta}{\bar{\sigma}}))(vx_i-q).$$

The equilibrium strategy for the supporters can be calculated using the same way in Proposition 4. Let $k_3^s(x_0,L)$ and $k_3^c(x_0,L)$ be the two thresholds which are consistent with the cutoff equilibrium for the two-cutoff strategy. For a given x_0 on the equilibrium path, $k_3^s(x_0,L)$, $k_3^c(x_0,L)$ and the optimal $L_3(x_0)$ chosen by the leader must satisfy the following equations:

$$(1-\Phi(\alpha k^c-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})-\frac{\psi}{\bar{\sigma}}x_0-\beta))((T-v)k^c+b+t+q)-b=0 \quad (24) \quad \texttt{\{eq8\}}$$

$$(1-\Phi(\alpha k^s-\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L})-\frac{\psi}{\bar{\sigma}}x_0-\beta))(v-s)k^s-I=0 \quad (25) \quad \texttt{\{eq9\}}$$

$$k^c(x_0,L)-\sigma_\epsilon\Phi^{-1}(1-\frac{C}{L})=k^s(x_0,L)-\sigma_\epsilon\Phi^{-1}(\frac{E}{L}). \quad (26) \quad \texttt{\{eq10\}}$$

Then we have that $k_3^s(x_0,L_3(x_0))>k^s(x_0,L'(x_0))$, $k_3^c(x_0,L_3(x_0))<k^c(x_0,L'(x_0))$ and $L_3(x_0)<L'(x_0)$.

Let $M_3^{\max}$ be the M such that $u_c'(k,M)$, $u_n''(k,M)\equiv(1-\Phi(\alpha k-M-\beta))(vk-q)$ and $u_b'(k,M)$ intersect at the same point. We have $M_3^{\max}>M^{med}$ because $u_n''(k,M)<u_n'(k,M)$. Therefore, when $x_0^{med'}$ satisfy $\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\sigma}x_0^{med'}=\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L})+\frac{\psi}{\sigma}x_0^{med'}=M_3^{\max}$, then we have $x_0^{med'}<x_0^{med}$. Furthermore, we have $x_0^{\min'}>x_{0_{TP}}^{\min}$, where $x_0^{\min'}$ is the minimal x_0 to make the finite threshold exist which is consistent with the cutoff equilibrium. It is easy to find when q increases, $x_0^{med'}$ will decrease, and $x_{0_{TP}}^{\min}$ will increase. Therefore, there exists a threshold $\bar{q}$ such that when $q>\bar{q}$, then there is no two cutoff thresholds which is consistent with the cutoff equilibrium. □

Proof of Proposition 10. The proof is very similar to Lemma 4, instead of repeating the whole procedure, I describe the intuition of the proof here. Since $\frac{t}{T-v}<\frac{q-I}{v-s}$. When $M\to+\infty$, the smaller root of $\Delta u_{nb}'(k,M)$ converge to $-\frac{q-I}{v-s}$ which is less than the smaller root of $\Delta u_{cn}'(k,M)$ converging to $-\frac{t}{T-v}$. It implies that when M is large, there only exist two-cutoff strategies, and when M is small, there only exist one-cutoff strategies.

1. Let M_{TR}^{med} satisfy $u_c'(k,M_{TR}^{med})=u_c'(k,M_{TR}^{med})=u_c'(k,M_{TR}^{med})$, which represent the leader's influence threshold between the one-cutoff strategy and two-cutoff strategy. Let $x_{0_{TR}}^{med}$ satisfies $M_{TR}^{med}=\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{E+C})+\frac{\psi}{\sigma}x_{0_{TR}}^{med}$.

Since $(1-\Phi(\alpha k-M-\beta))((T-v)k+t+b+q)-b-I$ is single peaked and increases with M, so there exist a $M_{TR}^{\min}$ such that $(1-\Phi(\alpha k-M-\beta))((T-v)k+t+b+q)-b-I=0$ has unique root for k. Let $x_{0_{TR}}^{min}$ satisfies $M_{TR}^{\min}=\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{E+C})+\frac{\psi}{\bar{\sigma}}x_{0_{TR}}^{min}$.

For a given x_0 and L, to guarantee the existence of one-cutoff strategy, we must have $M_{TR}^{\min}\le M<M_{TR}^{med}$. Where, $M=\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(1-\frac{C}{L})+\frac{\psi}{\bar{\sigma}}x_0$ when $L<E+C$ and $M=\frac{\sigma_\epsilon}{\bar{\sigma}}\Phi^{-1}(\frac{E}{L})+\frac{\psi}{\bar{\sigma}}x_0$ when

$L > E + C$. So, M increases with L when $L < E + C$, decreases with L when $L > E + C$, and increases with x_0. Therefore, when $x_0 < x_{0_{TR}}^{\min}$, M is less than $M_{TR}^{\min}$ for sure and no finite cutoff equilibrium exist.

2. When $x_{0_{TR}}^{\min} \leq x_0 < x_{0_{TR}}^{med}$, the influence M can never achieve M_{TR}^{med} regardless of L, therefore only one-cutoff strategy exists. Furthermore, in one-cutoff strategy, the optimal coalition size is $E + C$ which balance the participation and maintaining secrecy conditions.

3. When $x_{0_{TR}}^{med} \leq x_0$, the influence M can exceed M_{TR}^{med}, which lead to the existence of two-cutoff strategy. For any given x_0 in this case, if L is chosen such that k_{TR}^c and k_{TR}^s exist which satisfy (7)-(10), then we can enlarge L to solve k_{TR}^c and k_{TR}^s such that (24)-(26) are satisfied. Similarly, when a given x_0 and L leading to k_{TR}^c and k_{TR}^s satisfying (11)-(14), then we can reduce L to solve k_{TR}^c and k_{TR}^s such that (24)-(26) are satisfied. Part 3 is proved.

4. Using the method to prove Proposition 5, it is easy to find that optimal $L_{TR}^*(x_0)$ increases with x_0. When $x_0 > x_{0_{TR}}^{med}$, since the lower bound of $k^s(x_0, L)$ is $\frac{q-I}{v-s}$ and the lower bound of $k^c(x_0, L)$ is $\frac{t}{T-v}$, there exists a upper bound of the optimal coalition size which is L_{TR}^{max} and less than or equal to 1. $\square$

References

Acemoglu, D., Ticchi, D., and Vindigni, A. (2010). A theory of military dictatorships. *American Economic Journal: Macroeconomics*, 2(1):1–42.

Anderson, L. (2011). Demystifying the arab spring: parsing the differences between tunisia, egypt, and libya. *Foreign Aff.*, 90:2.

Angeletos, G.-M., Hellwig, C., and Pavan, A. (2006). Signaling in a global game: Coordination and policy traps. *Journal of Political Economy*, 114:452–484.

Angeletos, G.-M., Hellwig, C., and Pavan, A. (2007). Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks. *Econometrica*, 75(3):711–756.

Baliga, S. and Sjostrom, T. (2012). The strategy of manipulating conflict. *The American Economic Review*, 102(6):2897–2922.

Belkin, A. and Schofer, E. (2003). Toward a structural understanding of coup risk. *Journal of Conflict Resolution*, 47(5):594–620.

Benney, J. (2013). Stability maintenance at the grassroots: Chinas weiwen apparatus as a form of conflict resolution.

Besley, T. and Robinson, J. A. (2010). Quis custodiet ipsos custodes? civilian control over the military. *Journal of the European Economic Association*, 8(2-3):655–663.

Bharara, P. (2007). Corporations cry uncle and their employees cry foul: Rethinking prosecutorial pressure on corporate defendants. *Am. Crim. L. Rev.*, 44:53.

Boix, C. and Svolik, M. W. (2013). The foundations of limited authoritarian government: Institutions, commitment, and power-sharing in dictatorships. *The Journal of Politics*, 75(02):300–316.

Bueno de Mesquita, E. (2005). The quality of terror. *American Journal of Political Science*, 49(3):515–530.

Bueno de Mesquita, E. (2010). Regime change and revolutionary entrepreneurs. *American Political Science Review*, 104(3):446–466.

Carlsson, H. and Van Damme, E. (1993). Global games and equilibrium selection. *Econometrica: Journal of the Econometric Society*, pages 989–1018.

Casper, B. A. and Tyson, S. A. (2012). Popular uprisings and the strategic risk of a coup. In *American Political Science Association annual meeting paper*.

Chen, C.-C. (1975). The japanese adaptation of the pao-chia system in taiwan, 1895–1945. *The Journal of Asian Studies*, 34(02):391–416.

Chodakiewicz, M. J. (2004). *Between Nazis and Soviets: occupation politics in Poland, 1939-1947*. Lexington Books.

Chwe, M. S.-Y. (2000). Communication and coordination in social networks. *The Review of Economic Studies*, 67(1):1–16.

Chwe, M. S.-Y. (2013). *Rational ritual: Culture, coordination, and common knowledge*. Princeton University Press.

Decalo, S. (1976). *Coups and army rule in Africa: studies in military style*. yale university press New Haven.

DeNardo, J. (2014). *Power in numbers: The political strategy of protest and rebellion*. Princeton University Press.

Dewan, T. and Myatt, D. P. (2007). Leading the party: Coordination, direction, and communication. *American Political Science Review*, 101(04):827–845.

Dewan, T. and Myatt, D. P. (2008). The qualities of leadership: Direction, communication, and obfuscation. *American Political Science Review*, 102(03):351–368.

Echenique, F. and Edlin, A. (2004). Mixed equilibria are unstable in games of strategic complements. *Journal of Economic Theory*, 118(1):61–79.

Edmond, C. (2011). Information manipulation, coordination, and regime change. Technical report, National Bureau of Economic Research.

Egorov, G., Guriev, S., and Sonin, K. (2009). Why resource-poor dictators allow freer media: a theory and evidence from panel data. *American Political Science Review*, 103(04):645–668.

Finkel, S. E., Muller, E. N., and Opp, K.-D. (1989). Personal influence, collective rationality, and mass political action. *The American Political Science Review*, pages 885–903.

Galetovic, A. and Sanhueza, R. (2000). Citizens, autocrats, and plotters: A model and new evidence on coups d'etat. *Economics & Politics*, 12(2):183–204.

Galtung, J. (1967). On the effects of international economic sanctions, with examples from the case of rhodesia. *World Politics*, 19(03):378–416.

Gasiorowski, M. J. (1996). An overview of the political regime change dataset. *Comparative Political Studies*, 29(4):469–483.

Geddes, B. (1999). What do we know about democratization after twenty years? *Annual Review of Political Science*, 2(1):115–144.

Geddes, B., Wright, J., and Frantz, E. (2014). Autocratic breakdown and regime transitions: a new data set. *Perspectives on Politics*, 12(02):313–331.

Ginkel, J. and Smith, A. (1999). So you say you want a revolution a game theoretic explanation of revolution in repressive regimes. *Journal of Conflict Resolution*, 43(3):291–316.

Harden, B. (2008). Escapee tells of horrors in north korean prison camp. *Washington Post December*, 11:2008.

Howard, P. N., Duffy, A., Freelon, D., Hussain, M., Mari, W., and Mazaid, M. (2011). Opening closed regimes: what was the role of social media during the arab spring?

Hughes, L. (1985). Sofiya alekseyevna and the moscow rebellion of 1682. *The Slavonic and East European Review*, pages 518–539.

Jia, L. (2008). Qinhan shiqi zuxing lunkao [the experiment on family punishment in qin and han dynasty]. *Journal of Capital Normal University*, (In Chinese):11–17.

Kaempfer, W. H. and Lowenberg, A. D. (1988). The theory of international economic sanctions: A public choice approach. *The American Economic Review*, pages 786–793.

Katz, P. R. (2005). Governmentality and its consequences in colonial taiwan: A case study of the ta-pa-ni incident of 1915. *The Journal of Asian Studies*, 64(02):387–424.

Khawaja, M. (1993). Repression and popular collective action: Evidence from the west bank. In *Sociological Forum*, volume 8, pages 47–71. Springer.

Kim, N. K. (2014). Revisiting economic shocks and coups. *Journal of Conflict Resolution*, page 0022002713520531.

Levinson, D. J. (2003). Collective sanctions. *Stanford Law Review*, pages 345–428.

Li, H. and Ran, J. (2005). *Baojia System in Republic of China era*. Sichuan University Press.

Lichbach, M. I. (1998). *The rebel's dilemma*. University of Michigan Press.

Little, A. T. (2014). Coordination, learning, and coups.

Lohmann, S. (1994). The dynamics of informational cascades. *World politics*, 47(1):42–101.

Londregan, J. B. and Poole, K. T. (1990). Poverty, the coup trap, and the seizure of executive power. *World politics*, 42(02):151–183.

Luttwak, E. (1979). *Coup d'etat*. Duke Univ Press.

Major, S. and McGann, A. J. (2005). Caught in the crossfire innocent bystanders as optimal targets of economic sanctions. *Journal of Conflict Resolution*, 49(3):337–359.

Marinov, N. and Goemans, H. (2013). Coups and democracy. *British Journal of Political Science*, pages 1–27.

Migdal, J. S. (1974). *Peasants, politics, and revolution: pressures toward political and social change in the Third World*. Princeton University Press Princeton.

Morris, S. and Shin, H. S. (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review*, pages 587–597.

Morris, S. and Shin, H. S. (2001). Rethinking multiple equilibria in macroeconomic modeling. In *NBER Macroeconomics Annual 2000, Volume 15*, pages 139–182. MIT PRess.

Morris, S. and Shin, H. S. (2002). Social value of public information. *The American Economic Review*, 92(5):1521–1534.

Morris, S. and Shin, H. S. (2003). Global games: theory and applications. *Advances in Economics and Econometrics (Proceedings of the Eighth World Congress of the Econometric Society)*, 35:56–114.

Morris, S. and Shin, H. S. (2004). Coordination risk and the price of debt. *European Economic Review*, 48(1):133–153.

Olson Jr, M. (1965). Logic of collective action: public goods and the theory of groups, the.. *Harvard Economic Studies*.

Persson, T. and Tabellini, G. (2009). Democratic capital: The nexus of political and economic change. *American Economic Journal: Macroeconomics*, pages 88–126.

Popkin, S. L. (1979). *The rational peasant: The political economy of rural society in Vietnam*. Univ of California Press.

Powell, J. (2012). Determinants of the attempting and outcome of coups détat. *Journal of Conflict Resolution*, 56(6):1017–1040.

Powell, J. M. and Thyne, C. L. (2011). Global instances of coups from 1950 to 2010 a new dataset. *Journal of Peace Research*, 48(2):249–259.

Shadmehr, M. and Bernhardt, D. (2011). Collective action with uncertain payoffs: Coordination, public signals, and punishment dilemmas. *American Political Science Review*, 105(04):829–851.

Sinpeng, A. (2007). Democracy from above: Regime transition in the kingdom of bhutan. *Journal of Bhutan Studies*, 17:21–47.

Svolik, M. W. (2009). Power sharing and leadership dynamics in authoritarian regimes. *American Journal of Political Science*, 53(2):477–494.

Svolik, M. W. (2013). Contracting on violence the moral hazard in authoritarian repression and military intervention in politics. *Journal of Conflict Resolution*, 57(5):765–794.

Swartz, D. (2011). Jasmine in the middle kingdom: autopsy of chinas (failed) revolution. *American enterprise institute, April, http://www. aei. org/docLib/AO-2011-04-No-1-g. pdf. Accessed*, 15.

Tullock, G. (1971). The paradox of revolution. *Public Choice*, 11(1):89–99.

Tullock, G. (1974). *The social dilemma: The economics of war and revolution*. University publications Blacksburg.

Twitchett, D. C., Fairbank, J. K., Feuerwerker, A., Peterson, W. J., Liuv, K.-C., and MacFarquhar, R. (1978). *The Cambridge History of China*, volume 1991. Cambridge Univ Press.

Tyson, S. and Smith, A. (2013). Regime insiders and revolutions. In *Midwest Political Science Association Annual Meeting Paper*.

Wickham-Crowley, T. P. (1992). *Guerrillas and revolution in Latin America: A comparative study of insurgents and regimes since 1956*. Princeton University Press.

Xue, Y. (2005). The study on pao-chia system during japanese occupation period in taiwan and discuss the difference between cross-strait pao-chia systems.