The Power of Whispers:
A Theory of Rumor, Communication and Revolution*

HENG CHEN
University of Hong Kong
YANG K. LU
Hong Kong University of Science and Technology
WING SUEN
University of Hong Kong

December 31, 2013

Abstract. We study how rumors mobilize individuals who take collective action. Rumors may or may not be informative, but they create public topics on which people can exchange their views. Individuals with diverse private information rationally evaluate the informativeness of rumors about regime strength. A rumor against the regime can coordinate a larger mass of attackers if individuals can discuss its veracity than if they cannot. Communication can be so effective that a rumor can have an even greater impact on mobilization than when the same story is fully trusted by everybody. However, an extreme rumor can backfire and discourage mobilization.

Keywords. global game, public signals, swing population, mixture distribution, censorship

JEL Classification. D74, D83

*We thank Roland Bénabou, Matthias Doepke, Christian Hellwig, Satoru Takahashi, Yikai Wang, Fabrizio Zilibotti, and seminar participants at NBER Summer Institute, UBC-HKU Microeconomics Workshop, University of Hong Kong, Toulouse School of Economics, European University Institute and National University of Singapore for their suggestions. Guangyu Pei deserves praise for providing excellent research assistance to this project. Our research is partially funded by the Research Grants Council of Hong Kong (Project No. HKU 742112B).

1. Introduction

Collective actions, such as riots, currency attacks and bank runs, are often immersed in rumors. Perhaps the most dramatic place to witness rumors in action is a political revolution. Amid the recent Tunisian revolution, Ben Ali, the ex-Tunisian leader, was said to have fled his country. This was confirmed after conflicting rumors about his whereabouts, and finally led to the end of street protests. A while later in Egypt, it was widely reported that Mubarak's family had left for London, which was believed by many as a clear sign of fragility of the regime. Similar rumors about Qaddafi and his family appeared in Libya when the battle between the opposition and the regime intensified. Rumors are not unique to the series of revolutions in the Arab Spring. During the 1989 democracy movement in China, rumors repeatedly surfaced about the death of the leaders, Deng Xiaoping and Li Peng, as well as the divide among communist leaders.[1]

Are rumors just rumors? In many cases, yes. Rumors that spread during turmoils often quickly disappear without leaving a trace. That seems to be natural, as rational individuals may discount unreliable information they receive in those situations. However many historical incidents suggest that rumors often turn out to be particularly effective in mobilization. The Velvet Revolution in Czechoslovakia was described as a "revolution with roots in a rumor" (Bilefsky 2009). At the dawn of the revolution, a prominent (false) rumor that a 19-year old student was brutally killed by the police triggered many otherwise hesitant citizens to take to the streets. The revolution gained huge momentum right after that and the regime collapsed a few days later. In the Arab Spring, the news about Mubarak's family proved to be false, yet the opposition credited it for "mark[ing] a new phase" in their campaign.[2] Chinese history also offers many anecdotes in which rumors mobilized mass participation, including the Boxer Uprising, the Republican Revolution, and the May Fourth Movement (Zhang 2009). Similarly, riots are often amplified or even sparked by rumors as well: the 1921 Tulsa race riot, the 1967 Newark riot and the 2004 Rome riot provide dramatic examples.

A common interpretation of the role of rumors in mass movements is that individuals are just blindly herded by rumors. However we adopt the position that individuals are fully aware that rumors circulating in times of turmoil may or may not be well

[1]There were widespread rumors of many variants that Deng died of illness during the protest and that Li was gun shot to death. It was also widely rumored that some senior leaders in the Communist Party wrote an open letter to oppose taking any action against students. Some of these rumors were repeated in the print media. See, for example, the report in the daily newspaper *Ming Pao* on June 6, 1989.

[2]*World Tribune* reported on January 28, 2011 that "confirmed by a Western diplomat, . . . Mubarak's wife, Suzanne, son, Gamal, and granddaughter arrived in London on a private jet as Egypt's defense minister secretly flew to the United States."

founded, and that they update their beliefs in a Bayesian manner. Since rumors are widely circulated and commonly observed, they may serve as a coordination device just like a public signal in a coordination game. We explain why some rumors are effective in mobilizing participation in collective actions while others are not.

In this paper, we focus on two key aspects of rumors: that they may be true or not true, and that people talk about them. Individuals in times of uncertainty and crisis often seek others' opinions and discuss with peers about their judgment and evaluation of rumors. Information from fellow citizens can influence their beliefs and even actions. The core of our paper is to show that communication among individuals centering around a public topic can substantially change outcomes of collective actions.

Specifically, we model political revolution as a global game. Citizens are uncertain about the regime's strength and possess dispersed private information about it. In this model, global strategic complementarities arise because a citizen's incentive to revolt increases with the aggregate action of all other citizens. If the number of participants is sufficiently high, the regime collapses; otherwise it survives. Before citizens take actions, they hear a rumor about the regime. This rumor is a publicly observed message, which could be either an informative signal about the regime's strength or an uninformative noise unrelated to fundamentals. Citizens assess the informativeness of the rumor based on their private information. As a consequence of diverse private information among citizens, their assessments may also differ. Further, citizens communicate with one another and tell their peers whether they believe the rumor or not.

In this model, the degree of skepticism is endogenous: citizens whose private information differs more from the rumor are more skeptical of it. Due to this skepticism, rumors against the regime mobilize less attackers than when such news is known to be trustworthy. If a rumor is far different from the fundamental, it will also differ from most citizens' private information and therefore be heavily discounted by them in their belief updating. As a result, extreme rumors have little impact on equilibrium outcomes.

When citizens communicate, those whose private information is close to the rumor will tell their peers that it is informative. Recipients of confirmatory messages treat what their peers say as evidence for the truth being close to what the rumor suggests, and therefore become more responsive to the rumor. Consider, for example, a rumor against the regime. A fraction of the population (those with intermediate private information) will attack the regime if their peers tell them that the rumor is informative, and will not attack otherwise. If the rumor is indeed near the true strength, more citizens will receive confirmatory messages from their peers. Therefore, communication helps such a rumor to mobilize more attackers. By the same mechanism, if the rumor is far from the truth, most citizens will express disbelief to their peers, which discourages attacking. Therefore, communication overcomes or reinforces skepticism about a rumor, depending on whether the rumor is close to the fundamental or not.

Interestingly and surprisingly, we find that communication could make rumors even more powerful than trustworthy news in mobilizing individuals. Once citizens are allowed to communicate with each other, it is possible that the regime would survive if all of them believe that a rumor against the regime is trustworthy or informative, but would collapse if they believe the rumor may be uninformative.

Given that rumors can threaten the survival of the regime, it is natural to think the regime may increase its chance of survival by blocking any public information against it. We investigate this conjecture using our model and find that censorship does not necessarily help the regime to survive. That is because citizens would interpret the absence of rumors as a sign of bad news for the regime.

Our work enriches the global games literature (Morris and Shin 2003) in a couple of directions. We offer a specification of public signals using mixture distributions that allows us to capture people's skepticism. Specifically, by allowing an additional layer of Bayesian updating on the "quality" of the information source, our model captures the fact that individuals tend to discount information which is too different from their priors.[3] We will also show why this specification is qualitatively different from having a public signal with low precision.[4] In our model, the dispersion in private information is crucial, not only because it guarantees equilibrium uniqueness, but also because it generates diverse assessments on the informativeness of rumors, which provides a ground for the study of communication between citizens. It is the core of this paper.

In much of the global games literature, citizens are assumed to only respond to the signals they observe; any further interactions among citizens are often left out for simplicity. In reality, individuals do exchange information with each other before they make decisions and take actions. This is especially true in collective actions such as protests, demonstrations and revolutions. We model direct interaction between citizens by allowing them to communicate privately, rather than just observing a public signal of what others are doing.[5]

[3]Gentzkow and Shapiro (2006) show that individuals tend to believe that a news source is of high quality if it conforms to their prior expectations. Also see Suen (2010) for a model with similar features.

[4]A common implicit assumption in the literature is that citizens believe the public signal is informative. They assign a constant weight on the public signal based on its relative precision to private signals, and would not adjust the weight even though the public signal is remarkably different from what their own private information suggests.

[5]Angeletos and Werning (2006) explicitly acknowledge the importance of direct interaction between agents in coordination models. They allow agents to observe a public signal about the aggregate attack, which conveniently approximates the situation where agents could learn about the actions of others.

This paper should not be interpreted as contradicting the literature that stresses structural factors as root causes for a revolution (Skocpol 1979). Structural factors, such as the state of the economy and international pressure, are those that make a society "ripe" for revolution. However, it has been noted that structural factors are not sufficient for a successful revolution. In line with Bueno de Mesquita (2010), we argue that some random factors also play a role in determining the fate of a revolution. In our model, the realization of rumors serves as a source of randomness.

Our work is related to a small economics literature on rumors, e.g., Banerjee (1993) and Bommel (2003). Unlike their models, in which a rumor is passed on to others sequentially, we provide a model in a static setting, in which a rumor is heard by citizens simultaneously. We focus on the effect of communication among citizens about the rumor rather than the transmission of the rumor itself.

This paper also contributes to a growing literature on revolutions in economics. Edmond (2011) considers a coordination game where citizens' private information about the regime's strength is contaminated by the regime's propaganda. Our model differs in that private information is uncontaminated, but the public signal may be false and unrelated to fundamentals. Both Bueno de Mesquita (2010) and Angeletos and Werning (2006) study coordination games with two stages, where public signals arise endogenously in the first stage. In our model, the "attack stage" is preceded by a "communication stage," where a private message endogenously arises and enlarges citizens' information set. New development of this literature puts emphasis on uncertain payoffs from revolt (Bernhardt and Shadmehr 2011). Given that our focus is the effect of rumors and communication, we assume that only the regime strength is uncertain.

In other fields of social sciences, there is no lack of discussions on rumors (e.g., Allport and Postman 1947) and revolutions (e.g., Goldstone 1994). However there are few studies on the relationship between these two. The idea seems to have been "up in the air" that rumors motivate citizens to participate in social movements, but the precise mechanisms remain unspecified. Our model is a step toward formalizing one such mechanism to explain explicitly how rumors affect citizens' beliefs, actions, and therefore equilibrium outcomes in revolutions.

2. A Model of Rumors and Talk about Rumors

2.1. Players and payoffs

Consider a society populated by a unit mass of ex ante identical citizens, indexed by $i \in [0,1]$. Citizen i chooses one of two actions: revolt ($a_i = 1$) or not revolt ($a_i = 0$). The aggregate mass of population that revolt is denoted A. Nature selects the strength of the regime, θ, which is sometimes also referred to as the state. The regime survives if and only if $\theta > A$; otherwise it is overthrown. A citizen's payoff depends both on whether the regime is overthrown and on whether she chooses to revolt. A cost $c \in (0,1)$, has to be paid if she revolts. If the regime is overthrown, citizens who revolt receive a benefit $b = 1$, and those who do not participate receive no benefit.[6] A citizen's net utility is therefore:

$$u(a_i, A, \theta) = \begin{cases} 1 - c, & \text{if } a_i = 1 \text{ and } A \geq \theta; \\ -c, & \text{if } a_i = 1 \text{ and } A < \theta; \\ 0, & \text{if } a_i = 0. \end{cases}$$

2.2. Information structure

Citizens are ex ante identical and have improper prior on θ. They become ex post heterogeneous after each of them observes a noisy private signal,

$$x_i = \theta + \varepsilon_i,$$

where $\varepsilon_i \sim \mathcal{N}(0, \sigma_x^2)$ is independent of θ and is independently distributed across i. This assumption captures the situation that citizens have diverse assessments of the regime's strength before they hear any rumor and communicate. This seemingly standard assumption in the global games literature turns out to be crucial for our model, for people will not exchange information in a society where everyone shares the same beliefs.

Our next assumption is that all citizens hear a rumor, z, concerning the strength of the regime. The key issue that we focus on is how citizens evaluate and react to the rumor, when the rumor could be totally uninformative. Toward this end, the rumor is modeled as a public signal which could come from two alternative sources: either a source which offers an informative signal on the strength of the regime, or a source which only produces uninformative noise. Formally we model the random variable z as coming from a mixture distribution:

$$z \sim \begin{cases} I = \mathcal{N}(\theta, \sigma_z^2) & \text{with probability } \alpha, \\ U = \mathcal{N}(s, \sigma_U^2) & \text{with probability } 1 - \alpha; \end{cases}$$

[6] We abstract from free-riding issues, which have been carefully addressed by Edmond (2011) and Bernhardt and Shadmehr (2011). The benefits from regime change can be modeled as a public good that all citizens would enjoy. Edmond (2011) offers a general payoff structure to accommodate this concern. He shows that a condition can be derived such that citizens still have incentives to act against the regime, despite the free-riding problem. To avoid being repetitive, we adopt a simpler payoff structure in this paper.

where I indicates the informative source and U indicates the uninformative source. We assume that α, s, σ_z, and σ_U are commonly known to all citizens.

The parameter s can be interpreted as the "sentiment" of the public, which captures their perception of what uninformative messages would sound like.[7] For example, if the public is used to receiving propaganda materials telling that the regime is strong, then they may expect a high value of s.

We stress that our specification of rumor as a mixture distribution is different from an informative public signal with low precision. According to the linear Bayesian updating formula, all agents would react to an informative public signal in the same way regardless of their private information. In our specification, however, citizens make an inference on the informativeness of the rumor based on their own private information. They therefore have diverse opinions on the same rumor and react to it differently.

While not dismissing the relevance of rumormongers, we choose to model rumors as exogenous public signals in order to focus on their role in coordinating collective action. The origin and the content of rumors are assumed to be exogenous because of their sheer diversity and unpredictability. The possibility that rumors are cooked up strategically to influence individuals' beliefs is an important reason that people tend to be skeptical. But even in this case, the true source of rumors is often shrouded in obscurity, making it difficult to infer whether they are manufactured to defend the regime or to destabilize it.[8] Moreover, studies on rumors also show that they could be created unintentionally. For example, misunderstanding between individuals is a usual source of rumors (Allport and Postman 1947; Peterson and Gist (1951); and Buchner (1965)).

By modeling rumors as a public signal, we also abstract from the process of how rumors travel from one to another.[9] It is implicitly assumed that rumors could reach to every citizen in the game.[10] This assumption seems to be realistic for many revolutions in history: rumors against authorities did gain a substantial, even huge, amount of publicity under very repressive regimes.[11]

[7]The assumption that s is common to all the citizens is made to simplify the exposition. Allowing them to possess diverse sentiments will not change our results.

[8]See Knapp (1944), Nkpa (1977), Elias and Scotson (1994), and Gambetta (1994) for related analysis. The incentives of rumormongers could be unpredictable in the sense that they might be motivated by many different reasons (Zhang 2009; and Turner, et al. 1992).

[9]Acemoglu, Ozdaglar, and ParandehGheibi (2010) model how rumors spread in a network.

[10]We can also assume that a certain fraction of citizens do not hear any rumor. This would not affect the main results in our model.

[11]The rumor that a student was killed by the police, which ignited the Velvet Revolution in Czechoslovakia, was broadcast by Radio Free Europe. In 2009, the Iranian post-election protest intensified after a rumor surfaced in the internet that police helicopters poured acid and boiling water on protesters (Esfandiari 2010).

We maintain the following parameter restrictions throughout this paper:

$$\sigma_x < \sigma_z^2\sqrt{2\pi}; \tag{1}$$
$$\sigma_U^2 > \sigma_x^2 + \sigma_z^2 \equiv \sigma_I^2. \tag{2}$$

The first restriction is standard. When $\alpha = 1$, the model reduces to the standard global game model with public signal. Condition (1) is both sufficient and necessary for uniqueness of equilibrium in that model; see Angeletos and Werning (2006).

The second restriction captures the idea that uninformative noise exhibits greater variability than an informative signal.[12] This assumption is motivated by the fact that an informative signal is generated based on the true strength of the regime and its realization is anchored by the truth, while there are multiple possibilities that can generate an uninformative signal. Rumors may be made up by friends or enemies of the regime, or by people with unknown motives which are unrelated to regime survival. The possibility that rumors are often the result of mistakes also adds to this uncertainty. In other words, since uninformative noise is drawn from a distribution which is not anchored by facts or fundamentals, it tends to be more unpredictable. In Section 3, we will show that this assumption is responsible for the result that citizens tend to discount wild rumors.

2.3. Communication

After citizens observe their private signals and the rumor, they are randomly paired up to communicate with one another about the informativeness of the rumor. Specifically each citizen in a pair expresses her view on the likelihood that the rumor is drawn from an informative source, and hears her peer's view on the same matter. We assume that citizens can only convey their views in a binary fashion. Let y_i represent the message sent to citizen i by her peer k. The communication technology is characterized as follows:

$$y_i = \begin{cases} 1, & \text{if } \Pr[z \sim I | z, x_k] \geq \delta; \\ 0, & \text{if } \Pr[z \sim I | z, x_k] < \delta. \end{cases} \tag{3}$$

The parameter δ is common to all citizens and can be interpreted as their threshold for plausibility. Citizen k who sends the message $y_i = 1$ to citizen i can be interpreted as saying, "I believe the rumor is informative;" while the message $y_i = 0$ can be interpreted as "I don't believe it." A high value of δ means that citizens are unlikely to say they believe in the rumor unless they are sufficiently confident of their assessments. To rule out the possibility that agents will never say they believe in the rumor, we impose

[12] Since θ has a variance of σ_x^2 to agents with private signal x_i, the unconditional variance of the informative source is $\sigma_I^2 \equiv \sigma_x^2 + \sigma_z^2$.

an upper bound for the value of δ. Specifically, we assume:

$$\delta < \frac{\alpha\sigma_I^{-1}}{\alpha\sigma_I^{-1} + (1-\alpha)\sigma_U^{-1}} \equiv \bar{\delta}.$$

The communication rule (3) is non-strategic. Given that it is a game with a continuum of agents, and that each agent only talks to one other agent, there is no incentive for citizens to strategically manipulate their peers' belief by lying. In a sense, truthful revelation is probably a good description of the casual communication between acquaintances. However, we also acknowledge that people may have other concerns when they communicate, and as a result they may behave strategically. In Section 5.3, we study a model of strategic communication and show that the key features of the communication rule in the non-strategic setting still hold.

To simplify our analysis, we assume that conversations between citizens are conducted in a coarsened fashion, in which they express their assessments of an rumor in terms of whether they believe it or not.[13] However, what drives our results is not information coarsening, but the assumption that citizens talk about the *informativeness* of the rumor. Rumors usually anchor the topic of conversations, especially when people take collective action. It is reasonable to assume that people are interested in their veracity and are inclined to discuss about them. A model which allows for exchanging private signals only, without mentioning the rumor, will deliver very different results. Section 5.2 will elaborate on this point.

2.4. Posterior beliefs and decision rules

This model can be analyzed in two stages. In the communication stage, citizen i sends a message, $y_k \in \{0, 1\}$, to her peer k based on the information set $\{z, x_i\}$, and receives a private message $y_i \in \{0, 1\}$ from her peer k. In the attack stage, citizen i chooses to revolt or not, given the post-communication information set $\{z, x_i, y_i\}$.

Before communication, agents update their beliefs on the mixture weight:

$$\Pr[z \sim I | z, x_i] = \frac{\alpha\sigma_I^{-1}\phi\left(\sigma_I^{-1}(z - x_i)\right)}{\alpha\sigma_I^{-1}\phi\left(\sigma_I^{-1}(z - x_i)\right) + (1-\alpha)\sigma_U^{-1}\phi\left(\sigma_U^{-1}(z - s)\right)} \equiv w(z, x_i), \quad (4)$$

where ϕ is the standard normal density function.

The function $w(z, x_i)$ is single-peaked in x_i, reaching the maximum at $x_i = z$. This

[13] Information coarsening is a cost-effective way of exchanging casual information and binary communication is commonly used in real life. Therefore, we adopt this formulation in the benchmark case, since it offers both realism and tractability. In Section 5.2, we demonstrate that the main insights of our model continue to hold if we allow people to exchange the exact values of their probability assessment.

means that a citizen is more likely to believe the rumor is informative if her private information is closer to the rumor. There exists $\underline{x}(z)$ and $\overline{x}(z)$ such that $w(z, \underline{x}(z)) = w(z, \overline{x}(z)) = \delta$. Therefore, the communication decision rule can be written as:

$$y(z, x_i) = \begin{cases} 1, & \text{if } x_i \in [\underline{x}(z), \overline{x}(z)]\,; \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

After communication, agents update their beliefs with the message y_i received from their peers. Their beliefs will be re-weighted by the likelihood of receiving y_i for each state. Let $P(\cdot|z, x_i, y_i)$ be the cumulative distribution of the state, with the corresponding posterior density $p(\cdot|z, x_i, y_i)$. By Bayes' rule, agent i who receives the message $y_i = 1$ from her peer revises her belief about the state to:

$$P(\theta|z, x_i, 1) = \frac{\int_{-\infty}^{\theta} J(t, z) p(t|z, x_i)\,\mathrm{d}t}{\int_{-\infty}^{\infty} J(t, z) p(t|z, x_i)\,\mathrm{d}t}, \tag{6}$$

where $p(\cdot|z, x_i)$ is the density associated with the belief $P(\cdot|z, x_i)$ before communication, and $J(t, z)$ is the probability of receiving message $y_i = 1$ in state t:

$$J(t, z) = \Phi\left(\frac{\overline{x}(z) - t}{\sigma_x}\right) - \Phi\left(\frac{\underline{x}(z) - t}{\sigma_x}\right), \tag{7}$$

where Φ is the standard normal distribution function. See Figure 1 for illustration. Similarly, an agent i who receives the message $y_i = 0$ revises her belief to:

$$P(\theta|z, x_i, 0) = \frac{\int_{-\infty}^{\theta} (1 - J(t, z)) p(t|z, x_i)\,\mathrm{d}t}{\int_{-\infty}^{\infty} (1 - J(t, z)) p(t|z, x_i)\,\mathrm{d}t}.$$

It is worth noting that the posterior belief about θ is stochastically increasing in x_i.[14]

In the attack stage, agent i chooses a_i to maximize expected utility,

$$a(z, x_i, y_i) = \underset{a_i \in \{0,1\}}{\operatorname{argmax}} \left\{ \int_{-\infty}^{\infty} u(a_i, A(\theta, z), \theta) p(\theta|z, x_i, y_i)\,\mathrm{d}\theta \right\}, \tag{8}$$

where $A(\theta, z)$ is the mass of attackers when the state is θ and the rumor is z:

$$A(\theta, z) = \int_{-\infty}^{\infty} \left[J(\theta, z) a(z, x_i, 1) + (1 - J(\theta, z))\, a(z, x_i, 0) \right] \frac{1}{\sigma_x} \phi\left(\frac{x_i - \theta}{\sigma_x}\right) \mathrm{d}x_i. \tag{9}$$

[14]This follows from Milgrom (1981). Section B of the Technical Appendix provides an explanation.

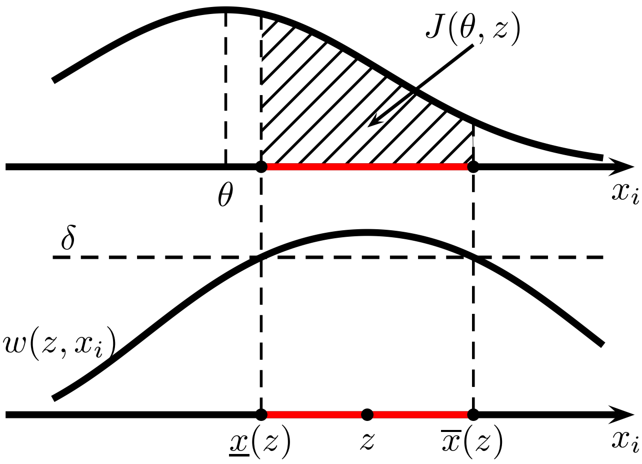


Figure 1. *Citizens with $x_i \in [\underline{x}(z), \overline{x}(z)]$ say they believe the rumor. In the aggregate, the mass of such citizens is $J(\theta, z)$.*

To summarize the timing of the game: nature first selects the strength of the regime; then citizens receive private signals and hear a rumor. In the communication stage, citizens are randomly matched and communicate. In the attack stage, they choose to revolt or not. The regime either survives or not, hinging on the mass of attackers.

2.5. Equilibrium

Definition 1. *An equilibrium is a set of posterior beliefs $p(\theta|z, x_i, y_i)$ that are derived from Bayes' rule, a message sending decision $y(z, x_i)$ and a revolt decision $a(z, x_i, y_i)$, such that (5) and (8) hold.*

In this paper, we focus on monotone equilibrium in which, for any z, there is a survival threshold $\theta^*(z)$ such that the regime collapses if and only if $\theta \leq \theta^*(z)$. For any fixed regime survival threshold $\hat{\theta}$, $P(\hat{\theta}|z, x_i, y_i) = \Pr[\theta \leq \hat{\theta}|z, x_i, y_i]$ gives the expected payoff from attacking the regime when the information set is $\{z, x_i, y_i\}$. It is decreasing in x_i since posterior belief about θ is stochastically increasing in x_i. This implies that, given a rumor z and a regime survival threshold $\hat{\theta}$, there exists a unique cutoff type $\hat{x}_I$ (when $y_i = 1$) or $\hat{x}_U$ (when $y_i = 0$) that is indifferent between attacking or not. The indifference conditions can be written as:

$$P(\hat{\theta}|z, \hat{x}_I, 1) = c, \tag{10}$$

$$P(\hat{\theta}|z, \hat{x}_U, 0) = c. \tag{11}$$

The decision rule adopted by citizens is monotone:

$$a(z, x_i, y_i) = \begin{cases} 1, & \text{if } x_i \leq \hat{x}_I \text{ and } y_i = 1, \\ & \text{or } x_i \leq \hat{x}_U \text{ and } y_i = 0, \\ 0, & \text{otherwise.} \end{cases}$$

Let $A(\theta; \hat{x}_I, \hat{x}_U, z)$ be the mass of attackers when the state is θ and when citizens adopt the cutoff rules $\hat{x}_I$ and $\hat{x}_U$:

$$A(\theta; \hat{x}_I, \hat{x}_U, z) = J(\theta, z)\Phi\left(\frac{\hat{x}_I - \theta}{\sigma_x}\right) + (1 - J(\theta, z))\Phi\left(\frac{\hat{x}_U - \theta}{\sigma_x}\right),$$

where the function J is given by (7). The regime survival threshold must satisfy the critical mass condition,

$$A(\hat{\theta}; \hat{x}_I, \hat{x}_U, z) = \hat{\theta}. \tag{12}$$

A monotone equilibrium can be characterized by a triple $(\theta^*(z), x_I^*(z), x_U^*(z))$ that solves equations (10), (11) and (12), simultaneously.

3. Rumors without Communication

Our model departs from the standard global game model in two respects: (1) the public signal may be uninformative; and (2) citizens can exchange messages concerning the informativeness of the public signal. To highlight the effects of these two separate features, we discuss in this section a model with feature (1) only. Such a model can be obtained as a special case of our model by setting $\delta = 0$, so that everyone always sends the same message, and communication becomes irrelevant. We refer to this special case as the "mute model," and let (θ_m^*, x_m^*) represent the equilibrium regime survival threshold and cutoff agent type.

The "mute model" nests two important benchmarks: $\alpha = 0$ and $\alpha = 1$. When $\alpha = 0$, citizens believe that the rumor is completely uninformative. The model reduces to a standard Morris and Shin (1998) model without public signal. We label it the "pure noise model" and use $(\theta_{ms}^*, x_{ms}^*)$ to denote the equilibrium pair. In this case, the equilibrium survival threshold is simply $\theta_{ms}^* = 1 - c$.

When $\alpha = 1$, the rumor is known to be trustworthy. We refer to this case as the "public signal model" and use $(\theta_{ps}^*, x_{ps}^*)$ to denote the equilibrium pair. In this case, $\theta_{ps}^*(z)$ monotonically decreases in z. A lower value of z indicates that the regime is more fragile. Therefore, agents are more aggressive in attacking, and the regime is more likely to collapse.[15]

[15] The subscript *ms* stands for "Morris-Shin," *ps* for "public signal," and *m* for "mute."

In the "mute model," the posterior belief about θ upon hearing a rumor z is a mixture of the posterior distribution in the "public signal model" and that in the "pure noise model," with weights given by the posterior belief that the rumor is informative or not. In other words,

$$P(\theta|z, x_i) = w(z, x_i)\Phi\left(\frac{\theta - X_i}{\sqrt{\beta}\sigma_x}\right) + (1 - w(z, x_i))\Phi\left(\frac{\theta - x_i}{\sigma_x}\right), \tag{13}$$

where the weight function $w(z, x_i)$ is given by (4). Note that X_i is citizen i's posterior mean of θ upon observing an informative public signal z. We have $X_i = \beta x_i + (1 - \beta)z$ with $\beta = \sigma_z^2/(\sigma_z^2 + \sigma_x^2)$, and the posterior variance is $\beta\sigma_x^2$. Since communication is ineffective, the critical mass condition reduces to the standard one in the Morris and Shin (1998) model.

Proposition 1. *Extreme rumors have no impact, i.e.,* $\lim_{z\to\pm\infty} \theta_m^*(z) = \theta_{ms}^*$*. There exists a unique neutral rumor* $\tilde{z}$ *such that* $\theta_m^*(\tilde{z}) = \theta_{ps}^*(\tilde{z}) = \theta_{ms}^*$*. In the "mute model,"* $\theta_m^*(z)$ *is always between* θ_{ms}^* *and* $\theta_{ps}^*(z)$*, but it is increasing then decreasing then increasing in* z*.*

When citizens hear an extreme rumor, i.e., $z \to \pm\infty$, they discard it as uninformative unless the true fundamental is also extreme and close to it.[16] Therefore, the equilibrium is identical to that in the "pure noise model." This property follows from our the parameter restriction (2), under which, for any finite x_i, $w(z, x_i)$ is single-peaked in z, and approaches 0 when z is extreme.[17]

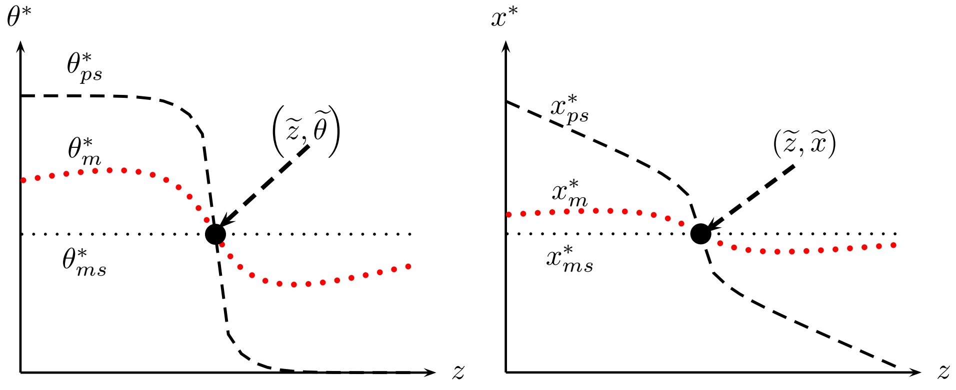
There exists a unique neutral rumor $\tilde{z}$, such that $\theta_{ps}^*(z) > \theta_{ms}^*$ for all $z < \tilde{z}$ and $\theta_{ps}^*(z) < \theta_{ms}^*$ for all $z > \tilde{z}$. We say that a rumor is "against the regime" if $z < \tilde{z}$, and that the rumor is "for the regime" if $z > \tilde{z}$. In the "mute model," upon hearing the neutral rumor $\tilde{z}$, the cutoff agent is indifferent as to whether the rumor is informative or not, because the likelihood of success is the same in both cases. This is why the equilibrium pair $(\theta_{ms}^*, x_{ms}^*)$ of the "pure noise model" also solves the "mute model" when $z = \tilde{z}$. See Figure 2.[18]

Skepticism is the key feature of the "mute model": citizens take into account the

[16] If the fundamental is extreme, the information structure does not affect the outcome of the revolution game. In this model, the survival threshold θ^* is between 0 and 1, because the mass of attacker is always between 0 and 1. For example, if the fundamental θ approaches $-\infty$, the regime collapses under any information structure.

[17] When $\sigma_I < \sigma_U$, an extreme z is infinitely more likely to be generated from an uninformative source than an informative one. But if we assume the opposite parameter restriction, then $w(z, x_i)$ has a single trough, and approaches 1 when z is extreme. In that case, citizens would regard rumors which are further away from their private information to be more plausible.

[18] To compute the numerical examples in the figures, unless otherwise specified, we use the following set of parameters: $c = 0.5$, $s = 0.5$, $\alpha = 0.5$, $\delta = 0$, $\sigma_U^2 = 1$, $\sigma_z^2 = 0.5$, and $\sigma_x^2 = 0.4$ for the "pure noise model," "public signal model" and "mute model." For the "communication model," we set $\delta = 0.5$.



(a) regime survival threshold (b) cutoff agent type

Figure 2. *θ^*_m is non-monotone and is between θ^*_{ps} and θ^*_{ms}. At $z = \tilde{z}$, the survival thresholds are the same in the three models. The qualitative properties of $x^*_m(z)$ are similar to those of $\theta^*_m(z)$.*

possibility that the rumor could just provide an uninformative noise. The effect of skepticism manifests itself in the fact that $\theta^*_m(z)$ is always between θ^*_{ms} and $\theta^*_{ps}(z)$ for any z. In the "mute model," citizens are less responsive to the rumor than when they are sure that it is informative, but are more responsive to it than when they are sure that it is not. This, however, does not mean that the "mute model" is simply a "public signal model" with a less precise public signal.

In fact, another aspect of skepticism is that $\theta^*_m(z)$ is not monotonic, in contrast to the monotonicity of $\theta^*_{ps}(z)$. To see the intuition, observe that from equation (13),

$$\frac{\partial P(\theta^*_m|z, x_i)}{\partial z} = -w\frac{\beta}{\sqrt{\beta}\sigma_x}\phi_I + (\Phi_I - \Phi_U)\frac{\partial w}{\partial z}, \tag{14}$$

where the subscripts I and U means that the functions are evaluated at the points $(\theta^*_m - X_i)/(\sqrt{\beta}\sigma_x)$ and $(\theta^*_m - x_i)/\sigma_x$, respectively. If the rumor is informative, an increase in z is an indication that the regime is strong, which lowers the probability that the regime will collapse. Hence the first term is negative. This is the standard "public signal effect." The second term captures the "skepticism effect," which is positive for extreme values of z. For instance, when z is sufficiently large, $\partial w/\partial z < 0$, which reflects greater skepticism toward the rumor when it moves further away from the agent's private signal. Moreover, $\Phi_I - \Phi_U < 0$, because the probability that the regime will collapse is lower if a rumor for the regime is informative than if it is not.

The magnitude of the "public signal effect" is small for z sufficiently large, whereas that of the "skepticism effect" is big. When the second term dominates the first term, an increase in the value of z actually raises the likelihood of regime collapsing for an agent. As a result, the marginal agent who is indifferent between attacking and not attacking must have a higher private information. Therefore, the survival threshold $\theta^*_m(z)$ increases in z. For z sufficiently close to $\tilde{z}$, the magnitude of $\Phi_I - \Phi_U$ is small, because it does not matter to the cut-off type agent whether the neutral rumor is informative or not. Therefore, the "public signal effect" dominates and the survival threshold decreases in z. This explains the non-monotonicity of $\theta_m^*(z)$.[19]

4. Rumors with Communication

In this section, we let $\delta \in (0, \overline{\delta})$ so that there is meaningful communication among the citizens. We focus on three key results in this "communication model." First, communication makes citizens even more skeptical of extreme rumors. When an extreme rumor against the regime circulates, communication discourages attackers. Second, communication makes citizens more responsive to rumors that are close to neutral. When a rumor mildly against the regime circulates, communication encourages attackers and can cause the regime to collapse, even though it could have survived had citizens remained silent. Third, the effect of communication can be so potent that a rumor can mobilize even more attackers than when the same story is regarded by everybody as fully trustworthy.

4.1. Extreme and Neutral Rumors

Proposition 2. *In the "communication model," the equilibrium triple (θ^*, x_I^*, x_U^*) has the following properties:*

1. $\lim_{z\to\pm\infty} \theta^*(z) = \theta_{ms}^*$, $\lim_{z\to\pm\infty} x_U^*(z) = x_{ms}^*$, *and* $\lim_{z\to\pm\infty} x_I^*(z) = \mp\infty$.
2. *There exists a z' such that $\theta^*(z') = \theta_m^*(z')$ and $x_I^*(z') = x_U^*(z') = x_m^*(z')$.*
3. *If $c = 0.5$, then z' coincides with $\tilde{z}$. Moreover, $x_I^*(z) > x_U^*(z)$ for all $z < z'$ and $x_I^*(z) < x_U^*(z)$ for all $z > z'$.*

When the rumor indicates that the regime is extremely weak, almost everyone says "I don't believe it" (because the limit of $J(\theta, z)$ is 0 for any finite θ). The skepticism of an agent (captured by the fact that the limit of $w(z, x_i)$ is 0) is reinforced by the skepticism of her peer. This explains why θ^* and x_U^* are the same as those in the "pure noise model." In this case, an agent i is very unlikely to receive a confirmatory message from her peer. But if such a rare event takes place, it must be that her peer has very negative private information about the regime strength. She will revise her belief so much so that, for any finite value of x_i, she becomes almost sure that the regime will collapse and will therefore revolt. That explains the limit behavior of x_I^*.

[19] The properties of $x_m^*(z)$ are similar to those of $\theta_m^*(z)$. This follows from the critical mass condition, which can be written as $x_m^*(z) = \theta_m^*(z) + \sigma_x \Phi^{-1}(\theta_m^*(z))$. Under parameter restriction (1), $x_m^*(z)$ is increasing in $\theta_m^*(z)$.

The limit behavior of x_I^* and x_U^* immediately implies that there exists z' such that $x_I^*(z') = x_U^*(z') = x'$. Such an x' turns out to be equal to $x_m^*(z')$ of the "mute model." This is because, when $x_I^* = x_U^*$, an agent's decision to revolt does not depend on what her peer says. In other words, the "communication model" and the "mute model" would produce the same outcomes when $z = z'$.

We show that z' coincides with $\tilde{z}$ when $c = 0.5$. This means that rumors for which cutoff type citizens do not care about what their peers say ($y = 1$ or $y = 0$) are also rumors for which they do not care about from which source the rumor is drawn ($z \sim I$ or $z \sim U$). This special case offers analytical tractability for properties of rumors close to the neutral one. Because equilibrium thresholds are the same across all the four models when $z = z' = \tilde{z}$, we study the role played by communication by letting z deviate from z' and comparing the equilibrium results with the other three benchmarks.[20] Therefore, we focus on this case to characterize the properties of non-extreme rumors.

4.2. Direct and Multiplier Effects

Part 3 of Proposition 2 establishes that, if a rumor is against the regime, citizens become more aggressive in attacking when receiving a confirmatory message than when receiving a contrarian message (i.e., $x_I^* > x_U^*$). In this case, citizens with private information $x_i < x_U^*$ attack the regime regardless of the message they receive. We label this group of citizens *revolutionaries*. Citizens with private information $x_i > x_I^*$ would not revolt regardless. This group is labeled *bystanders*. The group of citizens with $x_i \in [x_U^*, x_I^*]$ is called the *swing population*. Their revolt decisions are influenced by their peers' assessments of the rumor: they choose to revolt if they receive $y_i = 1$, and not to revolt otherwise.

To see why anyone from the swing population becomes more aggressive when receiving a confirmatory message of a negative rumor, we explicate the direct and multiplier effects. First, when the value of z deviates from the neutral rumor z', the direct effect comes from the response of individuals, without taking into account the resulting change in equilibrium survival threshold. Let $\hat{x}(z, \hat{\theta})$ denote the cutoff type who is indifferent between attacking and not attacking when the rumor is z and the regime survival threshold is $\hat{\theta}$. The size of the direct effect depends on the magnitude

[20] It is difficult to directly compare $\theta^*(z)$, $\theta_m^*(z)$ and $\theta_{ps}^*(z)$ because none of these equilibrium values admits a closed-form solution. In the case of extreme rumors, we analyze those models for z sufficiently large or small. By comparing the rates of convergence in each model, we can infer the levels of these functions. In the case of non-extreme rumors, we study their properties around the point $z = z' = \tilde{z}$ and compare their derivatives. Since the equilibrium thresholds in the four models are the same when the rumor is neutral, we can make conclusions about the levels of these functions for values of z near z'. Our results still hold when $z' \neq \tilde{z}$. We offer a discussion in Section C of the Technical Appendix and explain that the value of participation cost c does not affect the properties of equilibrium.

of $\partial\hat{x}/\partial z$. Second, the multiplier effect arises because of the complementarity in action among citizens, which amplifies the direct effect: when citizens become more aggressive, the regime survival threshold has to increase, which in turn raises the payoff from revolting and hence further increases the cutoff type. The magnitude of the multiplier effect depends on $\partial\hat{x}/\partial\hat{\theta}$.

Lemma 1. At $z = z' = \tilde{z}$ and $\hat{\theta} = \theta'$, the cutoff types who are indifferent between attacking and not attacking satisfy:

$$\frac{\partial\hat{x}_I}{\partial z} < \frac{\partial\hat{x}_m}{\partial z} < \frac{\partial\hat{x}_U}{\partial z} < 0, \tag{15}$$

$$\frac{\partial\hat{x}_I}{\partial\hat{\theta}} > \frac{\partial\hat{x}_m}{\partial\hat{\theta}} > \frac{\partial\hat{x}_U}{\partial\hat{\theta}} > 0. \tag{16}$$

When the rumor is neutral, the cutoff citizen is indifferent between receiving $y = 1$ or $y = 0$ from her peer, i.e., $x_I^* = x_U^* = x_m^*$. The ordering of $\partial\hat{x}/\partial z$ (direct effect) implies that a citizen who receives $y = 1$ reacts more to the rumor than she does without such communication. The reason is that a confirmatory message ($y = 1$) causes a citizen to believe that states close to the rumor are more likely to be the fundamental. In other words, the message leads to a more concentrated posterior density for states around $\theta = z'$. See the solid curves in Figure 3 for an illustration. This effect can be seen from equation (6), which implies that

$$\frac{p(\theta|z', x', 1)}{p(\theta|z', x')} = \frac{J(\theta, z')}{\int_{-\infty}^{\infty} J(t, z')p(t|z', x')\,\mathrm{d}t}. \tag{17}$$

The likelihood $J(\theta, z')$ in the numerator is increasing then decreasing in θ, with a peak at $\theta = z'$. The denominator is just the expected value of $J(\theta, z')$. Therefore, the ratio in (17) is greater than 1 when θ is close to z' and smaller than 1 when θ is far from it. As shown by the dashed curves in Figure 3, the same increase in the value of z shifts the density function $p(\cdot|z, x_i, 1)$ to the right by a greater amount than it does to the density function $p(\cdot|z, x_i)$. This explains why, in response to the same amount of change in z, $\hat{x}_I$ has to decrease by a larger amount than $\hat{x}_m$ in order to balance the indifference condition.

To understand the intuition behind the ordering of $\partial\hat{x}/\partial\hat{\theta}$ (multiplier effect), note that equation (17) implies that $p(\theta'|z', x', 1) > p(\theta'|z', x')$. The term $p(\theta'|z', x')$ reflects the increase in payoff from attacking in the "mute model" when the regime survival threshold is marginally raised above θ'. Since $p(\theta'|z', x', 1) > p(\theta'|z', x')$, this means that raising the survival threshold increases the incentive to attack by a greater amount

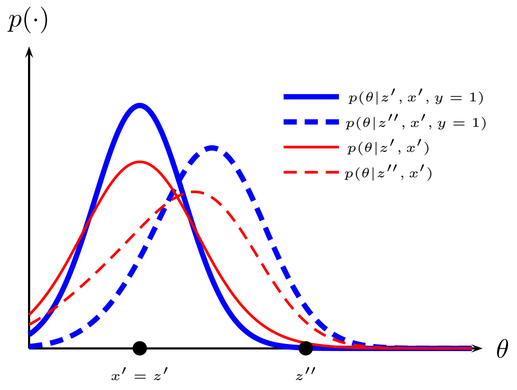


Figure 3. *The posterior density given $y = 1$ is more concentrated around $\theta = z'$ than its counterpart in the "mute model." When z increases from z' to z'', the density function given $y = 1$ shifts to the right by a greater amount than does the density function in the "mute model."*

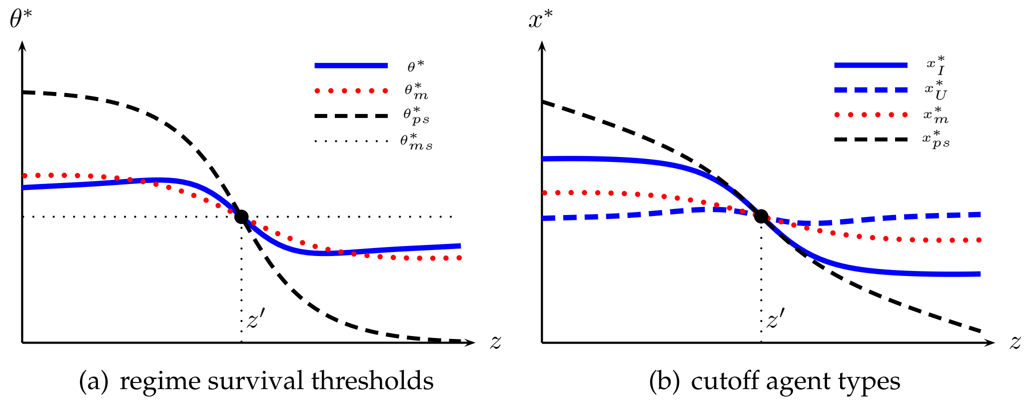


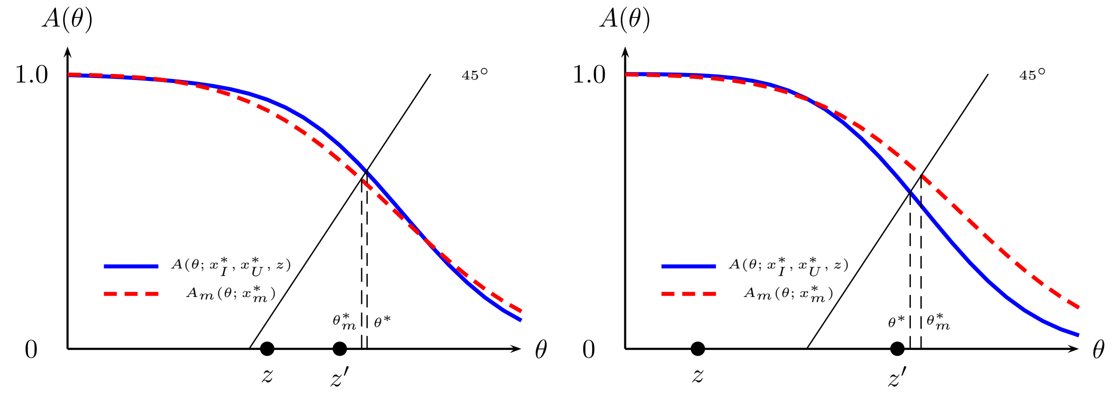
Figure 4. *Equilibrium regime survival thresholds and equilibrium cutoff agent types in the "communication model" in comparison to other models.*

for citizens who receive a confirmatory message. As a result, $\hat{x}_I$ also has to increase by a greater amount than $\hat{x}_m$ does to keep the cutoff types indifferent.

4.3. The Communication Effect: Does it Help or Hurt?

The first two key results of the "communication model" are established in this subsection: communication encourages mobilization when a negative rumor is close to neutral, but it discourages mobilization when a negative rumor is extreme.[21] See Figure 4 for a comparison of the regime survival thresholds under the "communication model" and the "mute model."

[21]The focus of our analysis is why some rumors can be so effective while others are not. Toward this end, we compare the equilibrium survival thresholds across models for given realizations of rumor z. An alternative question is which information structure—with or without communication—will be preferred by the regime ex ante, i.e., before the rumor is heard. In order to answer this question, we can compute the ex ante survival probability for each θ, by integrating over z using the mixture distribution. However, once we pool together all these effective and ineffective rumors as well as positive and negative rumors, the answer to this question depends on the parameters of the model and no general conclusion can be made.



(a) A mild negative rumor, i.e., $z < z'$ (b) An extreme negative rumor, i.e., $z << z'$

Figure 5. *Communication increases the mass of attackers in states close to z but lowers it in states far away from z.*

Interestingly, these two contrasting results are driven by the same mechanism. To explain, we start with how communication affects the total mass of attackers and then elaborate on how it translates into the ranking of equilibrium thresholds.

Communication among citizens allows them to be better informed regarding where the true state θ lies. They know better about whether θ is close to or far away from z than their counterparts in the "mute model." The key is that the fraction of citizens who *send* confirmatory messages, $J(\theta, z)$, will be larger if the true regime strength is closer to what the rumor suggests, and achieves a maximum when $\theta = z$. Therefore, more citizens will *hear* confirmatory messages when θ is near z and fewer when θ is far away from z.

This mechanism implies that if the rumor says that the regime is fragile and the regime is indeed weak, a larger fraction of the swing population will receive the message $y = 1$ from their peers and decide to join the revolutionaries and attack the regime, which results in a larger total mass of attackers $A(\theta)$ in the "communication model." On the other hand, the total mass of attackers will be smaller if the true regime strength is very different from what the rumor says.

Figure 5 plots the mass of attackers $A(\theta)$ against the regime strength θ, holding the cutoff rules constant. Figure 5(a) shows the situation for z being slightly below z': a larger mass of attackers are mobilized than that in the "mute model" when the regime strength θ is near z, and the mass is smaller when the true θ is far away from z. For the same reason, if the rumor indicates the regime is extremely fragile, it mobilizes more attackers when the regime is indeed extremely weak. But the mass of attackers falls relative to that in the "mute model" when the regime is indeed strong and far away from what the rumor suggests. Figure 5(b) illustrates this situation.

In Figure 5, the mass of attackers $A(\theta)$ in the "communication model" is larger than its counterpart in the "mute model" for states close to z. Since the equilibrium regime survival threshold is given by the intersection of $A(\theta)$ and the 45-degree line, this figure shows that $\theta^*(z) > \theta_m^*(z)$ for z slightly lower than z' and $\theta^*(z) < \theta_m^*(z)$ for z substantially lower than z'.

Proposition 3. *For z sufficiently negative, $\theta^*(z) < \theta_m^*(z)$; for z sufficiently large, $\theta^*(z) > \theta_m^*(z)$.*

Consider an extreme negative rumor. The cutoff citizens who receive $y = 0$ will be less aggressive than the cutoff type in the "mute model," who is in turn less aggressive than the cutoff citizens who receive $y = 1$. That is, $x_U^* < x_m^* < x_I^*$.[22] Because z is sufficiently negative, unless the true regime strength is extremely weak, the fraction of citizens who receive confirmatory messages, $J(\theta, z)$, will be very low. Therefore, only a negligible fraction of the swing population will join the revolutionaries and attack the regime. As a result, the mass of attackers in the "communication model" is smaller than in the "mute model." That is,

$$J(\theta, z)\left[\Phi\left(\frac{x_I^* - \theta}{\sigma}\right) - \Phi\left(\frac{x_U^* - \theta}{\sigma}\right)\right] + \Phi\left(\frac{x_U^* - \theta}{\sigma}\right) < \Phi\left(\frac{x_m^* - \theta}{\sigma}\right),$$

or $A(\theta) < A_m(\theta)$.[23] As a result, Proposition 3 is implied.

Proposition 4. *For z near z' (which equals $\tilde{z}$), $\theta^*(z) > \theta_m^*(z)$ if $z < z'$; and $\theta^*(z) < \theta_m^*(z)$ if $z > z'$.*

To understand this proposition, we study the change in total mass of attackers when z deviates from z'. In the "communication model," a slight decrease in z from z' leads to an increase in the size of swing population by $\sigma_x^{-1}\phi(\cdot)(\mathrm{d}x_I^*/\mathrm{d}z - \mathrm{d}x_U^*/\mathrm{d}z)$. A fraction J of the swing population would receive a confirmatory message from their peers and decide to join the revolutionaries and attack the regime. The change in the mass of revolutionaries is $\sigma_x^{-1}\phi(\cdot)\mathrm{d}x_U^*/\mathrm{d}z$. In the "mute model," the increase in the mass of attackers is $\sigma_x^{-1}\phi(\cdot)\mathrm{d}x_m^*/\mathrm{d}z$. For $A(\theta')$ to be greater than $A_m(\theta')$ when z is slightly below z', we must have

$$\frac{1}{\sigma_x}\phi(\cdot)\left[J(\theta', z')\left(\frac{\mathrm{d}x_I^*}{\mathrm{d}z} - \frac{\mathrm{d}x_U^*}{\mathrm{d}z}\right) + \frac{\mathrm{d}x_U^*}{\mathrm{d}z}\right] < \frac{1}{\sigma_x}\phi(\cdot)\frac{\mathrm{d}x_m^*}{\mathrm{d}z} < 0. \tag{18}$$

[22] The first inequality is established in the proof of Proposition 3, and the second follows from part 1 of Proposition 2.

[23] The proof of Proposition 3 shows that, as z gets sufficiently negative, the first term of the left-hand-side varnishes much quicker than the difference between the second term and the term on the right-hand-side.

To see why inequality (18) holds, it is sufficient to show that both direct and multiplier effects are larger in the "communication model" than those in the "mute model."[24] That is, at the point $\hat{\theta} = \theta'$ and $z = z'$,

$$J\frac{\partial \hat{x}_I}{\partial z} + (1-J)\frac{\partial \hat{x}_U}{\partial z} < \frac{\partial \hat{x}_m}{\partial z} < 0; \tag{19}$$

$$J\frac{\partial \hat{x}_I}{\partial \hat{\theta}} + (1-J)\frac{\partial \hat{x}_U}{\partial \hat{\theta}} > \frac{\partial \hat{x}_m}{\partial \hat{\theta}} > 1. \tag{20}$$

Inequalities (19) and (20) are stronger than Lemma 1. They hold because a large fraction of the population will receive confirmatory messages when the rumor is close to the true state. As a result, the population as a whole is better informed given the extra information from communication.

Further, consistent with inequality (18), θ^* is more responsive to the change in z around z' than is θ^*_m. Using the fact that $\hat{x}_I(\theta', z') = \hat{x}_U(\theta', z')$, we can differentiate the critical mass condition (12) to obtain:

$$\frac{\mathrm{d}\theta^*(z')}{\mathrm{d}z} = \frac{\frac{1}{\sigma_x}\phi(\cdot)\left(J\frac{\partial \hat{x}_I}{\partial z} + (1-J)\frac{\partial \hat{x}_U}{\partial z}\right)}{1 - \frac{1}{\sigma_x}\phi(\cdot)\left(-1 + J\frac{\partial \hat{x}_I}{\partial \hat{\theta}} + (1-J)\frac{\partial \hat{x}_U}{\partial \hat{\theta}}\right)}. \tag{21}$$

Comparing (21) with its counterpart in the "mute model,"[25] and using inequalities (19) and (20), we obtain:

$$\frac{\mathrm{d}\theta^*(z')}{\mathrm{d}z} < \frac{\mathrm{d}\theta^*_m(z')}{\mathrm{d}z} < 0. \tag{22}$$

Proposition 4 follows from inequality (22).

4.4. The Power of Whispers: Rumors vs. Trustworthy News

Strikingly, the effect of communication can be so large that, when a rumor against the regime is heard, the regime could survive when all agents believe that it is trustworthy, but could collapse when citizens know that the rumor may be uninformative.

In this comparison, communication is also allowed in the "public signal model," but it is completely ineffective. Given everybody believes that the rumor is informative (i.e., $\alpha = 1$), the posterior belief of the informativeness is always $w(z, x_i) = 1$. Therefore, for any plausibility threshold δ, everybody would send to and receive from her peers the message $y = 1$. Paradoxically, communication takes no effect in the

[24] We note that $\mathrm{d}x^*/\mathrm{d}z = \partial x^*/\partial z + (\partial x^*/\partial \hat{\theta})(\mathrm{d}\theta^*/\mathrm{d}z)$. Therefore, the magnitude of $\mathrm{d}x^*/\mathrm{d}z$ depends on both direct and multiplier effects. At the aggregate level, the direct effect in the "communication model" is the weighted average of $\partial \hat{x}_I/\partial z$ and $\partial \hat{x}_U/\partial z$, with the weights being J and $1-J$, respectively. The multiplier effect is the weighted average of $\partial \hat{x}_I/\partial \hat{\theta}$ and $\partial \hat{x}_U/\partial \hat{\theta}$.

[25] To obtain the counterpart of (21) in the "mute model," let $J = 1$ and replace $\partial \hat{x}_I/\partial z$ and $\partial \hat{x}_I/\partial \hat{\theta}$ with $\partial \hat{x}_m/\partial z$ and $\partial \hat{x}_m/\partial \hat{\theta}$.

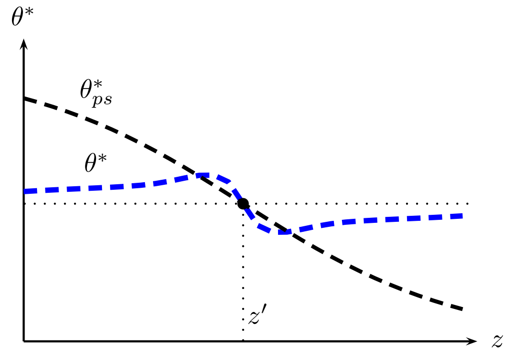


Figure 6. *Comparing the "communication model" to the "public signal model" when private information has moderately high precision.*

absence of skepticism.

Figure 6 shows that θ^* can be higher than θ^*_{ps} for z slightly below z'.[26] To see why this possibility could arise, we need to understand why it could hold that $A(\theta') > A_{ps}(\theta')$, when z deviates from z' to a slightly lower value. Similar to the previous analysis, the response of the mass of attackers to a change in z is governed by the direct and multiplier effects. In the following analysis, we just focus on the multiplier effect, i.e., $\partial\hat{x}/\partial\hat{\theta}$.[27]

We have shown the following two separate mechanisms: due to the effect of skepticism, $\partial\hat{x}_{ps}/\partial\hat{\theta} > \partial\hat{x}_m/\partial\hat{\theta}$; due to the effect of communication, $\partial\hat{x}_I/\partial\hat{\theta} > \partial\hat{x}_m/\partial\hat{\theta} > \partial\hat{x}_U/\partial\hat{\theta}$. Interestingly, the fraction of population who receive confirmatory messages can be large, such that the average response of the population is larger than that in the "public signal model." That is,

$$J(\theta', z')\frac{\partial\hat{x}_I}{\partial\hat{\theta}} + (1 - J(\theta', z'))\frac{\partial\hat{x}_U}{\partial\hat{\theta}} > \frac{\partial\hat{x}_{ps}}{\partial\hat{\theta}}. \tag{23}$$

This result is interesting in that skepticism provides the ground for communication, but its effect can also be undone by communication.

We show that inequality (23) holds when the threshold δ for sending a confirmatory message is neither too high nor too low and the relative precision of the private information is not too low.

Proposition 5. *For β larger than a threshold $\hat{\beta}$, there exists an interval $[\delta_1, \delta_2]$ such that, for any $\delta \in [\delta_1, \delta_2]$, $\theta^*(z) > \theta^*_{ps}(z)$ if $z < z'$, and $\theta^*(z) < \theta^*_{ps}(z)$ if $z > z'$, for z near z' (which equals $\tilde{z}$).*

[26]In plotting this figure, we set $\delta = 0.5$ and use the value of $\sigma_x^2 = 0.2$, which is lower than that used in Figure 4(a).

[27]Lemma 5 of the Technical Appendix establishes that the direct and multiplier effects sum up to 1. A stronger multiplier effect implies that the direct effect is also stronger.

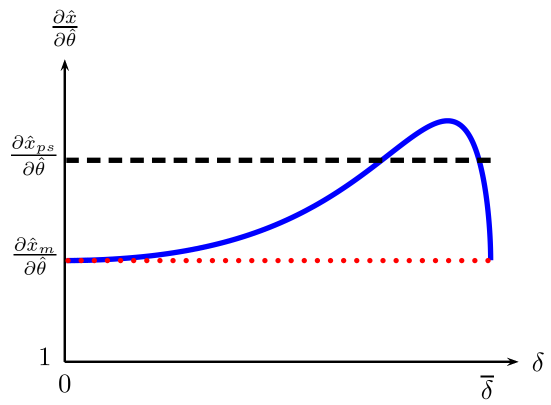


Figure 7. *The hump-shaped communication effect. The solid curve stands for the multiplier effect in the "communication model;" the dashed and dotted lines represent the multiplier effect in the "public signal model" and "mute model," respectively.*

To understand Proposition 5, note that communication is not very informative when the value of δ is very large or very small, because most citizens would be saying the same thing. Figure 7 shows that the multiplier effect in the "communication model" (the left-hand-side of inequality (23)) is the same as the multiplier effect in the "mute model" when δ is close to its bounds. Figure 7 also shows that the effect of communication is hump-shaped: for intermediate values of δ, the multiplier effect in the "communication model" is larger than that in the "mute model," and can even exceed that in the "public signal model."

Recall that δ determines the fraction of population J who send confirmatory messages. When the plausibility threshold δ is low, J is high, because citizens are less cautious in saying that they believe the rumor. Given that receiving a confirmatory message is so expected, recipients of such messages would not update their belief that much. Therefore, $\partial \hat{x}_I / \partial \hat{\theta}$ will be quite close to $\partial \hat{x}_m / \partial \hat{\theta}$.

When δ is larger, citizens becomes more careful about sending their peers confirmatory messages. In this case, recipients of confirmatory messages can be even more responsive than those in the "public signal model," i.e., $\partial \hat{x}_I / \partial \hat{\theta} > \partial \hat{x}_{ps} / \partial \hat{\theta}$. Because it is unlikely for citizens to receive $y = 1$ from their peers, such a confirmatory message will be strong evidence that θ is close to z'. This effect can be so strong that inequality (23) holds.

For δ very large, however, J becomes so small that only a negligible fraction of the population receives confirmatory messages. The weighted average, $J \partial \hat{x}_I / \partial \hat{\theta} + (1 - J) \partial \hat{x}_U / \partial \hat{\theta}$, would dip below $\partial \hat{x}_{ps} / \partial \hat{\theta}$ again.

Proposition 5 also says that the relative precision of private information has to be reasonably high for communication to take a strong effect. When the noise is smaller in private information, citizens' private signals will be on average closer to the true θ.

A confirmatory message will be stronger evidence for the true θ being close to what the rumor says. Therefore, recipients of confirmatory messages will be more responsive to the rumor, making the left-hand-side of (23) bigger. In the "public signal model," on the other hand, a higher precision of the private signals means that citizens will rely more on their private signal and be less responsive to the rumor, making the right-hand-side of (23) smaller.

Our analysis of rumors suggests that, for outcomes of regime changes, it matters little whether rumors reflect the truth or have no basis in fact. What matters is that rumors create public topics that people can talk about. By communication, people learn from what others believe regarding the rumor and can better coordinate their actions. That explains why some false rumors could mobilize citizens very effectively when collective actions take place.

5. Discussion

5.1. Prior Belief and Sentiment

Proposition 5 demonstrates that the effect of communication is larger when δ takes intermediate values and smaller when it is more extreme. A similar result can be formally established for the comparative statics of the prior α.[28] The effect of communication is smaller when α is closer to 1 or 0, and is bigger when α takes intermediate values. When agents have a relative more extreme prior, i.e., either they believe the rumor is very likely to be informative or uninformative, their peers' assessments about the rumor take little effect. On the other hand, if, ex ante, agents are more uncertain about the rumor's veracity, information exchange among agents has a stronger effect on the equilibrium. In other words, the absolute value of $\partial\theta^*/\partial z$ evaluated near z' increases and then decreases in α. Since the value of z' is independent of α, this in turn implies that $\theta^*(z)$ for z slightly below z' is increasing then decreasing in α.

The effect of changing α can be decomposed into three specific mechanisms. First, when the prior belief α is higher, the posterior belief ω that the rumor is informative is higher. Second, as a result, the faction of agents with $\omega \geq \delta$ is also higher. That is, J increases when ω increases. Third, the counteracting effect is that, when a larger fraction of agents send confirmatory messages to their peers, agents who receive confirmatory messages become less responsive to the change in z.

When α is sufficiently small and becomes slightly larger, the first two effects dominate the third and the equilibrium survival threshold is more responsive to the rumor. However, when α is large enough, both ω and J are close to 1 and the increments to

[28]For β larger than a threshold $\hat{\beta}'$, there exists an interval $[\alpha_1, \alpha_2]$ such that, for any α belongs to the interval, $\theta^*(z) > \theta^*_{ps}(z)$ if $z < z'$, and $\theta^*(z) < \theta^*_{ps}(z)$ if $z > z'$, for z near z' (which equals $\tilde{z}$).

both are very small when α increases further. Therefore, the third effect dominates and the absolute value of $\partial\theta^*/\partial z$ decreases as α increases beyond a certain point.

Further, the equilibrium survival threshold responds to changes in s in a similar fashion and the absolute value of $\partial\theta^*/\partial z$ does not respond to s monotonically. Consider the case where s increases from $s = z'$ to ∞. We label $s = z'$ a "neutral sentiment," since the equilibrium regime survival threshold is symmetric in s about the point $s = z'$. If s increases from the neutral value, the slope of $\theta^*(z)$ around z' will be steeper, indicating that the regime is less likely to survive when the rumor is negative. The reason is that when s is higher, citizens believe that a rumor against the regime ($z < z'$) is more likely to have come from an informative source. As a result, they believe the regime is weaker and are therefore more inclined to attack.

The aforementioned three mechanisms are still at work. If s is slightly increased from the neutral value, both ω and J increase and agents are more reactive to a negative rumor close to z', which dominates the third effect. If s is sufficiently large, both ω and J are close to 1. Since almost everybody thinks a negative rumor is informative and says so to their peers, information exchange becomes less useful for agents. Thus, the third effect dominates the first two and the slope of θ^* becomes flatter when s increases from a large value.

5.2. The Topic of Conversations Matters

In the baseline model, we have assumed that communication takes the form of exchanging coarse (binary) signals about the informativeness of the rumor. In this subsection, we demonstrate that information coarsening is not a crucial assumption. It is the exchange of views on the rumor's informativeness that drives our results. This mechanism is qualitatively different from the case where the content of conversation is unrelated to the rumor. To stress this point, we explore two alternative communication protocols, which allow citizens to exchange the exact value of $w(z, x_i)$ or the exact value of x_i, respectively. We then discuss the case where both protocols are allowed.

Exchange $w(z, x_i)$. Suppose citizens can tell each other the value of w (instead of just whether w is greater than or less than δ). Citizens i updates her belief in a Bayesian fashion when she receives the message $w_k = w(z, x_k)$ from her peer k:

$$\Pr[\theta < \hat{\theta} | z, x_i, w_k] = \frac{\int_{-\infty}^{\hat{\theta}} j(t, z, w_k) p(t|z, x_i)\, \mathrm{d}t}{\int_{-\infty}^{\infty} j(t, z, w_k) p(t|z, x_i)\, \mathrm{d}t},$$

where

$$j(t, z, w_k) = \frac{1}{\sigma_x}\phi\left(\frac{x^l - t}{\sigma_x}\right) + \frac{1}{\sigma_x}\phi\left(\frac{x^r - t}{\sigma_x}\right),$$

with x^l and x^r being the two values of x_k that solves the equation $w(z, x_k) = w_k$.

The function $j(t, z, w_k)$ reaches a local maximum at $t = z$, if w_k is high (x^l and x^r are near z), or a local minimum at $t = z$, if w_k is low (x^l and x^r are far from z). Moreover, at $z = z'$, there exists a $\hat{w}$ such that the ratio, $j(\theta', z', w_k) / \int_{-\infty}^{\infty} j(t, z', w_k) p(t|z', x') \, \mathrm{d}t$, is greater than 1 if $w_k > \hat{w}$. Therefore, citizens who receive a message $w_k > \hat{w}$ in this model are similar to citizens who receive a message $y = 1$ in the baseline model. Similarly, the ratio $j(\theta', z', w_k) / \int_{-\infty}^{\infty} j(t, z', w_k) p(t|z', x') \, \mathrm{d}t$ is less than 1 if $w_k < \hat{w}$. Therefore, citizens who receive a message $w_k < \hat{w}$ in this model are similar to citizens who receive a message $y = 0$ in the baseline model.

Intuitively, when citizen k conveys a high assessment that the rumor is likely to be informative, i.e., $w_k > \hat{w}$, her peer i assigns a higher probability weight (density) to states close to z. Similarly, when citizen k does not believe the rumor is sufficiently informative, i.e., $w_k < \hat{w}$, citizen i assigns lower weight to states close to z and higher weight to states far away from z. Recall this mechanism of probability re-weighting is precisely the key that drives our results in the baseline case. Therefore, the model of exchanging the exact probability assessments is not qualitatively different from our model of exchanging coarse information.

Exchange x_i. Now we turn to the case where citizens directly exchange their private information concerning the strength of the regime. To be more general, we assume that the communication process is noisy: each citizen receives her peer's private signal with an additive noise. That is, the message y_i received by the citizen i from her peer k is given by:

$$y_i = x_k + \xi_k,$$

where the noise $\xi_k \sim \mathcal{N}(0, \sigma_\xi^2)$ is independent of x_k and independently distributed across k. After communication, each citizen possesses an information set which consists of the rumor z and two private signals x_i and y_i. This information set is equivalent to $\{z, v_i\}$, where v_i is a private signal with higher precision than x_i. That is,

$$\begin{aligned} v_i &\equiv \frac{\sigma_x^2 + \sigma_\xi^2}{2\sigma_x^2 + \sigma_\xi^2} x_i + \frac{\sigma_x^2}{2\sigma_x^2 + \sigma_\xi^2} y_i, \\ \sigma_v^2 &\equiv \frac{\sigma_x^2 + \sigma_\xi^2}{2\sigma_x^2 + \sigma_\xi^2} \sigma_x^2 < \sigma_x^2. \end{aligned}$$

In other words, this setting is observationally equivalent to the "mute model," with private signals of higher quality. An increase in the precision of private information causes citizens to be less responsive to the rumor in the "mute model." See the dashed line in Figure 8 for the equilibrium survival threshold in the model with the modified

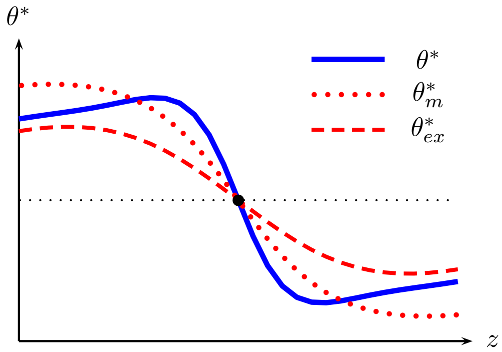


Figure 8. *Equilibrium survival threshold for regime strength is less responsive to rumors when citizens communicate with one another about their private signals.*

assumption concerning the exchange of private signals (labeled θ^*_{ex}).

The contrast between these two alternative models highlights that the topic of conversations matters: when citizens talk about what they privately know, they put less emphasis on public information; when they exchange views on the public signal that they commonly observe, they rely more on the public information.

Exchange both. Our analysis of communication would not be affected if we allow citizens to exchange their private signal, as long as conversation about the rumor is also allowed. Consider a hybrid model by allowing citizen i to exchange private information with a random peer k and to exchange views on informativeness of the rumor with another random peer k'. As the previous analysis suggests, the exchange of private information only improves its precision.

Because our analysis holds true for any precision of private information, changing its precision would not affect our results. One might conjecture that our results will be undermined once we allow for private-information exchange, but this is not necessarily the case either. In fact, Proposition 5 implies that more precise private information (higher β) can strengthen the effect of the rumor under certain conditions.

5.3. Strategic Communication

There are two key features of communication protocol in the benchmark model that drive our main results regarding the effect of rumors. First, agents follow an "interval" message sending rule. They send confirmatory messages when their private signals are within a certain interval. Second, this interval is not fixed. Instead, it is anchored by the rumor: the interval is symmetric around and shifts with the rumor's realization. The implication is that people confirm the veracity of the rumor when what they know is sufficiently close to it. We have seen how these two features contribute to making communication a potent mechanism to coordinate collective action.

Although non-strategic communication can be a realistic description of many conversations between acquaintances, especially when each individual has a negligible effect on the aggregate outcome, there are situations when strategic concerns also play a role. In times of turmoil, a citizen may not know who is a friend or an enemy, and she may have to be careful about what she tells her peer. In this subsection, we provide a model of strategic communication that captures some of these concerns. The resulting equilibrium communication rule is not mechanical as in the benchmark model, but still delivers the two key features of the benchmark communication protocol that drive our main results.

Assume that agents are randomly paired up and play a communication game. Each agent can choose to say "Yes, I believe it," or "No, I do not believe it." In contrast to the costless message sending setup in the benchmark model, agents may incur two types of cost. First, if one's message is subsequently contradicted by the facts, the sender may face penalties for spreading false rumors or for instilling skepticism toward legitimate news. This is especially relevant in autocratic regimes where the authority tends to punish people who express sympathetic views on unverified rumors. Alternatively, the sender may also incur a psychological cost when she realizes that she had misled her peer. We let c_1 be the expected cost to an agent when she says "yes" but the rumor turns out to be informative, and c_0 be the expected cost when she says "no" but the rumor turns out to be uninformative. Second, one does not fully know who one is talking to in a casual conversation. If a sender's assessment is at odds with that of her peer's, then she may face a chance that her peer may turn her in to the regime or to the rebels, so that she is singled out and punished.[29] Thus, in sending a message to her peer, an agent's cost depends on what the other agent's message is. We let d_1 be the expected cost to an agent if she says "yes" but her peer says "no," and let d_0 be the expected cost if she says "no" but her peer says "yes." The cost is normalized to 0 if both agents express the same assessment. The cost matrix to player i is summarized by Table 1, where $\mathbb{1}\{\cdot\}$ is the indicator function.[30]

[29]When they have different assessments on the same rumor, one of them can be considered as being unfriendly to the regime. There is a chance that the other agent betrays her and turns her in to the regime. Moreover, an agent who expresses assessments that are sympathetic to the regime can also be turned in by her peer and punished by the rebels for betraying the cause of the revolution.

[30]Note that the cost parameters c_l and d_l can be functions of z. But it does not affect the equilibrium results, since agents play this game after they observe the rumor.

Table 1. *Player i's cost matrix.*

		Player k "yes"	"no"
Player i	"yes"	$\mathbb{1}\{z \sim U\} \cdot c_1$	$\mathbb{1}\{z \sim U\} \cdot c_1 + d_1$
	"no"	$\mathbb{1}\{z \sim I\} \cdot c_0 + d_0$	$\mathbb{1}\{z \sim I\} \cdot c_0$

A message sending rule is a set S such that a player chooses to express confirmatory view if and only if her private information is in S. Given that player i expects her peer player k to adopt the message sending rule S, her assessment that player k will send a confirmatory message is

$$q(S; z, x_i) \equiv \omega(z, x_i) \Pr[x_k \in S | z, x_i, z \sim I] + (1 - \omega(z, x_i)) \Pr[x_k \in S | z, x_i, z \sim U].$$

Using the cost matrix in Table 1, player i chooses to say "yes" if and only if

$$q(S; z, x_i) \geq \frac{d_1 + c_1 - \omega(z, x_i)(c_0 + c_1)}{d_0 + d_1}. \tag{24}$$

The equilibrium message sending rule S^* is a fixed point: given agents follow the message sending rule S^*, inequality (24) hold if and only if $x_i \in S^*$.

Proposition 6. *If $c_1 > d_0 > c_1 - \omega(z, z)(c_0 + c_1)$, there exists an interval equilibrium message sending rule S^* and it is symmetric about the rumor z.*

When $d_0 \geq c_1$, it is an equilibrium that everybody endorses the rumor. If agent i expects that her peer will say "yes" and the cost c_1 of being punished for endorsing a false rumor is smaller relative to the cost d_0 of disagreeing with one's peer and being turned in to the regime or rebels, the agent i chooses to say "yes" as well, and no meaningful communication occurs. On the other hand, when $d_0 \leq c_1 - \omega(z, z)(c_0 + c_1)$, even the most "confident" agent whose private information is confirmed by the rumor (i.e., $x_i = z$) finds it costly to endorse its veracity. In that case, no citizen would say "yes" even if she expects her peer to say "yes."

If d_0 is neither too high nor too low, (i.e., it satisfies the restriction stated in Proposition 6), then it can support an equilibrium in which citizens who find the rumor sufficiently plausible would choose to say that they believe the rumor. To see this, suppose that the message sending rule takes the form of an interval: $S = [\underline{x}, \overline{x}]$. Then,

$$\begin{aligned} q(S; z, x_i) = \omega(z, x_i) & \left(\Phi\left(\frac{\overline{x} - \beta x_i - (1-\beta) z}{\sqrt{\beta + 1}\sigma_x} \right) - \Phi\left(\frac{\underline{x} - \beta x_i - (1-\beta) z}{\sqrt{\beta + 1}\sigma_x} \right) \right) \\ & + (1 - \omega(z, x_i)) \left(\Phi\left(\frac{\overline{x} - x_i}{\sqrt{2}\sigma_x} \right) - \Phi\left(\frac{\underline{x} - x_i}{\sqrt{2}\sigma_x} \right) \right). \end{aligned}$$

If S is symmetric about z, $q(S; z, x_i)$ increases in x_i from zero, peaks at $x_i = z$ and then decreases towards zero. Intuitively, when x_i is extreme, player i tends to believe that player k's private signal is also very extreme and the probability that x_k falls in any intermediate interval S is low. Thus the left-hand-side of inequality (24) is hump-shaped in x_i. Because of the property of $\omega(z, x_i)$, the right-hand-side of (24) is inverted hump-shaped in x_i. Proposition 6 shows that there exists an interval $[\underline{x}, \overline{x}]$ such that the left-hand-side exceeds the right-hand-side if and only if x_i belongs to this interval. Further, this interval must contain z and is symmetric about it. For any symmetric decision rule S, both $q(S; z, \cdot)$ and $w(z, \cdot)$ are symmetric about the rumor z. Given the symmetry of both sides of (24), the equilibrium message sending rule $S^*(z)$ is symmetric as well.

Depending on the cost parameters, the endogenous decision interval $S^*(z)$ may be larger or smaller than that specified in the benchmark case with exogenous plausibility threshold δ. Instead of reporting whether they believe a rumor mechanically, agents choose their messages depending on expected costs and on what they believe their peers will report. However, this equilibrium communication rule shares the same key features as the exogenous communication rule in the benchmark case. In terms of the effect of communication on equilibrium in the attack stage of the game, the two alternative communication protocols deliver the same qualitative results.

5.4. Robustness

Our baseline specification on communication is the simplest possible one that allows citizens to exchange views on the informativeness of the rumor. When we relax various simplifying assumptions, our main results still hold. We discuss three examples.

Finer messages. First, we investigate the case where citizens can exchange finer messages, rather than simply $y = 1$ or 0. Specifically, if citizens decide to send their peers contradictory messages $y = 0$, we allow them to justify and explain why they do not believe the rumor is informative. When $x_k < \underline{x}(z)$, citizen k can send her peer a message $y = 0L$, interpreted to mean "the rumor is not informative because it indicates that the regime is too much stronger than I believe;" she sends a message $y = 0R$ when $x_k > \overline{x}(z)$; and she sends $y = 1$ when $x_k \in [\underline{x}(z), \overline{x}(z)]$.

In the baseline model, the key result that rumors can be even more potent than trustworthy news is driven by the fact that the effect of communication can dominate skepticism when a confirmatory message is received and by the fact that the fraction of population who receives those messages can be sufficiently large. The same driving force can also dominate in this specification and lead to a similar result.

Meeting with like-minded citizens. Second, in our benchmark model, we assume that agents meet in a completely random manner. However, one may argue that in reality, people are more likely to meet someone who shares similar beliefs. To capture the spirit of such a semi-random meeting technology, we assume that citizens split into two groups after they hear the rumor: those whose private information is lower than the value of z form a group L and the rest form a group R. Then citizens are randomly matched in pairs in each group. Citizen k from L sends her peer a message $y = 0$ when $x_k < \underline{x}(z)$; or $y = 1$, when $\underline{x}(z) < x_k < z$. Similarly, citizen k' from R sends her peer $y = 0$ when $x_{k'} > \overline{x}(z)$; or $y = 1$ when $z < x_{k'} < \overline{x}(z)$.

This setup is quite similar to the specification of exchanging finer messages discussed above. The difference is that the group of agents who receive $y = 1$ in that case splits into two in this case. However, recipients of confirmatory messages in both groups consider those messages as evidence for θ being close to z. The same logic of the "finer messages" case continues to apply here and rumors can still be more potent than trustworthy news in this case.

Meeting with more than one citizen. Third, we can also extend our model to the case where agents meet up and communicate in a group. Specifically, N agents are randomly grouped together in the communication stage and then each of them sends one message to and receives $N - 1$ messages from his $N - 1$ group members simultaneously. Naturally, there are N types of agents after they communicate, depending on the number of confirmatory messages received, n, where $n \in \{0, 1, \ldots, N - 1\}$. It turns out that the posterior densities of agents that receive a larger number of confirmatory messages are more concentrated around the rumor and therefore become more responsive to it. It is also similar to the benchmark case in that the fraction of agents who receive a high number of confirmatory messages is larger if the rumor happens to be closer to the true fundamental. Therefore, our key results can still extend.[31]

5.5. Censorship: The Power of Silence

Our model shows that rumors against a regime could coordinate more citizens to attack, provided they are close to the truth. Thus autocratic governments may want to block negative rumors against it and to stop citizens from talking about these rumors. In reality, news censorship and the control of information flow are commonly adopted by many autocratic governments concerned about their survival.

However, does it always help the regime to survive if censorship is adopted to screen out all negative information? When citizens hear no rumor about the regime,

[31] We elaborate on the specification of this case in Technical Appendix D and present some results from the case with $N = 3$. It is worthwhile to point out that the case where infinitely many citizens can meet up and exchange information is different. If agents receive infinitely many messages, by the law of large numbers, they can back out the fundamental θ given the faction of confirmatory messages they receive and the message sending rule.

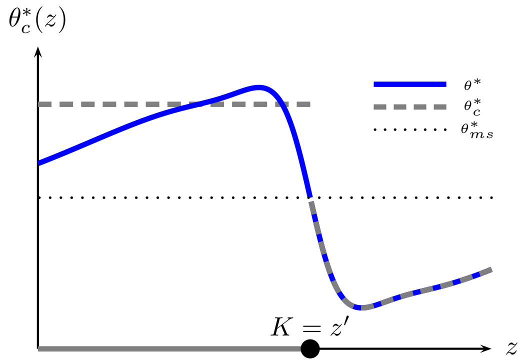


Figure 9. *The dashed line shows the equilibrium regime survival threshold in the "censorship model." Censorship can sometimes hurt the regime by raising the survival threshold relative to that in the "communication model."*

silence itself becomes a public signal about the regime strength, and it is not obvious that news censorship necessarily increases the chances of regime survival.

Assume that the regime blocks any rumor z if $z < K$. In other words, only rumors that suggest the regime is stronger than K can be heard and discussed by citizens. We assume that citizens are aware of this censorship rule. Therefore it is common knowledge that $z < K$ when citizens do not hear a rumor. In this case, citizens cannot communicate about the informativeness of z, since they do not observe it.

When z is observable, i.e., $z \geq K$, the equilibrium is exactly the same as in our "communication model." When citizens hear no rumor, they understand that the event $z < K$ has taken place but the authority has censored the news. Taking θ_c as the threshold for regime survival, they calculate the expected payoff of revolt in a Bayesian fashion,

$$\Pr[\theta \leq \theta_c | z < K, x_i] = \frac{\int_{-\infty}^{\theta_c} \left[\alpha\Phi\left(\frac{K-t}{\sigma_z}\right) + (1-\alpha)\Phi\left(\frac{K-s}{\sigma_U}\right)\right] \frac{1}{\sigma_x}\phi\left(\frac{t-x_i}{\sigma_x}\right) \mathrm{d}t}{\int_{-\infty}^{\infty} \left[\alpha\Phi\left(\frac{K-t}{\sigma_z}\right) + (1-\alpha)\Phi\left(\frac{K-s}{\sigma_U}\right)\right] \frac{1}{\sigma_x}\phi\left(\frac{t-x_i}{\sigma_x}\right) \mathrm{d}t}.$$

Now we return to the issue raised at the beginning of this subsection and examine the effect of censorship on regime survival in the case where $K = z'$, i.e., the regime blocks any negative rumors against it. See Figure 9 for illustration. The effect of censorship is represented by the difference between the solid line for the "communication model" and the dashed line for the model with censorship.

We note that the "censorship model" is a variant of the "mute model" when $z < K = z'$, since citizens cannot communicate when there is no rumor to talk about. In a sense, citizens pool the effects of negative rumors, both more and less dangerous ones, and use an average to make a decision, given they cannot observe the specific realization of z. Therefore, censorship does help the regime to survive when the realizations of rumors are the most dangerous ones (i.e., when $\theta^*(z)$ is around the peak value), but hurts the regime, when z is very far or close to z'.

Alternatively, the regime can also commit to a strategy that it blocks all the rumors, no matter they are good or bad, so that agents do not have any access to the public information. These two strategies have different implications for the ex ante survival probability of the regime. In the case of censoring all the news, the survival probability of the regime is 0 if the fundamental θ is smaller than θ^*_{ms}, and 1 if $\theta > \theta^*_{ms}$. In the case of censoring bad news only, the survival probability of the regime depends on θ. The stronger is the regime strength, the more likely does it survive. Once the strength θ is above θ^*_c, the regime can survive with probability 1.

We have shown that, for a negative rumor, $\theta^*_{ms} < \theta^*_c$. If the regime strength is very high (i.e., $\theta > \theta^*_c$), the censorship strategy does not make much difference, because the regime would survive in both cases. If the regime strength is intermediate (i.e., $\theta \in [\theta^*_{ms}, \theta^*_c]$), the regime is better off blocking all public information. The regime survives with probability 1 if all the public news are blocked, but has a chance to collapse if it blocks bad news only. The reason is that, in the latter case, when agents observe that the regime is suppressing rumors, the silence itself is a bad news and works against the regime. If the regime strength is too low (i.e., $\theta < \theta^*_{ms}$), the regime must collapse if agents cannot access any public information. But the regime can survive with a certain probability if it screens out bad news only, because there might be a chance that the rumor is in favor of the regime, which may help a weak one to sustain.

6. Conclusion

Social interaction is an important source of information for individuals, especially when the coordination motive is important. Our paper highlights the significance of this channel and contributes to this interesting but under-explored topic.

It is not news that revolutions in history are often intertwined with rumors. However, what strikes us is why some rumors, which often turned out to be false later, could be so effective for mobilization, while others were simply ignored. We offer an analysis of this phenomenon by focusing on two key aspects of rumors: that they may or may not be true, and that people talk about them. In this model, individuals' skepticism toward rumors arises as a rational response instead of a behavioral assumption. Moreover, we explicate a novel mechanism where the effect of citizens' skepticism can be undone or reinforced by communication among themselves.

To the best of our knowledge, our model is the first attempt to explicitly investigate the role of rumors in a regime change game. Our theory is interpreted in the context of political revolution, but it can also be extended to model rumors in bank runs, financial crises or currency attacks. We have not, however, addressed questions about how rumors originate or how they spread. Although we explore a number of communication protocols in this paper, our analysis is confined to decentralized communication with pairwise matching. The role played by social networks, mass media, and modern communication technologies in promulgating or abating rumors remains to be studied.

Appendix

Proof of Proposition 1. *Part 1.* Combine the indifference condition and the critical mass conditions in the "mute model," we get

$$\theta_m^* = (1-c) + \left[\Phi\left(\frac{\theta_m^* - (\beta x_m^* + (1-\beta)z)}{\sqrt{\beta}\sigma_x}\right) - c\right] \cdot \frac{w(z, x_m^*)}{1 - w(z, x_m^*)}.$$

For any finite x_m^*, $\lim_{z\to\infty} w(z, x_m^*) = 0$. Therefore,

$$\lim_{z\to\infty} \theta_m^*(z) = 1 - c = \theta_{ms}^*.$$

As z goes to infinity, $x_m^*(z)$ must remain finite. To see why this is the case, suppose z is sufficiently large and x^* goes to negative infinity, the equation above implies that $\theta_m^* = 1 - c$ and $A(\theta_m^*; z) = 0$. The critical mass condition is violated. Next suppose that z is sufficiently large and x^* goes to infinity. The equation above implies that $\theta_m^* < 1 - c$, but the fact that x^* is sufficiently large implies that $A(\theta_m^*; z) = 1$. The critical mass condition is again violated. A similar argument establishes that $\lim_{z\to-\infty} \theta_m^*(z) = \theta_{ms}^*$.

Part 2. Since $(\theta_{ms}^*, x_{ms}^*)$ satisfies the indifference condition of the "pure noise model," we have

$$\Phi\left(\frac{\theta_{ms}^* - x_{ms}^*}{\sigma_x}\right) = c.$$

Since $(\theta_{ms}^*, x_{ms}^*)$ also satisfies the indifference condition of the "public signal model" when $z = \tilde{z}$, we have

$$\Phi\left(\frac{\theta_{ms}^* - (\beta x_{ms}^* + (1-\beta)\tilde{z})}{\sqrt{\beta}\sigma_x}\right) = c.$$

The posterior belief $P(\cdot|\tilde{z}, x)$ in the "mute model" is just the weighted average of the left-hand-side of the two equations above. Hence $P(\theta_{ms}^*|\tilde{z}, x_{ms}^*) = c$. Furthermore, $(\theta_{ms}^*, x_{ms}^*)$ satisfies the critical mass condition of the "mute model." Therefore, it is an equilibrium for the "mute model" when $z = \tilde{z}$.

Part 3. We first show that if $z < \tilde{z}$, then $\theta_m^*(z) < \theta_{ps}^*(z)$. From the critical mass condition in the "mute model," we have $\Phi\left(\sigma_x^{-1}(\theta_m^* - x_m^*)\right) = 1 - \theta_m^*$. Therefore, the indifference condition in the "mute model" can be written as:

$$c = w(z, x_m^*)\Phi\left(\frac{\theta_m^* - (\beta x_m^* + (1-\beta)z)}{\sqrt{\beta}\sigma_x}\right) + (1 - w(z, x_m^*))(1 - \theta_m^*).$$

From the indifference condition of the "public signal model" and from the fact that

$1-\theta^*_{ms}=c$, we also have

$$c = w(z, x^*_m)\Phi\left(\frac{\theta^*_{ps} - (\beta x^*_{ps} + (1-\beta)z)}{\sqrt{\beta}\sigma_x}\right) + (1 - w(z, x^*_m))(1 - \theta^*_{ms}).$$

These two equations, together with the fact that $\theta^*_{ps} > \theta^*_{ms}$ when $z < \tilde{z}$, imply

$$\begin{aligned} & w(z, x^*_m)\Phi\left(\frac{q(\theta^*_m) - (1-\beta)z}{\sqrt{\beta}\sigma_x}\right) + (1 - w(z, x^*_m))(1 - \theta^*_m) \\ > & w(z, x^*_m)\Phi\left(\frac{q(\theta^*_{ps}) - (1-\beta)z}{\sqrt{\beta}\sigma_x}\right) + (1 - w(z, x^*_m))(1 - \theta^*_{ps}), \end{aligned}$$

where

$$q(\theta^*) \equiv \theta^* - \beta x^* = \theta^* - \beta(\theta^* + \sigma_x \Phi^{-1}(\theta^*)).$$

To show $\theta^*_m < \theta^*_{ps}$ from the above inequality, it suffices to show that $\mathrm{d}q(\theta^*)/\mathrm{d}\theta^* \leq 0$. We have

$$\frac{\mathrm{d}q(\theta^*)}{\mathrm{d}\theta^*} = 1 - \beta - \frac{\beta\sigma_x}{\phi\left(\Phi^{-1}(\theta^*)\right)} \leq 1 - \beta - \beta\sigma_x\sqrt{2\pi},$$

which is negative by assumption (1). Hence $\theta^*_m < \theta^*_{ps}$ when $z < \tilde{z}$.

Next, we show that $\theta^*_m(z) > \theta^*_{ms}$ if $z < \tilde{z}$. Suppose this is not true. Then $\Phi(\sigma_x^{-1}(\theta^*_m - x^*_m)) = 1 - \theta^*_m \geq 1 - \theta^*_{ms} = c$, which implies

$$\Phi\left(\frac{\theta^*_m - (\beta x^*_m + (1-\beta)z)}{\sqrt{\beta}\sigma_x}\right) \leq c.$$

Moreover, the fact that $(\theta^*_{ms}, x^*_{ms})$ satisfies the indifference condition of the "public signal model" at $z = \tilde{z}$ implies

$$\Phi\left(\frac{\theta^*_{ms} - (\beta x^*_{ms} + (1-\beta)\tilde{z})}{\sqrt{\beta}\sigma_x}\right) = c.$$

These two conditions can be combined to give

$$q(\theta^*_m) - q(\theta^*_{ms}) \leq (1-\beta)(z - \tilde{z}) < 0.$$

Since $\mathrm{d}q(\theta^*)/\mathrm{d}\theta^* < 0$, this inequality implies $\theta^*_m > \theta^*_{ms}$, a contradiction. Thus, when $z < \tilde{z}$, we must have $\theta^*_m(z) \in (\theta^*_{ms}, \theta^*_{ps}(z))$. When $z \geq \tilde{z}$, the proof is symmetric.

The following is to show that (a) $\theta^*_m(z)$ is increasing then decreasing for $z \in (-\infty, \tilde{z})$;

and (b) $\theta_m^*(z)$ is decreasing then increasing for $z \in (\tilde{z}, \infty)$. Fix a $z_0 \in (-\infty, \tilde{z})$. Define

$$f(z) \equiv P(\theta_m^*(z_0)|z, x_m^*(z_0)).$$

We show that f is single-peaked in z for $z \in (-\infty, \tilde{z})$. It suffices to verify that $\mathrm{d}f(z)/\mathrm{d}z = 0$ implies $\mathrm{d}^2 f(z)/\mathrm{d}z^2 < 0$. To simplify the notation, we use the subscript I or U to denote the posterior distribution (or density) when z is known to be informative or uninformative, respectively. We have

$$\begin{aligned}
\frac{\mathrm{d}f(z)}{\mathrm{d}z} &= (\Phi_I - \Phi_U)\frac{\partial w}{\partial z} - w\frac{\partial \Phi_I}{\partial z} \\
&= \left[(1-w)\left(\frac{z-s}{\sigma_U^2} + \frac{x_m^*(z_0) - z}{\sigma_I^2}\right)\left(1 - \frac{\Phi_U}{\Phi_I}\right) - \frac{1-\beta}{\sqrt{\beta}\sigma_x}\frac{\phi_I}{\Phi_I}\right] w\Phi_I.
\end{aligned}$$

When $\mathrm{d}f(z)/\mathrm{d}z = 0$, the second derivative is given by:

$$\begin{aligned}
\frac{1}{w\Phi_I}\frac{\mathrm{d}^2 f(z)}{\mathrm{d}z^2} = &-w(1-w)\left(\frac{z-s}{\sigma_U^2} + \frac{x_m^*(z_0) - z}{\sigma_I^2}\right)^2\left(1 - \frac{\Phi_U}{\Phi_I}\right) \\
&+ (1-w)\left(\frac{1}{\sigma_U^2} - \frac{1}{\sigma_I^2}\right)\left(1 - \frac{\Phi_U}{\Phi_I}\right) \\
&- (1-w)\left(\frac{z-s}{\sigma_U^2} + \frac{x_m^*(z_0) - z}{\sigma_I^2}\right)\left(\frac{\partial\left(\Phi_U/\Phi_I\right)}{\partial z}\right) \\
&- \left(\frac{1-\beta}{\sigma_x\sqrt{\beta}}\right)\frac{\partial\left(\phi_I/\Phi_I\right)}{\partial z}.
\end{aligned}$$

(i) The first term is negative because $\Phi_I > \Phi_U$ for $z < \tilde{z}$. To see this, suppose the contrary is true. Then $\Phi_I \leq \Phi_U$ implies

$$\theta_m^*(z_0) + \sqrt{\beta}x_m^*(z_0) \leq (1 + \sqrt{\beta})z.$$

But for $z_0 \leq \tilde{z}$ the left-hand side is greater than $\theta_{ms}^* + \sqrt{\beta}x_{ms}^*$, which is equal to $(1+\sqrt{\beta})\tilde{z}$, a contradiction. (ii) The second term is negative because $\sigma_U^2 > \sigma_I^2$ from parameter restriction (2). (iii) The third term is negative because $(z-s)/\sigma_U^2 + (x_m^*(z_0) - z)/\sigma_I^2 > 0$ whenever $\mathrm{d}f(z)/\mathrm{d}z = 0$ and because Φ_U/Φ_I is increasing in z. (iv) The fourth term is negative because the function ϕ_I/Φ_I is increasing in z.

The single-peakedness of $f(z)$ for $z \in (-\infty, \tilde{z})$ implies that in this range there can be at most one $z_1 \neq z_0$ such that $\theta_m^*(z_1) = \theta_m^*(z_0)$. Suppose otherwise. Let $z_1 \neq z_2 \neq z_0$ be such that $\theta_m^*(z_1) = \theta_m^*(z_2) = \theta_m^*(z_0)$. By the critical mass condition, this implies $x_m^*(z_1) = x_m^*(z_2) = x_m^*(z_0)$. Since $(\theta_m^*(z_0), x_m^*(z_0))$ satisfies the equilibrium conditions for $z \in \{z_1, z_2, z_0\}$, the equation $f(z) = c$ has at least three solutions, which contradicts the single-peakedness of f.

In parts 1 and 2 of the proposition, we already established that $\theta^*_m(z)$ is higher than θ^*_{ms} for $z \in (-\infty, \tilde{z})$, and approaches it when z goes to minus infinity or to $\tilde{z}$. Thus, $\theta^*_m(z)$ must be increasing for z sufficiently negative, and decreasing for z sufficiently close to $\tilde{z}$. Together with the fact that for any z_0 in this range, there can be at most one z_1 such that $\theta^*_m(z_0) = \theta^*_m(z_1)$, this implies that $\theta^*_m(z)$ must be increasing then decreasing in this range.

For the case $z \in (\tilde{z}, \infty)$, write:

$$\frac{\mathrm{d}f(z)}{\mathrm{d}z} = \left[(1-w)\left(\frac{z-s}{\sigma_U^2} + \frac{x^*_m(z_0)-z}{\sigma_I^2}\right)\left(\frac{1-\Phi_U}{1-\Phi_I} - 1\right) - \frac{1-\beta}{\sqrt{\beta}\sigma_x}\frac{\phi_I}{1-\Phi_I}\right] w(1-\Phi_I).$$

We can show that the bracketed term is increasing when it is equal to zero, because $\Phi_I < \Phi_U$ when $z > \tilde{z}$ and because $\phi_I/(1-\Phi_I)$ is decreasing in z. Hence $f(z)$ must be decreasing then increasing in z in this range. Following similar reasoning as in the earlier case, this implies that $\theta^*_m(z)$ is decreasing then increasing for $z \in (\tilde{z}, \infty)$. ■

Proof of Proposition 2. *Part 1.* We only prove the limiting properties of $x^*_U(z)$ and $\theta^*(z)$ here. The limiting properties of $x^*_I(z)$ require comparing the rates of convergence of different functions, and they are formally established in Lemma 2 in the Technical Appendix.

Suppose the limit values of both $x^*_U(z)$ and $\theta^*(z)$ are finite. The indifference condition (11) in "communication model" requires

$$\lim_{z\to\infty} \int_{-\infty}^{\theta^*} \frac{1-J(t,z)}{\Pr[y_i = 0|z, x^*_U]} p(t|z, x^*_U)\,\mathrm{d}t = c.$$

By Lemma 2 (Claim 1) in the Technical Appendix, both $\underline{x}(z)$ and $\overline{x}(z)$ go to infinity as z goes to infinity. Therefore, for any $t \leq \theta^*$, the probability that x_j does not belong to $[\underline{x}(z), \overline{x}(z)]$ goes to one. We thus have

$$\lim_{z\to\infty} \frac{1-J(t,z)}{\Pr[y_i = 0|z, x^*_U]} = 1.$$

The indifference condition for type x^*_U becomes:

$$\lim_{z\to\infty} \int_{-\infty}^{\theta^*} p(t|z, x^*_U)\,\mathrm{d}t = \Phi\left(\frac{\theta^* - x^*_U}{\sigma_x}\right) = c,$$

where the first equality follows because $w(z, x^*_U)$ goes to 0 as z goes to infinity.

When z goes to infinity, $J(\theta^*, z)$ goes to zero. Therefore the critical mass condition (12) for the "communication model" becomes,

$$\lim_{z \to \infty} J(\theta^*, z)\Phi\left(\frac{x_I^* - \theta^*}{\sigma_x}\right) + (1 - J(\theta^*, z))\Phi\left(\frac{x_U^* - \theta^*}{\sigma_x}\right) = \Phi\left(\frac{x_U^* - \theta^*}{\sigma_x}\right) = \theta^*.$$

Given (θ^*, x_U^*) solves the same equation system as that in the "pure noise model," we conclude that $\lim_{z\to\infty} x_U^*(z) = x_{ms}^*$ and $\lim_{z\to\infty} \theta^*(z) = \theta_{ms}^*$. The proof of the case for the limit as z goes to minus infinity is analogous.

Part 2. From part 1, $x_I^*(z) > x_U^*(z)$ for z sufficiently negative and $x_I^*(z) < x_U^*(z)$ for z sufficiently large. Both $x_I^*(z)$ and $x_U^*(z)$ are continuous. Therefore there exists a z' such that $x_I^*(z') = x_U^*(z')$.

Let $\theta^*(z') = \theta'$ and $x_I^*(z') = x_U^*(z') = x'$. We proceed to establish that (θ', x') solves the "mute model" as well. To see this, we first note that $x_I^* = x_U^*$ implies that the critical mass condition (12) of the "communication model" reduces to its counterpart in the "mute model." Next, note that for any value of z', x' and θ', we have

$$P(\theta'|z', x') = \Pr[y_i = 1|z', x']P(\theta'|z', x', 1) + \Pr[y_i = 0|z', x']P(\theta'|z', x', 0).$$

Therefore, if z', x', and θ' satisfy the indifference conditions $P(\theta'|z', x', 1) = c$ and $P(\theta'|z', x', 0) = c$ in the "communication model," then they must satisfy the indifference condition $P(\theta'|z', x') = c$ in the "mute model" as well.

Part 3. We first show that z' coincides with $\tilde{z}$ when $c = 0.5$. Let $\theta_m^*(\tilde{z}) = \tilde{\theta}$ and $x_m^*(\tilde{z}) = \tilde{x}$. We need to show that $(\tilde{\theta}, \tilde{x}, \tilde{x})$ solves the "communication model" at $z = \tilde{z}$, when $c = 0.5$. Toward this end, it suffices to show that $P(\tilde{\theta}|\tilde{z}, \tilde{x}, 1) = c$. Given that $(\tilde{\theta}, \tilde{x})$ solves the "mute model" at $z = \tilde{z}$, we have $P(\tilde{\theta}|\tilde{z}, \tilde{x}) = c$. These two conditions would imply that $P(\tilde{\theta}|\tilde{z}, \tilde{x}, 0) = c$. Hence the indifference condition (11) for the "communication model" is satisfied. Given that $x_I^*(\tilde{z}) = x_U^*(\tilde{z}) = \tilde{x}$, the critical mass condition (12) holds as well. To see this, note that both $J(t, \tilde{z})$ and $p(t|\tilde{z}, \tilde{z})$ are symmetric about the point $t = \tilde{z}$, which gives

$$\int_{-\infty}^{\tilde{z}} J(t, \tilde{z})p(t|\tilde{z}, \tilde{z})\,\mathrm{d}t = 0.5\int_{-\infty}^{\infty} J(t, \tilde{z})p(t|\tilde{z}, \tilde{z})\,\mathrm{d}t$$

Hence, $P(\tilde{z}|\tilde{z}, \tilde{z}, 1) = 0.5$. For $c = 0.5$, $\tilde{\theta} = \tilde{x} = \tilde{z}$. Therefore, we have $P(\tilde{\theta}|\tilde{z}, \tilde{x}, 1) = c$. This establishes that $\theta^*(\tilde{z}) = \theta_m^*(\tilde{z})$ and $x_I^*(\tilde{z}) = x_U^*(\tilde{z}) = x_m^*(\tilde{z})$.

Next we show that $x_I^*(z) > x_U^*(z)$ for $z < z'$. Suppose otherwise. Then, since $\lim_{z\to-\infty} x_I^*(z) > \lim_{z\to-\infty} x_U^*(z)$, there must exist some $z_0 < z'$ such that $x_I^*(z_0) = x_U^*(z_0)$. According to Lemma 4 in the Technical Appendix, for any z_0, there exists a unique pair $(\hat{\theta}_0, x_0)$ such that $P(\hat{\theta}_0|z_0, x_0, 1) = P(\hat{\theta}_0|z_0, x_0, 0) = c$. It is easy to verify that such a pair is $(\hat{\theta}_0, x_0) = (z_0, z_0)$. However, such a pair does not satisfy the critical mass condition, because $A(z_0; z_0, z_0, z_0) \neq z_0$, a contradiction. A symmetric argument shows that $x_I^*(z) < x_U^*(z)$ for $z > z'$. ■

Proof of Proposition 3. We prove the second half of this proposition; the proof of the first part is analogous. We proceed by constructing a contradiction. Suppose $\theta^*(z) \leq \theta_m^*(z)$ for z sufficiently large. By Lemma 6 in the Technical Appendix, $0 < \partial \hat{x}_m / \partial \hat{\theta} < 1$ when z is sufficiently large. This implies

$$x_m^*\left(\theta_m^*(z)\right) - \theta_m^*(z) \leq x_m^*\left(\theta^*(z)\right) - \theta^*(z).$$

Since Φ is monotone, the above inequality together with the critical mass condition for the "mute model" imply:

$$\theta_m^*(z) = \Phi\left(\frac{x_m^*\left(\theta_m^*(z)\right) - \theta_m^*(z)}{\sigma_x}\right) \leq \Phi\left(\frac{x_m^*\left(\theta^*(z)\right) - \theta^*(z)}{\sigma_x}\right).$$

Lemma 7 in the Technical Appendix establishes that, for z sufficiently large and for any $\hat{\theta}$,

$$\left(1 - J(\hat{\theta}, z)\right)\Phi\left(\frac{\hat{x}_U(\hat{\theta}) - \hat{\theta}}{\sigma_x}\right) > \Phi\left(\frac{\hat{x}_m(\hat{\theta}) - \hat{\theta}}{\sigma_x}\right),$$

which implies that

$$J(\hat{\theta}, z)\Phi\left(\frac{\hat{x}_I(\hat{\theta}) - \hat{\theta}}{\sigma_x}\right) + \left(1 - J(\hat{\theta}, z)\right)\Phi\left(\frac{\hat{x}_U(\hat{\theta}) - \hat{\theta}}{\sigma_x}\right) > \Phi\left(\frac{\hat{x}_m(\hat{\theta}) - \hat{\theta}}{\sigma_x}\right).$$

Evaluating both sides at $\hat{\theta} = \theta^*(z)$, it follows that

$$\theta^*(z) > \Phi\left(\frac{x_m^*\left(\theta^*(z)\right) - \theta^*(z)}{\sigma_x}\right) \geq \theta_m^*(z),$$

a contradiction. ■

Proof of Proposition 4. By Claim 2 in the proof of Lemma 1,

$$
\begin{aligned}
& J\frac{\partial \hat{x}_I}{\partial \hat{\theta}} + (1-J)\frac{\partial \hat{x}_U}{\partial \hat{\theta}} \\
=& -p(\theta'|z',x')\left[J(\theta',z')\frac{1}{\int_{-\infty}^{\theta'} \frac{J(t,z')}{J(\theta',z')}\frac{\partial p(t|z',x')}{\partial x}\,\mathrm{d}t} + (1-J(\theta',z'))\frac{1}{\int_{-\infty}^{\theta'} \frac{1-J(t,z')}{1-J(\theta',z')}\frac{\partial p(t|z',x')}{\partial x}\,\mathrm{d}t} \right] \\
>& -p(\theta'|z',x')\frac{1}{J(\theta',z')\int_{-\infty}^{\theta'} \frac{J(t,z')}{J(\theta',z')}\frac{\partial p(t|z',x')}{\partial x}\,\mathrm{d}t + (1-J(\theta',z'))\int_{-\infty}^{\theta'} \frac{1-J(t,z')}{1-J(\theta',z')}\frac{\partial p(t|z',x')}{\partial x}\,\mathrm{d}t} \\
=& \frac{-p(\theta|z',x')}{\int_{-\infty}^{\theta'} \frac{\partial p(t|z',x')}{\partial x}\,\mathrm{d}t} \\
=& \frac{\partial \hat{x}_m}{\partial \hat{\theta}},
\end{aligned}
$$

where the inequality follows from Jensen's inequality and the fact that the function $1/t$ is concave for $t<0$. This establishes the first inequality of (20). Moreover, since

$$
P_x = -p + \frac{(1-\beta)w}{\sqrt{\beta}\sigma_x}\phi\left(\frac{\theta'-X'}{\sqrt{\beta}\sigma_x}\right) < 0,
$$

we have $\partial \hat{x}_m/\partial\hat{\theta} = -p/P_x > 1$. This establishes the second inequality of (20).

Inequality (19) follows from (20) and from Lemma 5 in the Technical Appendix, which shows that $\partial\hat{x}/\partial z = 1 - \partial\hat{x}/\partial\hat{\theta}$. Finally, inequalities (20) and (19), together with the comparison of the decomposition equation (21) of the "communication model," and its counterpart in the "mute model," establish (22). The proposition follows. ∎

Proof of Proposition 5. We write the difference between $\partial\hat{x}_{ps}/\partial\hat{\theta}$ and $\partial\hat{x}_m/\partial\hat{\theta}$ as:

$$
\frac{\partial \hat{x}_{ps}}{\partial \hat{\theta}} - \frac{\partial \hat{x}_m}{\partial \hat{\theta}} = \left(\frac{1-\beta}{\beta}\frac{1-w}{w\frac{1}{\sqrt{\beta}} + (1-w)}\right)\frac{\partial \hat{x}_m}{\partial \hat{\theta}} \equiv L_{ps}\frac{\partial \hat{x}_m}{\partial \hat{\theta}};
$$

and we write

$$
\left(J\frac{\partial \hat{x}_I}{\partial \hat{\theta}} + (1-J)\frac{\partial \hat{x}_U}{\partial \hat{\theta}}\right) - \frac{\partial \hat{x}_m}{\partial \hat{\theta}} = \left(\frac{D^2}{(JP_x - D)\left((1-J)P_x + D\right)}\right)\frac{\partial \hat{x}_m}{\partial \hat{\theta}} \equiv L\frac{\partial \hat{x}_m}{\partial \hat{\theta}},
$$

where

$$
D = \int_{-\infty}^{\theta'} \frac{\partial J(t,z')}{\partial t}\frac{\partial P(t|z',x')}{\partial x}\,\mathrm{d}t.
$$

Lemma 8 in the Technical Appendix shows that there always exists δ such that the lower bound of L is increasing in β and is bounded above 0 when β approaches 1. On the other hand, L_{ps} is bounded from above with an upper bound $(1-\beta)/\beta$. Observe that $(1-\beta)/\beta$ is decreasing in β and approaches 0 when β approaches 1. There must exist a threshold $\hat{\beta}$ such that, for any $\beta > \hat{\beta}$, there exists δ so that $L > L_{ps}$. By continuity, for any $\beta > \hat{\beta}$, there exists an interval $[\delta_1, \delta_2]$ so that for any $\delta \in [\delta_1, \delta_2]$, $L > L_{ps}$, which implies that inequality (23) holds. Using the same logic as in Proposition 4, we obtain $\mathrm{d}\theta^*(z')/\mathrm{d}z < \mathrm{d}\theta^*_{ps}(z')/\mathrm{d}z < 0$. The proposition follows. ■

Proof of Proposition 6. Consider a symmetric communication rule $S = [z-\Delta, z+\Delta]$. It is straightforward to see that both $\omega(z,\cdot)$ and $q(S;z,\cdot)$ are symmetric about z. Therefore, if $x_i = \underline{x} < z$ solves (24) with equality, then $x_i = \overline{x} = z + (z - \underline{x}) > z$ also does.

To establish the existence of interval decision rules, we denote $\tilde{q}(\Delta) = q(S;z,z-\Delta)$ and $\tilde{\omega}(\Delta) = \omega(z, z-\Delta)$. We need to establish that there exists $\Delta > 0$, such that

$$\tilde{q}(\Delta) = \frac{d_1 + c_1 - \tilde{\omega}(\Delta)(c_0 + c_1)}{d_0 + d_1}.$$

Note that $\tilde{\omega}(0) = \omega(z,z)$ and $\lim_{\Delta\to\infty} \tilde{\omega}(\Delta) = 0$, while $\lim_{\Delta\to 0} \tilde{q}(\Delta) = 0$ and $\lim_{\Delta\to\infty} \tilde{q}(\Delta) = 1$. Therefore, given the condition stated in the proposition, there exists $\Delta^* > 0$ such that the equation holds. Moreover, given $S^* = [z-\Delta^*, z+\Delta^*]$, $q(S^*;z,x_i)$ is increasing then decreasing in x_i, with a peak at $x_i = z$. Thus,

$$q(S^*;z,x_i) \geq q(S^*;z,\underline{x}) = \frac{d_1 + c_1 - \omega(z,\underline{x})(c_0+c_1)}{d_0+d_1} \geq \frac{d_1 + c_1 - \omega(z,x_i)(c_0+c_1)}{d_0+d_1}$$

if and only if $x_i \in S^*$. ■

References

Acemoglu, D., A. Ozdaglar, and A. ParandehGheibi (2010). Spread of (mis)information in social networks. *Games and Economic Behavior 70*(2), 194–227.

Allport, G.W. and L.J. Postman (1947). *The Psychology of Rumor*. New York,: H. Holt and Company.

Angeletos, G.M. and I. Werning (2006). Crises and crices: information aggregation, multiplicity, and volatility. *American Economic Review 96*(5), 1720–1736.

Banerjee, A.V. (1993). The economics of rumours. *Review of Economic Studies 60*(2), 309–327.

Bernhardt, D. and M. Shadmehr (2011). Collective action with uncertain payoffs: Coordination, public signals and punishment dilemmas. *American Political Science Review 105*(4), 829–851.

Bilefsky, D. (2009). Celebrating revolution with roots in a rumor. *The New York Times*, November 18, p. 16.

Bommel, J.V. (2003). Rumors. *Journal of Finance 58*(4), 1499–1519.

Buchner, H.T. (1965). A theory of rumor transmission. *Public Opinion Quarterly 29*(1), 54–70.

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

Edmond, C. (2011). Information manipulation, coordination, and regime change. Working Paper 17395, National Bureau of Economic Research.

Elias, N. and J.L. Scotson (1994). *The Established and The Outsiders: A Sociology Enquiry into Community Problems*, 2d ed. SAGE Publications.

Esfandiari, G. (2010). The Twitter devolution. *Foreign Policy*, June 7.

Gambetta, D. (1994). Godfather's gossip. *European Journal of Sociology 35*(2), 199–223.

Gentzkow, M. and J.M. Shapiro (2006). Media bias and reputation. *Journal of Political Economy 114*(2), 280–316.

Goldstone, J.A. (1994). *Revolutions: Theoretical, Comparative, and Historical Studies*, 2d ed. Fort Worth: Harcourt Brace.

Knapp, R.H. (1944). A psychology of rumor. *The Public Opinion Quarterly 8*(1), 22–37.

Milgrom, P.R. (1981). Good news and bad news: Representation theorems and applications. *Bell Journal of Economics 12*(2), 380–391.

Morris, S. and H.S. Shin (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review 88*(3), 587–597.

Morris, S. and H.S. Shin (2003). Global games: Theory and applications. In M. Dewatripont, L.P. Hansen, and S.J. Turnovsky, eds., *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, Vol. 1. Cambridge: Cambridge University Press.

Nkpa, N.K.U. (1977). Rumors of mass poisoning in Biafra. *Public Opinion Quarterly 41*(3), 332–346.

Peterson, W.A. and N.P. Gist (1951). Rumor and public opinion. *American Journal of Sociology 57*(2), 159–167.

Skocpol, T. (1979). *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. New York: Cambridge University Press.

Suen, W. (2010). Mutual admiration clubs. *Economic Inquiry 48*(1), 123–132.

Turner, M.E., A.R. Pratkanis, P. Probasco, and C. Leve (1992). Threat, cohesion, and group effectiveness: Testing a social identity maintenance perspective on group think. *Journal of Personality and Social Psychology 63*(5), 781–796.

Zhang, N. (2009). Rumors and mobilization in China's contentions politics. Ph.D. dissertation, Chinese University of Hong Kong.