Why Do Authoritarian Regimes Allow Citizens to Voice Opinions Publicly?*

Jidong Chen† Yiqing Xu‡

Journal of Politics, forthcoming

Abstract

Why would an authoritarian regime allow citizens to voice opinions publicly if the exchange of information among citizens spurs social instability as has been often alleged? We show that an authoritarian regime can strengthen its rule by allowing citizens to communicate with each other publicly. From the government's perspective, such communication has two interrelated functions. First, if public communication reveals a shared feeling of dissatisfaction towards the government among citizens, the government will detect the danger and improve policies accordingly. Second, and perhaps more interestingly, public communication disorganizes citizens if they find themselves split over policies. We show that the government allows public communication if and only if it perceives sufficient preference heterogeneity among citizens. The model also illustrates that public communication could serve as a commitment device ensuring government responsiveness when it faces high dissatisfaction, which in turn makes the government better off than with private polling.

Keywords: Authoritarian Governance, Public Communication, Horizontal Communication, China

*This version: September 1st, 2016. This paper was previously circulated as "On Deliberative Authoritarian Governance" and "Authoritarian Governance with Public Communication." We thank the editor and the two anonymous referees for their helpful comments and suggestions. We are indebted to Carles Boix, Scott Gehlbach, Bob Gibbons, Jens Hainmueller, Haifeng Huang, Matias Iaryczower, John Londregan, Peter Lorentzen, Nolan McCarty, Adam Meirowitz, Kris Ramsay, Tom Romer, Jim Sndyer, Leonard Wantchekon, as well as seminar participants at Florida State University, LSE, MIT, Princeton, Stanford, Tsinghua SEM, University of Illinois Urbana Champaign, Harvard-MIT-BU China Workshop, New Faces in China Studies at Duke, Political Economy Graduate Conference at NYU, Biennial Conference of China Development Studies at Shanghai Jiao Tong University, the Conference in Political Economy at Renmin University, 2014 Conference of Theoretical Research in Development Economics at Barcelona GSE, 2014 SPSA and MPSA Conferences for helpful comments. We are also grateful for the Q-APS program for its generous support.

†Jidong Chen, Assistant Professor, Business School, Beijing Normal University. Rear Main Building 1720, Beijing Normal University, 19 Xinjiekouwai St., Beijing 100875, China. Email: gdongchen@gmail.com.

‡Yiqing Xu, Assistant Professor, Department of Political Science, University of California, San Diego. Social Science Building 377, 9500 Gilman Drive #0521, La Jolla, CA 92093. Email: yiqingxu@ucsd.edu.

To silence the populace is as grim a task as preventing flood. A blocked river would eventually inundate and cause great catastrophe; the same can be said of a stifled people. It is therefore wise to dredge the river to let it run free, and to enable the people to speak its mind.

——Discourses of the States (*Guo Yu*), around 500 BC

1 Introduction

In order to maintain social stability and stay in power, an authoritarian incumbent has to find a way to please or repress citizens under its rule (Svolik 2012). Sophisticated authoritarian rulers provide citizens with economic benefits, and prevent them from coordinating with each other against the government (Bueno de Mesquita and Downs 2005). Some argue that an authoritarian regime must severely limit citizens' freedom to criticize the government (Levitsky and Way 2002), and dislikes information exchange among citizen (Hollyer, Rosendorff and Vreeland 2011, 2015*a*).

These characterizations are largely correct. However, in order to survive, authoritarian regimes also need information from the citizens. A growing literature in comparative politics has investigated how authoritarian regimes use various measures to learn from the citizens. For example, Botero, Ponce and Shleifer (2013) report that citizens in many authoritarian countries regularly complain to the authorities about poor government services and the misconduct of officials, in the light of which they argue that the complaints of citizens help sustain the quality of governance in these counties. Particularly, in China, the world's most populated authoritarian country, a non-trivial proportion of citizens in both urban and rural areas very frequently make public complaints on a set of policy issues (Lorentzen 2013; Tsai and Xu 2016). These findings echo some scholars' earlier claim that the Chinese state sets up institutions that aim to solicit information from the citizens in an orderly and peaceful manner (Oi 2003; Nathan 2003; Lorentzen 2013).

On the other hand, authoritarian regimes sometimes allow citizens to voice their opinions on public platforms, such as the social media, concerning issues that seem politically sensitive. A recent, vivid example came from Hong Kong, one of China's special administrative regions. At the end of 2014, a group of Hong Kong students initiated a protest against the Communist Party's decision on Hong Kong's electoral reform, which was joined later by older Hong Kong residents.[1] The Hong Kong protest was certainly a sensitive political issue in the eyes of the Communist Party of China (CCP). But surprisingly, after some initial hesitation, the Chinese government allowed its citizens to discuss this event almost freely online. It came to pass, however, that Chinese online commentators were deeply split on the Hong Kong protest, for a large proportion of them actually voiced support for the government's position. Why did an authoritarian government allow citizens to communicate with each other publicly on such sensitive issues?

To answer this question and to better understand an authoritarian government's incentives and strategies, we developed a game theoretic model. In the benchmark model, we assume that a citizen is either satisfied or dissatisfied with the *status quo* policy. The preferences of the citizens are potentially correlated. They do not know each other's preference and are not fully aware of how their preferences are correlated. At the same time, the government is not fully informed of the citizens' preferences. However, it does receive a private signal indicating the overall level of dissatisfaction towards the *status quo* policy. Public communication is a process through which citizens publicly express their preferences. Unlike democracies where the people's freedom of speech is protected by law, in authoritarian regimes the possibility for citizens to speak their mind freely is dependent upon the strategic choice of the government. If permitted, each citizen sends a message at no cost (i.e., cheap-talk), which becomes public information. The government then chooses a policy accordingly.

[1] For media coverage of the protest, see, for examples, The New York Times (http://goo.gl/TkUUJW) and The Washington Post (https://goo.gl/gMXP3A). Mainland Chinese's mixed feelings towards the protest were also reported by the media, e.g. The New York Times (http://goo.gl/3cjzZx), NPR (http://goo.gl/gZr7JL).

After learning the policy, the citizens will simultaneously decide whether to participate in collective action demanding their desired policy.

We argue that such communication generates not only vertical information flows from the citizens to the government, but also horizontal information flows across the citizens.[2] Horizontal communication, by making public the private information about individual preferences, can either coordinate or discourage citizens in collective action. We call these the effects the *coordination* effect and the *discouragement* effect respectively. As horizontal information flows take place, vertical information flow enable the government to respond efficiently to the fluctuating public opinion and reduce the risk of collective action by meeting the policy wishes of the citizens. By allowing the government to condition policies on the vertical information flows, public communication allows the government to preempt collective action caused by horizontal learning. Thus, the strategic response to vertical information flows mitigates the cost of horizontal learning by the citizens. This sometimes tilts the cost and benefit of citizens' communication in favor of openness. The value of vertical information flow offers a direct justification for the principles asserted in *Guo Yu* from our first the epigraph: a stifled people is like a blocked river; it becomes extremely dangerous if the ruler does not know that the people are angry after they are forbidden to speak out.

Thus we identify three driving forces that jointly determine the government's net gain from opening up public communication. The first is the policy-adjustment effect through vertical learning. The effect is positive because the government can always make good use of the citizens' private information without its constraint being tightened. The second one, which is positive, is the discouragement effect through horizontal information flows. The third one, which is negative, is the coordination effect in horizontal communication.

[2] This paper shares a feature with Farrell and Gibbons (1989), who investigate "cheap talk with two audiences." A key difference is that, in this paper, the two receivers (the government and the other citizen) take action sequentially rather than simultaneously. Our model can also be understood as a veto bargaining game with pre-bargaining communication. The proposer is able to control the information inflow while the two citizens need to coordinate on exercising the "veto power."

On the basis of the interaction of the three driving forces, we characterize the government's equilibrium strategy, and show that it allows public communication if and only if it perceives sufficient preference heterogeneity among the citizens. The government mainly makes a trade-off between the potential benefit from the discouragement effect and the potential cost from the coordination effect mitigated by the policy-adjustment effect. This result illustrates how government's incentive of allowing public communication is tied with its perception of the chances of collective action.[3] It also suggests that even when the policy-adjustment effect is small, the government may still allow public communication in order to take advantage of the discouragement effect. As a result, public communication does not necessarily imply policy changes.

We show furthermore that paradoxically, owing to the presence of the discouragement effect, the government may strictly prefer public communication to private polling, in which each citizen privately communicates her preference to the government. Specifically, when the government knows that there is a high probability that the citizens have opposite policy preferences, allowing public communication will make it better off. This is because public communication serves as a commitment device, ensuring that the government fully responds to problems that spur popular anger and thereby benefiting the government. Under public communication, because citizens' preferences are publicly revealed, they are discouraged from joining a protest when they find themselves split over the policies. With private polling, however, when a dissatisfied citizen fails to see any policy changes, she is likely to over-estimate the probability that others are as equally dissatisfied with the government policies as she is. The chances of collective action are therefore higher than under public communication.[4]

[3]This finding is consistent with recent evidence that the Chinese government allows citizens to express their opinions with much freedom, but actively censors information that may spur collective action (King, Pan and Roberts 2013).

[4]We also specify a condition under which the government strictly prefers private polling to public communication, that is, when citizens are not fully aware of their homogeneous preferences while the government has such knowledge. In this case, the government would be strictly better off by using private polling to prevent

As an extension, we also analyze how the government reacts to new informational technology, which enables citizens to privately communicate with each other without being controlled by the government. Specifically, we consider the chance of an effective private horizontal communication across citizens that does not rely on government-provided platforms. We demonstrate that even a tiny probability like this can make the government more inclined to supply public platforms allowing citizens to speak.

The main results of the paper are closely related to work that investigates the government's control of information and citizens' collective action (e.g., Casper and Tyson 2014; Dimitrov 2014; Edmond 2013; Egorov and Sonin 2011; Gehlbach and Sonin 2008; Hollyer, Rosendorff and Vreeland 2015*a*,*b*; Huang, Boranbay and Huang 2016; Landa and Tyson 2016; Little 2012, 2015, 2016*a*,*b*; Lohmann 1993; Lorentzen 2014, 2016; Shadmehr and Bernhardt 2011, 2015, 2016; Smith and Tyson 2014). For example, Magaloni (2006, 2010), Gandhi and Lust-Okar (2009), Little (2012) and Egorov and Sonin (2011) study how an authoritarian regime can use controlled elections to signal its strength and deter collective action. Little (2016*a*) further argues that the benefit of signaling strength in elections does not prevent the incumbent from learning from voters and making compromises accordingly. Lorentzen (2014) shows that a regime can use media to monitor bureaucratic agents, thus reaping the benefits of vertical information flows without being endangered by horizontal information flows.

In line with Lorentzen (2014), our model incorporates the trade-off between the benefit of vertical information flowing and the cost of horizontal information flowing that may coordinate citizens. Substantively, this paper is different from Egorov and Sonin (2011) and Lorentzen (2014) in that the value of vertical information flow in our framework is directly derived from policy adjustment rather than from bureaucratic monitoring. Perhaps more importantly, we demonstrate that even in the absence of vertical learning, an authoritarian government can still benefit from horizontal information flows because of the discourage-

dissatisfaction from spreading.

ment effect, and that public communication is a commitment device that can lend greater advantage to the government than merely conducting private polling.

Our paper is also closely related to the literature in comparative politics and positive political theory that studies the effect of information disclosure on citizens' collective action. For example, in common-value models of collective action, Shadmehr and Bernhardt (2011) and Little (2016*b*) show that allowing citizens to know more information does not necessarily increase their incentive to protest. Different from their models, in our model, the government uses dual strategies: it influences citizens' choice-making both directly—by changing the policy—and indirectly—by shaping their beliefs on the fundamental. It is therefore helpful to analyze the interaction between vertical and horizontal information flows. Moreover, contrary to common-value models, our framework takes account of private values. It is precisely preference heterogeneity among the citizens that makes the discouragement effect possible.

2 Model

In this section, we introduce the basic framework and discuss the mechanics of public communication using a simplified benchmark model. We provide proofs for a generalized model in the Appendix.

2.1 Setup

Players and policy preferences. There are three players: a government, citizen 1 and citizen 2.[5] Without loss of generality, we focus on a general policy issue. For this particular issue, the government has two policy options to choose from, $x \in \{Q, R\}$. We call Q the *status quo* policy and R the reform policy. The government strictly prefers the status quo to reform, as if it is costly to implement the reform policy. Its preference over policy x is

[5] We can also interpret them as two groups of citizens.

written as:

$$u_G(x) = \begin{cases} 0 & \text{if } x = Q \\ -\mu & \text{if } x = R \end{cases}, \tag{1}$$

where $\mu > 0$ captures the government's interest in keeping the *status quo*. The two citizens can be one of two types, an *anti-reform* type, who prefers the *status quo* to reform on this issue, or a *pro-reform* type, who strictly prefers reform to the *status quo* in this issue. Citizen i's type is denoted as $\omega_i \in \{\underline{\omega}, \overline{\omega}\}$, with $\underline{\omega}$ and $\overline{\omega}$ representing an anti-reform type and a pro-reform type, respectively. We normalize a citizen's policy gain from the status quo to zero, no matter what type she is, i.e.,

$$u_i(Q; \omega_i) = 0, \quad i = 1, 2. \tag{2}$$

A citizen with an anti-reform preference likes the status quo better than reform:

$$u_i(R; \omega_i = \underline{\omega}) = -L < 0, \quad i = 1, 2. \tag{3}$$

A pro-reform citizen gets a strictly positive payoff from the reform policy:

$$u_i(R; \omega_i = \overline{\omega}) = L, \quad i = 1, 2, \tag{4}$$

where L is common knowledge.[6] We say reform policy R is the *desired policy* of the pro-reform type, and the status quo is the *desired policy* of the anti-reform type.

Information. Because the level of citizen dissatisfaction fluctuates over time, the two citizens' types ω_1 and ω_2 are unknown to each other and to the government. However, both the government and the citizens share a common prior that a citizen prefers reform with

[6] In the generalized model of the Appendix, we consider a more general preference $u_i(R; \omega_i = \underline{\omega}) = -\underline{L} \leq 0$, where $\underline{L}$ can be different from L.

probability one-half:[7]

$$\Pr(\omega_i = \overline{\omega}) = p = \frac{1}{2}, \quad i = 1, 2. \tag{5}$$

Intuitively, p measures the conflict of interest between the citizens and the government.

The two citizens' preferences over the issue (i.e., types) are potentially correlated. If one of the citizens is pro-reform, with probability γ, then the other one is also pro-reform, i.e.,

$$\omega_j|_{\omega_i=\overline{\omega}} = \begin{cases} \overline{\omega} & \text{with probability} \quad \gamma \\ \underline{\omega} & \text{with probability} \quad 1-\gamma \end{cases}, \quad i, j \in \{1, 2\},\ i \neq j. \tag{6}$$

γ represents the preference correlation of the citizens,[8] who face uncertainty about the distribution of public opinion, and only know that γ follows a distribution on $[0, 1]$ according to a cumulative distribution function $G(\cdot)$. In other words, the citizens merely have a rough idea about γ, such as its expected value $\overline{\gamma} = E(\gamma)$. The government faces uncertainty about the realization of the citizens' preferences, yet gets a private signal about their distribution γ. For the sake of simplicity, we assume that the government directly observes the preference correlation γ.[9] γ captures preference homogeneity among the citizens. A smaller γ means that the preferences of the citizens are more heterogeneous.

Timing and actions. The timing of actions is as follows.

Period (0) *Institutional design.*The government chooses whether or not to allow public com-

[7]In the Appendix, we characterize the equilibrium for an arbitrary prior p.

[8] The distribution of $\omega_j|_{\omega_i=\underline{\omega}}$ is characterized in Equation (A2) in the Appendix. Under the assumption that $p = \frac{1}{2}$, $\Pr(\omega_j = \underline{\omega}|\omega_i = \underline{\omega}) = \gamma$.

[9] When the government does not directly observe the preference correlation γ, the results will be exactly the same, as long as it gets a private signal $\widetilde{t}$ with $E(\gamma|t)$ strictly increasing, continuously differentiable in the signal realization t, and $\lim_{t\to\underline{t}} E(\gamma|t) = \max\{0, 1 - \frac{1-p}{p}\}$, $\lim_{t\to\overline{t}} E(\gamma|t) = 1$, where $[\underline{t}, \overline{t}]$ is the support of the marginal distribution of $\widetilde{t}$. In the Supplementary Appendix, we investigate an extension when the government does not have more information than the citizens. We make a relevant discussion and comparison with the benchmark model.

munication among the citizens, $\alpha \in \{0,1\}$. When $\alpha = 0$, no citizen's voice will be heard by the government or her fellow citizen. On the contrary, when $\alpha = 1$, a citizen is allowed to send messages about her preferences over the policy issue, which will be heard by both the government and the other citizen.

Period (1) *Public communication.* If allowed (i.e., $\alpha = 1$), each citizen sends a message $m_i \in \{0,1\}$ to the government and to each other at no cost. The message is publicly observable. We interpret $m_i = 1$ as complaining (dislike the *status quo* policy) and $m_i = 0$ as abstaining. If $\alpha = 0$, this period is skipped.

Period (2) *Policy adjustment.* The government chooses an effort $e \in [0,1]$ to adjust the policy. With probability e, the reform policy R will be implemented; otherwise, the status quo Q will be kept. The citizens observe only the policy outcome x, but do not observe the government's effort e.

Period (3) *Collective action.* Each citizen simultaneously decides whether to participate in a popular protest ($a_i = 1$) or not ($a_i = 0$).

Collective action and payoffs. Each player's payoff consists of two parts: a payoff at the policy-adjustment stage, and a payoff at the collective-action stage. We denote them as the "policy payoff" and the "collective-action payoff," respectively, although the latter can also be seen as policy-driven. For the sake of simplicity, we assign equal weight to the two payoffs.[10]

If both citizens participate, the collective action will succeed for sure. In this case, a new policy $y \neq x$ gets implemented and the government suffers $\rho_2 > 0$.[11] If only one of the citizens participates, the government suffers a cost $\rho_1 > 0$. Throughout the paper we

[10] The results are qualitatively the same if we assign different weights.

[11] Depending on the magnitude of ρ_2, we can interpret collective action in different ways. It can be a small-scale protest demanding that the government change a particular policy or punish a misbehaved local official, as often happens in China. It can also be a social movement aiming at a regime change, after which the citizens or the new government implement the reform policy.

assume $\rho_2 > \mu > \rho_1$. Moreover, with probability $\lambda \in (0, \min\{\frac{1}{2}, \frac{1}{L}\})$, the individual protest is successful and the new policy $y \neq x$ is implemented; with probability $(1-\lambda)$, it is not successful and the original policy x remains unchanged. If neither citizen participates, no policy change happens and the government suffers no cost. Citizen i's collective-action payoff is represented by:

	participate (j)	abstain (j)
participate (i)	$u_i(y;\omega_i) - k_i$	$\lambda u_i(y;\omega_i) + (1-\lambda)u_i(x;\omega_i) - k_i$
abstain (i)	$\lambda u_i(y;\omega_i) + (1-\lambda)u_i(x;\omega_i)$	$u_i(x;\omega_i)$

where k_i is citizen i's cost of participating in collective action. k_i is i's private information and is only known to her after she observes the government's policy x.[12] k_i is assumed to be independent and identically distributed between 0 and 1 with a cumulative distribution function $F(\cdot)$. We assume $F(\cdot)$ is weakly concave; and $f(k) = F'(k) > 0, \forall k \in [0,1]$.[13] The distribution of the cost of participating in collective action captures the repression technology of the government.

The government's total payoff also consists of two parts: the utility from its policy preference $u_G(x)$ and the cost of collective action, which we summarize as follows:

	participate (j)	abstain (j)
participate (i)	$-\rho_2$	$-\rho_1$
abstain (i)	$-\rho_1$	0

[12] We can show that, even if the citizens know their private costs of collective action in the communication stage, the outcome induced in any symmetric cut-point equilibrium will be the same as when they do not know the costs. The intuition is that a citizen with a high collective-action cost always has an incentive to pretend to be of low cost so as to persuade the other citizen to join the protest. As a result, in any symmetric cut-point equilibrium, cheap-talk produces no information on the collective-action cost that would change the equilibrium outcome.

[13] We can verify that the uniform distribution and any distribution with a cumulative distribution function $F(k) = k^\delta$ $(0 < \delta < 1)$ satisfy this property. The concavity of the distribution is used merely to guarantee the unique prediction in the collective-action stage. Without this assumption, we may need to deal with the problem of multiple equilibria, although the properties in the equilibrium we focus on are still valid.

Denote $A \equiv (1-\lambda)L$, which is the payoff gain of joining a protest (excluding the protest cost) provided that the other citizen also participates. Similarly, $B \equiv \lambda L$, is the payoff gain when the other citizen does not participate. Hence, we have $0 < B < \min\{1, A\}$.[14]

The equilibrium notion is *Perfect Bayesian Equilibrium.* Because multiple equilibria may exist as in other cheap-talk/signaling games, we focus on equilibria in which citizens truthfully reveal their preferences (types) when allowed. Such equilibria satisfy the following requirements:

(i) $\alpha^*(\gamma)$, the probability of the government allowing public communication, maximizes its expected payoff given the citizens' equilibrium strategies, beliefs, and the government equilibrium policy-adjustment strategies;

(ii) when public communication is (not) allowed, $\alpha = 1$ ($\alpha = 0$), $e_1^*(\gamma, m_1, m_2)$ ($e_0^*(\gamma)$), the probability of the government making an effort to adjust the policy, maximizes its expected payoff given the citizens' equilibrium strategies and beliefs;

(iii) when public communication is not allowed ($\alpha = 0$) and the citizen observes the final policy x, $\widehat{\gamma}_x^*$, the citizens' belief about γ for x that takes place with positive probability in the equilibrium, is formed based on Bayes' rule, given the government's equilibrium strategies;[15]

(iv) each citizen in the collective-action stage maximizes her own welfare, given the equilibrium belief about γ and the other one's equilibrium strategy.

We first pin down the equilibrium features at the collective-action stage based on the conjecture that such a citizen-truth-telling equilibrium exists. Then, we use those properties to check the incentive compatibility (IC) constraints of the citizens at the communication stage. In the end, we investigate the government's optimal choice of whether to allow public

[14]The condition $A > B$ captures an important feature of authoritarian regimes: complementarity in collective action, that is, when knowing that others are less likely to join the collective action, one's incentive to join it also decreases.

[15]When public communication is allowed, the citizens will directly observe each other's preference because we focus on the citizen-truth-telling equilibrium.

communication.

2.2 Equilibrium Characterization

The collective-action stage. The following lemma, as well as its more technical version (Lemma 4 in the Appendix), shows that in any citizen-truth-telling equilibrium, the citizens' strategy at the collective-action stage is uniquely determined.[16] Specifically, if her desired policy is implemented, a citizen never protests; otherwise she protests if and only if her realized cost of joining the protest is relatively small.

Since neither the government nor the other citizen observes a citizen's cost of collective action, both can only gauge her probability of participation based on the prior of her cost of collective action and the equilibrium cut-point, which is characterized by Lemma 4 in the Appendix. Equation A18 in the Appendix defines $p_0(\widehat{\gamma})$ as the endogenous probability that a discontent citizen protests, where $\widehat{\gamma}$ is the probability with which she believes that the other citizen is of the same type.

Lemma 1 (Characterizing the collective-action stage) *In any equilibrium in which both citizens truthfully reveal their types when they are allowed to speak, each citizen in the collective-action stage uses a cut-point strategy characterized in Lemma 4 in the Appendix. Specifically, (1) a citizen never protests if her desired policy is implemented; (2) otherwise, the probability that she protests equals* $p_0(\widehat{\gamma})$*, where* $\widehat{\gamma}$ *is her perception whether the other citizen shares the same preference; and (3)* $p_0(\widehat{\gamma})$ *is an increasing function characterized by Equation A18 in the Appendix.*

[16] Without the standard *common-value global games* setup, we still get the uniqueness feature at the collective-action stage. As implied by Morris and Shin (2006), whether the prediction for collective action in both common value and private value games is unique depends on how we technically parameterize the payoffs and the uncertainty, both of which capture assumptions of common knowledge. The technical approach we use, i.e., incorporating private costs, is similar to Palfrey and Rosenthal (1985), that Morris and Shin (2006) call a *private-value interaction/global game*.

With public communication, when the other citizen also claims to be dissatisfied (pro-reform), a dissatisfied citizen joins a protest with the highest possible probability $p_0(1)$ since she knows that the other one has similar incentive, and the collective action is likely to be successful. If a dissatisfied citizen infers that the other one is satisfied, she then chooses to join a protest with probability $p_0(0)$. However, if public communication is not allowed, a dissatisfied citizen can only condition her behavior on her perception of preference homogeneity $\widehat{\gamma}$ (which may depend on the observed policy x). The probability of her joining a protest is therefore $p_0(\widehat{\gamma})$. Lemma 1 helps us understand the role of horizontal communication in facilitating or impeding collective action. Compared with the case of no communication, with public communication, a dissatisfied citizen increases her probability of joining a protest from $p_0(\widehat{\gamma})$ to $p_0(1)$ when she finds the other one is also dissatisfied and decreases her probability of protest from $p_0(\widehat{\gamma})$ to $p_0(0)$ when she finds that the other citizen is satisfied. We illustrate the two scenarios in Figure 1.

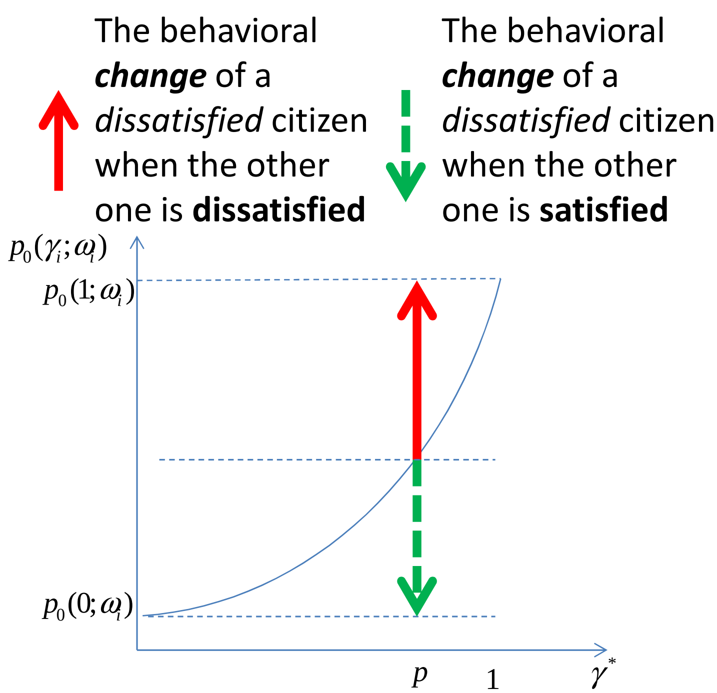

Figure 1. Horizontal communication and participation in collective action

The government's expected payoff. To formally understand the government's incentive of allowing public communication, we write down its expected payoff as follows. $G_1(e, \omega_1, \omega_2)$ is the government's expected payoff in every possible situation (ω_1, ω_2) when it allows public communication and chooses an effort $e \in [0, 1]$; $G_0(e, \omega_1, \omega_2)$ is its expected payoff in every possible situation (ω_1, ω_2) when it forbids communication and chooses an effort $e \in [0, 1]$.

For example, if both citizens are pro-reform, i.e., $\omega_1 = \omega_2 = \overline{\omega}$, with public communication they know each other's type and each independently protests against the status quo policy with probability $p_0(1)$. The government's payoff therefore is:

$$G_1(e, \omega_1, \omega_2)|_{\omega_1=\omega_2=\overline{\omega}} = -(1-e)W(p_0(1)) - e\mu, \tag{7}$$

where

$$W(x) = \rho_2 x^2 + 2\rho_1 x(1-x). \tag{8}$$

The expected payoffs under the other situations can be similarly pinned down.[17] We summarize the government's payoffs under the two circumstances ($\alpha = 0$ and $\alpha = 1$) and their differences in Table 1.[18] In the table, $\widehat{\gamma}_1^*$ ($\widehat{\gamma}_0^*$) is denoted as the dissatisfied citizens' perceived preference homogeneity when she does (not) observe the reform when public communication is not allowed. According to the last row of the table, without considering the learning effect of its own, the government benefits from public communication when the two citizens are of different types, and loses from public communication when they share the same interest. In other words, public communication may be beneficial to the government even if it does not adjust policy according to what it learns from the citizens.

The government's incentive to allow public communication. In this section we formally define the three effects that shape the government's incentive. Notice that the com-

[17] See the Appendix for detailed derivations.

[18] For the moment, we use the same notation e for policy adjustment when the public communication is allowed or not allowed, implicitly assuming that the government's efforts under the two circumstances are the same.

Table 1. Net Gain from Opening Public Communication Holding Government's Effort Constant

	$\omega_1 = \omega_2 = \underline{\omega}$	$\omega_1 = \omega_2 = \overline{\omega}$	$\omega_1 \neq \omega_2$
$\alpha = 1$	$-e[\mu + W(p_0(1))]$	$-(1-e)W(p_0(1)) - e\mu$	$-(1-e)\rho_1 p_0(0)$ $-e[\mu + \rho_1 p_0(0)]$
$\alpha = 0$	$-e[\mu + W(p_0(\widehat{\gamma}_1^*))]$	$-(1-e)W(p_0(\widehat{\gamma}_0^*)) - e\mu$	$-(1-e)\rho_1 p_0(\widehat{\gamma}_0^*)$ $-e[\mu + \rho_1 p_0(\widehat{\gamma}_1^*)]$
$G_1 - G_0$	$-eZ(\widehat{\gamma}_1^*) \leq 0$	$-[1-e]Z(\widehat{\gamma}_0^*) \leq 0$	$(1-e)\rho_1[p_0(\widehat{\gamma}_0^*) - p_0(0)]$ $e\rho_1[p_0(\widehat{\gamma}_1^*) - p_0(0)] > 0$

Note: $Z(x) \triangleq [W(p_0(1)) - W(p_0(x))]$. Equivalent expressions of the government payoffs are Equations A25 and A26 in the Appendix.

munication generates a combination of horizontal information flows and vertical information flows. Horizontal communication among the citizens has two possible effects on the government's net gain from allowing public communication: the *coordination* effect, when the citizens find out that they share the same preference, i.e., $\omega_1 = \omega_2$, and the *discouragement* effect, when the citizens realize that they are of different types, i.e., $\omega_1 \neq \omega_2$.

Suppose $e_0^*(\gamma)$ is the government's equilibrium policy choice when public communication is not allowed, i.e., $e_0^*(\gamma) \in \arg\max_e E[G_0(e, \omega_1, \omega_2)|\gamma]$. The coordination effect is formally defined as the government's net gain from citizens' horizontal learning when they are of the same type, assuming the government sticks to $e_0^*(\gamma)$, i.e.,

$$\sum_{w_1 = w_2} \Pr(\omega_1, \omega_2|\gamma)[G_1(e_0^*, \omega_1, \omega_2) - G_0(e_0^*, \omega_1, \omega_2)], \tag{9}$$

which equals $-\frac{1}{2}\gamma\{e_0^*(\gamma)[W(p_0(1)) - W(p_0(\widehat{\gamma}_1^*))] + (1 - e_0^*(\gamma))[W(p_0(1)) - W(p_0(\widehat{\gamma}_0^*))]\}$ after simplification and is always non-positive.

Similarly, the discouragement effect is formally defined as a net gain from horizontal learning when citizens are different types, assuming the government sticks to $e_0^*(\gamma)$, i.e.,

$$\sum_{w_1 \neq w_2} \Pr(\omega_1, \omega_2|\gamma)[G_1(e_0^*, \omega_1, \omega_2) - G_0(e_0^*, \omega_1, \omega_2)], \tag{10}$$

which equals $(1-\gamma)\rho_1[(1 - e_0^*(\gamma))(p_0(\widehat{\gamma}_0^*) - p_0(0)) + e_0^*(\gamma)(p_0(\widehat{\gamma}_1^*) - p_0(0))]$ after simplification and is always non-negative.

The government also gains from vertical information flows since it learns the citizens' preferences and adjusts policy when it finds that the citizens pose a *real* threat to its rule. The *policy-adjustment* effect from vertical communication is formally defined as the government's net gain from learning the citizens' preferences through public communication, assuming that the citizens already know each other's preference, i.e.,

$$E[\max_{e} G_1(e, \omega_1, w_2) - G_1(e_0^*(\gamma), \omega_1, \omega_2)|\gamma]. \tag{11}$$

It is always non-negative because there is no loss to obtain additional information that does not change the government's constraints.

Hence we get the following **Hierarchical Communication Identity**, that shows that the government's payoff difference between allowing and forbidding public communication can be decomposed into the above three driving forces, i.e.,

$$\begin{aligned} & E[\max_{e} G_1(e, \omega_1, \omega_2) - G_0(e_0^*(\gamma), \omega_1, \omega_2)|\gamma] \\ = \ & \underbrace{E[\max_{e} G_1(e, \omega_1, w_2) - G_1(e_0^*(\gamma), \omega_1, \omega_2)|\gamma]}_{\text{the policy-adjustment effect from vertical communication}} \\ + & \underbrace{\sum_{w_1 = w_2} \Pr(\omega_1, \omega_2)[G_1(e_0^*(\gamma), \omega_1, \omega_2) - G_0(e_0^*(\gamma), \omega_1, \omega_2)|\gamma]}_{\text{the coordination effect from horizontal communication}} \\ + & \underbrace{\sum_{w_1 \neq w_2} \Pr(\omega_1, \omega_2)[G_1(e_0^*(\gamma), \omega_1, \omega_2) - G_0(e_0^*(\gamma), \omega_1, \omega_2)|\gamma]}_{\text{the discouragement effect from horizontal communication}} \end{aligned} \tag{12}$$

We summarize the above results in the following lemma.

Lemma 2 ***(The Hierarchical Communication Identity)*** *The government's net gain from opening public communication can be decomposed into three driving forces: (1) the policy-adjustment effect (a direct informational gain from vertical communication, provided that the citizens already know each other's type) that is non-negative; (2) the coordination effect (a net gain from horizontal communication when the citizens are of the same type,* *provided that the government sticks to the effort without communication) that is non-positive; and (3) the discouragement effect (a net gain from horizontal communication when citizens are of different types, provided that the government sticks to the effort without communication) that is non-negative.*

The above communication identity illustrates that the government's incentive for allowing public communication depends on the three effects: coordination effect, discouragement effect and the policy-adjustment effect. When making the decision whether to allow public communication platforms, the government mainly makes a trade-off between the potential benefit from the discouragement effect and the potential cost from the coordination effect, which is however mitigated by the policy-adjustment effect. Before we fully characterize the equilibrium, the following observation is helpful for simplifying the analysis.

Lemma 3 (No effort without public communication) *In any equilibrium, the government chooses not to make any effort to adjust the policy when public communication is not allowed, i.e.,* $e_0^*(\gamma) \equiv 0$.

(See the proof of the generalized version Lemma 5 in the Appendix.)

To understand the intuition behind this lemma, suppose that the government chooses to reform when public communication is not allowed. In this case, it exerts effort μ to change the policy. If it allows public communication, however, it suffers μ when both citizens are pro-reform, but as long as one of the citizens is anti-reform, the government suffers less than μ. Therefore, allowing public communication offers a profitable deviation for the government. Conversely, the government will not to make any effort to adjust policy when it shuts down public communication; otherwise it would simply allow public communication to reap the potential benefits.

Since the government does not make any effort without public communication, at the policy-adjustment stage, no additional private information of the government will be revealed. Thus, when public communication is not allowed, the government gets $-M(\gamma, p_0(\widehat{\gamma}_0^*))$, where $\widehat{\gamma}_0^* = E(\gamma|\alpha = 0; \alpha^*(\cdot))$ is the citizens' equilibrium belief about preference homogeneity when public communication is not allowed, and

$$M(\gamma, x) = p\gamma x^2 \rho_2 + 2px(1 - \gamma x)\rho_1 \tag{13}$$

is the government's expected loss from pro-reform citizens' collective action when each of them protests with probability x. When public communication is allowed, the government gets $\max\{-W(p_0(1)), -\mu\}$ if it sees two pro-reform citizens, where $W(p_0(1))$ is the government's cost from citizens' collective action if it sticks to the status quo. The maximum function represents the government's choice whether to make an effort to reform. It makes no effort and gets $-p_0(0)\rho_1$ if it sees only one pro-reform citizen as making an effort brings a higher cost $\mu + p_0(0)\rho_1$. The government makes no effort and gets 0 if the two citizens are anti-reform. The difference in the government's payoffs between allowing and not allowing public communication therefore is:

$$G_{diff}(\gamma, \widehat{\gamma}_0^*) = M(\gamma, p_0(\widehat{\gamma}_0^*)) - \frac{1}{2}\gamma \min\{W(p_0(1)), \mu\} - (1 - \gamma)p_0(0)\rho_1, \tag{14}$$

which is a function of its private information about preference homogeneity γ. The following existence result characterizes the government's equilibrium decision regarding opening public communication.

Proposition 1 (Equilibrium characterization) *(1) in any equilibrium, the government allows public communication if and only if its private signal indicates that the citizens' preferences are sufficiently heterogeneous, i.e.,*

$$\alpha^* = \begin{cases} 1 & \text{if } \gamma < \gamma^* \\ 0 & \text{if } \gamma \geq \gamma^* \end{cases}, \tag{15}$$

where $\gamma^* \in (0, 1]$*;*

(2) when the public communication is shut down, citizens will think that they are more homogeneous than in the case when public communication is allowed, i.e., $E(\gamma|\alpha = 0) > E(\gamma|\alpha = 1)$*; and*

(3) if $\min_x G_{diff}(x, E(\gamma|\gamma \geq x)) < 0$*, there exists an equilibrium with interior cut-point equilibrium* $\gamma^* \in (0, 1)$.[19]

(See the proof of the generalized version Proposition 4 in the Appendix.)

The first part of Proposition 1 suggests that the government's private signal of the preference homogeneity γ serves as a measurement for the likelihood of collective action. The government allows public communication if and only if the likelihood of collective action is small. To understand the intuition, consider a special case when $W(p_0(1)) \leq \mu$, because reforming is too costly, the government does not benefit from improving the policy even it knows both citizens are pro-reform. Thus, the government makes the trade-off only between the coordination effect and the discouragement effect. When the γ is low, and the government knows that discouragement is more likely to take place, it will choose to allow public communication. Conversely, when γ is high, the government will choose to shut it down. This special case also implies that the appearance of robust public discussion on policies in the public domain does not necessarily mean that the government will listen to the citizens.

The second part of the proposition suggests that whenever public discussion over an issue is shut down, a dissatisfied citizen will tend to believe that more people demand a policy change than when public discussion is allowed. This is consistent with the insight that an authoritarian incumbent can use publicly observable action to signal its strength (Little 2016*a*; Egorov and Sonin 2011; Huang 2015).

As a corollary, the following statement performs a measure of the government's decision on openness when its cost of policy adjustment varies.

[19] An equilibrium with $\gamma^* = 1$ always exists. In this equilibrium, whenever the government shuts down public communication, the citizens believe that their preferences are perfectly correlated.

Corollary 1 (Effect of Policy-Adjustment Cost on Openness) *Provided that the cost the government bears when only one citizen protests ρ_1 is sufficiently small,*

(1) if the cost of policy adjustment is not tiny, i.e., so long as $\mu > W(p_0(E(\gamma), \overline{\omega}))$, there is an equilibrium $\gamma^* \in (0, 1)$ with partial openness; and

(2) if the cost of policy adjustment is sufficiently small, i.e., so long as $\mu < W(p_0(0, \overline{\omega}))$, in the unique equilibrium $\gamma^* = 1$, that is, the government always allows public communication. (See the Appendix for the proof.)

Consistent with our intuition, the above corollary suggests that a higher willingness (or less cost) to adjust the policy of the government makes it more willing to allow citizens to discuss the issue in public.

As multiple equilibria may exist, the following proposition helps to select the unique equilibrium that maximizes the government's welfare, which also minimizes openness, i.e., $\Pr(\gamma : \alpha^*(\gamma) = 1)$.

Proposition 2 (Equilibrium selection) *There exists a unique equilibrium γ^{**} that offers the government the highest possible payoff among all equilibria. $\gamma^{**} = \inf_{\gamma^* \text{ is an equilibrium}} \gamma^* > 0$ also has the minimum level of openness (i.e., the smallest probability of allowing public communication) among all equilibria.*

(See the Appendix for the proof.)

Proposition 2 suggests that the government will keep the level of openness at the minimum level in order to maximize its payoff. To understand the intuition, consider two possible equilibria γ_a^*, γ_b^* with $0 < \gamma_a^* < \gamma_b^* < 1$. When preference homogeneity is relatively low, i.e., $\gamma < \gamma_a^*$, the government allows public communication and therefore gets the same expected payoff under the two equilibria. When preference homogeneity is moderate, i.e., $\gamma_a^* < \gamma < \gamma_b^*$, in equilibrium γ_a^* the government's payoff under no communication is higher than that under public communication, which is the same as the government's payoff in the equilibrium γ_b^*. When preference homogeneity is relatively high, i.e., $\gamma > \gamma_b^*$, neither equilibria involves public communication. However, in equilibrium γ_a^*, perceived preference homogeneity, $E(\gamma|\gamma > \gamma_a^*)$ is lower than that under equilibrium γ_b^*, i.e., $E(\gamma|\gamma > \gamma_b^*)$. As a result, citizens have less incentives to join a collective action and the government suffers less cost from citizens' collective action in equilibrium γ_a^*. Hence, the degree of openness negatively correlates with the government's payoff among all possible equilibria.

Note that the fact that the government can maximize its payoff by selecting the equilibrium that has the minimum level of openness does not rule out the possibility that it actually allows public communication in certain equilibria. We want to emphasize that the former result is about equilibrium selection while the decision of opening public communication channels $\alpha^*(\gamma)$ is a strategic choice made by the government in each equilibrium. In the following section we assume that an interior equilibrium $\gamma^* < 1$ always exists. We will also focus on the particular equilibrium with the minimum level of openness γ^{**}.

3 Private Polling

To prevent citizens from coordinating with each other in the public space, the government can shut down public communication channels and privately elicit information from individuals (say, by conducting private polls), and then they can decide how to respond to citizens' policy demands. In a private polling setup, citizens' messages are directly observed by the government, but not by each other. Is private polling a better option than public communication for an authoritarian government? In this section, we will compare the government's welfare in this private polling game and in the benchmark model where it chooses between public communication and no communication.

When the government conducts private polls, the citizens only learn each other's preference based on the observed government policies. When the government faces one or no pro-reform citizen, there is no need to make an effort to reform because in the absence of a policy change, the government's cost is at most ρ_1, which is smaller than the policy-adjustment cost μ. We denote ε^* as the probability of the government implementing the reform policy when it faces two pro-reform citizens. Upon two pro-reform citizens, the government's choice ε^* balances the tradeoff between reform and no reform while bearing the cost of collective action given the two citizens' belief about preference homogeneity $q(\varepsilon^*)$. $q(\varepsilon^*)$ is calculated by Bayes' rule and depends on both ε^* and $\overline{\gamma}$, which is perceived preference homogeneity.

$$q(\varepsilon^*) \equiv \Pr(w_j = \overline{\omega} | w_i = \overline{\omega}, x = Q) = \frac{\overline{\gamma}(1-\varepsilon^*)}{\overline{\gamma}(1-\varepsilon^*) + (1-\overline{\gamma})}. \tag{16}$$

Its first derivative $q'(\varepsilon^*) = -\overline{\gamma}(1-\overline{\gamma})(1-\overline{\gamma}\varepsilon^*)^{-2} < 0$, hence, $q(\varepsilon^*)$ is strictly decreasing in ε^*. By Lemma 1, a pro-reform citizen protests against the *status quo* policy with probability $p_0(q)$, which is also decreasing in ε^*. To understand this, when a pro-reform citizen believes that the government is making a higher effort, she is more likely to believe that the lack of reform is a result of the lack of support for the reform policy from her fellow citizen. We leave the equilibrium characterization in the private polling game and its comparison with the equilibrium under public communication to the Appendix, and summarize the result as follows.

Proposition 3 (Public communication vs. private polling) *Provided that it is not sufficiently costless for the government to adjust the policy, i.e., so long as $W(p_0(0, \overline{\omega})) < \mu$,*

(1) when the government's private signal indicates that the preferences of the citizens are relatively heterogeneous (i.e., there exists a $\widehat{\gamma} > 0$, such that, for any $\gamma < \widehat{\gamma}$), it strictly prefers public communication to private polling; and

(2) when the government's private signal indicates that the preferences of the citizens are relatively homogeneous ($\gamma \geq \gamma^{**}$) and knows that the citizens believe that their preferences are relatively heterogeneous ($W(p_0(\overline{\gamma})) \leq \mu$), it strictly prefers private polling to public communication — and in the benchmark game, it prefers no public communication.

(See the Appendix for the proof.)

The intuition is straightforward. Under public communication, because citizens directly observe each other's preference, there is a chance for a pro-reform citizen to realize that the lack of reform is a result of the lack of support for reform from her fellow citizen. Under private polling, they cannot directly observe each other's preference. When the government understands they are of different opinions, it is better for the government to let the citizens discover that fact through public communication. Conversely, when citizens do not understand that they are of the same opinion, the government does not want them to find it out when communicating with each other, hence, private polling is preferred.

4 Extension: Private Horizontal Communication

In reality, citizens may have private channels of communicating with each other that is not controlled by the regime. For examples, people can directly talk to each other without using public communication platforms; they can also use modern communication technologies that circumvent state regulations. In this section, we extend the model to capture the possibility of private horizontal communication among citizens and its impact on the government's strategic decision..

We assume that when public communication is not allowed, with probability h, citizens can directly learn each other's preference through private communication channels; with probability $1-h$, the communication is not successful and, as a result, they do not know each other's preference. h measures the effectiveness of citizens' horizontal communication outside government-regulated channels. Proposition 6 in the Appendix demonstrates that such a possibility makes the government more inclined to open platforms for public communication. In other words, if citizens are able to privately communicate their preferences with each other, the government would rather make the conversation public. This proposition complements one of Proposition 2's implications, namely that the government will utilize resources and powers at its disposal to minimize the level of openness for its own benefit. In reality, however, certain societal and technological factors are not directly controlled by the government. Proposition 6, therefore, emphasizes how these factors may affect its strategic choice.

In the Supplementary Appendix, we offer a variety of additional extensions to the benchmark model and illustrate how our main results are robust to alternative assumptions. Specifically, we consider three cases: (1) when the government can use a more sophisticated way of information management; (2) when the citizens are *ex ante* asymmetric; (3) when the citizens cannot Bayesian-update information. In addition, we also analyze the case when the government does not have more information than the citizens.

5 Conclusion

By introducing both vertical and horizontal communication into a collective-veto bargaining structure, this paper develops a tractable model to explain why an authoritarian state may use public communication to strengthen its rule. Public communication in authoritarian regimes allow citizens to publicly express their opinions and communicate policy preferences. We show that the government can gain from vertical communication as it learns citizens' policy preferences and adjusts policy accordingly. Meanwhile, the government can either gain or lose from horizontal information flows, depending on whether it impedes or encourages citizens' collective action by informing them of one another's preferences.

In general, a sophisticated authoritarian government opens public communication when the coordination effect from horizontal communication is dominated by the policy-adjustment effect from vertical communication and the discouragement effect from horizontal communication. Our model suggests that the government is more likely to open public communication when it perceives a low likelihood of collective action, or when the citizens are more likely to have private communication or other forms of horizontal interactions.

Since public communication helps reshape citizens' beliefs regarding one another's policy preference, we show that the government prefers public communication to private polling when its private signal indicates that citizen preferences are sufficiently heterogeneous. This is because with private polls, the government loses the chance to convince a pro-reform citizen that her fellow citizens are anti-reform and "blame" them for the lack of reform. When the preferences of the citizens are homogeneous and unknown to the citizens, the government prefers private polling to public communication so as to prevent collect action from taking place.

Nathan (2003) observes that the availability of new information and communication technology is not likely to lead to a regime change in China because routinized protests cannot send strong enough signals to trigger a large mass movement that is needed for a fundamental change. This paper offers an alternative explanation for the lack of a regime change. Citizens' communication through public platforms not only enables the government to learn from the citizens effectively and change policies that spur popular anger, but also informs the citizens about the perhaps conflicting interests among themselves and thereby discourages them from coordinating with each other.

References

Botero, Juan, Alejandro Ponce and Andrei Shleifer. 2013. "Education, Complaints, and Accountability." *Journal of Law and Economics* 56(4):959–996.

Bueno de Mesquita, Bruce and George W. Downs. 2005. "Development and Democracy." *Foreign Affairs* September–October(84):77–86.

Casper, Brett Allen and Scott A Tyson. 2014. "Popular Protest and Elite Coordination in a Coup détat." *The Journal of Politics* 76(02):548–564.

Dimitrov, Martin. 2014. "Dictatorship and Information: Autocratic Resilience in Communist Europe and China." Presented at the Annual Meeting of the Southern Political Science Association, New Orleans.

Edmond, Chris. 2013. "Information Manipulation, Coordination, and Regime Change." *The Review of Economic Studies* 80(4):1422–1458.

Egorov, Georgy and Konstantin Sonin. 2011. "Incumbency Advantage in Non-Democratic Elections." Presented at the Annual Meeting of the Midwest Political Science Association, Chicago.

Farrell, Joseph and Robert Gibbons. 1989. "Cheap Talk with Two Audiences." *The American Economic Review* 79(5):1214–1223.

Gandhi, Jennifer. 2008. *Political Institutions under Dictatorship.* Cambridge University Press Cambridge.

Gandhi, Jennifer and Ellen Lust-Okar. 2009. "Elections under authoritarianism." *Annual Review of Political Science* 12:403–422.

Gehlbach, Scott and Konstantin Sonin. 2008. "Government Control of the Media." Presented at the Annual Meeting of the American Political Science Association, Boston.

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. 2011. "Democracy and Transparency." *The Journal of Politics* 73(04):1191–1205.

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. 2015*a*. "Transparency, Protest and Political (In)Stability." *American Political Science Review* (forthcoming).

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. 2015*b*. "Why Do Autocrats Disclose?" Presented at the Annual Meeting of the Midwest Political Science Association, Chicago.

Huang, Haifeng. 2015. "Propaganda as Signaling." *Comparative Politics* 47(4):419–444.

Huang, Haifeng, Serra Boranbay and Ling Huang. 2016. "Media, Protest Diffusion, and Authoritarian Resilience." *Political Science Research and Methods* (forthcoming).

Kalathil, Shanthi and Taylor C. Boas. 2003. *Open Networks, Closed Regimes: The Impact of the Internet on Authoritarian Rule*. Washington, D.C.: The Brookings Institution Press.

King, Gary, Jennifer Pan and Margaret Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *The American Political Science Review* 107(2):1–18.

Landa, Dimitri and Scott Tyson. 2016. "Enforcement and Accomplishment: A Theory of Leadership in Organizations." Mimeo.

Levitsky, Steven and Lucan Way. 2002. "The Rise of Competitive Authoritarianism." *Journal of Democracy* 13(2):51–65.

Little, Andrew. 2012. "Elections, Fraud, and Election Monitoring in the Shadow of Revolution." *Quarterly Journal of Political Science* 7(3):249–283.

Little, Andrew. 2015. "Propaganda and Credulity." Mimeo, Cornell University.

Little, Andrew. 2016*a*. "Are Non-competitive Elections Good for Citizens?" *Journal of Theoretical Politics* (forthcoming).

Little, Andrew. 2016*b*. "Communication Technology and Protest." *The Journal of Politics* 78(1):152–166.

Lohmann, Susanne. 1993. "A Signaling Model of Informative and Manipulative Political Action." *The American Political Science Review* 87(2):319–333.

Lorentzen, Peter. 2013. "Regularizing Rioting: Permitting Public Protest in an Authoritarian Regime." *Quarterly Journal of Political Science* 8(2):127–158.

Lorentzen, Peter. 2014. "China's Strategic Censorship." *American Journal of Political Science* 58(2):402–414.

Lorentzen, Peter. 2016. "Chinas Controlled Burn: Information management and state-society relations under authoritarianism." Book Manuscript.

Magaloni, Beatriz. 2006. *Voting for Autocracy: Hegemonic Party Survival and Its Demise in Mexico.* New York: Cambridge University Press.

Magaloni, Beatriz. 2010. "The Game of Electoral Fraud and the Ousting of Authoritarian Rule." *American Journal of Political Science* 54(3):751–765.

Morris, Stephen and Hyun Song Shin. 2006. Heterogeneity and Uniqueness in Interaction Games. In *The Economy as an Evolving Complex System III: Current Perspectives and Future Directions*, ed. Lawrence E. Blume and Steven N. Durlauf. New York: Oxford University Press. 207–242.

Nathan, Andrew. 2003. "Authoritarian Resilience." *Journal of Democracy* 14(1):6–17.

Oi, Jean. 2003. "Bending without Breaking, The Adaptability of Chinese Political Institutions". In *How Far Across the River? Chinese Policy Reform at the Millennium*, ed.

Nicholas C. Hope, Dennis Tao Yang and Mu Yang Li. Palo Alto: Stanford University Press. 450–468.

Palfrey, Thomas and Howard Rosenthal. 1985. "Voter Participation and Strategic Uncertainty." *The American Political Science Review* 79(1):62–78.

Shadmehr, Mehdi and Dan Bernhardt. 2011. "Collective Action with Uncertain Payoffs: Coordination, Public Signals and Punishment Dilemmas." *American Political Science Review* 105(4):829–51.

Shadmehr, Mehdi and Dan Bernhardt. 2015. "State Censorship." *American Economic Journal: Microeconomics* 7(2):280–307.

Shadmehr, Mehdi and Dan Bernhardt. 2016. "Can Citizen Communication Hinder Successful Revolution?" Mimeo.

Smith, Alastair and Scott Tyson. 2014. "Inside the Palace: Collection Action between Elites." Prepared for the Annual Meeting of the Midwest Political Science Association, Chicago.

Stockmann, Daniela and Mary E Gallagher. 2011. "Remote control: How the media sustain authoritarian rule in China." *Comparative Political Studies* 44(4):436–467.

Svolik, Milan. 2012. *The Politics of Authoritarian Rule*. Cambridge: Cambridge University Press.

Truex, Rory. 2013. "Representation Within Bounds." Presented at the Annual Meeting of the American Political Science Association, Chicago.

Tsai, Lily L. and Yiqing Xu. 2016. "Outspoken Insiders: Political Connections and Citizen Participation in Authoritarian China." Mimeo, Massachusetts Institute of Technology.

Appendix for *Why Do Authoritarian Regimes Allow Citizens to Voice Opinions Publicly?*

Jidong Chen (Beijing Normal) Yiqing Xu (UCSD)

This appendix supplements the paper *Why Do Authoritarian Regimes Allow Citizens to Voice Opinions Publicly?*. Specifically, we prove most of the results with a more generalized model, of which the benchmark model in the main text is a special case. First, we present the model.

The model

The two citizens' *policy preferences* in the policy-adjustment stage are $u_i(x, \omega_i, \omega_j), i = 1, 2$.

$Pr(\omega_i = \overline{\omega}) = p, i = 1, 2$

$$\omega_j|_{\omega_i=\overline{\omega}} = \begin{cases} \overline{\omega} & \text{with probability} \quad \gamma \\ \underline{\omega} & \text{with probability} \quad 1-\gamma \end{cases}, \quad i \neq j, \tag{A1}$$

where $\gamma \in (\max\{0, 1 - \frac{1-p}{p}\}, 1)$, and

$$\omega_j|_{\omega_i=\underline{\omega}} = \begin{cases} \overline{\omega} & \text{with probability} \quad \frac{p}{1-p}(1-\gamma) \\ \underline{\omega} & \text{with probability} \quad 1 - \frac{p}{1-p}(1-\gamma) \end{cases}, \quad i \neq j. \tag{A2}$$

Throughout the Appendix, we assume $\mu > 2p\rho_1$.

To incorporate the analysis on the government's policy-adjustment capacity, in this generalized model, we assume that in the policy-adjustment stage, the reform policy R will be implemented with probability $e\sigma$, otherwise the status quo policy Q remains. $\sigma \in [0, 1]$ is exogenous and measures the government's ability to adjust the policy. Our benchmark model is a special case when $\sigma = 1$.

Collective-action payoff is characterized by:

	participate (j)	abstain (j)
participate (i)	$V_i^{11}(x;\omega_i,\omega_j)-k_i$	$V_i^{10}(x;\omega_i)-k_i$
abstain (i)	$V_i^{01}(x;\omega_i,\omega_j)$	$V_i^{00}(x;\omega_i)$

where k_i is the private cost of participating in collective action, which is private information. k_i is independently and identically distributed with a cumulative distribution function $F(\cdot)$ and support $[\underline{k},\overline{k}]$. k_i is only observed by citizen i after she observes government policy x. We also define

$$A_i(x;\omega_i,\omega_j)=V_i^{11}(x;\omega_i,\omega_j)-V_i^{01}(x;\omega_i,\omega_j); \tag{A3}$$

$$B_i(x;\omega_i)=V_i^{10}(x;\omega_i)-V_i^{00}(x;\omega_i). \tag{A4}$$

We call $D(\omega_i)$ is the desired policy of the type ω_i and denote $D(\underline{\omega})=Q, D(\overline{\omega})=R$.

It can be verified that the assumptions in the benchmark model are special cases of the following assumptions.

Assumption 1 *$F(\cdot)$ is weakly concave; $f(k)=F'(k)>0,\forall k\in[\underline{k},\overline{k}]$.*

Assumption 2 *Whenever the desired policy is chosen by the government, type ω_i never has an incentive to protest, that is, $\max\{A_i(D(\omega_i);\omega_i,\omega_j),B_i(D(\omega_i),\omega_i)\}\leq\underline{k},\forall\omega_j\in\{\underline{\omega},\overline{\omega}\}$.*

Assumption 3 *If $x\neq D(\omega_i), A(\omega_i)\triangleq A_i(x;\omega_i=\omega_j), B(\omega_i)\triangleq B_i(x;\omega_i)$ do not depend on i. $min\{A(\overline{\omega}),\overline{k}\}>B(\overline{\omega})>\underline{k}$; for the type $\underline{\omega}$, we either have $min\{A(\underline{\omega}),\overline{k}\}>B(\underline{\omega})>\underline{k}$ or $A(\underline{\omega})=B(\underline{\omega})=\underline{k}$.*

Under the assumptions that $u_i(R;\omega_i=\overline{\omega})=L, u_i(R;\omega_i=\underline{\omega})=-\underline{L}\leq 0,\quad i=1,2$, we have: $A(\overline{\omega})=(1-\lambda)L, B(\overline{\omega})=\lambda L, A(\underline{\omega})=(1-\lambda)\underline{L}, B(\underline{\omega})=\lambda\underline{L}$.

Lemma 4 (Characterizing the equilibrium in the above collective-action game) *(1) Whenever the government implements its desired policy $D(\omega_i)$, type ω_i never joins the collective action;*

(2) if $x \neq D(\omega_i)$, type ω_i protests according to a cut-point strategy:

$$a_i(\omega_0) = \begin{cases} 1 & \textit{if} \quad k_i \leq k^* \\ 0 & \textit{if} \quad k_i > k^* \end{cases} . \tag{A5}$$

$k^* = T_0(\widehat{\gamma}; \omega_i)$, where $\widehat{\gamma}$ is the probability with which she believes that the other player is of the same type; $T_0(\widehat{\gamma}; \omega_i)$ is uniquely and well defined by $T_0 = \min\{\widehat{\gamma}(A(\omega_i) - B(\omega_i))F(T_0) + B(\omega_i), \overline{k}\}$; and $T_0(\widehat{\gamma}; \omega_i)$ is weakly increasing, and it is strictly increasing whenever $\widehat{\gamma} \leq \min\{\gamma_0, 1\}$, where $\gamma_0 = \frac{\overline{k} - B(\omega_i)}{A(\omega_i) - B(\omega_i)}$. Hence $p_0(\widehat{\gamma}; \omega_i) = F(T_0(\widehat{\gamma}; \omega_i))$ is the probability with which a citizen protests if her desired policy is not implemented.

Proof *of* **Lemma** 4

Without loss of generality, suppose ω_0 is the dissatisfied type and her desired policy is not implemented. Define $A = A(\omega_0), B = B(\omega_0)$.

(a) Assumption 2 implies that it is a dominant strategy for type $\omega_i \neq \omega_0$ not to protest.

(b) Thus the only uncertainty is to what extent a type ω_0 citizen will join the protest.

(b.1) We first claim that in equilibrium citizen i of type ω_0 uses a cut-point strategy

$$a_i(\omega_0) = \begin{cases} 1 & \text{if} \quad k_i \leq k^* \\ 0 & \text{if} \quad k_i > k^* \end{cases} , \tag{A6}$$

because her payoff gain in protest is: $\widehat{\gamma}\Pr(j \text{ protest}|\omega_i = \omega_0)A + (1 - \widehat{\gamma}\Pr(j \text{ protest}|\omega_i = \omega_0))B - k_i$. Now suppose i's cut-point is $k_i^*, i = 1, 2$.

(b.2) According to (b.1) the payoff gain of player i is therefore $\widehat{\gamma}F(k_j^*)(A - B) + B - k_i$. It then can be verified that the equilibrium condition is equivalent to

$$k_1^* = \min\{\widehat{\gamma}F(k_2^*)(A - B) + B, \overline{k}\}, \tag{A7}$$

$$k_2^* = \min\{\widehat{\gamma}F(k_1^*)(A - B) + B, \overline{k}\}. \tag{A8}$$

(b.3) Suppose $\gamma_0 = \frac{\overline{k}-B}{A-B} < 1$, and focus on the situation where $\widehat{\gamma} \geq \frac{\overline{k}-B}{A-B}$.

We claim that the unique solution to Equation (A7) and Equation (A8) is $k_1^* = k_2^* = \overline{k}$. We can easily verify that $k_1^* = k_2^* = \overline{k}$ is a solution to the above equations and hence is an equilibrium. We need to check the other two possibilities.

Possibility 1: If at least one cut-point $k_i^* = \overline{k}$, then according to Equation (A7) and Equation (A8), the other cut-point automatically becomes the corner solution $\overline{k}$.

Possibility 2: Both cut-points are interior $k_1^*, k_2^* \in [B, \overline{k})$. Without loss of generality, let's assume $k_1^* \leq k_2^*$, so we get:

$$\widehat{\gamma} F(k_2^*)(A-B) + B \leq \widehat{\gamma} F(k_1^*)(A-B) + B, \tag{A9}$$

therefore $k_1^* \geq k_2^*$ so that $k_1^* = k_2^* \in [B, \overline{k})$. Let's denote them as k^*, we then have:

$$k^* = \widehat{\gamma} F(k^*)(A-B) + B. \tag{A10}$$

Because of Assumption 1, $\psi(x) \triangleq \widehat{\gamma} F(x)(A-B) + B - x$ is also weakly concave. In addition we have $\psi(\underline{k}) = B - \underline{k} > 0, \psi(\overline{k}) \geq 0$.

$\forall k \in (\underline{k}, \overline{k})$ can be represented by $k = \theta \underline{k} + (1-\theta)\overline{k}$ for some $\theta \in (0,1)$. So $\psi(k) \geq \theta \psi(\underline{k}) + (1-\theta)\psi(\overline{k}) > 0$. As a result, $k_1^* = k_2^* = \overline{k}$ is the unique equilibrium.

(b.4) When $0 < \widehat{\gamma} < \min\{\frac{\overline{k}-B}{A-B}, 1\}$, we first claim that any equilibrium $k_1^*, k_2^* \in [B, \overline{k})$. If not, we must have:

$$\overline{k} = \min\{\widehat{\gamma} F(k_i^*)(A-B) + B, \overline{k}\}, \tag{A11}$$

for some i. However the right-hand side $\min\{\widehat{\gamma} F(k_i^*)(A-B)+B, \overline{k}\} = \widehat{\gamma} F(k_i^*)(A-B)+B < \overline{k}$ because $\widehat{\gamma} < \min\{\frac{\overline{k}-B}{A-B}, 1\}$. So we get a contradiction. Hence, $k_1^*, k_2^* \in [B, \overline{k})$. Similarly as in (b.3), without loss of generality, let's assume $k_1^* \leq k_2^*$, so we get:

$$\widehat{\gamma} F(k_2^*)(A-B) + B \leq \widehat{\gamma} F(k_1^*)(A-B) + B, \tag{A12}$$

therefore $k_1^* \geq k_2^*$ so that $k_1^* = k_2^* \in [B, \overline{k})$. Let's denote them as k^*, we then have:

$$k^* = \widehat{\gamma} F(k^*)(A - B) + B. \tag{A13}$$

Because of Assumption 1 $\psi(x) \triangleq \widehat{\gamma} F(x)(A - B) + B - x$ is also weakly concave. In addition we have $\psi(\underline{k}) = B - \underline{k} > 0, \psi(\overline{k}) = \gamma(A - B) - (\overline{k} - B) < 0$.

Because of continuity of $\psi(x)$, $\exists$ a solution $k^* \in (\underline{k}, \overline{k})$ such that $k^* = \widehat{\gamma} F(k^*)(A-B)+B$.

Because of concavity of $\psi(x)$, applying the same logic in (b.3), $\forall k \in (\underline{k}, k^*)$, $\psi(k) > 0$ and $\forall k \in (k^*, \overline{k})$, $\psi(k) < 0$. As a result, k^* is the unique cut-point equilibrium.

(b.5) Therefore the equilibrium cut-point of type $\omega_i = \omega_0$ is uniquely determined by $k^* \triangleq T_0(\widehat{\gamma})$, where $T_0(\widehat{\gamma})$ is uniquely and well defined by

$$T_0(\widehat{\gamma}) = \min\{\widehat{\gamma} F(T_0(\widehat{\gamma}))(A - B) + B, \overline{k}\}. \tag{A14}$$

(c) When $\gamma_0 = \frac{\overline{k}-B}{A-B} < 1$ and $\widehat{\gamma} \geq \frac{\overline{k}-B}{A-B}, T_0(\gamma) = \overline{k}$. When $\gamma < \min\{\frac{\overline{k}-B}{A-B}, 1\}$, $T_0(\widehat{\gamma})$ is uniquely and well defined by Equation (A13). To rewrite the above equation, we get:

$$\widehat{\gamma} = \frac{T_0 - B}{(A - B)F(T_0)}. \tag{A15}$$

To show that $T_0(\widehat{\gamma})$ is strictly increasing when $\widehat{\gamma} < \min\{\frac{\overline{k}-B}{A-B}, 1\}$, we only need to show that above well-defined function is strictly increasing in T_0 when $T_0 \geq B$. It is obvious that $\frac{T_0-B}{(A-B)F(T_0)}$ is differentiable, thus we have:

$$\frac{d\frac{T_0-B}{(A-B)F(T_0)}}{dT_0} = \frac{F(T_0) - (T_0 - B)f(T_0)}{(A - B)(F(T_0))^2}. \tag{A16}$$

We only need to show $F(T_0) - (T_0 - B)f(T_0) > 0$. Because $F(T_0)$ is differentiable, $\exists \xi \in [\underline{k}, T_0]$ s.t. $F(T_0) = F(\underline{k}) + f(\xi)(T_0 - \underline{k}) > f(\xi)(T_0 - B) \geq (T_0 - B)f(T_0)$. The last inequality comes from concavity. ■

According to the above lemma, in our benchmark model, $T_0(\widehat{\gamma};\omega_i)$ is uniquely and well defined by $T_0 = \min\{\widehat{\gamma}(1-2\lambda)ZF(T_0)+\lambda Z, 1\}$ and $\gamma_0 = \frac{1-\lambda Z}{(1-2\lambda)Z}$, where

$$Z = \begin{cases} \underline{L} & \text{if} \quad \omega_i = \underline{\omega} \\ L & \text{if} \quad \omega_i = \overline{\omega} \end{cases} . \tag{A17}$$

Therefore $p_0(\widehat{\gamma};\omega_i) = F(T_0(\widehat{\gamma};\omega_i))$ is determined by:

$$F^{-1}(p_0) = \min\{\widehat{\gamma}(1-2\lambda)Zp_0 + \lambda Z, 1\}. \tag{A18}$$

Deriving of the Government's Expected Payoffs

If both citizens are pro-reform, i.e., $\omega_1 = \omega_2 = \overline{\omega}$, with public communication they know each other's type and each independently protests against the *status quo* policy with probability $p_0(1)$. The government's payoff therefore is:

$$G_1(e,\omega_1,\omega_2)|_{\omega_1=\omega_2=\overline{\omega}} = -(1-e\sigma)W(p_0(1)) - e\sigma\mu, \tag{A19}$$

where $W(x) = \rho_2 x^2 + 2\rho_1 x(1-x)$.

Without public communication, a dissatisfied citizen believes that the other citizen is also dissatisfied with probability $\widehat{\gamma}_0^*$, so that she will protest with probability $p_0(\widehat{\gamma}_0^*)$. Hence without public communication the government's payoff is:

$$G_0(e,\omega_1,\omega_2)|_{\omega_1=\omega_2=\overline{\omega}} = -(1-e\sigma)W(p_0(\widehat{\gamma}_0^*)) - e\sigma\mu. \tag{A20}$$

If both citizens are anti-reform types, i.e., $\omega_1 = \omega_2 = \underline{\omega}$, they do not protest against the status quo policy so that the government suffers 0. If reform is launched, under public communication, each of them independently protests with probability $p_0(1)$. The government's payoff is:

$$G_1(e,\omega_1,\omega_2)|_{\omega_1=\omega_2=\underline{\omega}} = -e\sigma[\mu + W(p_0(1))]. \tag{A21}$$

Similarly, without public communication, the government's payoff is

$$G_0(e,\omega_1,\omega_2)|_{\omega_1=\omega_2=\underline{\omega}} = -e\sigma[\mu + W(p_0(\widehat{\gamma}_1^*))]. \tag{A22}$$

The only difference between the above two payoffs is in a dissatisfied citizen's belief of the other citizen's type.

If one citizen is pro-reform and the other is anti-reform, i.e., $\omega_1 \neq \omega_2$, with public communication, whoever is dissatisfied protests with probability $p_0(0)$. Hence, the government's payoff is:

$$G_1(e,\omega_1,\omega_2)|_{\omega_1\neq\omega_2} = -(1-e\sigma)\rho_1 p_0(0) - e\sigma(\mu + \rho_1 p_0(0)). \tag{A23}$$

Similarly when public communication is not allowed, the government gets

$$G_0(e,\omega_1,\omega_2)|_{\omega_1\neq\omega_2} = -(1-e\sigma)\rho_1 p_0(\widehat{\gamma}_0^*) - e\sigma(\mu + \rho_1 p_0(\widehat{\gamma}_1^*)). \tag{A24}$$

The government's payoffs in equations are summarized as follows.

$$G_1(e,\omega_1,\omega_2) = \begin{cases} -e\sigma[\mu + W(p_0(1,\underline{\omega}))] & \text{if} \quad \omega_1=\omega_2=\underline{\omega} \\ -(1-e\sigma)W(p_0(1,\overline{\omega})) - e\sigma\mu & \text{if} \quad \omega_1=\omega_2=\overline{\omega} \\ -(1-e\sigma)\rho_1 p_0(0,\overline{\omega}) - e\sigma(\mu+\rho_1 p_0(0,\underline{\omega})) & \text{if} \quad \omega_1=\underline{\omega},\ \omega_2=\overline{\omega} \\ -(1-e\sigma)\rho_1 p_0(0,\overline{\omega}) - e\sigma(\mu+\rho_1 p_0(0,\underline{\omega})) & \text{if} \quad \omega_1=\overline{\omega},\ \omega_2=\underline{\omega} \end{cases}. \tag{A25}$$

$$G_0(e,\omega_1,\omega_2) = \begin{cases} -e\sigma[\mu + W(p_0(\underline{\widehat{\gamma}_1^*},\underline{\omega}))] & \text{if} \quad \omega_1=\omega_2=\underline{\omega} \\ -(1-e\sigma)W(p_0(\widehat{\gamma}_0^*,\overline{\omega})) - e\sigma\mu & \text{if} \quad \omega_1=\omega_2=\overline{\omega} \\ -(1-e\sigma)\rho_1 p_0(\widehat{\gamma}_0^*,\overline{\omega}) - e\sigma(\mu+\rho_1 p_0(\underline{\widehat{\gamma}_1^*},\underline{\omega})) & \text{if} \quad \omega_1=\underline{\omega},\ \omega_2=\overline{\omega} \\ -(1-e\sigma)\rho_1 p_0(\widehat{\gamma}_0^*,\overline{\omega}) - e\sigma(\mu+\rho_1 p_0(\underline{\widehat{\gamma}_1^*},\underline{\omega})) & \text{if} \quad \omega_1=\overline{\omega},\ \omega_2=\underline{\omega} \end{cases}, \tag{A26}$$

where

$$\underline{\widehat{\gamma}_1^*} = 1 - \frac{p}{1-p}(1 - \widehat{\gamma}_1^*). \quad \text{(A27)}$$

We summarize the government's payoffs in the generalized model under the two circumstances ($\alpha = 0$ and $\alpha = 1$) and their differences in the following table. Without considering the learning effect of its own, the government benefits from public communication when the two citizens are of different types, and loses from public communication when both citizens share the same interest. In other words, public communication might be beneficial to government even if it does not adjust policy based on what it learns from the citizens.

Table 2. Net Gain from Opening Public Communication Holding Government's Effort Constant

	$\omega_1 = \omega_2 = \underline{\omega}$	$\omega_1 = \omega_2 = \overline{\omega}$	$\omega_1 \neq \omega_2$
$\alpha = 1$	$-e\sigma[\mu + W(p_0(1))]$	$-(1-e\sigma)W(p_0(1)) - e\sigma\mu$	$-(1-e\sigma)\rho_1 p_0(0)$ $-e\sigma[\mu + \rho_1 p_0(0)]$
$\alpha = 0$	$-e\sigma[\mu + W(p_0(\underline{\widehat{\gamma}_1^*}))]$	$-(1-e\sigma)W(p_0(\widehat{\gamma}_0^*)) - e\sigma\mu$	$-(1-e\sigma)\rho_1 p_0(\widehat{\gamma}_0^*)$ $-e\sigma[\mu + \rho_1 p_0(\widehat{\gamma}_1^*)]$
$G_1 - G_0$	$-e\sigma Z(\underline{\widehat{\gamma}_1^*}) \leq 0$	$-[1-e\sigma]Z(\widehat{\gamma}_0^*) \leq 0$	$(1-e\sigma)\rho_1[p_0(\widehat{\gamma}_0^*) - p_0(0)]$ $e\sigma\rho_1[p_0(\widehat{\gamma}_1^*) - p_0(0)] > 0$

Note: $Z(x) \triangleq [W(p_0(1)) - W(p_0(x))]$.

Lemma 5 (No effort without public communication) *Given one of the following conditions,*

(a) $\rho_2 \geq 2\rho_1, A(\overline{\omega}) = A(\underline{\omega}), B(\overline{\omega}) = B(\underline{\omega}), p \leq \frac{1}{2}$,

(b) $pW(p_0(1, \overline{\omega})) \leq \mu$, *where* $W(x)$ *is defined by Equation 8,*

(c) $\rho_2 \leq 2\rho_1$,

(d) σ *is sufficiently close to 1 (including* $\sigma = 1$*),*

(e) $\sigma = 0$,

in any equilibrium, the government chooses not to make any effort to adjust the policy when public communication is not allowed, i.e., $e_0^*(\gamma) \equiv 0$.

Proof *of* **Lemma** 5

When $\sigma = 0$, any equilibrium is equivalent to the case with no effort. Without loss of generality, let's assume $\sigma > 0$. We have

$$e_0^*(\gamma) \in \arg\min_e [(1-e\sigma)M(\gamma, p_0(\widehat{\gamma}_0^*, \overline{\omega})) + e\sigma(\mu + \underline{M}(\gamma, p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega}))], \tag{A28}$$

where $M(\gamma, x)$ is defined by Equation 13, and

$$\underline{M}(\gamma, p_0(\underline{\widehat{\gamma}_1^*})) = (1 - 2p + p\gamma)p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega})^2(\rho_2 - 2\rho_1) + 2(1-p)\rho_1 p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega}), \tag{A29}$$

where

$$\underline{\widehat{\gamma}_1^*} = 1 - \frac{p}{1-p}(1 - \widehat{\gamma}_1^*). \tag{A30}$$

Notice that the payoff gain $M(\gamma, p_0(\widehat{\gamma}_0^*, \overline{\omega})) - \underline{M}(\gamma, p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega})) - \mu$ is linear in γ, hence, $e_0^*(\gamma)$ must follow a cut-point rule.

(a) Since $A(\overline{\omega}) = A(\underline{\omega}), B(\overline{\omega}) = B(\underline{\omega})$, we have $p_0(x, \underline{\omega}) = p_0(x, \overline{\omega})$ for any x.

The slope of γ in the payoff gain is $pp_0(\widehat{\gamma}_0^*, \overline{\omega})^2(\rho_2 - 2\rho_1) - pp_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega})^2(\rho_2 - 2\rho_1)$.

(a.1) Suppose $e_0^*(\gamma)$ involves with an interior cut-point.

If $\widehat{\gamma}_0^* > \underline{\widehat{\gamma}_1^*}$, the payoff gain is strictly increasing in γ. So the government makes an effort only if γ is higher. Therefore $\widehat{\gamma}_1^* > \widehat{\gamma}_0^*$, so that $\underline{\widehat{\gamma}_1^*} \geq \widehat{\gamma}_1^* > \widehat{\gamma}_0^*$. It is a contradiction.

So we must have $\widehat{\gamma}_0^* \leq \underline{\widehat{\gamma}_1^*}$ and the payoff gain is strictly decreasing in γ.

The payoff gain at $\gamma = 0$ equals $2pp_0(\widehat{\gamma}_0^*, \overline{\omega})\rho_1 - 2(1-p)\rho_1 p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega}) - (1-2p)p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega})^2(\rho_2 - 2\rho_1) - \mu < 0$. It is a contradiction given the fact that the payoff gain is decreasing and $e_0^*(\gamma)$ follows an interior cut-point rule.

(a.2) As a result, $e_0^*(\gamma)$ should be independent of γ. Because $M(\gamma, p_0(\widehat{\gamma}_1^*, \overline{\omega})) \leq \underline{M}(\gamma, p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega}))$, $e_0^*(\gamma) > 0$ is ruled out, and $e_0^*(\gamma) \equiv 0$ is the only possible equilibrium strategy. As long as off-the-equilibrium-path citizens believe they are weakly more homogeneous than in the case without reform, the government never has an incentive to deviate from this pooling strategy.

(b) When $\rho_2 \geq 2\rho_1$, we have $M(\gamma, p_0(\widehat{\gamma}_0^*, \overline{\omega})) < M(1, p_0(1, \overline{\omega})) = pW(p_0(1, \overline{\omega})) \leq \mu$, thus it is the strictly dominant strategy not to make any effort. When $\rho_2 < 2\rho_1$, the proof is the same as in part (c).

(c) Similarly $M(\gamma, p_0(\widehat{\gamma}_0^*, \overline{\omega})) < 2p\rho_1 p_0(\widehat{\gamma}_0^*, \underline{\omega}) \leq \mu$, thus it is the strictly dominant strategy not to make any effort.

(d) Without loss of generality, assume that $W(p_0(1, \overline{\omega})) > \mu$ so that the government always makes an effort to adjust the policy under public communication when it sees two dissatisfied citizens; otherwise, we can directly apply the results in part (b). Suppose there is a set of γ with a positive measurement, under which the government chooses not to allow public communication but chooses to reform. Under these γ's, the government gets $-\sigma[\mu + \underline{M}(\gamma, p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega}))] - (1-\sigma)M(\gamma, p_0(\widehat{\gamma}_0^*, \overline{\omega})) = -\sigma\mu - [\sigma\underline{M}(\gamma, p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega})) + (1-\sigma)M(\gamma, p_0(\widehat{\gamma}_0^*, \overline{\omega}))]$. If the government allow public communication, it gets $-p\gamma[\sigma\mu + (1-\sigma)W(p_0(1, \overline{\omega}))] - 2p(1-\gamma)p_0(0, \overline{\omega})\rho_1$. In the following we will show that it makes the government strictly better off to allow public communication.

$$
\begin{aligned}
&-\sigma\mu - [\sigma\underline{M}(\gamma, p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega})) + (1-\sigma)M(\gamma, p_0(\widehat{\gamma}_0^*, \overline{\omega}))] \\
&\leq -\sigma\mu \\
&< min\{-p[\sigma\mu + (1-\sigma)W(p_0(1, \overline{\omega}))], -2pp_0(0, \overline{\omega})\rho_1\} \\
&\leq -p\gamma[\sigma\mu + (1-\sigma)W(p_0(1, \overline{\omega}))] - 2p(1-\gamma)p_0(0, \overline{\omega})\rho_1
\end{aligned}
$$

The second inequality is valid when σ is sufficiently close to 1. As a result, there is a profitable deviation for the government, so that the government chooses not to adjust policy without public communication.

(e) The result is obvious when $\sigma = 0$. ■

Proposition 4 *Provided either $p \leq \frac{1}{2}$, or σ is sufficiently close to 1, or any other condition in Lemma 5 is satisfied, (1) in any equilibrium, the government allows public communication if and only if its private signal indicates that citizen preferences are sufficiently heterogeneous, i.e.,*

$$
\alpha^* = \begin{cases} 1 & if\ \gamma < \gamma^* \\ 0 & if\ \gamma \geq \gamma^* \end{cases}, \tag{A31}
$$

where $\gamma^* \in [0,1]$*; when* $p \leq \frac{1}{2}, \gamma^* > 0$;

(2) when public communication is shut down, citizens will think that their preferences are more homogeneous than when public communication is allowed, i.e., $E(\gamma|\alpha = 0) > E(\gamma|\alpha = 1)$*, provided that* $\gamma^* > \max\{0, 1 - \frac{1-p}{p}\}$.

Below we assume $p \leq \frac{1}{2}$ *and at least one condition in Lemma 5 is satisfied.*

(3) If $\min_x G_{diff}(x, E(\gamma|\gamma \geq x)) < 0$*, an equilibrium exists with interior cut-point* $\gamma^* \in (0,1)$*;*[1]

(4) if one of the following conditions is satisfied, for sufficiently small σ*, we have:* $\min_x G_{diff}(x, E(\gamma|\gamma \geq x)) < 0$*:*

(4.a) $p_0(E(\gamma), \overline{\omega}) < 1$*, and* ρ_2/ρ_1 *is sufficiently large,*

(4.b) $\rho_2 \geq 2\rho_1, \exists \gamma_0' \in (\max\{0, 1 - \frac{1-p}{p}\}, 1)$ *such that* $p_0(E(\gamma|\gamma \geq \gamma_0'), \overline{\omega}) \leq \gamma_0' p_0(1, \overline{\omega}) + (1 - \gamma_0')p_0(0, \overline{\omega})$*; and*

(5) if one of the following conditions is satisfied, for σ *that is sufficiently close to 1 (including* $\sigma = 1$*), we have* $\min_x G_{diff}(x, E(\gamma|\gamma \geq x)) < 0$*:*

(5.a) $W(p_0(E(\gamma), \overline{\omega})) < \mu$*, and* ρ_1 *is sufficiently small,*

(5.b) $\rho_2 \leq 2\rho_1$.

Proof *of* **Proposition** 4

The difference in the government's payoff is $G_{diff}(\gamma, \widehat{\gamma}_0^*) = (1 - e^*(\gamma)\sigma)M(\gamma, p_0(\widehat{\gamma}_0^*, \overline{\omega})) + e^*(\gamma)\sigma(\mu + \underline{M}(\gamma, p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega})) - p\gamma \min\{W(p_0(1, \overline{\omega})), \sigma\mu + (1-\sigma)W(p_0(1, \overline{\omega}))\} - 2p(1-\gamma)p_0(0, \overline{\omega})\rho_1$. We know that the payoff gain as a function of γ is either linear or piecewise linear. When it is piecewise linear, it is inverse V-shaped.

(1) (a) First we show that $G_{diff}(0, \widehat{\gamma}_0^*) \geq 0$. We have

$$G_{diff}(0, \widehat{\gamma}_0^*) = -2pp_0(0, \overline{\omega})\rho_1 + (1 - e^*(0))2pp_0(\widehat{\gamma}_0^*, \overline{\omega})\rho_1 + e^*(0)[\sigma((1-2p)p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega})^2(\rho_2 - 2\rho_1) + 2(1-p)\rho_1 p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega}) + \mu) + (1-\sigma)p_0(\widehat{\gamma}_0^*, \overline{\omega})].$$

[1] An equilibrium with $\gamma^* = 1$ always exists. In this equilibrium, whenever the government shuts down the platform of public communication, citizens believe that their preferences are perfectly correlated.

When σ is sufficiently close to 1, or any other condition in Lemma 5 is satisfied, $e^*(0) = 0$ so that $G_{diff}(0, \widehat{\gamma}_0^*) = 2p\rho_1[p_0(\widehat{\gamma}_0^*, \overline{\omega}) - p_0(0, \overline{\omega})] \geq 0$.

When $p \leq \frac{1}{2}$, without loss of generality we assume that $\rho_2 > 2\rho_1$, otherwise we can directly apply the proof above. So we have

$(1-2p)p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega})^2(\rho_2 - 2\rho_1) + 2(1-p)\rho_1 p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega}) + \mu$

$\geq (1-2p)p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega})^2\rho_2 + 2(1-p-1+2p)\rho_1 p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega}) + \mu$

$= (1-2p)p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega})^2\rho_2 + 2p\rho_1 p_0(\underline{\widehat{\gamma}_1^*}, \underline{\omega}) + \mu$

$\geq \mu$

$\geq 2pp_0(\widehat{\gamma}_0^*, \overline{\omega})\rho_1$.

Therefore $G_{diff}(0, \widehat{\gamma}_0^*) \geq 2p\rho_1[p_0(\widehat{\gamma}_0^*, \overline{\omega}) - p_0(0, \overline{\omega})] \geq 0$.

(b) Because of the linearity or piecewise linearity ("inverse V-shaped") of the payoff gain function, the government's equilibrium strategy can be either the strategy that is independent of γ or a cut-point strategy with the following form

$$\alpha^* = \begin{cases} 1 & \text{if } \gamma < \gamma^* \\ 0 & \text{if } \gamma \geq \gamma^* \end{cases}, \tag{A32}$$

where $\gamma^* \in [0, 1]$.

(c) When $p \leq \frac{1}{2}$, suppose a pooling equilibrium exists with $\gamma^* = 0$, the citizens' equilibrium belief $\widehat{\gamma}_0^* = E(\gamma) > 0$. The payoff gain at $\gamma = 0$ equals $2p\rho_1(p_0(E(\gamma), \overline{\omega}) - p_0(0, \overline{\omega})) > 0$, which is a contradiction. Therefore, in any equilibrium under the assumption that $p \leq \frac{1}{2}$, $\gamma^* > 0$.

(2) When the citizens observe $\alpha = 0$, their belief toward the social homogeneity is $E(\gamma|\gamma \geq \gamma^*)$. When the citizens observe $\alpha = 0$, their belief toward the preference homogeneity is $E(\gamma|\gamma < \gamma^*)$. It is obvious that $E(\gamma|\gamma \geq \gamma^*) > E(\gamma|\gamma < \gamma^*)$.

(3) Under the assumptions, we have

$$G_{diff}(\gamma,\widehat{\gamma}_0^*)=M(\gamma,p_0(\widehat{\gamma}_0^*,\overline{\omega}))-p\gamma\min\{W(p_0(1,\overline{\omega})),\sigma\mu+(1-\sigma)W(p_0(1,\overline{\omega}))\}-2p(1-\gamma)p_0(0,\overline{\omega})\rho_1. \tag{A33}$$

$E(\gamma|\gamma\geq x)=\int_{\gamma\geq x}\gamma d\frac{G(\gamma)-G(x)}{1-G(x)}=\frac{\int_{\gamma\geq x}\gamma dG(\gamma)}{1-G(x)}$ is continuous in x, so that $G_{diff}(x,E(\gamma|\gamma\geq x))$ is continuous in x.

$G_{diff}(x,E(\gamma|\gamma\geq x))|_{x=0}=2p\rho_1(p_0(E(\gamma),\overline{\omega})-p_0(0,\overline{\omega}))>0$, thus as long as $\min_x G_{diff}(x,E(\gamma|\gamma\geq x))<0$, there always exists a fixed point $\gamma^*\in(0,1)$ such that $G_{diff}(\gamma,E(\gamma|\gamma\geq\gamma^*))\geq 0$ with $\gamma\leq\gamma^*$, and $G_{diff}(\gamma,E(\gamma|\gamma\geq\gamma^*))\leq 0$ with $\gamma>\gamma^*$.

(4) Assume $\sigma=0$ and fix a γ close to the lower bound, the payoff gain is $G_{diff}(\gamma,\widehat{\gamma}_0^*)=M(\gamma,p_0(\widehat{\gamma}_0^*,\overline{\omega}))-p\gamma W(p_0(1,\overline{\omega}))-2p(1-\gamma)p_0(0,\overline{\omega})\rho_1$

$=p\gamma p_0(\widehat{\gamma}_0^*,\overline{\omega})^2(\rho_2-2\rho_1)+2pp_0(\widehat{\gamma}_0^*,\overline{\omega})\rho_1-p\gamma[p_0(1,\overline{\omega})^2(\rho_2-2\rho_1)+2p_0(1,\overline{\omega})\rho_1]-2p(1-\gamma)p_0(0,\overline{\omega})\rho_1$

$=-p\gamma[p_0(1,\overline{\omega})^2-p_0(\widehat{\gamma}_0^*,\overline{\omega})^2](\rho_2-2\rho_1)+2p\rho_1[p_0(\widehat{\gamma}_0^*,\overline{\omega})-\gamma p_0(1,\overline{\omega})-(1-\gamma)p_0(0,\overline{\omega})]$.

(4.a) As $\frac{\rho_2}{\rho_1}$ goes to infinity, the payoff gain is close to $p\gamma p_0^2(\widehat{\gamma}_0^*,\overline{\omega})\rho_2-p\gamma p_0^2(1,\overline{\omega})\rho_2<0$.

(4.b) When $\rho_2\geq 2\rho_1$ and a γ exists such that $p_0(\widehat{\gamma}_0^*,\overline{\omega})-\gamma p_0(1,\overline{\omega})-(1-\gamma)p_0(0,\overline{\omega})\leq 0$, the payoff gain is also negative at this γ.

When we let σ get close to 0, we will therefore always find an x such that $G_{diff}(x,E(\gamma|\gamma\geq x))<0$.

(5) First assume that $W(p_0(1,\overline{\omega}))>\mu$, so that

$$G_{diff}(\gamma,\widehat{\gamma}_0^*)=M(\gamma,p_0(\widehat{\gamma}_0^*,\overline{\omega}))-p\gamma[\sigma\mu+(1-\sigma)W(p_0(1,\overline{\omega}))]-2p(1-\gamma)p_0(0,\overline{\omega})\rho_1. \tag{A34}$$

We need to show that there is an interior $\widehat{\gamma}$ such that $M(\widehat{\gamma},p_0(\widehat{\gamma}_0^*,\overline{\omega}))-2p(1-\gamma)p_0(0,\overline{\omega})\rho_1<p\widehat{\gamma}\mu$ when ρ_1 is sufficiently small, where $\widehat{\gamma}_0^*=E(\gamma|\gamma\geq\widehat{\gamma})$.

Since $W(p_0(E(\gamma),\overline{\omega}))<\mu$, there is an interior $\widehat{\gamma}$ such that $W(p_0(\widehat{\gamma}_0^*,\overline{\omega}))<\mu$, therefore $p\widehat{\gamma}W(p_0(\widehat{\gamma}_0^*,\overline{\omega}))+2p(1-\widehat{\gamma})[p_0(\widehat{\gamma}_0^*,\overline{\omega})-p_0(0,\overline{\omega})]\rho_1<p\widehat{\gamma}\mu$ when ρ_1 is sufficiently small. So we have

$M(\widehat{\gamma}, p_0(\widehat{\gamma}_0^*, \overline{\omega})) - 2p(1-\widehat{\gamma})p_0(0, \overline{\omega})\rho_1$

$= p\widehat{\gamma}(\rho_2 - 2\rho_1)p_0(\widehat{\gamma}_0^*, \overline{\omega})^2 + 2p\rho_1 p_0(\widehat{\gamma}_0^*, \overline{\omega}) - 2p(1-\widehat{\gamma})p_0(0, \overline{\omega})\rho_1$

$= p\widehat{\gamma}[(\rho_2 - 2\rho_1)p_0(\widehat{\gamma}_0^*, \overline{\omega})^2 + 2p\rho_1 p_0(\widehat{\gamma}_0^*, \overline{\omega})] + 2p(1-\widehat{\gamma})[p_0(\widehat{\gamma}_0^*, \overline{\omega}) - p_0(0, \overline{\omega})]\rho_1$

$= p\widehat{\gamma}W(p_0(\widehat{\gamma}_0^*, \overline{\omega})) + 2p(1-\widehat{\gamma})[p_0(\widehat{\gamma}_0^*, \overline{\omega}) - p_0(0, \overline{\omega})]\rho_1$

$< p\widehat{\gamma}\mu$.

As a result, when σ is sufficiently close to 1, we have $M(\widehat{\gamma}, p_0(\widehat{\gamma}_0^*, \overline{\omega})) - 2p(1-\widehat{\gamma})p_0(0, \overline{\omega})\rho_1 < p\widehat{\gamma}[\sigma\mu + (1-\sigma)W(p_0(1, \overline{\omega}))]$, i.e., $G_{diff}(\widehat{\gamma}, E(\gamma|\gamma \geq \widehat{\gamma})) < 0$.

When $W(p_0(1, \overline{\omega})) \leq \mu$, the proof is similar as above.

When $\rho_2 \leq 2\rho_1$, we have $W(p_0(E(\gamma), \overline{\omega})) < \mu$ so that $\min_x G_{diff}(x, E(\gamma|\gamma \geq x))$.

(6) Finally to complete the proof we also need to check the IC conditions of the citizens so that they have incentives to fully reveal their preferences when public communication is allowed. The IC conditions are checked in the following lemma. ∎

Lemma 6 *Provided the conditions below, citizens' truth-telling incentives are satisfied:*

(I) sufficient conditions for incentive compatibility of type $\underline{\omega}$*:*

$u_i(R, \underline{\omega}) \leq u_i(Q, \underline{\omega})$,

$V_i^{01}(Q, \underline{\omega}, \overline{\omega}) \geq \max\{V_i^{00}(R, \underline{\omega}), V_i^{10}(R, \underline{\omega})\}, V_i^{00}(Q, \underline{\omega}) \geq \max\{V_i^{00}(R, \underline{\omega}), V_i^{10}(R, \underline{\omega}), V_i^{01}(Q, \underline{\omega}, \overline{\omega})\}$;

and

(II) sufficient conditions for incentive compatibility of type $\overline{\omega}$*:*

$V_i^{00}(R, \overline{\omega}) \geq \max\{V_i^{00}(x, \overline{\omega}), V_i^{10}(x, \overline{\omega}), V_i^{00}(x, \overline{\omega}), V_i^{01}(x, \overline{\omega}, \overline{\omega})\}, \forall x \in \{R, Q\}$;

$u_i(R, \overline{\omega}) \geq u_i(Q, \overline{\omega})$,

$V_i^{11}(Q, \overline{\omega}, \overline{\omega}) \geq V_i^{10}(Q, \overline{\omega}), V_i^{01}(Q, \overline{\omega}, \overline{\omega}) \geq V_i^{00}(Q, \overline{\omega})$.

Proof *of* **Lemma** 6

(a) Denote q_t as the probability that reform will be implemented when the government observes t number of pro-reform citizens.

First notice that in any citizen truth-telling equilibrium, $q_2 \geq q_1 = q_0 = 0$.

(b) We check the payoff gain of the $\underline{\omega}$ type between claiming $\overline{\omega}$ and $\underline{\omega}$.

When $\omega_j = \underline{\omega}$, by claiming $\overline{\omega}$ she gets $u_i(Q, \underline{\omega}) + V_i^{00}(Q; \underline{\omega})$, by claiming $\underline{\omega}$ she gets $u_i(Q,\underline{\omega})+V_i^{00}(Q,\underline{\omega})$, which makes her weakly better than claiming $\overline{\omega}$.

When $\omega_j=\overline{\omega}$, by claiming $\overline{\omega}$ she gets $q_2[u_i(R,\underline{\omega})+\max\{V_i^{00}(R,\underline{\omega}),V_i^{10}(R,\underline{\omega})-k_i\}]+(1-q_2)[u_i(Q,\underline{\omega})+\delta_jV_i^{01}(Q,\underline{\omega},\overline{\omega})+(1-\delta_j)V_i^{00}(Q,\underline{\omega})]$;

by claiming $\underline{\omega}$, she gets $u_i(Q,\underline{\omega})+\underline{\delta}_jV_i^{01}(Q,\underline{\omega},\overline{\omega})+(1-\underline{\delta}_j)V_i^{00}(Q,\underline{\omega},\overline{\omega})$, where $\underline{\delta}_j\leq\delta_j$ according to Lemma 4.

As $V_i^{01}(Q,\underline{\omega},\overline{\omega})\geq\max\{V_i^{00}(R,\underline{\omega}),V_i^{10}(R,\underline{\omega})\}$, $V_i^{00}(Q,\underline{\omega})\geq\max\{V_i^{00}(R,\underline{\omega}),V_i^{10}(R,\underline{\omega}),V_i^{01}(Q,\underline{\omega},\overline{\omega})\}$, claiming $\overline{\omega}$ does not provide a profitable deviation.

(c) We check the payoff gain of type $\overline{\omega}$ between claiming $\overline{\omega}$ and $\underline{\omega}$. When $\omega_j=\underline{\omega}$, the policy will never be changed, therefore, the other person will never protest. Thus, type $\overline{\omega}$ is indifferent between claiming $\overline{\omega}$ and $\underline{\omega}$.

When $\omega_j=\overline{\omega}$, by claiming $\overline{\omega}$, she gets $q_2[u_i(R,\overline{\omega})+V_i^{00}(R,\overline{\omega})]+(1-q_2)[u_i(Q,\overline{\omega})+VV]$,

where $VV=\max\{\delta'V_i^{11}(Q,\overline{\omega},\overline{\omega})+(1-\delta')V_i^{10}(Q,\overline{\omega})-k_i,\delta'V_i^{01}(Q,\overline{\omega},\overline{\omega})+(1-\delta')V_i^{00}(Q,\overline{\omega})\}$;

By claiming $\underline{\omega}$, she gets $[u_i(Q,\overline{\omega})+ZZ]$,

where $ZZ=\max\{\underline{\delta'}V_i^{11}(Q,\overline{\omega},\overline{\omega})+(1-\underline{\delta'})V_i^{10}(Q,\overline{\omega})-k_i,\underline{\delta'}V_i^{01}(Q,\overline{\omega},\overline{\omega})+(1-\underline{\delta'})V_i^{00}(Q,\overline{\omega})\}$,

$\underline{\delta'}\leq\delta'$ according to Lemma 4.

Because $V_i^{11}(Q,\overline{\omega},\overline{\omega})\geq V_i^{10}(Q,\overline{\omega})$, $V_i^{01}(Q,\overline{\omega},\overline{\omega})\geq V_i^{00}(Q,\overline{\omega})$, we have

$\delta'V_i^{11}(Q,\overline{\omega},\overline{\omega})+(1-\delta')V_i^{10}(Q,\overline{\omega})\geq\underline{\delta'}V_i^{11}(Q,\overline{\omega},\overline{\omega})+(1-\underline{\delta'})V_i^{10}(Q,\overline{\omega})$,

$\delta'V_i^{01}(Q,\overline{\omega},\overline{\omega})+(1-\delta')V_i^{00}(Q,\overline{\omega})\geq\underline{\delta'}V_i^{01}(Q,\overline{\omega},\overline{\omega})+(1-\underline{\delta'})V_i^{00}(Q,\overline{\omega})$.

Hence, $VV\geq ZZ$, and claiming $\underline{\omega}$ does not offer a profitable deviation.

Thus, IC constraint is satisfied. ■

Proof *of* **Corollary** 1

(1) The result directly comes from the last part of Proposition 4.

(2) When $\mu<W(p_0(0,\overline{\omega}))$, it's easy to verify that the government's payoff gain is strictly positive, hence, it always allows public communication. ■

Proof *of* **Proposition** 2

In the generalized version, we need to assume that at least one of the conditions in Lemma 5 is satisfied.

Define $\gamma^{**} = \inf_{\gamma^* \text{ is an equilibrium}} \gamma^*$. First of all it is well defined. We then show that

(a) it is an equilibrium;

(b) $\gamma^{**} > 0$; and

(c) it maximizes the government's welfare and minimizes openness.

Recall that the payoff gain of the government is

$$G_{diff}(\gamma, \widehat{\gamma}_0^*) = M(\gamma, p_0(\widehat{\gamma}_0^*, \overline{\omega})) - p\gamma \min\{W(p_0(1, \overline{\omega})), \sigma\mu + (1-\sigma)W(p_0(1, \overline{\omega}))\} - 2p(1-\gamma)p_0(0, \overline{\omega})\rho_1, \tag{A35}$$

(d) By the definition of γ^{**}, a series of equilibria γ_t^* exist such that $G_{diff}(\gamma_t^*, E(\gamma|\gamma \geq \gamma_t^*)) = 0$, and $\gamma_t^* \to \gamma^{**}$. By the continuity of $G_{diff}(x, E(\gamma|\gamma \geq x))$, we know that $G_{diff}(\gamma^{**}, E(\gamma|\gamma \geq \gamma^{**})) = 0$ so that γ^{**} is also an equilibrium.

(e) $G_{diff}(0, E(\gamma|\gamma \geq 0) = 2p\rho_1(p_0(E(\gamma), \overline{\omega}) - p_0(0, \overline{\omega})) > 0$, therefore $\gamma^{**} > 0$.

(f) We need to show the following property: if $\gamma_a^* > \gamma_b^*$ and they are both equilibria, then the government's welfare is strictly higher under γ_b^* than under γ_a^*. Denote $G(\gamma; \gamma^*)$ as the government's welfare under the equilibrium γ^*.

When $\gamma < \gamma_b^*$, the government opens public communication under both equilibria, therefore, $G(\gamma; \gamma_b^*) = G(\gamma; \gamma_a^*)$.

When $\gamma \in [\gamma_b^*, \gamma_a^*)$, under the equilibrium γ_b^*, the government prefers forbidding public communication. Thus its welfare $G(\gamma; \gamma_b^*)$ is strictly higher than the welfare under openness which equals to $G(\gamma; \gamma_a^*)$.

When $\gamma \geq \gamma_a^*, G(\gamma; \gamma_a^*) = -M(\gamma, p_0(E(\gamma|\gamma \geq \gamma_a^*), \overline{\omega})), G(\gamma; \gamma_b^*) = -M(\gamma, p_0(E(\gamma|\gamma \geq \gamma_b^*), \overline{\omega}))$. Since $\gamma_a^* > \gamma_b^*$, $E(\gamma|\gamma \geq \gamma_a^*) \geq E(\gamma|\gamma \geq \gamma_b^*)$. Because $M(\gamma, x)$ is increasing in x, we get $G(\gamma; \gamma_a^*) \leq G(\gamma; \gamma_b^*)$.

Since γ^{**} is the "smallest" equilibrium cut-point, it is obvious that it allows minimum level of openness among all equilibria. ■

Equilibrium characterization in the private polling game

Upon two pro-reform citizens, the government's payoff is $-\mu$ when the reform policy is launched and $-W(p_0(q(\varepsilon^*)))$ when the status quo is kept, where $W(\cdot)$ is defined by Equation 8. The government's equilibrium choice ε^* must solve the following problem:

$$\max_{\varepsilon \in [0,1]} \varepsilon\sigma(-\mu) + (1 - \varepsilon\sigma)[-W(p_0(q(\varepsilon^*))]. \tag{A36}$$

There are two scenarios to be considered. An equilibrium with $\varepsilon^* > 0$ exists, if $W(p_0(\overline{\gamma})) > \mu$. An equilibrium with $\varepsilon^* = 0$ exists, if $\mu \geq W(p_0(\overline{\gamma}))$. As $W(p_0(\cdot))$ is an increasing function, $W(p_0(\overline{\gamma})) > \mu$ if and only if the public perceived preference homogeneity is sufficiently high. The following proposition compares the government's expected payoffs in the equilibrium of the private-polling game and in the equilibrium of the benchmark model with the choice of allowing public communication and not.

Proposition 5 (Public communication v.s. private polling) *(1) Provided that $p \leq \frac{1}{2}$, and $W(p_0(0, \overline{\omega})) < \mu$ whenever $\sigma = 1$, $\exists \widehat{\gamma} \in (0, 1]$, such that, for any $\gamma < \widehat{\gamma}$, the government's payoff in the public communication game is strictly higher than its payoff in the private polling game;*

(2) Provided the conditions in Proposition 1 and $\rho_2 > \mu$, when the government's private signal indicates that the citizens' preferences are relatively homogeneous ($\gamma \geq \gamma^{**}$) and it knows that the citizens believe their preferences are heterogeneous ($W(p_0(\overline{\gamma}, \overline{\omega})) \leq \mu$), the government's payoff in the private polling game is strictly higher than its payoff in the public communication game.

Proof *of* **Proposition** 5

Observe that: an equilibrium with $\varepsilon^* > 0$ exists, if $W(p_0(\overline{\gamma}, \overline{\omega}))) > \mu$; an equilibrium with $\varepsilon^* = 0$ exists, if $\mu \geq W(p_0(\overline{\gamma}, \overline{\omega}))$.

Without loss of generality, assume $\sigma > 0$.

(1) The government's expected payoff under public communication ($\varepsilon = 1$) is $-p\gamma \min\{[\sigma\mu + (1-\sigma)W(p_0(1, \overline{\omega}))], W(p_0(1, \overline{\omega}))\} - 2p(1-\gamma)[p_0(0, \overline{\omega})\rho_1]$.

The government's expected payoff in the private polling equilibrium (ε^*) is $-p\gamma \min\{[\sigma\mu + (1-\sigma)W(p_0(q(\varepsilon^*), \overline{\omega}))], W(p_0(q(\varepsilon^*), \overline{\omega}))\} - 2p(1-\gamma)p_0(q(\varepsilon^*), \overline{\omega})\rho_1$.

Under the assumption that $W(p_0(0, \overline{\omega})) < \mu$ whenever $\sigma = 1$, we always have $q(\varepsilon^*) > 0$. Therefore, for sufficiently small γ, i.e., $\gamma \to 0^+$, the government strictly prefers public communication ($\varepsilon = 1$) to private polling equilibrium.

(2) Suppose $\mu \geq W(p_0(\overline{\gamma}, \overline{\omega}))$, we must have $\varepsilon^* = 0$. Thus the government never adjusts policy, and the the citizens' belief about preference homogeneity is $\overline{\gamma} = E(\gamma)$, which is lower than the belief they hold in the public communication game when the communication platform is shut down, that is $E(\gamma | \gamma \geq \gamma^{**})$. Therefore, collective action is less likely to happen and the government gets a strictly higher payoff in the private polling game than in the public communication game. ■

Private channels of horizontal communication

We model the private channel of citizens' horizontal communication in the following way. When public communication is not allowed, with probability h, through certain private channels of communication, citizens can directly learn each other's preference; with probability $1 - h$, their communication is not successful and they do not know each other's preference. Thus h captures the effectiveness of citizens' horizontal interaction outside government regulated communication platforms. Proposition 6 shows that such channels will push the government to become willing to allow public communication.

Proposition 6 (Private channels of horizontal communication) *Assume $W(p_0(1, \overline{\omega})) > \mu > pW(p_0(1, \overline{\omega}))$, $p \leq \frac{1}{2}$, and conditions in Proposition 4 are satisfied,*

(1) in any equilibrium, the government allows public communication if and only if its

private signal indicates that citizens are sufficiently heterogeneous, i.e.,

$$\alpha^* = \begin{cases} 1 & \textit{if } \gamma < \gamma^*(h,\sigma) \\ 0 & \textit{if } \gamma \geq \gamma^*(h,\sigma) \end{cases} ; \tag{A37}$$

(2) the possibility of a successful private communication among citizens forces the government to become more open; namely, $\gamma^*(h,\sigma) > \gamma^{**}(0,\sigma)$, *and* $\gamma^{**}(h,\sigma)$ *is strictly increasing in* h *when* $\gamma^{**}(h,\sigma) < 1$, *provided* $\forall h < 1$, *where* $\gamma^{**}(h,\sigma)$ *is the equilibrium that offers the government the highest payoff among all equilibria; and*

(3) when the likelihood of a successful private communication among citizens is large (h *is sufficiently close to* 1, *including* $h = 1$*), the government always allows public communication.*

Proof *of* **Proposition** 6

First we show that whenever public communication is not allowed, the government never makes effort, namely $e_0^*(\gamma; h, \sigma) = 0$.

When the citizens successfully communicate with each other, a successful reform costs the government at least μ, and the status quo policy costs the government: $p\gamma W(p_0(1,\overline{\omega})) + 2p(1-\gamma)p_0(0,\overline{\omega})\rho_1 < \mu$.

When the citizens do not successfully communicate with each other, a successful reform costs the government at least μ, and the *status quo* policy costs the government: $p\gamma p_0(\widehat{\gamma}_0,\overline{\omega})^2(\rho_2 - 2\rho_1) + 2p\rho_1 p_0(\widehat{\gamma}_0,\overline{\omega})$. When $\rho_2 > 2\rho_1$, $p\gamma p_0(\widehat{\gamma}_0,\overline{\omega})^2(\rho_2 - 2\rho_1) + 2p\rho_1 p_0(\widehat{\gamma}_0,\overline{\omega}) \leq pW(p_0(1,\overline{\omega})) < \mu$. When $\rho_2 \leq 2\rho_1$, $p\gamma p_0(\widehat{\gamma}_0,\overline{\omega})^2(\rho_2 - 2\rho_1) + 2p\rho_1 p_0(\widehat{\gamma}_0,\overline{\omega}) \leq 2p\rho_1 p_0(\widehat{\gamma}_0,\overline{\omega}) < \mu$. As a result, $e_0^*(\gamma; h, \sigma) = 0$.

(1) The government's payoff gain, therefore, is

$G_{diff}(\gamma, \widehat{\gamma}_0^*; h) = (1-h)M(\gamma, p_0(\widehat{\gamma}_0^*,\overline{\omega})) + h[p\gamma W(p_0(1,\overline{\omega})) + 2p(1-\gamma)p_0(0,\overline{\omega})\rho_1] - p\gamma[\sigma\mu + (1-\sigma)W(p_0(1,\overline{\omega}))] - 2p(1-\gamma)p_0(0,\overline{\omega})\rho_1$. which is a linear function in γ, and the intercept is $(1-h)2pp_0(\widehat{\gamma}_0^*,\overline{\omega})\rho_1 + 2hpp_0(0,\overline{\omega})\rho_1 - 2pp_0(0,\overline{\omega})\rho_1 \geq 0$. Thus any equilibrium must follow the cut-point rule (including the degenerate one):

$$\alpha^* = \begin{cases} 1 & \text{if } \gamma < \gamma^*(h,\sigma) \\ 0 & \text{if } \gamma \geq \gamma^*(h,\sigma) \end{cases} . \tag{A38}$$

(2) $G_{diff}(x, E(\gamma|\gamma \geq x); h)|_{x=0} = (1-h)2p\rho_1 p_0(\overline{\gamma}, \overline{\omega}) + h \cdot 2p\rho_1 p_0(0, \overline{\omega}) - 2pp_0(0, \overline{\omega})\rho_1 > 0$, provided $h < 1$.

So we know that $\forall x \in [0, \gamma^*(h,\sigma)), G_{diff}(x, E(\gamma|\gamma \geq x); h) > 0$. We can also verify that

$G_{diff}(\gamma, \widehat{\gamma}_0^*; h) = (1-h)G_{diff}(\gamma, E(\widetilde{\gamma}|\widetilde{\gamma} \geq \gamma); h = 0) + hV_{vertical}(\gamma)$, where $V_{vertical} = \sigma p\gamma[W(p_0(1, \overline{\omega})) - \mu] > 0$.

$\forall \gamma \in [0, \gamma^{**}(0,\sigma)], G_{diff}(\gamma, E(\widetilde{\gamma}|\widetilde{\gamma} \geq \gamma); h = 0) \geq 0, V_{vertical}(\gamma) > 0$. Therefore $G_{diff}(\gamma, \widehat{\gamma}_0^*; h) > 0$, thus we must have $\gamma^*(h,\sigma) > \gamma^{**}(0,\sigma)$ for $h > 0$.

Next we will show that $G_{diff}(\widehat{\gamma}, E(\gamma|\gamma \geq \widehat{\gamma}); h_1) > 0$ implies $G_{diff}(\widehat{\gamma}, E(\gamma|\gamma \geq \widehat{\gamma}); h_2) > 0$ whenever $h_2 > h_1$. Suppose $G_{diff}(\widehat{\gamma}, E(\gamma|\gamma \geq \widehat{\gamma}); h_1) > 0$. If $G_{diff}(\widehat{\gamma}, E(\gamma|\gamma \geq \widehat{\gamma}); h = 0) > 0$, $G_{diff}(\widehat{\gamma}, E(\gamma|\gamma \geq \widehat{\gamma}); h_2) > 0$. If $G_{diff}(\widehat{\gamma}, E(\gamma|\gamma \geq \widehat{\gamma}); h = 0) \leq 0$, $G_{diff}(\widehat{\gamma}, E(\gamma|\gamma \geq \widehat{\gamma}); h_2) = G_{diff}(\widehat{\gamma}, E(\gamma|\gamma \geq \widehat{\gamma}); h_1) - (h_1 - h_2)[G_{diff}(\widehat{\gamma}, E(\gamma|\gamma \geq \widehat{\gamma}); h = 0) - V_{vertical}(\widehat{\gamma})] > 0$. As a result, we always have $\gamma^{**}(h_2,\sigma) > \gamma^{**}(h_1,\sigma)$, provided $\gamma^{**}(h_1,\sigma) < 1$.

(3) When h is sufficiently close to 1, the government's payoff gain is sufficiently close to the value of vertical information $V_{vertical}(\gamma) > 0$, hence it always wants to allow public communication. ■

Supplementary Appendix for *Why Do Authoritarian Regimes Allow Citizens to Voice Opinions Publicly?*

Jidong Chen (Beijing Normal) Yiqing Xu (UCSD)

In this Supplementary Appendix, we offer various extensions of our benchmark model in the paper *Why Do Authoritarian Regimes Allow Citizens to Voice Opinions Publicly?*. We illustrate how our main results are robust to some alternative assumptions. Specifically, we consider the following cases: (1) when the government can use a more sophisticated way of information management; (2) when citizens are *ex ante* asymmetric; (3) when citizens cannot Bayesian-update information. In addition, we also analyze the case when the government does not have more information than the citizens.

More sophisticated information management

As implied by Proposition 2, the government needs to keep the degree of openness to the minimum in order to achieve its maximum payoff. In reality, an authoritarian may use propaganda or internet censorship to manipulate citizens' belief and persuade them that the government is popularly supported by at least a proportion of citizens so as to distort their incentives to join the protest (Stockmann and Gallagher 2011; Shadmehr and Bernhardt 2011; Dimitrov 2014). For example, the Chinese government not only deletes "negative" news online that can potentially raise popular anger (King, Pan and Roberts 2013), but also hires Internet commentators to post favorable comments toward government policies as a way to sway public opinion (Kalathil and Boas 2003).[1] Even in authoritarian legislatures, an incumbent could manipulate voices in a very sophisticated way (such as personnel control) to his own favor(Gandhi 2008; Truex 2013).

[1] The commentators are commonly called "fifty-cent party" by netizens as a satire since they are said to be paid RMB fifty cents for each post.

In Proposition 7 (below), we extend the benchmark model by allowing the government to flexibly manipulate information. This extension not only demonstrates that the benchmark model is useful in capturing the more sophisticated information management by an authoritarian regime, but also illustrates that the logic of censorship and propaganda relies on the key trade-off between coordination effect and the discouragement effect introduced in the benchmark model. When the citizens are likely to think they are both pro-reform, the coordination effect is more likely to take place. Thus, the government may want to change such perception by producing discouraging information to keep the coordination effect at minimum. We show that the government does not fully allow information disclosure when both citizens are dissatisfied with the *status quo* policy, and that when they are of different types, the government truthfully discloses it.

Specifically we consider the following signal-jamming technology:

> *When both citizens claim they are pro-reform, the government truthfully discloses it (as "popular anger") to the public with probability (1-c) and reports "mixed opinions" with probability c; otherwise it always reports "mixed opinions."*

Proposition 7 (A signal-jamming technology) *Consider the above signal-jamming technology. Provided* $W(p_0(1,\overline{\omega})) > \mu > W(p_0(0,\overline{\omega}))$,

(1) the government always has an incentive to censor the information indicating that citizens are homogeneously against it; namely, in any equilibrium $c^* > 0$*; and*

(2) provided $(1-\lambda)L > 1, \gamma_0 \triangleq \frac{1-\lambda L}{(1-2\lambda)L}$*, any* $c^* \in [\frac{1}{1-\sigma}\frac{\gamma_0}{1-\gamma_0}\frac{1-\overline{\gamma}}{\overline{\gamma}}, 1]$ *is an equilibrium.*

Proof *of* **Proposition** 7

Because $W(p_0(1,\overline{\omega})) > \mu$, after disclosing the information *"popular anger," the regime makes effort* $e = 1$, and gets $\sigma(-\mu) + (1-\sigma)[-W(p_0(1,\overline{\omega}))]$.

(1) Suppose an equilibrium exists with $c^* = 0$. Upon the news *"mixed opinions"* and the failure of reform, a pro-reform citizen believes that the other citizen is anti-reform, thus upon two pro-reform citizens, if the government chooses to disclose it as *"mixed opinions"*,

it will get: $\max\{-W(p_0(0,\overline{\omega})), \sigma(-\mu) + (1-\sigma)[-W(p_0(0,\overline{\omega}))]\}$. Because $W(p_0(0,\overline{\omega})) < \mu$, we have:

$$\max\{-W(p_0(0,\overline{\omega})), \sigma(-\mu) + (1-\sigma)[-W(p_0(0,\overline{\omega}))]\}$$
$$= -W(p_0(0,\overline{\omega})) > \sigma(-\mu) + (1-\sigma)[-W(p_0(1,\overline{\omega}))].$$

Thus, there is always a profitable deviation. As a result, $c^* = 0$ cannot be an equilibrium.

(2) Let's focus on the equilibrium where the government always makes an effort to adjust its policy when facing two pro-reform citizens. The government's incentive compatibility constraint is therefore

$$W(p_0(1,\overline{\omega})) \geq \mu, W(p_0(q^*,\overline{\omega})) \geq \mu,$$

where $q^* = \frac{\overline{\gamma}c^*(1-\sigma)}{\overline{\gamma}c^*(1-\sigma)+(1-\overline{\gamma})}$ is the probability with which each pro-reform citizen thinks the other citizen has the same preference upon the failure of reform and the news "*heterogeneous opinions.*"

We also need incentive compatibility constraint on disclosing the information

$$W(p_0(q^*,\overline{\omega})) = W(p_0(1,\overline{\omega})).$$

As long as $q^* \geq \gamma_0$, we get both $W(p_0(q^*,\overline{\omega})) \geq \mu$ and $W(p_0(q^*,\overline{\omega})) = W(p_0(1,\overline{\omega}))$. So we only need to check if $q^* \geq \gamma_0$.

And $q^* \geq \gamma_0$ is equivalent to $c^* \geq \frac{1}{1-\sigma}\frac{\gamma_0}{1-\gamma_0}\frac{1-\overline{\gamma}}{\overline{\gamma}}$. As a result, any $c^* \geq \frac{1}{1-\sigma}\frac{\gamma_0}{1-\gamma_0}\frac{1-\overline{\gamma}}{\overline{\gamma}}$ is an equilibrium. ■

Asymmetric citizens

In our benchmark model, the citizens are *ex ante* the same and have the same prior distribution p. In reality, however, preferences of certain groups of citizens may already be known to the public and the uncertainty is only about other citizens' opinions. In Proposition 8, we investigate such a possibility in a variant version of the model. We show that the basic insight in the benchmark model directly applies to the case with *ex ante* asymmetric citizens, and the government allows public communication if and only if its private signal indicates that the two citizens share different preferences.

We consider the following environment. Citizen 1 is pro-reform type $\overline{\omega}$ for sure, she and the government face the uncertainty about citizen 2's preference. With probability θ, citizen 2 supports reform and with probability $1-\theta$, she strictly prefers the *status quo*. The government directly observes θ whereas citizen 1 only knows that θ is determined according to a cumulative distribution function $G(\theta)$. A higher θ implies a higher likelihood that citizens are both discontent with the status quo policy.[2] Similarly as the main result, Proposition 8 shows that the government allows public communication if and only if its private signal indicates that the two citizens share divergent preferences, i.e., θ is low.

Proposition 8 *Assume k_i is i.i.d. uniform distribution on [0,1] and $\rho_2 \geq 2\rho_1$. $\exists L_0 > 0$ such that $\forall L = \underline{L} \in (0, L_0]$, the government allows public communication if and only if government's private signal indicates that citizen 2 is likely to prefer the status quo policy (i.e., θ is low).*

Proof *of* **Proposition** 8

(1) Suppose $\widehat{\theta}$ is citizen 1's equilibrium perception toward θ.

According to the same logic of Lemma 4's proof, we know that the probability that citizen i joins a protest (upon the *status quo* policy) $P_i(\widehat{\theta})$ is pinned down by

$\widehat{\theta}(1-2\lambda)LP_2(\widehat{\theta}) + \lambda L = P_1(\widehat{\theta})$ and $(1-2\lambda)LP_1(\widehat{\theta}) + \lambda L = P_2(\widehat{\theta})$.

Hence, $P_1(\widehat{\theta}) = \frac{\widehat{\theta}(1-2\lambda)L+\lambda L}{1-\widehat{\theta}[(1-2\lambda)L]^2}$, $P_2(\widehat{\theta}) = (1-2\lambda)LP_1(\widehat{\theta}) + \lambda L$, both of which are smaller than 1 (provided that L is sufficiently small) and strictly increasing in $\widehat{\theta}$.

As L is sufficiently small, the cost that the government suffers when the *status quo* policy remains is also sufficiently small. As a result, the government will never make an effort. Thus the government's payoff gain from allowing public communication is:

$$\theta P_1(\widehat{\theta})P_2(\widehat{\theta})\rho_2 + P_1(\widehat{\theta})(1-\widehat{\theta}P_2(\widehat{\theta}))\rho_1 + \theta P_2(\widehat{\theta})(1-P_1(\widehat{\theta}))\rho_1$$
$$-\theta[P_1(1)P_2(1)\rho_2 + P_1(1)(1-P_2(1))\rho_1 + P_2(1)(1-P_1(1))\rho_1]$$
$$-(1-\theta)P_1(0)\rho_1$$

[2]What θ also captures is preference divergence between the government and citizen 2.

It is a linear function in θ, whose slope is:

$P_1(\widehat{\theta})P_2(\widehat{\theta})\rho_2 - P_1(\widehat{\theta})P_2(\widehat{\theta})\rho_1 - P_2(\widehat{\theta})P_1(\widehat{\theta})\rho_1$

$-[P_1(1)P_2(1)\rho_2 + P_1(1)(1 - P_2(1))\rho_1 + P_2(1)(1 - P_1(1))\rho_1]$

$+P_1(0)\rho_1$

$= -[P_1(1)P_2(1) - P_1(\widehat{\theta})P_2(\widehat{\theta})](\rho_2 - 2\rho_1)$

$-[P_1(1) + P_2(1) - P_1(0)]\rho_1$

$< 0.$

As a result, the government allows public communication if and only if θ is sufficiently low.

■

When citizens cannot Bayesian-update information

In the benchmark model, we implicitly assume that citizens are purely rational and can update information based on the government's action. In Proposition 9, we relax this assumption by assuming that citizens cannot update any information based on the government's action (especially the action regarding allowing public communication). We show that our main result still holds. However, as citizens become less sophisticated, the government will become less open.

Suppose that citizens cannot infer γ from the government's decisions. Without public communication, their belief towards the preference homogeneity is always $\overline{\gamma} = E(\gamma)$ and is unaffected even when they see the government's action. Under the assumptions in Lemma 5, we can show that the government never makes effort when communication is not allowed by the similar logic of its proof. Accordingly, we write down the government's payoff gain function as follows:

$$G_{diff}(\gamma) = M(\gamma, p_0(\overline{\gamma}, \overline{\omega})) - p\gamma \min\{W(p_0(1, \overline{\omega})), \sigma\mu + (1-\sigma)W(p_0(1, \overline{\omega}))\} - 2p(1-\gamma)p_0(0, \overline{\omega})\rho_1, \quad \text{(S 1)}$$

Under the assumptions of Proposition 4 and by the similar logic of its proof, we can show that the payoff gain function is strictly decreasing, therefore in any equilibrium, the government allows public communication if and only if γ is small. Suppose γ^{***} is one of the new cut-point equilibria. When $\gamma = \gamma^{**}$, which is the equilibrium that maximizes the government's payoff when the citizens are purely rational, the government's payoff without public communication in the new equilibrium γ^{***} should be higher than its payoff under the equilibrium γ^{**} without public communication. This is because in the new equilibrium γ^{***} citizens' belief $\overline{\gamma}$ is lower than their belief about preference homogeneity $E(\gamma|\gamma \geq \gamma^{**})$ in the equilibrium γ^{**}. In other words, in the new equilibrium when the government shuts down public communication, it gets a higher degree of social stability than in the old case when citizens can update information. Hence, the government's payoff without public communication when $\gamma = \gamma^{**}$ under the new equilibrium γ^{***} should be strictly higher than its payoff with pubic communication. As a result, we must have $\gamma^{***} < \gamma^{**}$. This implies that the less sophisticated the citizens are, the less open the government becomes. We summarize the result as follows.

Proposition 9 *Provided the conditions in Lemma 5 and Proposition 4, when citizens cannot Bayesian-update information, in any equilibrium, the government allows communication if and only if its private signal indicates that citizens are sufficiently heterogeneous, that is*

$$\alpha^* = \begin{cases} 1 & \text{if } \gamma < \gamma^{***} \\ 0 & \text{if } \gamma \geq \gamma^{***} \end{cases} ; \tag{S 2}$$

Furthermore, any equilibrium γ^{***} induces a lower degree of openness than the welfare-maximizing equilibrium γ^{**} when citizens are purely rational, i.e., $\gamma^{***} < \gamma^{**}$.

When the government does not have more information than the citizens

In this part, we present a variant version of the model where both the government and the citizens observe the realization of γ.[3] Because the citizens know more than in the benchmark model, they can coordinate their behavior based on the realization of the preference homogeneity γ even without public communication.

We make several additional assumptions in this part to simplify the analysis. First, each player's gain from the policy is larger than the upper bound of the collective-action cost, i.e., $L = \underline{L} > 1$. Second, the probability that an individual challenge succeeds is relatively small:

$$\lambda < \min\{\frac{1}{L}, 1 - \frac{1}{L}\}. \quad \text{(S 3)}$$

Furthermore, we assume $\rho_2 \geq 2\rho_1$.

Denote $A \equiv (1-\lambda)L$, which is the net payoff gain from joining a protest provided that the other citizen also participates. Similarly, $B \equiv \lambda L$ is the payoff gain when the other citizen does not participate. Hence, we have $0 < B < 1 < A$.

$$M = p\gamma p_0(\gamma)^2\rho_2 + 2pp_0(\gamma)\rho_1(1-\gamma p_0(\gamma)) \quad \text{(S 4)}$$

is the government's expected loss from citizens' collective action. Without public communication, if the government makes an effort, it gets a cost $\sigma(\mu + M) + (1-\sigma)M$, thus it never makes an effort without public communication and gets a payoff $-M$. For simplicity, we assume $\sigma = 1$.

When public communication is allowed, the government will get $-\mu$ if it sees two pro-reform citizens, get $-p_0(0)\rho_1$ if it sees only one pro-reform citizen and get 0 if it receives no

[3] It is equivalent to assume that both the government and citizens only know the expected preference correlation, $E(\gamma)$.

complaints. The difference in the government's payoff therefore is:

$$M - p\gamma\mu - 2p(1-\gamma)p_0(0)\rho_1. \tag{S 5}$$

Proposition 10 (Citizens' preference homogeneity increases regime openness) *Provided that* $\frac{F^{-1}(y)-B}{y}$ *is concave,*[4] *when* λ *is sufficiently small, in the truth telling equilibrium, the government allows the citizens to speak if and only if the preference correlation of the citizens* γ *is sufficiently high. Specifically,* $\exists\gamma^* \in (0, \gamma_0)$*, where* $\gamma_0 = \frac{1-B}{A-B}$*, such that*

$$\alpha^* = \begin{cases} 1 & \textit{if } \gamma > \gamma^* \\ 0 & \textit{if } \gamma \leq \gamma^* \end{cases}. \tag{S 6}$$

This proposition seems contradictory to the main result in Proposition 4 that the government allows public communication when its privately observed γ is small. Under the assumption that the government does not know more about γ than the citizens, γ serves as a piece of public information that can coordinate citizens' action. The common knowledge on γ allows the citizens to have "implicit" communication even if the government does not allow them to communicate publicly. Namely, since γ is publicly observable, even without public communication, the two citizens can coordinate their actions based on γ through logic of "high order beliefs" naturally imbedded in the game structure. Thus γ here not only serves as the signal of the government, but also serves as the coordination device that horizontally connects the citizens. From this point of view, this comparative statics is consistent with Proposition 6 that investigates the effect of citizens' horizontal communication on the government's openness, as the publicly observable γ captures how citizens are horizontally connected even without the platform of public communication provided by the government. More broadly, if we interpret γ as how sophisticated the citizens can privately connect with each other, this comparative statics is also consistent with Proposition 9 that shows less

[4]The uniform distribution naturally satisfies this condition.

sophistication of citizens will make the government less open. To prove the proposition, we will use the following two lemmas.

Lemma 7 *Suppose $f : I \to f(I)$ is a real-value function, I is an inertial on $\mathbb{R}$. $f(x)$ is concave and strictly increasing, then $f^{-1}(y)$ is convex.*

Proof *of* **Lemma** 7

$\forall y_1, y_2 \in f(I)$ and $y_1 < y_2, \forall \theta \in [0,1]$, suppose $x_1 = f^{-1}(y_1), x_2 = f^{-1}(y_2)$. We need to show $f^{-1}(\theta y_1 + (1-\theta)y_2) \leq \theta f^{-1}(y_1) + (1-\theta)f^{-1}(y_2)$.

It is equivalent to $f[f^{-1}(\theta y_1 + (1-\theta)y_2)] \leq f[\theta f^{-1}(y_1) + (1-\theta)f^{-1}(y_2)]$

i.e., $\theta y_1 + (1-\theta)y_2 \leq f(\theta x_1 + (1-\theta)x_2)$

i.e., $\theta f(x_1) + (1-\theta)f(x_2) \leq f(\theta x_1 + (1-\theta)x_2)$.

It is exactly the concavity of $f(x)$. ■

Lemma 8 *Suppose $\frac{F^{-1}(y)-B}{y}$ is concave, then $p_0(\gamma)$ is convex in γ when $\gamma \leq \gamma_0$.*

Proof *of* **Lemma** 8

By Lemma 4, we know that $T_0(\gamma)$ is strictly increasing in γ when $\gamma \leq \gamma_0$, so $p_0(\gamma) = F(T_0(\gamma))$ is also strictly increasing in γ when $\gamma \leq \gamma_0$.

To apply Lemma 7, we only need to check that $p_0^{-1}(\cdot)$ is concave. When $\gamma \leq \gamma_0$, according to Lemma 1, $T_0(\gamma)$ is determined by $T_0(\gamma) = \gamma F(T_0(\gamma))(A-B) + B$. Thus p_0 is determined by:

$$F^{-1}(p_0) = \gamma p_0(A-B) + B, \tag{S 7}$$

which is equivalent to $\gamma = \frac{F^{-1}(p_0)-B}{(A-B)p_0}$.

Because $\frac{F^{-1}(y)-B}{y}$ is concave by assumption, $p_0^{-1}(\cdot)$ is concave, therefore by Lemma 7, $p_0(\gamma)$ is convex in γ when $\gamma \leq \gamma_0$. ■

Proof *of* **Proposition** 10

First we provide the more generalized conditions, of which the conditions stated in Proposition 10 are a special case.

$\rho_2 \geq 2\rho_1$,

$\rho_2 > \mu \geq \max\{F(B)\rho_1, 2pp_0(0)\rho_1\}$

$\mu > F(B)[2\rho_1 f(F(B)) + (\rho_2 - 2\rho_1)F(B) + 2\rho_1]$

(a) Given Equation (S 4), whenever $\gamma \neq \gamma_0$,

$$\frac{dM}{d\gamma} = \zeta_1 p_0'(\gamma) + \zeta_2 \gamma 2 p_0(\gamma) p_0'(\gamma) + \zeta_2 p_0(\gamma)^2, \tag{S 8}$$

where $\zeta_1 = 2p\rho_1, \zeta_2 = p(\rho_2 - 2\rho_1)$. Therefore $M(\gamma)$ is strictly increasing in γ.

(b) According to Lemma 2, whenever $\gamma \geq \gamma_0$, the coordination effect is 0, therefore the government's payoff gain is always positive, provided $\rho_2 > \mu$, so that in the following we will only consider the government's payoff gain when $\gamma < \gamma_0$.

(c) When $\gamma < \gamma_0$,

$$\frac{d^2M}{d\gamma^2} = \zeta_1 p_0''(\gamma) + 2\zeta_2(p_0(\gamma)p_0''(\gamma) + (p_0'(\gamma))^2)\gamma + 4\zeta_2 p_0(\gamma)p_0'(\gamma). \tag{S 9}$$

So M is strictly increasing and convex in γ given we already know that $p_0''(\gamma)$ is convex by Lemma 8, when $\gamma \leq \gamma_0$. Thus the payoff gain function when $\gamma \leq \gamma_0$ is convex.

(d) payoff gain$|_{\gamma=0} = 2pp_0(0)\rho_1 - 2pp_0(0)\rho_1 = 0$.

(e)

$$\frac{d\text{payoff gain}}{d\gamma}|_{\gamma=0} = \frac{dM}{d\gamma}|_{\gamma=0} - p\mu + 2pp_0(0)\rho_1. \tag{S 10}$$

We have

$$\frac{dM}{d\gamma}|_{\gamma=0} = \zeta_1 p_0'(0) + \zeta_2 p_0(0)^2. \tag{S 11}$$

Recall that p_0 is determined by Equation (S 7), we have:

$$\frac{1}{f(p_0(0))}p_0'(0) = p_0(0)(A - B). \tag{S 12}$$

That is, $p_0'(0) = f(p_0(0))p_0(0)(A - B)$. Put it into Equation (S 11) and Equation (S 11) we have:

$$\frac{d\text{payoff gain}}{d\gamma}|_{\gamma=0} = p\{F(B)[2\rho_1 f(F(B))(A - B) + (\rho_2 - 2\rho_1)F(B) + 2\rho_1] - \mu\}. \tag{S 13}$$

Because $\mu > F(B)[2\rho_1 f(F(B)) + (\rho_2 - 2\rho_1)F(B) + 2\rho_1]$, we have $\frac{d\text{payoffgain}}{d\gamma}|_{\gamma=0} < 0$.

(f) Because $\frac{d\text{payoffgain}}{d\gamma}|_{\gamma=0} < 0$, payoff gain$|_{\gamma=0} = 0$, payoff gain$|_{\gamma\geq\gamma_0} > 0$, the payoff gain has a unique zero point in $(0, \gamma_0)$. ■