WORKING PAPER · NO. 2023-67

The Dictator's Dilemma: A Theory of Propaganda and Repression

A. Arda Gitmez and Konstantin Sonin

AUGUST 2023

Becker Friedman Institute | Political Economics Initiative

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

The Dictator's Dilemma: A Theory of Propaganda and Repression*

A. Arda Gitmez[†]   Konstantin Sonin[‡]

August 2, 2023

Repression and information manipulation are two main tools of any modern authoritarian regime. Our theoretical model demonstrates how repression and propaganda complement each other: when the regime's opponents are facing stricter punishment, the effect of persuasion is stronger, and propaganda is used by the regime more heavily. Similarly, when repression eliminates those citizens who are relatively more skeptical about the regime, the rest can be more heavily influenced. Finally, we show that when citizens self-select into receiving information from individual sources, the dictator cannot do better than resorting to public messaging.

Keywords: Authoritarian regimes, repression, propaganda, Bayesian persuasion.

JEL Classification: D85, L82.

*We are grateful to Georgy Egorov, Scott Gehlbach, Anton Kolotilin, Aleksandr Levkun, Zhaotian Luo, Roger Myerson, Tymofiy Mylovanov, Torsten Persson, Maria Titova, and Andy Zapechelnyuk for many helpful comments.

[†]Bilkent University. E-mail: arda.gitmez@bilkent.edu.tr.

[‡]University of Chicago. E-mail: ksonin@uchicago.edu

Introduction

Repression and propaganda have always been considered the primary tools of autocratic control (Svolik, 2012). In the 20$^{\text{th}}$ century, information manipulation was a central focus in the study of totalitarian dictatorships such as Hitler's Germany, Stalin's Russia, and Mao's China, in which the state tried to control all aspects of subjects' lives (Arendt, 1951; Friedrich and Brzezinski, 1956; Cassinelli, 1960). With the demise of totalitarian dictatorships, propaganda is no longer considered a means of ideological indoctrination, but rather as a tool used by a leader to maintain his reputation as a strong and competent hand (Guriev and Treisman, 2019; Gratton and Lee, 2023; Harrison, 2023). At the same time, repression remains a critical instrument in the dictator's arsenal. In the year following the protests of the summer of 2020, Belarus' Alexander Lukashenko had more than 30,000 people arrested and hundreds given long jail terms, a more than ten-fold increase over the average number of political prisoners during the previous decade (de Vogel, 2022).

The tradition to consider repression and information manipulation as two tools in the autocrat's disposal goes back centuries. In *The Prince*, Niccolo Machiavelli writes on whether it is better to be feared than loved: "The answer is that one would like to be both the one and the other." (Machiavelli, 2019). The early formal theory of nondemocratic government (Wintrobe, 1990, 1998) focused on a simple trade-off: the dictator was deciding how to optimally allocate resources between "repression" and "benefits" to population aimed to make the dictator more popular. In a sweeping recent study of modern authoritarian regimes, Guriev and Treisman (2022) argue that sophisticated propaganda might be a substitute for brutal repression, persuading citizens in the leader's competence. They describe the phenomenon of "spin dictators" who rely primarily on propaganda as a substitute for terror and violence against the political opposition. Citing dozens of examples, Guriev and Treisman (2022) document the change in the relative weights of repression and propaganda from the 20$^{\text{th}}$ century to the 21$^{\text{st}}$. Modern dictators do not arrest hundreds of thousands as Stalin, Hitler, or Mao did or carry public executions as Iraq's Saddam Hussein, Cambodia's Pol Pot, or Equatorial Guinea's Macias Nguema. They primarily rely on selective censorship, digital surveillance, and sophisticated propaganda.[1]

Yet when circumstances change, the dictator's spin is no longer sufficient. In a stark recent example, since the start of Russia's full-scale invasion of Ukraine in February 2022, Putin's regime has doubled down on both tools of authoritarian control. Though information manipulation has

[1] See Gehlbach et al. (2022) for a model that distinguishes between different forms of information manipulation, and Shadmehr (2014) for a model where censorship is used to hide the extent of repression.

long been Putin's weapon of choice against the domestic opposition (Gehlbach, 2010; Rozenas and Stukal, 2019; Guriev and Treisman, 2022), it has dramatically increased after the invasion. The government closed down independent media outlets and implemented draconian punishments for journalists and ordinary people sharing information on social media (Gould-Davies, 2023). Simultaneously, it increased repression – with thousands of people opposing war and the country's course being arrested or forced out of the country, and opposition politicians getting prison terms of 8-25 years (Stoner, 2022; Treisman, 2022).

In this paper, we study how information manipulation naturally complements repression.[2] To this end, we present a theoretical model of an autocratic leader who attempts to minimize the number of citizens who oppose the leader through repression and propaganda. The basic logic behind the complementarity is as follows. At a given level of repression, there is a marginal citizen who is indifferent between opposing the leader and not. Suppose the repression (defined as the punishment for opposing the leader) increases. The previously marginal citizen is now inframarginal, and is unwilling to oppose. When the leader chooses the information manipulation strategy, he is trading off the *intensity* versus *credibility* of propaganda: a higher level of information manipulation contains more intense, but less credible propaganda (i.e., it moves the citizens' beliefs less). When the repression is higher, citizens are more compliant, meaning they would be willing to support even with less credible propaganda. As a result, the leader chooses a more intense propaganda level. It should be noted that, in our basic setup, repressions happen *ex post*, punishing those who oppose the dictator. This does not imply that every non-supporter is punished; the repression may only work in expectation and the results would go through.

Our main result (Proposition 2) reveals the complementarity between propaganda and *ex post* repression. In Section 4, we consider the possibility of *ex ante* repression, targeting those who are inclined to oppose. Ex ante repression occurs before propaganda and contains the purging of citizens who are identified to be skeptics of the regime based on their initial beliefs.[3] Our insights extend to this case: When citizens who are most skeptical about the regime are singled out and repressed, the rest tend to have more favorable opinions. Under the propaganda level tailored for the whole population, they would have been "underpersuaded" from the perspective of the leader. Therefore, the leader increases the intensity of propaganda once the most disloyal elements of the

[2] We use the terms *propaganda* and *information manipulation* interchangeably. In other contexts, *propaganda* might be used more restrictively, referring to a particular type of information manipulation, or, vice versa, more loosely, covering techniques that are not modeled as information manipulation by economic theorists.

[3] Arendt (1951) makes a distinction between the *dictatorial* terror, aimed against the well-identified opponents of the regime, from an all-pervasive *totalitarian* terror of purges, mass executions, and concentration camps. Modern theories of repressions with strategic targeting and selection include Myerson (2015), Tyson (2018), and Dragu and Przeworski (2019).

society are taken out. Thus, the dictator does not face the choice of repression versus propaganda, but rather benefits from them reinforcing each other.

Repressions have been shown to change citizens' behavior. Montagnes and Wolton (2019) and Rozenas (2020) use communist purges in Stalin's Russia and Mao's China to demonstrate this effect. Physical elimination as in Esteban, Morelli and Rohner (2015) changes the composition of the society; other forms of political disenfranchisement might have the same effect as well. In addition to mass executions, Stalin relocated hundreds of thousands from the places where they were a political threat to distant regions of Russia. In most cases, Stalin's mass repression campaigns were organized around broad ethnic or social categories (Gregory, Schröder and Sonin, 2011); in our model, this would correspond to the leader repressing citizens based on imperfect information about their initial beliefs. In the realm of democratic politics, Glaeser and Shleifer (2005) show that the incumbent politician might deliberately choose policies that drive voters who oppose him out of the district. Our theory applies to such situations as well. After repression, the rest of the society will be exposed to more information manipulation.

To model information manipulation, we use the basic model of Bayesian persuasion (Kamenica and Gentzkow, 2011; Bergemann and Morris, 2019) on an audience with heterogeneous priors (Alonso and Câmara, 2016*a*; Laclau and Renou, 2017; Gitmez and Molavi, 2022).[4] Compared to other communication protocols, the model of Bayesian persuasion assumes fuller commitment on behalf of the sender.[5] This makes sense in an applied model: dictators do not edit news in real time. Instead, they pass laws, establish institutions of censorship, and appoint editors to control the flow of information. The choice of an institutional bias or an editor of known ideological preferences corresponds to the choice of the main control parameter in the model.

Yet there are theoretical advantages of using the Bayesian persuasion model as well. First and most importantly, the model allows one to study the *maximum propaganda:* it provides the upper limit on the amount of persuasion that can be done via any information exchange between a sender and a receiver. At the same time, our qualitative results easily translate to other information-exchange models such as such as cheap talk in Crawford and Sobel (1982), verifiable messaging in Milgrom (1981) (see also Titova, 2022), and signaling in Spence (1973). Though the machinery of a model of propaganda and repression based on other communication protocols

[4] Also see Kosterina (2022), who studies Bayesian persuasion with unknown priors and a worst-case-maximizing sender, and Shimoji (2022), who develops a linear programming approach to Bayesian persuasion with heterogeneous priors.

[5] Though our model assumes full commitment, this assumption is not necessary; the results still hold if the probability that the dictator can renege on his commitment is not too high. See Section 6.

would be different, the main intuition is the same. (In Section 6, we discuss why our basic intuition carries through for other communication protocols and the role of other assumptions.)

An important element of our model is that the leader organizes information manipulation as public communication: he establishes an institution which learns the true state of the world and makes a public report. Then, a natural question is whether the leader could do better if it were possible to target different citizens with different messages. In Section 5, we demonstrate that the possibility of private persuasion does not add to the leader's persuasion power. As a technical matter, Proposition 4 justifies our assumption that the leader sticks to the public persuasion mechanism. This result mirrors the main result in Kolotilin et al. (2017).[6] The difference is that while in Kolotilin et al. (2017) the receivers have heterogeneous preferences, in our model they have heterogeneous priors. Substantively, the result explains why in many circumstances authoritarian regimes use blank, one-size-fits-all messaging rather than target groups with different attitudes individually.

The rest of the paper is organized as follows. Section 2 sets up our model. Section 3 studies the main case, when the leader optimally combines repression and persuasion. Section 4 deals with the case when the leader can repress the opposition ex ante. Section 5 deals with targeted propaganda. Finally, Section 6 discusses robustness of our results to alternative technical assumptions.

2 The Setup

There is a sender s (the *leader*) and a continuum of receivers $I = [0,1]$ (the *citizens*, where a generic citizen is denoted with i). A state of the world is denoted by $\omega \in \{0,1\}$. Here, $\omega = 1$ is the state of the world in which the citizens' preferences align with the leader, i.e., the state in which the leader is good), and $\omega = 0$ is the state where there is a misalignment. Citizen i's prior on the leader's being good is

$$\mu_i \equiv \Pr_i\{\omega = 1\}$$

Citizens have *heterogeneous priors*; the distribution of priors among citizens has density $f(\mu)$. Throughout the paper, we maintain the following assumption on the distribution of priors.

[6] Wang (2015) and Chan et al. (2019) also compare public and private persuasion, where the receivers' heterogeneous preferences are known by the sender. Alonso and Câmara (2016*b*) and Inostroza and Pavan (2022) study public persuasion towards heterogeneous receivers with known preferences, whereas Bardhi and Guo (2018), Arieli and Babichenko (2019), Taneva (2019) and Mathevet, Perego and Taneva (2020) study private persuasion towards heterogeneous receivers with known preferences. On private persuasion towards heterogeneous receivers with private preferences, Guo and Shmaya (2019) characterize the optimal information structure, and Heese and Lauermann (2021) study a voting setting.

Assumption 1. Probability density function f has full support on $[0,1]$, is continuously differentiable with $f(0) = f(1) = 0$, and is strictly log-concave, that is, $\frac{\partial^2}{\partial \mu^2} \log f(\mu) < 0$ for all $\mu \in [0,1]$.

The assumption that the density function f is log-concave is a mild one; see Bagnoli and Bergstrom (2005) for a list of commonly used log-concave distributions, which include, among others, uniform, (truncated) normal, and beta (with both parameters ≥ 1) distributions. The leader's prior is $\mu_s \in (0,1)$, and $F(\mu)$ denotes the cdf associated with $f(\mu)$. Both μ_s and f are common knowledge.

Given her information about ω, each citizen $i \in I$ takes an action $a_i \in \{0,1\}$. Here, $a_i = 1$ corresponds to supporting the leader, and $a_i = 0$ corresponds to opposing the leader. Citizen i's payoff is $u(a_i, \omega)$, which is given in Table 1.

	$\omega = 0$	$\omega = 1$
$a_i = 0$	$q - r$	$-r$
$a_i = 1$	0	$1 - q$

Table 1. Citizen Payoffs with Repression level r.

Parameter $q \in (0,1)$ captures the citizens' willingness to protest against the leader, or, equivalently, the aversion to supporting the leader. In addition, $r \in [0, q)$ is the *repression level*, set by the leader before citizens choose their actions. The mechanics of r is simple: repressing the citizens by $r \in [0, q)$ reduces the citizens' payoffs from opposing the leader by r, thereby making it less attractive to oppose.

Interpretation-wise, any act by the leader that increases the cost of opposing the leader, effectively, is *repression* – this includes harassment, intimidation, imprisonment, torture, or execution. The incidence of repression is determined by a citizen's choice of action and, therefore, could be thought of as the expected cost of opposing the leader. This is important as even in totalitarian dictatorships only a small share of population is actually punished. (The Great Terror, 1937-1939, the most intensive years of Stalin's repressions, had about 700,000 people executed and 800,000 more sent to jails and labor camps. These are staggering numbers and yet only about 1% of the USSR population at the time, Conquest, 2008.[7])

Even though the repressions increase the cost of opposing the leader, imposing repressions are costly for the ruler. In particular, if the leader chooses a repression level $r \in [0, q]$, he pays the cost

[7]In line with the logic of our model, Stalin's Great Terror was accompanied by a massive propaganda campaign (Conquest, 2008; Kotkin, 2017). In another example, propaganda has played a critical role ever since the Chinese Communist Party took over in 1949, yet it became even more ferocious during the early years of the Cultural Revolution, 1966-76, which was a combination of elite purges by Mao's faction and mass terror (MacFarquhar and Schoenhals, 2006).

of $\phi \cdot c(r)$. Here, $\phi > 0$ parametrizes the cost of repression. We impose the following assumption on the cost of repression, which ensures that the solution is well-behaved and interior:

Assumption 2. $c : [0, q] \to \mathbb{R}$ is strictly convex, with $c(0) = c'(0) = 0$ and $\lim_{r \to q} c'(r) = \infty$.

The leader's payoff is the total support he receives from the citizens, minus the cost of repression:

$$u_s\left(\{a_i\}_{i \in I}\right) = \int_{i \in I} a_i \cdot di - c(r) \tag{1}$$

The agents do not have any information about ω beyond the information conveyed by the information structure, chosen by the leader. The leader uses a public persuasion mechanism that sends messages from M. That is, the leader commits to an information structure $\{\sigma(\cdot|\omega)\}_{\omega \in \{0,1\}}$ where

$$\sigma(\cdot|\omega) \in \Delta(M) \qquad \text{for all } \omega \in \{0, 1\},$$

and the message drawn, $m \in M$, is publicly observable to each citizen and the leader. We assume that $|M|$ is large enough, so that there is a sufficient number of action recommendations for each receiver. As we will show later, under Assumption 1, the leader uses at most two messages.

The timing of the game is as follows: first, the leader chooses repression level $r \in [0, q]$. Then, the leader commits to an information structure. Finally, the message is drawn, the citizens take their actions, and payoffs are realized.

3 Propaganda and Repression

In this section, we present our main results on complementarity between repression and propaganda. We proceed by backward induction. So, we start with the following question: given a repression level $r \in [0, q)$, what is the leader's optimal information structure? We construct the **value function** of the leader as a function of his own posterior belief $\mu = \Pr_s(\omega = 1|m)$, and then use the concavification approach of Kamenica and Gentzkow (2011).

Suppose the leader's posterior is $\mu \in [0, 1]$. Because the leader and the citizens observe (and update their beliefs based on) the same public message m, as long as we know the priors of the leader and the citizen i, we can back out the posterior of citizen i from the leader's posterior. By Proposition 1 of Alonso and Câmara (2016*a*), citizen i has the following posterior:

$$\mu_i' = \frac{\mu \frac{\mu_i}{\mu_s}}{\mu \frac{\mu_i}{\mu_s} + (1 - \mu) \frac{1 - \mu_i}{1 - \mu_s}} \tag{2}$$

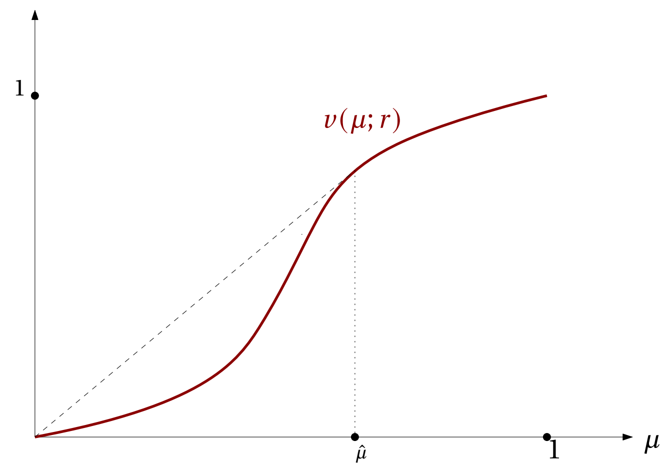

Figure 1. The Leader's Value Function and its Concavification.

Consider the citizen i with posterior μ_i'. By Table 1, she supports the leader if and only if

$$\mu_i' \cdot (1-q) \geq (1-\mu_i') \cdot q - r,$$

which simplifies to $\mu_i' \geq q - r$.[8] Substituting (2) into this inequality and rearranging terms, any citizen i with prior

$$\mu_i \geq \frac{1-\mu}{1-\mu+\mu\frac{1-(q-r)}{q-r}\frac{1-\mu_s}{\mu_s}}$$

supports the leader. Given (1), the leader's value function, as a function of leader's posterior μ, is

$$v(\mu; r) \equiv 1 - F\left(\frac{1-\mu}{1-\mu+\mu\frac{1-(q-r)}{q-r}\frac{1-\mu_s}{\mu_s}}\right). \tag{3}$$

The optimal solution relies on the characterization of the concave closure of $v(\mu; r)$. An inspection of (3) immediately reveals that $v(\mu; r)$ is increasing in μ, with $v(0; r) = 0$ and $v(1; r) = 1$. The next Lemma 1 shows that the value function has a specific shape.

Lemma 1. *$v(\mu; r)$ is strictly S-shaped, i.e., there is some $\tilde{\mu} \in [0,1]$ such that $v(\mu; r)$ is strictly convex for $\mu \in [0, \tilde{\mu}]$ and strictly concave for $\mu \in [\tilde{\mu}, 1]$.*

Figure 1 illustrates the leader's value function in light of Lemma 1. Visual inspection reveals that the leader's optimal information structure invokes two posteriors.[9] This can be achieved by using

[8] Our notation suggests that a receiver takes the sender's favorite action when indifferent. Since f is continuously differentiable by Assumption 1, it has no mass points, so the measure of citizens with priors equal to q is zero. Therefore, this choice of notation is inconsequential for the analysis.

[9] When $\mu_s \geq \hat{\mu}$ in Figure 1, the optimal policy does not reveal any information. In this case, any policy with two messages where $\sigma(m = good|\omega = 0) = \sigma(m = good|\omega = 1)$ is optimal. Among many optimal policies, we choose the one where $\sigma(m = good|\omega = 0) = 1$, so that the "bad" message is never sent and the beliefs following $m = bad$ are free. Consequently, one can choose the posterior following $m = bad$ to be zero, to remain consistent with the discussion.

two messages about the leader's quality: $m \in \{bad, good\}$. Moreover, one of the posteriors in the support is $\mu = 0$, i.e., one of the messages perfectly reveals $\omega = 0$. This can be achieved by setting $\sigma(m = good|\omega = 1) = 1$ in the optimal policy. Therefore, the optimal policy is characterized by a single-dimensional object:

$$\beta = \sigma(m = good|\omega = 0) \in [0, 1].$$

The most natural interpretation of β is the level of *propaganda*: it is the likelihood that the message is "the leader is good" when the leader is, in fact, bad. Since the message $m = bad$ perfectly reveals $\omega = 0$, it stands as an admission that the leader is bad. When the leader is bad, the admission occurs with probability $1 - \beta$. This probability can also be interpreted as the media freedom; for instance, this is the probability that the leader's restraint on media allows a scandal involving the leader to be released. Following this interpretation, β is a measure of leader's control over media. When β is high, the citizens rarely hear any scandals about the leader – instead, they hear positive messages about the leader more and more frequently. This is the sense in which β corresponds to the propaganda level.

Let $\beta^*(r)$ denote the propaganda level chosen by the leader under the optimal policy as a function of r. Let

$$V^*(r) \equiv (\mu_s + (1 - \mu_s)\beta^*(r)) \cdot v\left(\frac{\mu_s}{\mu_s + (1 - \mu_s)\beta^*(r)}; r\right) \tag{4}$$

denote the leader's (subjective) payoff under the optimal information manipulation policy. Equation (4) reveals the main trade-off in the leader's problem. A higher level of $\beta^*(r)$ leads to a higher frequency of good news, and therefore more *intense* propaganda – this is observed through the first appearance of $\beta^*(r)$ in Equation (4). On the other hand, a higher frequency of good news, by definition, means that they are less effective in shifting citizens' beliefs – this is seen through the second appearance of $\beta^*(r)$ in Equation (4). Therefore, more intense propaganda is inevitably less *credible*. The leader chooses $\beta^*(r)$ to optimally resolve the intensity versus credibility trade-off.

3.1 Comparative Statics with Respect to Repression

Before we move on to the leader's choice of repression, we take a brief detour to present a comparative statics result, which will be a critical stepping stone in our analysis. The next Proposition 1 deals with comparative statics with respect to r, the repression level.

Proposition 1. *The propaganda level chosen by the leader, $\beta^*(r)$, is increasing in the repression level r.*

Intuitively, a higher q means that the threshold belief for supporting the leader, $q - r$, is lower: for any posterior belief, citizens are more inclined to support the leader. That is, propaganda is more effective even when it is less credible. This tilts the balance in favor of intensity in the leader's trade-off, and the leader chooses more intense but less credible propaganda.

In a recent working paper, Curello and Sinander (2022) derive a similar comparative statics result regarding the extent of information manipulation. Their comparative statics are based on a partial order on value functions $v(\mu; r)$;[10] in contrast, our comparative statics are on a more primitive parameter of the model r. Proposition 1 of Curello and Sinander (2022) implies that as r changes, the value functions are ranked by their partial order: given $r_1 \geq r_2$, $v(\mu; r_1)$ is *ordinally less convex* than $v(\mu; r_2)$.

3.2 Optimal Combination of Repression and Propaganda

We now proceed with the leader's choice of repression level. Given the analysis in Subsection 3.1, the leader knows that repression level r will be accompanied by propaganda level $\beta^*(r)$ and the payoff $V^*(r)$ from information manipulation. Then, the leader's optimal level of repression, r^*, is the solution to the following problem:

$$r^* \in \arg\max_{r \in [0,q]} \{V^*(r) - \phi \cdot c(r)\}, \tag{5}$$

We are interested in how the leader's chosen repression and propaganda levels are affected by the cost of repression, which is parameterized by ϕ. Our next result shows, naturally, that when repression is cheaper, it is used more intensely.

Lemma 2. *The leader's chosen repression level, r^*, is decreasing in the cost of repression, ϕ.*[11]

Proof. By Assumption 2, the objective function $V^*(q - r) - \phi \cdot c(r)$ is strictly submodular in ϕ and r. The result follows from usual supermodularity arguments (e.g., Theorem 5 of Milgrom and Shannon, 1994). □

The following Proposition 2, our main result, summarizes the above analysis.

Proposition 2. *The leader's optimal propaganda level, β^*, is decreasing in the cost of repression, ϕ.*

Proof. By Lemma 2, r^* is decreasing in ϕ. By Proposition 1, $\beta^*(q - r^*)$ is increasing in r^*, and therefore it is decreasing in ϕ. □

[10] Similarly, Proposition 2 of Kolotilin, Mylovanov and Zapechelnyuk (2022) presents a comparative statics result based on the *shifts* of value functions.

[11] Our treatment does not exclude the possibility that there are multiple optimal levels of r^*. In that case, Lemma 2 and Proposition 2 should be interpreted as decreasing in the sense of *strong set order* (Milgrom and Shannon, 1994).

Proposition 2 shows that, *when the cost of repression is lower, a higher intensity of repression is accompanied by a higher level of propaganda*. Propaganda and repression are complements.

4 *Ex ante* Repressions

One critical feature of repression as defined in Sections 2-3 is that the punishment applies to those who act, in equilibrium, against the dictator. In this section, we consider another type of repression – when citizens are targeted *ex ante*, rather than *ex post*. For instance, the dictator might identify, perhaps with some noise, those who are hostile to the regime and eliminate them. In this case, repression alters the distribution of attitudes towards the regime, rather than the incentives for citizens. The analysis below demonstrates that the effect of *ex ante* repression is qualitatively the same: the more "pro-regime" is the resulting distribution, the higher is the level of propaganda for those who are not eliminated.

Let f_2 be an initial density of priors that satisfies Assumption 1. The leader has access to some informative signal about citizens' priors. As in the model of identity-based Stalinist repression in Gregory, Schröder and Sonin (2011), suppose that there is an institution, e.g., a secret police, that assigns a label $\ell_i \in \{Skeptic, Supporter\}$ to each citizen i with prior μ_i. The secret police's labeling technology is noisy:

$$\Pr\{\ell_i = Supporter \mid \mu_i\} = \rho(\mu_i),$$
$$\Pr\{\ell_i = Skeptic \mid \mu_i\} = 1 - \rho(\mu_i),$$

where $\rho(\mu_i) \in [0,1]$. Suppose $\rho(\mu)$ is continuously differentiable, increasing and log-concave in $\mu \in [0,1]$: this formulation allows for a variety of labeling technologies, such as the linear ($\rho(\mu) = \mu$) or sigmoid $\left(\rho(\mu) = \frac{1}{1+e^{-\mu}}\right)$ functions.

It is straightforward to see that the leader will prefer to purge only those who are labeled $Skeptic$. Suppose for simplicity that the leader purges all citizens labeled $\ell_i = Skeptic$ from the society, and only those with label $\ell_i = Supporter$ remain. Denote the distribution of priors of the remaining citizens by f_1:

$$f_1(\mu) = \frac{\rho(\mu) f_2(\mu)}{\int_0^1 \rho(\tilde{\mu}) f_2(\tilde{\mu}) d\tilde{\mu}}.$$

As long as $\rho(\mu)$ is continuously differentiable and log-concave, $f_1(\mu)$ satisfies Assumption 1.[12]

[12] Most importantly, $\log f_1(\mu) = \log \rho(\mu) + \log f_2(\mu) - \log \int_0^1 \rho(\tilde{\mu}) f_2(\tilde{\mu}) d\tilde{\mu}$, which is strictly concave in μ, and therefore $f_1(\mu)$ is strictly log-concave.

Moreover,

$$\frac{f_1(\mu)}{f_2(\mu)} = \frac{\rho(\mu)}{\int_0^1 \rho(\tilde{\mu}) f_2(\tilde{\mu}) d\tilde{\mu}}$$

which is increasing in μ because $\rho(\mu)$ is increasing. That is, f_1 is larger than f_2 in the *likelihood ratio order*. The following Proposition shows that the leader engages in more propaganda following a purge of skeptics.

Proposition 3. *Consider two distributions of priors, f_1 and f_2, that both satisfy Assumption 1 and where f_1 is larger than f_2 in the likelihood ratio order:*

$$\frac{f_1(\mu)}{f_2(\mu)} \text{ is increasing in } \mu.$$

Then, the propaganda level chosen by the leader under f_1 is larger than the propaganda level chosen under f_2: for any $r \in [0, q]$,

$$\beta_1^*(r) \geq \beta_2^*(r).$$

Proposition 3 shows that *ex ante* repression and propaganda complement each other as well: with skeptics purged, the rest face mor intense propaganda. Intuitively, the leader with distribution f_1 is a "more universally loved" leader than a leader with distribution f_2, because priors under f_1 tend to be larger compared to priors under f_2. In equilibrium, a more popular leader *rides his popularity* and provides less information to citizens, resulting in more propaganda.

5 Targeted Propaganda

In the analysis of *ex ante* repression in Section 4, the crucial precondition for the complementarity of propaganda and repression was reliance on purges being *targeted* towards citizens who have more pessimistic beliefs. A natural question to ask is whether the leader also benefits from targeted information manipulation. In this section, we argue that the answer is no. In particular, allowing for *private persuasion,* i.e., the opportunity to design type-specific propaganda, does not actually expand the menu of the leader's tools. We show that it is never (strictly) optimal for the leader to create different information sources that appeal to different citizens. Therefore, under the optimal policy, the leader sticks with public propaganda.

The first observation, almost trivial from the economic theory standpoint, is that if the dictator has the ability to address each citizen's type individually, then there will be maximum possible persuasion and, ultimately, more support than in the case we are going to focus on. Critically, the dictator would need to keep each citizen unaware of messages that other citizens receive – otherwise, we would be in the realm of public persuasion analyzed above. With individualized messages to each citizens the problem is, of course, that the real dictators cannot possibly know each individual's type. In an authoritarian regime, deep skeptics have all the incentives to conceal their type out of fear of being singled out for repression. In such a situation, the dictator who tries to increase information manipulation by individualizing the messages has to rely on citizens self-selection in receiving these messages. To analyze such a situation, we need to introduce the formalism of *incentive compatibility* (Kolotilin et al., 2017; Bergemann and Morris, 2019).

Consider the setup in Section 2, but suppose that instead of being restricted to use a public persuasion mechanism, the leader can use *private persuasion mechanisms*. Each citizen i receives a message $m_i \in \{0,1\}$ from the news source designed for her. To this end, consider the following setup: each citizen reports her prior $\mu \in [0,1]$, and given the reported prior, the mechanism sends an action recommendation $m \in \{0,1\}$.[13]

Definition 1. A **persuasion mechanism** σ is $\sigma = \{\sigma(\mu,0), \sigma(\mu,1)\}_{\mu\in[0,1]}$ where

$$\sigma(\mu,\omega) = \Pr(m=1|\mu,\omega) \in [0,1] \qquad \text{for } \mu \in [0,1], \omega \in \{0,1\}$$

Consider a receiver i with prior $\mu_i = \mu$. Her (subjective) payoff from reporting a prior $\tilde{\mu}$, taking action $d_0 \in \{0,1\}$ following message $m=0$, and taking action $d_1 \in \{0,1\}$ following $m=1$ is

$$\begin{aligned} U_\sigma(\mu,\tilde{\mu},d_0,d_1) \equiv \mu \cdot \Big(&\sigma(\tilde{\mu},1)\cdot d_1 \cdot (1-q) + \sigma(\tilde{\mu},1)\cdot(1-d_1)\cdot(-r) \\ &+ (1-\sigma(\tilde{\mu},1))\cdot d_0 \cdot (1-q) + (1-\sigma(\tilde{\mu},1))\cdot(1-d_0)\cdot(-r)\Big) \\ &+ (1-\mu)\cdot(\sigma(\tilde{\mu},0)\cdot(1-d_1)\cdot(q-r) + (1-\sigma(\tilde{\mu},0))\cdot(1-d_0)\cdot(q-r)) \end{aligned} \tag{6}$$

The critical element of private persuasion is *incentive compatibility*. The essence is that if a private persuasion mechanism is in place, then each citizen observes and follows her news source, not others' news sources. So if σ is an incentive compatible private persuasion mechanism, citizen i voluntarily chooses to get the messages designed exclusively for her, drawn according to $(\sigma(\mu_i, 0), \sigma(\mu_i,1))$.

The assumption that each citizen follows only one news source requires an explanation of its own. One standard explanation is cognitive constraints; another explanation is a budget constraint in terms of opportunity costs. If citizens were able to follow more than one source, then of course they would follow all possible news sources. In this case, any persuasion mechanism will be equivalent to a public persuasion mechanism; we analyzed this case in Section 3.

[13]The sufficiency of two messages is immediate from a revelation principle argument.

Definition 2. A persuasion mechanism σ is **incentive compatible** if, for all $\mu \in [0,1]$,

$$U_\sigma(\mu,\mu,0,1) \geq U_\sigma(\mu,\tilde{\mu},d_0,d_1) \qquad \text{for all } \tilde{\mu} \in [0,1], d_0, d_1 \in \{0,1\}.$$

Let the subjective payoff of a truthful and obedient citizen (i.e., a citizen who follows the news source designed for her and takes $d_0 = 0$, $d_1 = 1$) with prior μ be $U_\sigma(\mu) \equiv U_\sigma(\mu,\mu,0,1)$. Standard mechanism design arguments yield the following result:

Lemma 3. *If σ is an incentive compatible persuasion mechanism, then $U_\sigma(\mu)$ is convex, with $U_\sigma(0) = q - r$ and $U_\sigma(1) = 1 - q$. Moreover, both $\sigma(\mu,0)$ and $\sigma(\mu,1)$ are increasing in μ, with*

$$U'_\sigma(\mu) = (1-(q-r))\sigma(\mu,1) - (q-r)(1-\sigma(\mu,0)) - r,$$
$$U_\sigma(\mu) - \mu U'_\sigma(\mu) = (q-r)(1-\sigma(\mu,0))$$

for all $\mu \in [0,1]$.

We are going to prove that for any incentive compatible private persuasion mechanism, there exists a direct public mechanism that achieves the same outcome. Formally, a public persuasion mechanism $\pi : \{0,1\} \to \Delta(M)$ is a distribution of messages in each state. Let $M = [0,1]$, and consider direct public mechanisms for the sender: for each $\mu \in M$, $\Pr_s\{\omega = 1|\mu\} = \mu$. That is, under a direct public mechanism, the sender's posterior following a message is the message itself.[14] Given a state ω, let $\pi(\mu|\omega)$ denote the cdf of messages. Consequently, the ex ante distribution of messages (from the sender's perspective) is $\pi_s(\mu) = \mu_s \pi(\mu|\omega = 1) + (1-\mu_s)\pi(\mu|\omega = 0)$. For a mechanism to be direct, it needs to satisfy $\mathbb{E}_s[\mu] = \mu_s$.

We say that a direct public mechanism π **achieves the same outcome** as a private persuasion mechanism σ if it induces the same distribution of actions for each receiver in each state. Note that because of heterogeneous priors, citizens expect different states to realize with different probabilities, so the *ex ante* distributions of actions still differ.

Proposition 4. *For any incentive compatible private persuasion mechanism σ, there exists a direct public mechanism π that achieves the same outcome.*

Proposition 4 implies that the leader can achieve any outcome via a public persuasion mechanism, where he offers the same information structure for each citizen. This justifies our focus on public mechanisms in Section 3, and it implies that the optimal policy we characterized yields a

[14]Because we are working with heterogeneous priors, the posterior following a message is different across receivers. It is, therefore, impossible to define a *direct* public mechanism for everyone. We opt for defining a direct public mechanism for a particular agent, the sender.

higher payoff to the leader than any incentive compatible private persuasion mechanism. Intuitively, this is because the incentive compatibility constraints are extremely binding for the leader, to the extent that a public mechanism (which satisfies incentive compatibility trivially) can yield the same payoff.

This result is closely related to the "impossibility of private persuasion" result in Kolotilin et al. (2017); the difference is that our result is in a setup with heterogeneous priors, rather than with heterogeneous preferences. We cannot rely on the garbling result of Blackwell (1953) as Kolotilin et al. (2017) does, as Blackwell's famed result requires common posteriors following a public message. However, in our proof, we modify the proof of Proposition 2 in the Online Appendix to Kolotilin et al. (2017) to obtain a similar result in our setting.

Substantively, Proposition 4 provides a rationale for why an authoritarian regime might prefer a standardized approach to censorship and mass propaganda. The reality is of course more complex. Vladimir Lenin, the founder of the communist state after the Russian Revolution, stated in 1921: "We need full and truthful information. And the truth should not depend upon whom it has to serve. We can accept only the division between the unofficial information (for the Comintern Executive Committee only) and official information (for everybody)" (Egorov, Guriev and Sonin, 2009).[15] In a famous example of information parsed differently for different types, the 1956 speech by Soviet leader Nikita Khrushchev denouncing Stalin's repressions and cult of personality was distributed, in a written form, to local communist party secretaries, read aloud to ordinary party members at specially arranged meetings, and kept secret from the rest of the population (Taubman, 2003). Critically, Khrushchev relied on the existing designation of citizens' types, with the communist party members being presumably more loyal than citizens at large. Proposition 4 implies that absent such a division – that is, if the leader had to rely on citizens' self-selection – he would not be able to do better than public persuasion.

6 Robustness

In this section, we briefly discuss alternatives to our main assumptions and robustness of our core results. Specifically, we focus on the assumptions of the full commitment on behalf of the sender of information, the binary action space, and heterogeneous priors. In each case, our central result, the complementarity between information manipulation and repression, and main comparative

[15] The Lenin's formula was employed by other communist regimes as well; see Gehlbach and Keefer (2011), King, Pan and Roberts (2013) and Lorentzen (2014) for the analysis of China's information policy.

statics results, are robust to alternative assumptions. Thus, the choice of assumptions for the main model is dictated by our desire to keep the model as tractable as possible and the real-world context that we analyze.

6.1 Partial Commitment to Information Design

Assuming full commitment to information design by the sender (the dictator) simplifies the technical analysis, yet is not necessary. Consider a more general model that does not assume full commitment on the sender's part. Let us introduce probability p, $0 \leq p \leq 1$ and give the dictator an opportunity to manipulate the outcome *ex post* with probability p. In other words, we allow the commitment to the information design to fail with some probability. In the particular case when $p = 0$, this is our main case, a Bayesian persuasion model (Kamenica and Gentzkow, 2011). When $p > 0$, it is a more general model of information manipulation, in which there is no full commitment. The particular case of $p = 1$ corresponds to the Crawford-Sobel "no commitment" communication protocol (Crawford and Sobel, 1982).

To make our point that the general patterns would hold at this model, fix a distribution of beliefs f, the benefit of protesting q and a repression level r. The leader's benefit from inducing a posterior of μ on a representative citizen i who shares the same prior as the leader (i.e., a citizen i with $\mu_i = \mu_s$) is $v(\mu; r)$ defined in (3). Suppose the leader chooses an information strategy with two messages, $m \in \{good, bad\}$, with

$$\sigma(m = good|\omega = 1) = 1$$
$$\sigma(m = good|\omega = 0) = \widehat{\beta}_p(r) \in [0, 1]$$

Given Theorem 1 of Lipnowski, Ravid and Shishkin (2022) and Lemma 1, the sender optimal equilibrium of the game with partial commitment indeed contains such an information strategy. In equilibrium, the message is drawn according to σ, and whenever the leader has the opportunity to manipulate the message, he manipulates it such that the citizens observe $m = 1$. Therefore, the ex ante probability that the leader assigns to the citizens observing $m = 1$ is

$$(1 - p)\left(\mu_s + (1 - \mu_s)\widehat{\beta}_p(r)\right) + p$$

with the complementary probability, citizens observe $m = 0$. The representative citizen's posterior assigned to $\omega = 1$ after observing $m = 1$ is

$$\frac{\mu_s}{(1 - p)\left(\mu_s + (1 - \mu_s)\widehat{\beta}_p(r)\right) + p}$$

and the posterior assigned to $\omega = 1$ after observing $m = 0$ is zero. Overall, the leader's chosen propaganda level solves the following optimization problem.

$$\widehat{\beta}_p(r) = \arg\max_{\beta\in[0,1]} \left((1-p)\left(\mu_s + (1-\mu_s)\beta\right) + p\right) \cdot v\left(\frac{\mu_s}{(1-p)\left(\mu_s + (1-\mu_s)\beta\right) + p}; r\right) \tag{7}$$

Note that the objective function is a modified version of (4), and the two functions coincide when $p = 0$.

Our next result characterizes the propaganda level chosen under partial commitment in relation to the full commitment propaganda level.

Proposition 5. *Given $p \in [0,1]$, the propaganda level $\widehat{\beta}_p(r)$ satisfies:*

$$\widehat{\beta}_p(r) = \max\left\{\frac{\beta^*(r) - p}{1-p}, 0\right\}$$

An implication of Proposition 5 is that $\widehat{\beta}_p(r) \leq \beta^*(r)$ for any p. As in the running example of Lipnowski, Ravid and Shishkin (2022), the leader commits to a more informative structure (less propaganda) to compensate for the fact that beliefs following $m = good$ will be distorted downwards, because the citizens realize that the leader may have manipulated the message. This is also the reason why $\widehat{\beta}_p(r)$ is decreasing in p: lower commitment means the leader has to compensate more.

For our purposes, Proposition 5 reveals that the comparative statics result under full commitment (Propositions 1 and 3) carry over to the case of partial commitment. That is, even under partial commitment, the propaganda level is increasing in repression, and the model maintains the complementarity between propaganda and repression.

6.2 Binary Action Space

Our main motivation for the use of the binary action space comes from political economy considerations. Though the actual country leaders might have various motives, the standard assumption that a leader maximizes their chances to stay in power allows us to consider democratic and authoritarian leaders within the same analytical framework. For an authoritarian leader, this means preventing a coup, a revolution, or a massive protest that would lead to an ouster. (Svolik, 2012 provides statistics that these types of exit cover the overwhelming majority of autocrats' exits; see Dorsch and Maarek, 2018, and Egorov and Sonin, 2022, for updated statistics.) Following the early theories of autocrats' critical moments (Kuran, 1989; Lohmann, 1993), nearly all modern models assume binary choice "support" ("abstain") vs. "no support" ("rebel") for citizens (Persson and

Tabellini, 2009; Bueno de Mesquita, 2010; Shadmehr and Bernhardt, 2011; Little, 2012; Edmond, 2013; Tyson and Smith, 2018; Shadmehr, 2019; Barbera and Jackson, 2020; Egorov and Sonin, 2021).

While the binary action space for citizens is a standard assumption in models of authoritarian control, it grossly simplifies our analysis of persuasion. Specifically, the proof of Proposition 4 that deals with private vs. public persuasion does not readily extend to a larger action space. (See also a discussion of the binary vs. non-binary action space in Kolotilin et al., 2017.) Still, our main results, Propositions 2-3 on complementarity between information manipulation and repression, would hold if we use a larger action space. Whenever propaganda affects the action choice in favor of an action leader prefers, repression will have the complementary effect by enhancing the incentives to take the said action.

6.3 Heterogeneous priors

The main reason to include heterogeneous priors in our model is our willingness to be as realistic as possible. As discussed above, in an authoritarian polity, the change at the top requires collective action by citizens, which justifies the assumption of a binary action space. From a theoretical standpoint, assuming heterogeneous priors with the binary action space is almost equivalent to assuming heterogeneous utilities of citizens. (There is still a subtle difference as with heterogeneous priors one cannot use the garbling result of Blackwell, 1953, as Kolotilin et al., 2017 do, so we rely on an appropriate modification in Section 5.) With heterogeneous utilities, citizens have the same priors about the leader's competence, yet have different individual payoffs when the leader is competent. With heterogeneous priors, the payoff is the same, but the subjective probabilities about the leader's competence differ. Empirical literature on authoritarian transitions points out to heterogeneity of people's beliefs about the leader's quality (e.g., Kuran, 1989). In a recent study on impact of propaganda amid increased repressions on the Russians' attitude towards the Russia-Ukraine war, Alyukov (2022) notes the heterogeneity of attitudes.

A further advantage of having heterogeneous priors instead of heterogeneous preferences is that it allows us to study comparative statics with respect to repression (which, in our main setup in Section 2, affects preferences) without placing too much structure on the model. If we had heterogeneous preferences instead of heterogeneous priors, we would need to specify how repression r translates into the preferences of different citizens differently. With heterogeneous priors, we can maintain a uniform repression level r for every citizen and still obtain heterogeneous effects on citizens' behavior.

7 Conclusion

We offer a model of information manipulation and repression, two main tools in any autocrat's arsenal, considering both public and private persuasion and different types of repression. With a higher level of repression, the leader's marginal supporter is more disposed towards support and, therefore, can be more heavily manipulated. In George Orwell's Oceania, people are forced to use the *newspeak*, a special language designed to limit their ability to articulate anti-government concepts, cannot switch off radio that translates propaganda, and are forced to participate in ideological indoctrination meetings. Yet the ultimate message of *1984* is that it is the physical torture, applied to some, that makes citizens believe what the government wants them to believe.

References

Alonso, Ricardo and Odilon Câmara. 2016*a*. “Bayesian Persuasion with Heterogeneous Priors.” *Journal of Economic Theory* 165:672–706.

Alonso, Ricardo and Odilon Câmara. 2016*b*. “Persuading Voters.” *American Economic Review* 106(11):3590–3605.

Alyukov, Maxim. 2022. “Propaganda, Authoritarianism and Russia’s Invasion of Ukraine.” *Nature Human Behaviour* 6(6):763–765.

Arendt, Hannah. 1951. *The Origins of Totalitarianism.* New York: Harcourt and Brace.

Arieli, Itai and Yakov Babichenko. 2019. “Private Bayesian Persuasion.” *Journal of Economic Theory* 182:185–217.

Bagnoli, Mark and Ted Bergstrom. 2005. “Log-Concave Probability and Its Applications.” *Economic Theory* 26(2):445–469.

Barbera, Salvador and Matthew O. Jackson. 2020. “A Model of Protests, Revolution, and Information.” *Quarterly Journal of Political Science* 15(3):297–335.

Bardhi, Arjada and Yingni Guo. 2018. “Modes of Persuasion Toward Unanimous Consent.” *Theoretical Economics* 13(3):1111–1149.

Bergemann, Dirk and Stephen Morris. 2019. “Information Design: A Unified Perspective.” *Journal of Economic Literature* 57(1):44–95.

Blackwell, David. 1953. “Equivalent Comparisons of Experiments.” *The Annals of Mathematical Statistics* pp. 265–272.

Bueno de Mesquita, Ethan. 2010. “Regime Change and Revolutionary Entrepreneurs.” *American Political Science Review* 104(3):446–466.

Cassinelli, C. W. 1960. “Totalitarianism, Ideology, and Propaganda.” *The Journal of Politics* 22(1):68–95.

Chan, Jimmy, Seher Gupta, Fei Li and Yun Wang. 2019. “Pivotal Persuasion.” *Journal of Economic Theory* 180:178–202.

Conquest, Robert. 2008. *The Great Terror: A Reassessment.* Oxford University Press on Demand.

Crawford, Vincent P. and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50(6):1431–1451.

Curello, Gregorio and Ludvig Sinander. 2022. "The Comparative Statics of Persuasion." *Working Paper* .

de Vogel, Sasha. 2022. "Anti-Opposition Crackdowns and Protest: the Case of Belarus, 2000–2019." *Post-Soviet Affairs* 38(1-2):9–25.

Dorsch, Michael and Paul Maarek. 2018. "Rent Extraction, Revolutionary Threat, and Coups in Non-democracies." *Journal of Comparative Economics* 46(4):1082–1103.

Dragu, Tiberiu and Adam Przeworski. 2019. "Preventive Repression: Two Types of Moral Hazard." *American Political Science Review* 113(1):77–87.

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

Egorov, Georgy and Konstantin Sonin. 2021. "Elections in Non-Democracies." *Economic Journal* 131(636):1682–1716.

Egorov, Georgy and Konstantin Sonin. 2022. "The Political Economics of Non-Democracy." *Journal of Economic Literature* , forthcoming.

Egorov, Georgy, Sergei Guriev and Konstantin Sonin. 2009. "Why Resource-Poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." *American Political Science Review* 103(4):645–668.

Esteban, Joan, Massimo Morelli and Dominic Rohner. 2015. "Strategic Mass Killings." *Journal of Political Economy* 123(5):1087–1132.

Friedrich, Carl J. and Zbigniew K. Brzezinski. 1956. *Totalitarian Dictatorship and Autocracy.* Cambridge, MA: Harvard University Press.

Gehlbach, Scott. 2010. "Reflections on Putin and the Media." *Post-Soviet Affairs* 26(1):77–87.

Gehlbach, Scott and Philip Keefer. 2011. "Investment without Democracy: Ruling-Party Institutionalization and Credible Commitment in Autocracies." *Journal of Comparative Economics* 39(2):123–139.

Gehlbach, Scott, Zhaotian Luo, Anton Shirikov and Dmitriy Vorobyev. 2022. "A Model of Censorship, Propaganda, and Repression." *Working Paper* .

Gitmez, A. Arda and Pooya Molavi. 2022. "Polarization and Media Bias." *Working Paper* .

Glaeser, Edward L. and Andrei Shleifer. 2005. "The Curley Effect: The Economics of Shaping the Electorate." *Journal of Law, Economics, and Organization* 21(1):1–19.

Gould-Davies, Nigel. 2023. "How the War Has Changed Russia." *Survival* 65(2):23–26.

Gratton, Gabriele and Barton E. Lee. 2023. "Liberty, Security, and Accountability: The Rise and Fall of Illiberal Democracies." *Review of Economic Studies* , forthcoming.

Gregory, Paul R., Philipp J. H. Schröder and Konstantin Sonin. 2011. "Rational Dictators and the Killing of Innocents: Data from Stalin's Archives." *Journal of Comparative Economics* 39(1):34–42.

Guo, Yingni and Eran Shmaya. 2019. "The Interval Structure of Optimal Disclosure." *Econometrica* 87(2):653–675.

Guriev, Sergei and Daniel Treisman. 2019. "Informational Autocrats." *Journal of Economic Perspectives* 33(4):100–127.

Guriev, Sergei and Daniel Treisman. 2022. *Spin Dictators: The Changing Face of Tyranny in the 21st Century.* Princeton, New Jersey: Princeton University Press.

Harrison, Mark. 2023. *Secret Leviathan: Secrecy and State Capacity Under Soviet Communism.* Stanford University Press.

Heese, Carl and Stephan Lauermann. 2021. "Persuasion and Information Aggregation in Elections." *Working Paper* .

Inostroza, Nicolas and Alessandro Pavan. 2022. "Adversarial Coordination and Public Information Design." *Working Paper* .

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review* 101(6):2590–2615.

King, Gary, Jennifer Pan and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107(2):326–343.

Kolotilin, Anton, Timofiy Mylovanov and Andriy Zapechelnyuk. 2022. "Censorship as Optimal Persuasion." *Theoretical Economics* 17(2):561–585.

Kolotilin, Anton, Tymofiy Mylovanov, Andriy Zapechelnyuk and Ming Li. 2017. "Persuasion of a Privately Informed Receiver." *Econometrica* 85(6):1949–1964.

Kosterina, Svetlana. 2022. "Persuasion with Unknown Beliefs." *Theoretical Economics* 17(3):1075–1107.

Kotkin, Stephen. 2017. *Stalin, Vol. II: Waiting for Hitler, 1929–1941*. Penguin UK.

Kuran, Timur. 1989. "Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution." *Public Choice* 61(1):41–74.

Laclau, Marie and Ludovic Renou. 2017. "Public Persuasion." *Working Paper* .

Lipnowski, Elliot, Doron Ravid and Denis Shishkin. 2022. "Persuasion via Weak Institutions." *Journal of Political Economy* 130(10):2705–2730.

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

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

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

MacFarquhar, Roderick and Michael Schoenhals. 2006. *Mao's Last Revolution*. Harvard University Press.

Machiavelli, Niccolo. 2019. *Machiavelli: The Prince*. Cambridge Texts in the History of Political Thought 2nd ed. Cambridge: Cambridge University Press.

Mathevet, Laurent, Jacopo Perego and Ina Taneva. 2020. "On Information Design in Games." *Journal of Political Economy* 128(4):1370–1404.

Milgrom, Paul and Chris Shannon. 1994. "Monotone Comparative Statics." *Econometrica* 62(1):157–180.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *The Bell Journal of Economics* pp. 380–391.

Montagnes, Pablo and Stephane Wolton. 2019. "Mass Purges: Top-Down Accountability in Autocracy." *American Political Science Review* 113(4):1045–1059.

Myerson, Roger B. 2015. "Moral Hazard in High Office and the Dynamics of Aristocracy." *Econometrica* 83(6):2083–2126.

Persson, Torsten and Guido Tabellini. 2009. "Democratic Capital: The Nexus of Political and Economic Change." *American Economic Journal: Macroeconomics* 1(2):88–126.

Rozenas, Arturas. 2020. "A Theory of Demographically Targeted Repression." *Journal of Conflict Resolution* 64(7-8):1254–1278.

Rozenas, Arturas and Denis Stukal. 2019. "How Autocrats Manipulate Economic News: Evidence from Russia's State-Controlled Television." *Journal of Politics* 81(3):982–96.

Shadmehr, Mehdi. 2014. "Why Don't Regimes Publicize Repression?" *Working Paper* .

Shadmehr, Mehdi. 2019. "Multiplicity and Uniqueness in Regime Change Games." *The Journal of Politics* 81(1):303–308.

Shadmehr, Mehdi and Dan Bernhardt. 2011. "Collective Action with Uncertain Payoffs: Coordination, Public Signals, and Punishment Dilemmas." *American Political Science Review* 105(4):829–851.

Shaked, Moshe and J. George Shanthikumar. 2007. *Stochastic Orders*. New York, NY: Springer Science & Business Media.

Shimoji, Makoto. 2022. "Bayesian Persuasion in Unlinked Games." *International Journal of Game Theory* 51(3-4):451–481.

Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 87(3):355–374.

Stoner, Kathryn. 2022. "The War in Ukraine: How Putin's War in Ukraine Has Ruined Russia." *Journal of Democracy* 33(3):38–44.

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

Taneva, Ina. 2019. "Information Design." *American Economic Journal: Microeconomics* 11(4):151–85.

Taubman, William. 2003. *Khrushchev: The Man and His Era.* WW Norton & Company.

Titova, Maria. 2022. "Persuasion with Verifiable Information." *Working Paper* .

Treisman, Daniel. 2022. "Putin Unbound: How Repression at Home Presaged Belligerence Abroad." *Foreign Affairs* 101(3):40–53.

Tyson, Scott A. 2018. "The Agency Problem Underlying Repression." *The Journal of Politics* 80(4):1297–1310.

Tyson, Scott and Alastair Smith. 2018. "Dual-Layered Coordination and Political Instability: Repression, Co-optation, and the Role of Information." *Journal of Politics* 80(1):44–58.

Wang, Yun. 2015. "Bayesian Persuasion with Multiple Receivers." *Working Paper* .

Wintrobe, Ronald. 1990. "The Tinpot and the Totalitarian: An Economic Theory of Dictatorship." *American Political Science Review* 84(3):849–872.

Wintrobe, Ronald. 1998. *The Political Economy of Dictatorship.* Cambridge: Cambridge University Press.

Appendix: Proofs

Proof of Lemma 1. As f is continuously differentiable by Assumption 1, $v''(\mu; r) = \frac{\partial^2 v(\mu;r)}{\partial \mu^2}$ exists. We will show that $v''(\mu; r)$ satisfies the *strict single-crossing-from-above property*:

$$\text{If } v''(\mu_1; r) \geq 0 \text{ for some } \mu_1 \in [0,1], \text{ then } v''(\mu_2; r) > 0 \text{ for all } \mu_2 < \mu_1,$$

which implies that there exists some $\tilde{\mu}$ such that $v''(\mu; r) > 0$ for $\mu \leq \tilde{\mu}$ and $v''(\mu; r) < 0$ for $\mu \geq \tilde{\mu}$.

By (3), $v(\mu; r) = 1 - F(g(\mu))$ where

$$g(\mu) \equiv \frac{1-\mu}{1-\mu+\gamma\mu}, \qquad \gamma \equiv \frac{1-(q-r)}{q-r}\frac{1-\mu_s}{\mu_s} > 0$$

Then,

$$v''(\mu; r) = -f'(g(\mu)) \cdot (g'(\mu))^2 - f(g(\mu)) \cdot g''(\mu)$$

Suppose $v''(\mu_1; r) \geq 0$ for some $\mu_1 \in [0,1]$. This implies

$$\frac{f'(g(\mu_1))}{f(g(\mu_1))} \leq -\frac{g''(\mu_1)}{(g'(\mu_1))^2} = \frac{2}{\gamma}\left((1-\gamma) - (1-\gamma)^2\mu_1\right) \tag{8}$$

Take any $\mu_2 < \mu_1$. Because $g(\mu)$ is strictly decreasing in μ, $g(\mu_2) > g(\mu_1)$. Because $f(\mu)$ is strictly log-concave by Assumption 1,

$$\frac{f'(g(\mu_2))}{f(g(\mu_2))} < \frac{f'(g(\mu_1))}{f(g(\mu_1))} \tag{9}$$

Moreover, $\left((1-\gamma) - (1-\gamma)^2\mu\right)$ is decreasing in μ, so equations (8)-(9) yield that

$$\frac{f'(g(\mu_2))}{f(g(\mu_2))} < \frac{2}{\gamma}\left((1-\gamma) - (1-\gamma)^2\mu_2\right) = -\frac{g''(\mu_2)}{(g'(\mu_2))^2}$$

Therefore, $v''(\mu_2; r) > 0$. $\square$

Proof of Proposition 1. Consider $0 < r_1 \leq r_2 < q$. For $k \in \{1,2\}$, the value function under r_k is $v(\mu; r_k)$. Let $co\ v(\mu; r_k)$ denote the concave closure of $v(\mu; r_k)$. By Lemma 1, $v(\mu; r_k)$ is strictly S-shaped, and therefore the set where $v(\mu; r_k)$ coincides with its concave closure takes the following form:

$$\{\mu \in [0,1] : v(\mu; r_k) = co\ v(\mu; r_k)\} = \{0\} \cup [\hat{\mu}_k, 1] \text{ for some } \hat{\mu}_k \in [0,1].$$

By Corollary 2 of Kamenica and Gentzkow (2011), when $\mu_s < \hat{\mu}_k$, the optimal policy generates two posteriors for the leader: $\mu \in \{0, \hat{\mu}_k\}$. When $\mu_s \geq \hat{\mu}_k$, the optimal policy does not reveal any information. Moreover, $\hat{\mu}_k$ satisfies the following properties:

- $v'(\mu; r_k) \cdot \mu < v(\mu; r_k)$ for all $\mu \in (0,1)$ if and only if $\hat{\mu}_k = 0$.
- $v'(\mu; r_k) \cdot \mu > v(\mu; r_k)$ for all $\mu \in (0,1)$ if and only if $\hat{\mu}_k = 1$.
- When $\hat{\mu}_k \in (0,1)$,

$$v'(\hat{\mu}_k; r_k) \cdot \hat{\mu}_k = v(\hat{\mu}_k; r_k). \tag{10}$$

Let

$$y_k(\mu) \equiv v'(\mu; r_k) \cdot \mu - v(\mu; r_k) \qquad \text{for all } \mu \in [0,1] \tag{11}$$

Note that $y_k'(\mu) = v''(\mu; r_k)\mu$. By Lemma 1, then, $y_k'(\mu)$ satisfies single-crossing from above. This means $y_k(\mu)$ is first increasing and then decreasing. Furthermore, $v'(0; r_k) = v'(1; r_k) = 0$ because $f(1) = f(0) = 0$. Then, $y_k(0) = 0$ and $y_k(1) = -1$.

Define the set $\mathcal{U}_k \equiv \{\mu \in [0,1] : \quad y_k(\mu) \geq 0\}$. Based on our discussion so far, we conclude that $\mathcal{U}_k = [0, \hat{\mu}_k]$, for $k \in \{1,2\}$.

Next, we show that $\hat{\mu}_1 \geq \hat{\mu}_2$. As in Lemma 1,

$$y_k(\mu) = f(g_k(\mu))\,(1 - g_k(\mu))\,\frac{1 - g_k(\mu) + \gamma_k g_k(\mu)}{\gamma_k} - (1 - F(g_k(\mu)))\,,$$

where

$$g_k(\mu) \equiv \frac{1-\mu}{1-\mu+\gamma_k\mu}, \qquad \gamma_k \equiv \frac{1-(q-r_k)}{q-r_k}\frac{1-\mu_s}{\mu_s}.$$

Note that $g_k(\mu)$ is strictly decreasing in μ, with $g_k(0) = 1$ and $g_k(1) = 0$. Therefore, the range of $g_k(\mu)$ is $[0,1]$. For any $t \in [0,1]$ define

$$z_k(t) \equiv f(t)(1-t)\frac{1-t+\gamma_k t}{\gamma_k} - (1 - F(t))$$

Then, $y_k(\mu) = z_k(g_k(\mu))$. Because $y_k(\mu)$ is first increasing and then decreasing in μ, and $g_k(\mu)$ is monotonic in μ, we conclude that $z_k(t)$ is first increasing and then decreasing in t. Moreover, $z_k(0) = y_k(1) = -1$, $z_k(1) = y_k(0) = 0$. As a result, the set $\mathcal{V}_k \equiv \{t \in [0,1] : \quad z_k(t) \geq 0\}$ has the form $\mathcal{V}_k = [\hat{t}_k, 1]$, where $\hat{t}_k = g_k(\hat{\mu}_k)$.

Based on our discussion so far, characterizing $\hat{\mu}_k$ is a two-step procedure:

1. First, construct $\mathcal{V}_k$ and find $\hat{t}_k$.
2. Second, calculate $\hat{\mu}_k = g^{-1}(\hat{t}_k) = \frac{1-\hat{t}_k}{1-\hat{t}_k+\gamma_k\hat{t}_k}$.

Recall that $r_1 \le r_2$, which implies $\gamma_1 \le \gamma_2$. Then, $\frac{1-t+\gamma_2 t}{\gamma_2} \le \frac{1-t+\gamma_1 t}{\gamma_1}$ for all $t \in [0,1]$. This, in turn, implies that $z_1(t) \ge z_2(t)$. Thus, for any $t \in \mathcal{V}_2$, $z_2(t) \ge 0$, and, therefore, $z_1(t) \ge 0$ and $t \in \mathcal{V}_1$. That is, $\mathcal{V}_2 \subseteq \mathcal{V}_1$ and $\hat{t}_1 \le \hat{t}_2$. Finally,

$$\hat{\mu}_1 = \frac{1-\hat{t}_1}{1-\hat{t}_1+\gamma_1\hat{t}_1} \ge \frac{1-\hat{t}_1}{1-\hat{t}_1+\gamma_2\hat{t}_1} \ge \frac{1-\hat{t}_2}{1-\hat{t}_2+\gamma_2\hat{t}_2} = \hat{\mu}_2$$

where the first inequality follows from $\gamma_1 \le \gamma_2$ and the second inequality follows from $\hat{t}_1 \le \hat{t}_2$.

To conclude the proof, consider three cases. If $\mu_s \ge \hat{\mu}_1$, the optimal policy does not reveal any information in either case. Given that we already set $\sigma(m = good|\omega = 1) = 1$, the optimal policy includes $\sigma(m = good|\omega = 0) = 1$. Therefore, $\beta^*(r_1) = \beta^*(r_2) = 1$. If $\hat{\mu}_1 > \mu_s \ge \hat{\mu}_2$, the optimal policy under $v(\mu; r_2)$ does not reveal any information. In this case, $\beta^*(r_2) = 1$ and $\beta^*(r_1) < 1$. Finally, if $\mu_s > \hat{\mu}_1$, the optimal propaganda levels of propaganda $\beta^*(r_1)$ and $\beta^*(r_2)$ satisfy

$$\frac{\mu_s}{\mu_s + (1-\mu_s)\beta^*(r_k)} = \hat{\mu}_k, \qquad \text{for } k \in \{1,2\}.$$

Then, $\hat{\mu}_1 \ge \hat{\mu}_2$ implies $\beta^*(r_1) \le \beta^*(r_2)$. □

Proof of Proposition 3. For $k \in \{1,2\}$, let $v_k(\mu; r)$ denote the value function under $f_k(\mu)$. Denoting $y_k(\mu) \equiv v_k'(\mu; r) \cdot \mu - v_k(\mu; r)$ as in the proof of Proposition 1, we show show that $\hat{\mu}_1 \le \hat{\mu}_2$.

Consider some $\mu \in [0, \hat{\mu}_1]$, that is, $y_1(\mu) \ge 0$. Then, $v_1'(\mu; r) \cdot \mu \ge v_1(\mu; r)$. By (3), it implies that

$$\frac{f_1\left(\frac{1-\mu}{1-\mu+\mu\frac{1-(q-r)}{q-r}\frac{1-\mu_s}{\mu_s}}\right)}{1-F_1\left(\frac{1-\mu}{1-\mu+\mu\frac{1-(q-r)}{q-r}\frac{1-\mu_s}{\mu_s}}\right)} \frac{\mu\frac{1-(q-r)}{q-r}\frac{1-\mu_s}{\mu_s}}{\left(1-\mu+\mu\frac{1-(q-r)}{q-r}\frac{1-\mu_s}{\mu_s}\right)^2} \ge 1$$

Because the likelihood ratio order implies the hazard rate order (see, e.g., Theorem 1.C.1 of Shaked and Shanthikumar, 2007), $\frac{f_2(x)}{1-F_2(x)} \ge \frac{f_1(x)}{1-F_1(x)}$ for all $x \in [0,1]$, we have

$$\frac{v_2'(\mu; r)}{v_2(\mu; r)}\mu \ge 1,$$

which is equivalent to $y_2(\mu) \ge 0$. Therefore, $\mu \in [0, \hat{\mu}_2]$ and $\hat{\mu}_2 \ge \hat{\mu}_1$. Replicating the argument in the Proof of Proposition 1 yields $\beta_2^*(r) \le \beta_1^*(r)$. □

Proof of Proposition 4. Consider the public persuasion mechanism defined as

$$\pi(\mu|\omega) = 1 - \sigma(g(\mu), \omega) \qquad \text{for all } \mu \in [0,1], \omega \in \{0,1\}$$

where again

$$g(\mu) \equiv \frac{1-\mu}{1-\mu+\gamma\mu}, \qquad \gamma \equiv \frac{1-(q-r)}{q-r}\frac{1-\mu_s}{\mu_s}.$$

Note that $g(\mu)$ is strictly decreasing in μ, with $g(0) = 1$ and $g(1) = 0$. Moreover, $g(g(\mu)) = \mu$ for all $\mu \in [0,1]$.

Take μ, $\mu' \in [0,1]$ with $\mu \geq \mu'$. Since g is strictly decreasing, $g(\mu) \leq g(\mu')$. By Lemma 3, $\sigma(\mu, \omega)$ is increasing in μ, which implies: $\sigma(g(\mu), \omega) \leq \sigma(g(\mu'), \omega)$. Therefore, $\pi(\mu|\omega) \geq \pi(\mu'|\omega)$. We conclude that $\pi(\mu|\omega)$ is increasing in μ and it is a cdf.

Next, we show that $\mathbb{E}_s[\mu] = \mu_s$, so that this is indeed a direct mechanism. Substituting $U'_\sigma(g(\mu))$ and $U_\sigma(g(\mu)) - g(\mu)U'_\sigma(g(\mu))$ by Lemma 3, we obtain

$$\begin{aligned}
\pi_s(\mu) &= \mu_s \pi(\mu|\omega = 1) + (1-\mu_s)\pi(\mu|\omega = 0) \\
&= \mu_s(1-\sigma(g(\mu), 1)) + (1-\mu_s)(1-\sigma(g(\mu), 0)) \\
&= \mu_s - \frac{\mu_s}{1-(q-r)}\left[(1-\gamma)U_\sigma(g(\mu)) + (1-g(\mu)+\gamma g(\mu))\, U'_\sigma(g(\mu))\right]
\end{aligned}$$

Multiplying the second term by $\frac{-\gamma}{(1-g(\mu)+\gamma g(\mu))^2} g'(\mu) = 1$ and re-arranging terms yields

$$\pi_s(\mu) = \mu_s \frac{1-q}{1-(q-r)} + \frac{1-\mu_s}{q-r}\left[\frac{(1-\gamma)}{(1-g(\mu)+\gamma g(\mu))^2} U_\sigma(g(\mu)) + \frac{1}{1-g(\mu)+\gamma g(\mu)} U'_\sigma(g(\mu))\right] g'(\mu).$$

Taking the integral and changing variables with $v = g(\mu)$ gives

$$\int_0^1 \pi_s(\mu)d\mu = \mu_s \frac{1-q}{1-(q-r)} + \frac{1-\mu_s}{q-r}\int_{v=g(0)}^{v=g(1)}\left[\frac{1-\gamma}{(1-v+\gamma v)^2} U_\sigma(v) + \frac{1}{1-v+\gamma v} U'_\sigma(v)\right] dv.$$

Since $g(0) = 1$ and $g(1) = 0$, then

$$\begin{aligned}
\int_0^1 \pi_s(\mu)d\mu &= \mu_s \frac{1-q}{1-(q-r)} - \frac{1-\mu_s}{q-r}\int_0^1 \left(\frac{d}{d\mu}\frac{U_\sigma(\mu)}{1-v+\gamma v}\right) dv \\
&= \mu_s \frac{1-q}{1-(q-r)} - \frac{1-\mu_s}{q-r}\left(\frac{U_\sigma(1)}{\gamma} - U_\sigma(0)\right)
\end{aligned}$$

By Lemma 3, $U_\sigma(0) = q - r$ and $U_\sigma(1) = 1 - q$, so $\int_0^1 \pi_s(\mu)d\mu = 1 - \mu_s$. Therefore,

$$\mathbb{E}_s[\mu] = \mu_s,$$

and we conclude that the public persuasion mechanism π is direct.

Finally, we show that π induces the same distribution of actions as σ for each receiver in each state. This is true because for each receiver i with prior μ_i and each state ω, the probability of taking action $a_i = 1$ is

$$\begin{aligned}
\Pr\{a_i = \omega\} &= \Pr\left\{\frac{\mu\frac{\mu_i}{\mu_s}}{\mu\frac{\mu_i}{\mu_s} + (1-\mu)\frac{1-\mu_i}{1-\mu_s}} \geq q - r \middle| \omega\right\} \\
&= \Pr\left\{\mu \geq \frac{1-\mu_i}{1-\mu_i+\gamma\mu_i} \middle| \omega\right\}
\end{aligned}$$

$$= \Pr\{\mu \geq g(\mu_i)|\omega\}$$

$$= 1 - \pi(g(\mu_i)|\omega).$$

Since $\pi(\mu|\omega) = 1 - \sigma(g(\mu), \omega)$, $\Pr\{a_i = \omega\} = \sigma(g(g(\mu_i)), \omega)$. Moreover, since $g(g(\mu)) = 1$ for each $\mu \in [0,1]$, $\Pr\{a_i = \omega\} = \sigma(\mu_i, \omega)$, and the result follows. □

Proof of Proposition 5. Consider the optimization problem in (7) with a change of variables, where:

$$\mu_p \equiv \frac{\mu_s}{(1-p)\left(\mu_s + (1-\mu_s)\beta\right) + p} \tag{12}$$

Note that μ_p is strictly decreasing in β, with $\mu_p = \mu_s$ when $\beta = 1$ and $\mu_p = \frac{\mu_s}{(1-p)\mu_s + p}$ when $\beta = 0$. Therefore, $\mu_p \in [\mu_s, \frac{\mu_s}{(1-p)\mu_s+p}]$. With the change of variables, the optimization problem is

$$\max_{\mu_p \in [\mu_s, \frac{\mu_s}{(1-p)\mu_s+p}]} \frac{\mu_s}{\mu_p} v(\mu_p; r) \tag{13}$$

with a maximizer μ_p^* such that, by (12),

$$\mu_p^* = \frac{\mu_s}{(1-p)\left(\mu_s + (1-\mu_s)\widehat{\beta}_p(r)\right) + p} \tag{14}$$

Let $h(\mu_p) \equiv \frac{\mu_s}{\mu_p} v(\mu_p; r)$ denote the objective function. Then,

$$h'(\mu_p) = -\frac{\mu_s}{(\mu_p)^2} v(\mu_p; r) + \frac{\mu_s}{\mu_p} v'(\mu_p; r)$$
$$= \frac{\mu_s}{\mu_p}\left(-\frac{v(\mu_p; r)}{\mu_p} + v'(\mu_p; r)\right)$$

Therefore, $h'(\mu_p) \geq 0$ if and only if $v'(\mu_p; r)\mu_p - v(\mu_p; r) \geq 0$. As discussed in the proof of Proposition 1, $v'(\mu_p; r)\mu_p - v(\mu_p; r)$ satisfies single-crossing from above, which means $h(\mu_p)$ is quasiconcave with its peak at some $\hat{\mu} \in [0,1]$.

To proceed with the proof, we consider three exhaustive cases:

1. If $\hat{\mu} < \mu_s$, in the relevant range of μ_p in (13), $h(\mu_p)$ is decreasing. Therefore, $\mu_p^* = \mu_s$. By (14), $\widehat{\beta}_p(r) = 1$.

   At the end of the proof of Proposition 1, we showed that if $\hat{\mu} < \mu_s$, the optimal full commitment policy involves $\beta^*(r) = 1$. Therefore, in this case, $\widehat{\beta}_p(r) = 1 = \frac{1-p}{1-p} = \frac{\beta^*(r)-p}{1-p} \geq 0$.

2. If $\hat{\mu} \in [\mu_s, \frac{\mu_s}{(1-p)\mu_s+p}]$, the peak of $h(\mu_p)$ remains in the relevant range of μ_p. Therefore, $\mu_p^* = \hat{\mu}$. By (14),

$$\frac{\mu_s}{(1-p)\left(\mu_s + (1-\mu_s)\widehat{\beta}_p(r)\right) + p} = \hat{\mu} \tag{15}$$

At the end of the proof of Proposition 1, we showed that if $\mu_s \geq \hat{\mu}$, then:

$$\frac{\mu_s}{\mu_s + (1-\mu_s)\beta^*(r)} = \hat{\mu} \tag{16}$$

Combining (15) and (16), we have:

$$\frac{\mu_s}{(1-p)\left(\mu_s + (1-\mu_s)\widehat{\beta}_p(r)\right) + p} = \frac{\mu_s}{\mu_s + (1-\mu_s)\beta^*(r)} \implies \widehat{\beta}_p(r) = \frac{\beta^*(r) - p}{1-p}$$

where, by the fact that $\hat{\mu} \leq \frac{\mu_s}{(1-p)\mu_s + p}$ and by (16), we have: $\beta^*(r) - p \geq 0$.

3. If $\hat{\mu} > \frac{\mu_s}{(1-p)\mu_s+p}$, in the relevant range of μ_p in (13), $h(\mu_p)$ is increasing. Therefore, $\mu_p^* = \frac{\mu_s}{(1-p)\mu_s+p}$. By (14), $\widehat{\beta}_p(r) = 0$.

   Since $\mu_s > \frac{\mu_s}{(1-p)\mu_s+p} \geq \hat{\mu}$, in this case, $\beta^*(r)$ is defined by (16). But then, the fact that $\hat{\mu} > \frac{\mu_s}{(1-p)\mu_s+p}$, combined with (16), implies: $\beta^*(r) < p$. Therefore, in this case, $\widehat{\beta}_p(r) = 0 \geq \frac{\beta^*(r)-p}{1-p}$.

In any case, we have shown that $\widehat{\beta}_p(r) = \max\left\{\frac{\beta^*(r)-p}{1-p}, 0\right\}$. □