

NBER WORKING PAPER SERIES

INCUMBENCY ADVANTAGE IN NON-DEMOCRACIES

Georgy Egorov
Konstantin Sonin

Working Paper 20519
<http://www.nber.org/papers/w20519>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2014

We thank Daron Acemoglu, Alberto Alesina, Alex Debs, Scott Gehlbach, Mike Golosov, John Londregan, Santiago Oliveros, Carlo Prato, Alvaro Sandroni, Mehdi Shadmehr, participants of Harvard-MIT Positive Political Theory seminar, MOVE Workshop on Selection of Politicians in Barcelona and MPSA 2011, NASM of the Econometric Society 2011, EEA/ESEM 2012, BCEP conference on institutions and political conflict, and AEA and IEA 2014 conferences for helpful comments and suggestions, and Egor Abramov and Pavel Andreyanov for excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Georgy Egorov and Konstantin Sonin. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Incumbency Advantage in Non-Democracies
Georgy Egorov and Konstantin Sonin
NBER Working Paper No. 20519
September 2014
JEL No. D72,D82,H00

ABSTRACT

In elections that take place in a less-than-perfect democracy, incumbency advantages are different from those in mature democracies. The incumbent can prevent credible challengers from running, organize vote fraud, or even physically eliminate his main opponents. At the same time, formally winning the election does not guarantee staying in power. We present a unified model of elections and mass protests where the purpose of competitive elections is to reveal information about the relative popularity of the incumbent and the opposition. Citizens are heterogenous in their attitudes toward the dictator, and these individual preferences serve as private signals about the aggregate distribution of preferences; this ensures a unique equilibrium for any information the incumbent may reveal. We show that the most competent or popular dictators run in competitive elections, mediocre ones prevent credible opponents from running or cancel elections, and the least competent ones use outright repressions. A strong opposition makes competitive elections more likely but also increases the probability of repression. A totalitarian regime, where repression is cheaper, will have more repression, but even in the absence of repression, competitive elections will be rarer. A crueler, say, military, regime, where protesting is costly, makes repression less likely and, surprisingly, competitive elections more likely.

Georgy Egorov
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208
and NBER
g-egorov@kellogg.northwestern.edu

Konstantin Sonin
Higher School of Economics
20 Myasnitskaya ul.
Moscow, 101000, Russia
ksonin@hse.ru

1 Introduction

In 1987, 1993, and 1999, Hosni Mubarak, the president of Egypt, held ‘elections’ in which no other candidate was allowed to run. In 2005, Mubarak agreed to allow other candidates to be on the ballot, but the election was still marred with an uneven playing field, extensive fraud, and jailing of his opponents. In early 2011, Mubarak faced mass protests, was abandoned by his key supporters, and ended up under house arrest on corruption charges. In Europe, Alexander Lukashenko, the president of Belarus, held elections in 2001 and 2006 which were not recognized by the international community as free and fair. In 2010, he allowed multiple opposition candidates to be on the ballot yet had most of them jailed on the election night. Following the events in Tunisia and Egypt in late 2010-early 2011, Nursultan Nazarbayev, Kazakhstan’s ruler since independence in 1991, suddenly announced that he will stand for re-election even though his term from the previous election in 2005 would expire in 2012. On April 3, 2011, with no opposition candidates on the ballot, Nazarbayev was reported the winner with 95.5 percent of the vote.

The above examples allow us to make several important observations. In a non-democratic regime, having an election or not, allowing a serious opponent to take part or forbidding opposition leaders from running, and choosing the extent to which the population is informed about the outcome are all parts of the incumbent’s strategy set. At the same time, unlike in democratic countries, winning an election does not guarantee staying in power: the dictator may still be vulnerable to mass protests, as the example of Hosni Mubarak or a more recent case of Viktor Yanukovich in Ukraine in 2014 demonstrate. Not every non-democratic election is just a staged spectacle, however. In Azerbaijan in 2003, the opposition leader Isa Qambar, a former speaker of the parliament and acting president, was allowed to run against Ilham Aliyev, heir apparent and son of the outgoing president Heydar Aliyev, and finished with 15 percent of the vote to Aliyev’s 75 percent.¹ In Yugoslavia in 2000, the incumbent Slobodan Milosevic finished second in the first round with 39 percent, losing to the opposition leader Vojislav Kostunica with 49 percent. He resigned following the mass protests before the scheduled run-off.

In this paper, we study an environment where the incumbent is able to use strategies which are not available in democratic regimes, such as eliminating opposition leaders physically, imprisoning them, preventing them from running, and using vote fraud to get the desired result. In our model, the dictator faces possible mass protests and tries to minimize their scope so as to maximize his chances of staying in power. He chooses his actions, such as allowing the opposition leader

¹The election was more competitive than the final tally may suggest; see Cornell and Ismailzade (2004) for details.

to participate in fair or semi-fair elections, or, on the contrary, killing him, with this goal in mind.² We thus contribute to the literature on nondemocratic politics by modeling elections as means of communication with citizens while facing potential anti-government protests. In doing so, we distinguish between competitive elections, which, even if not perfectly fair and fraudless, are informative about the relative popularity of the dictator and the oppositions, and uncontested elections, the results of which are predetermined and uninformative.

Our baseline model is simple: the dictator faces a threat of mass protests against him and his goal is to make these protests as limited and contained as possible. There is an opposition leader, who is the single most credible alternative to the incumbent. The incumbent is able to eliminate this challenger; we model this as an effective yet costly option because of the need to buy military support or of possible international rebuke. Alternatively, he can challenge him in a free and fair election (we introduce the possibility of fraud as an extension in Section 4) or prevent him from running.

Individual citizens have heterogeneous preferences for the challenger vis-a-vis the incumbent and are more likely to protest if they hate the dictator more or if they are more sure of the chance of success. Importantly, these individual preferences also serve as private signals about the aggregate distribution of other citizens' preferences, whereas the dictator's actions whether to repress, whether to have elections, as well as the vote tally, are all public signals about this distribution. The dictator has an informational advantage (e.g., he is able to observe economic performance and deduce the aggregate), and while he cannot directly manipulate citizens' attitude towards him (at least in the short run), he is able to manipulate citizens' beliefs about his popularity through public signals. Our model of protests, where private signals come from individuals' preferences and are not erased by revelation of public information, differentiates our model from alternatives that use global games. In particular, we get a unique equilibrium for any public signal that the dictator may make available.

We show that the most popular dictators will choose competitive elections, dictators with intermediate popularity will do nothing (this may take the form of no elections, or phoney elections, depending on the historic context), and the least popular dictators will repress the opposition. These thresholds for competitive election and for repression depend on the parameters of the model in simple and intuitive ways. For example, consider a totalitarian regime, where repression is particularly easy, or an oil-rich regime, where staying in power is very attractive because of

²For models of incumbency advantage in democracies, see, e.g., Ashworth (2005) and Ashworth and E. Bueno de Mesquita (2008).

high rents from office. Not surprisingly, these regimes will repress more often because it is cheaper relative to the benefits. Our model predicts that even when repression is not used, competitive elections are less likely, and this is precisely because repression is cheaper; if such regime does not use repression, citizens infer that the incumbent is sure of his popularity, which makes them more skeptical of the chances to overthrow him and limits the scope of protests in the absence of further information. Thus, it is not surprising that the incumbent would not want to provide this information by participating in a competitive election.

Alternatively, consider a regime where protesting on the streets is particularly costly, say a military regime where the army is an active participant of protests and riots. These regimes will, not surprisingly, use less repression against the opposition, because protests are expected to be limited. Competitive elections will be more likely, however: with fewer protesters, the marginal protester is more averse to the regime, and it takes better public information to dissuade him from protesting. Thus, our model predicts that totalitarian regimes and military regimes will be polar opposites in terms of using repression and competitive elections, even though we do not model them explicitly as such.

Extending our model, we demonstrate that a challenger may want to pass on an opportunity to participate in a competitive election since such an election is organized by the incumbent when it benefits him more than the challenger. This gives rise to further interesting comparative statics results. Suppose, for example, that a country had democratic traditions before falling into dictatorship, so the challenger is expected to run in an election rather than ignore it and try to come to power through mass protests. He could similarly face competition for the position of the opposition's leader. He will then be more likely to run, but the incumbent will also be more likely to organize competitive elections because failure to do so will be interpreted as a signal of his weakness. Thus, unambiguously, countries which previously had a strong democratic tradition are more likely to keep having competitive elections when a dictator is in control. But they are also more likely to have repression of opposition leaders: as we show, if the dictator fails to organize a competitive election, he will be perceived as even weaker, and thus more dictators will resort to repression if the country had democracy before.

There is a substantial literature in political science that strives to explain elections held by autocrats (see, e.g., Miller, 2010, for a recent survey). The two most common explanations are: first, information-gathering on national and local issues;³ second, sending a signal about the strength of

³In Londregan and Vindigni (2006), election outcome is a signal that helps to arrange a stable agreement with the

the incumbent to deter entry (Geddes, 2009).⁴ Przeworski (2009) describes ‘plebiscitary elections,’ which the regime uses to demonstrate that it can “force everyone to appear in a particular place on a particular day and perform the act of throwing a piece of paper into a designated box.” (See Geddes, 2006; both Blaydes, 2006, and Magaloni, 2006 find evidence of this motive in their respective countries of study.) Howard and Roessler (2006) provide a classification of regimes based on their decisions to have elections and to allow opposition to participate. Fearon (2011) allows democratically elected leaders to ban further elections and analyzes when democracy is self-enforcing. Other models, e.g., Shadmehr and Bernhardt (2013) do not consider elections explicitly but rather emphasize the dictator’s ability to control availability of information, e.g., through media censorship (see also Egorov, Guriev, and Sonin, 2009, for empirical evidence on media freedom in nondemocracies). For a broader treatment of strategies of political survival see, e.g., B. Bueno de Mesquita et al. (2003), Myerson (2008), and Svoboda (2008).

Mass protests, one of the main causes of loss of power by dictators, are studied in a number of recent contributions. Acemoglu and Robinson (2001, 2006) assume that from time to time potential dissidents (‘the poor’) are able to overcome collective action problem and coordinate on protests, while Acemoglu, Robinson, and Verdier (2004) model protests as a coordination game between two groups of citizens modeled as unitary actors. Shadmehr and Bernhardt (2011) model protests as a coordination game in which participants are uncertain about the relative payoffs of a successful revolution and the status-quo. Limiting the amount of information available to citizens might increase the likelihood of protests as each individual citizen is forced to rely on others’ information to a larger extent. In Persson and Tabellini (2009), citizens’ decisions to defend democracy by protesting against a potential coup are modeled using a global game.

Several papers study dynamics of mass protest and the role of information. In Fearon (2001), citizens protest each period, and the number of current protesters contains useful information for future protesters. In E. Bueno de Mesquita (2010), protests are modeled as a coordination game dictators’ opponents. In Gandhi and Przeworski (2006), Magaloni (2010), Boix and Svoboda (2011), and Miller (2010) power-sharing arrangements are based on signals provided by election outcomes. Martinez-Bravo et al. (2014) show that local elections on the village level in China mattered for economic outcomes (see also Brandt and Turner, 2007), while Lorentzen (2013) describes how, in the absence of elections on the provincial level, China uses local riots to gather information about corruption of local officials.

⁴In an important case study of Mexican party dictatorship, Magaloni (2006) analyses both motives. Blaydes (2006, 2008) focuses on Egypt to demonstrate how local elections are used as a monitoring device to identify optimal, from the dictator’s standpoint, allocation of offices and distribution of patronage. Gandhi and Lust-Okar (2009) find similar patterns in Jordan.

with multiple equilibria, and the vanguard of revolution provides information to other potential protesters, thus altering the focal point (see also Shadmehr and Bernhardt, 2014, and Kricheli et al., 2011). Of particular interest are the papers that link mass protests to elections, both theoretical and empirical. Simpser (2013) collects evidence of electoral manipulation around the world and suggests that fraud may be used to signal strength of the regime. Little (2013) studies electoral fraud with rational voters; in his paper, however, the dictator does not possess superior information and his decisions do not have informational value to the citizens (see also Little, 2012). Kuhn (2011) presents a model where both fraud and protests are decisions made by unitary actors; he argues that protests are only possible if the election is won by the incumbent by a narrow margin and there is evidence of fraud. In Little, Tucker, and LaGatta (2013), the results of an election convey the same information to the dictator and the citizens, and the main question is whether the dictator will agree to step down voluntarily after losing. Gehlbach and Simpser (2014) study dictators' incentives to manipulate election results in a two-person 'sender-receiver' model (see also Rozenas, 2013). Our paper contributes to the literature on mass protests by introducing a simple and tractable model of protests with a unique equilibrium for any public information made available, which makes it particularly suitable for models involving manipulation of information.

Our paper is also related to the literature on violence and political repressions in nondemocracies (see Wintrobe, 1990, 1998 for early models). Acemoglu, Robinson, and Verdier (2004) and Padró i Miquel (2007) show how the use of force and fear helps to extract rents. (See Acemoglu, Golosov, and Tsyvinski, 2008, for a general theory of political incentives; Padró i Miquel and Yared, 2012, analyze politics of indirect control under the threat of using violence.) In Egorov and Sonin (2005) and Debs (2010), the winner of a power contest decides the fate of the loser and may execute the latter in order to prevent him from challenging his position again. In Acemoglu, Egorov, and Sonin (2008) powerful coalitions are able to eliminate political opponents until a stable coalition is formed. In this paper, we simultaneously study the use of repression and communication strategies by the incumbent dictator.⁵

The rest of the paper is organized as follows. Section 2 introduces a formal model. Section 3 analyzes this model. We first study the citizens' decisions to protest, then their voting decisions,

⁵Other formal models of nondemocratic regimes look specifically at economic performance of these regimes; see, e.g., Gehlbach and Keefer (2011) on investment, Acemoglu, Egorov, and Sonin (2010) and Egorov and Sonin (2011) on selection of members of the government, Besley and Kudamatsu (2010) and Svoboda (2009) on autocrat's accountability to a narrow selectorate.

and then the dictator’s incentives to use elections and to repress. In Section 4, we endogenize the opposition’s decision to participate in elections, explore the possibility of electoral fraud, and argue that our results are robust. Section 5 concludes.

2 Model

We study a simple two-period model of political selection where citizens decide whether to keep the incumbent dictator (D) or bring the challenger (C) to power after the first period. There is no discounting, and each of the politicians gets utility A every period he is in office. Unlike models of democratic politics, election is not the only way the office-holder may change; transition of power may also be the result of mass protests.

Both politicians are characterized by their abilities, a_D and a_C , which are drawn from the same normal distribution $\mathcal{N}(a_0, \sigma_a^2)$,⁶ and neither is observed by the citizens. However, citizens observe their economic well-being, which is a signal about the incumbent leader’s ability: if in some period $j \in \{1, 2\}$ politician $P \in \{D, C\}$ is in office, citizen i gets a payoff

$$r_i^P = a_P + \delta_i^P. \quad (1)$$

This specification captures several important points. First, citizens observe their own well-being, but not that of their fellow citizens. Thus, by the end of the first period, they hold heterogeneous beliefs about the dictator’s ability a_D .⁷ Second, they expect their utility to remain the same as long as the incumbent stays in office: for each citizen, the individual shock δ_i^P depends on the politician in power, but not on the period;⁸ all δ_i^P are assumed to be independent and distributed as $\mathcal{N}(0, \sigma_\delta^2)$. Third, citizens do not get any extra information about the challenger’s ability a_C by the end of the first period, and expect it to be a_0 .

Let us denote the expected net gain of citizen i from regime change by b_i :

$$b_i = \mathbb{E}a_C - r_i^D = a_0 - a_D - \delta_i^D. \quad (2)$$

⁶The assumption that they are drawn from the same distribution is not important and is done to save on notation. All the results would go through if $a_D \sim \mathcal{N}(\mu_D, \sigma_D^2)$ and $a_C \sim \mathcal{N}(\mu_C, \sigma_C^2)$.

⁷In a more realistic model, citizens would observe the well-being of their friends and neighbors, but not of those living far away. Our results go through as long as citizens are heterogeneous in their preferences.

⁸This assumption captures the idea that each politician pursues some specific policies which create winners and losers in the society. In the model, this creates conflict of interest, which will guide citizens’ voting and protesting behavior (see also E.Bueno de Mesquita, 2010). All the results will go through if instead of holding heterogeneous expectations about their payoffs should the incumbent stay in power, citizens received heterogeneous taste shocks which would affect their decisions to vote and to protest.

The citizen benefits if the dictator is replaced with the challenger if $b_i > 0$, and is worse off otherwise.

The dictator can lose power to the challenger in two ways: either by running against him in an election and losing it or, even if he runs unopposed and wins, he may still lose power as a result of mass protests. We follow Persson and Tabellini (2009) in assuming that the probability of dictator leaving the office, π , equals the share of population protesting. Each individual in the society makes the decision to protest independently and simultaneously. A citizen who decided to protest gets a disutility of $-c$ (where $c > 0$), which accounts for the probability of being shot, wounded, fired from the job, etc.; we expect c to be higher in more cruel regimes. At the same time, a citizen gets an extra “warm glow” utility, which reflects personal satisfaction from protesting against the hated regime and personally contributing to the dictator’s departure.⁹ We make the simple assumption that the warm glow utility of citizen i is proportional to his economic dissatisfaction with the regime b_i introduced in (2). More precisely, if citizen i protests and the dictator leaves, i gets an extra utility of αb_i . Citizen i gets some part of this warm glow, γb_i with $\gamma < \alpha$, even if he protested unsuccessfully; this captures the possibility that a sufficiently dissatisfied citizen may protest even if he does not expect the protest to bring him immediate benefits (e.g., the “Arab Spring” in Tunisia started with a young merchant self-immolating; there were similar episodes following the failure of the Prague Spring).¹⁰ Clearly, these preferences and intuitions are reversed for a person who strongly supports the dictator (i.e., if b_i is negative and large in absolute value); such a person would never protest as there is no benefit from protesting, only cost.

The payoffs from protesting are summarized in the following matrix:

	Dictator leaves	Dictator stays	
Citizen protests	$\alpha b_i - c$	$\gamma b_i - c$	(3)
Citizen stays home	0	0	

We summarize our assumptions on the payoff matrix (3) here:

⁹Persson and Tabellini (2009) introduced this parameter to capture the “warm glow” that an individual may experience from (successfully) defending the idea he/she firmly believes in, such as defending democracy or overthrowing a much-hated dictator. In this paper, we assume that some “warm glow” from protests may be experienced even if the uprising ultimately fails (γb_i in the top-right cell). This is in line with the recent (and growing literature) about ethical actions and warm glow in voting (e.g., Cherepanov, Feddersen, and Sandroni, 2013; see Feddersen, Gailmard, and Sandroni, 2009, for experimental evidence).

¹⁰Kricheli, Magaloni, and Livne (2011) build a two-period model of protests, where protests in the first period serve as a signal for other citizens to participate in the second period. There, a citizen who dislikes the dictator enough will protest so as to make his discontent with the dictator public information.

Assumption 1 $c > 0, \alpha > \gamma > 0$.

The assumption that $\gamma > 0$ is important: it implies that there is always an agent who protests (notice that for b_i high enough, protesting is a dominant strategy). The last assumption, $\alpha > \gamma$, captures the increasing differences intuition: If a citizen wants the dictator to leave ($b_i > 0$), his propensity to protest is higher if the dictator leaves than if the dictator stays. Indeed, this is equivalent to

$$\alpha b_i - c > \gamma b_i - c, \tag{4}$$

which simplifies to $\alpha > \gamma$.

We also maintain the following simple assumption, which ensures the existence and uniqueness of an equilibrium, throughout the paper. The assumption provides that there is a sufficient variation in citizens' idiosyncratic payoffs from the incumbent's rule. In other words, an individual's attitude toward the dictator, b_i , is not a too good predictor of other citizens' attitudes.

Assumption 2 *The variance of individual shocks is sufficiently large:*

$$\sigma_\delta > \frac{1}{2\sqrt{2 \ln 2}} \frac{c(\alpha - \gamma)}{\gamma^2}. \tag{5}$$

The dictator knows his own competence, and thus the actual distribution of people's attitudes. He has several strategies to choose from in order to maximize the chance of staying in power. One decision he has to make is whether or not to have elections with the challenger on the ballot. If the incumbent decides to have such elections, each citizen votes either for him or for the challenger, and the incumbent wins if the share of votes he gets, τ , is at least $\tilde{\tau} \in (0, 1)$.

In the main model, we assume that once the challenger is on the ballot, votes are counted fairly; in an extension in Section 4, we consider the case where the dictator cannot commit to fair counting and show that our results are robust. But to have competitive elections, the incumbent needs some cooperation from the challenger; at least, the challenger must agree to run. There are various reasons why a potential challenger might prefer not to run: e.g., he might not believe that votes are going to be counted fairly or that the voting outcome will show that the opposition is not popular; we endogenize the challenger's decision to participate in Section 4. For now, we assume that there is an exogenous probability $\eta \in (0, 1)$ that no credible opponent participates in elections even if the dictator tries to organize them. What is critical is that the voters do not observe the reason why the challenger does not participate in elections: they either see him on the ballot and

decide that the outcome is informative, or they do not, in which case the outcome of the election is not informative, but absence of the challenger on the ballot is.

Another option in the incumbent's arsenal is to repress the challenger; in the real world, it might be imprisonment, exile, or even assassination. This option is costly; using it costs some $k > 0$. Without a challenger, not only fair elections are impossible, but citizens become more averse to protests: the citizen's expectation of the next leader's competence is $a_0 - \Delta$ where $\Delta > 0$, and thus the benefit of citizen i from removing the dictator from power is $b_i^\Delta = b_i - \Delta$. For simplicity, we focus on the extreme case $\Delta = +\infty$, which says that if the challenger is repressed, the opposition is effectively decapitated and there are no credible alternatives left; as a result, nobody protests against the dictator and he is certain to stay in power. In other words, we assume that at cost k , the dictator is able to ensure his tenure in the near future.¹¹

To summarize, the timing of the game is as follows.

1. The competences of the incumbent and challenger, a_D and a_C , are realized.
2. Each citizen i gets utility r_i^D ; the incumbent leader learns his competence a_D .
3. The incumbent decides whether or not to repress. If he represses, the game goes to Step 7.
4. If there was no repression, the incumbent decides whether to have competitive elections; he succeeds in doing so with probability η , which is exogenous in Section 3 and is endogenized in Section 4.
5. Each citizen votes, the votes are counted, and the tally is announced. If $\tau < \tilde{\tau}$, the dictator is removed from office, and the game moves to Step 7 with the challenger in power for the second period.
6. Each citizen decides whether or not to protest with their payoffs given by (3). With probability π , where π is the share of those who protest, the challenger becomes the new leader, and with probability $1 - \pi$, the incumbent stays in power.
7. Each citizen i gets their second-period utility r_i^P , the dictator gets A if he stays in power ($A - k$ if he repressed), and the game ends.

We use the Perfect Bayesian equilibria in pure strategies for the equilibrium concept, so both politicians and citizens act fully rationally. We restrict ourselves to equilibria where citizens vote

¹¹Shadmehr (2013) considers repressions that provoke more, rather than less, protests.

sincerely (we discuss below why such strategies are natural in this game where votes serve as a signal relevant for protests). Throughout the paper, F and f are the c.d.f. and p.d.f. of a standard normal distribution.

3 Analysis

Our analysis proceeds as follows. First, we study citizens' decisions to revolt for any public information they may have and show that under Assumption 2 there exists a unique equilibrium which takes the threshold form: citizens with low b_i do not protest, and citizens with high b_i protest. This is true both in the case where citizens know the value of b and when they only know its distribution, regardless of what this distribution is. Second, we study citizens' voting decisions and show that sincere voting is an equilibrium. Third, we analyze the incumbent's decision to have competitive elections. We find that there is a unique threshold such that less popular dictators do not have competitive elections and more popular dictators try to do so. Finally, we introduce the strategic decision to repress and complete the analysis of the model.

3.1 Decisions to protest

We start with characterizing individuals' decisions to protest. Denote $b = a_0 - a_D$ and $\delta_i = -\delta_i^D$; with this notation,

$$b_i = b + \delta_i, \tag{6}$$

where δ_i is distributed as $\mathcal{N}(0, \sigma_\delta^2)$. This represents b_i , which is known to citizen i , as a sum of the common component b and a zero-mean idiosyncratic shock δ_i . Suppose that by the time of protests (Stage 6), after taking all public information into account, i.e., whether or not there was repression, competitive election, and the outcome of the election τ , b is believed to be taken from some distribution G . (This G will depend on all the prior actions by the dictator, but we keep the notation simple for now.) Thus, we study the decision of citizen i to protest if he thinks that b is taken from distribution G and he also observes his b_i . One can think about b_i as a signal about b , which is relevant to citizen i because it will determine the distribution of signals of other citizens, on which they base their decisions to protest; this, in turn, will determine the probability of success, and this is valuable information for citizen i making the decision.¹²

¹²The protesting game has a lot in common with global games (Carlsson and van Damme, 1993), which are often used to get unique equilibria in games with strategic complementarities such as currency attacks or mass protests. In

Each citizen i , knowing b_i , updates her priors on the distribution of b , thus getting distribution $G_{b_i} = G \mid b_i$.¹³ Because of the simple equation (6) that links b and b_i , we can prove that G_x first-order stochastically dominates G_y whenever $x > y$: for any $\xi \in \mathbb{R}$ such that $0 < G(\xi) < 1$, we have $G_x(\xi) < G_y(\xi)$.

In a pure-strategy equilibrium, each individual i decides whether or not to protest. Consider the set of protesters: let R_G denote the set of $x \in \mathbb{R}$ such that a citizen who got $b_i = x$ decides to protest. It is natural to expect (see Appendix for the full proof) that this set takes a form $R_G = [z_G, +\infty)$, where individual i with $b_i = z_G$ is indifferent.¹⁴ If so, the share of protesters, and thus the chance of the incumbent losing office, is equal to

$$\hat{\pi}_G = \hat{\pi}_G(b) = \Pr(b + \delta_i > z_G) = 1 - F\left(\frac{z_G - b}{C \sigma_\delta}\right). \quad (7)$$

However, someone who does not know b (e.g., citizen i with $b_i = x$) needs to integrate over all possible values of b ; for this person, the perceived probability of success is

$$\pi_{G_x} = \Pr(b + \delta_j > z_G \mid b + \delta_i = x) = 1 - \int_{-\infty}^{+\infty} F\left(\frac{z_G - b}{\sigma_\delta}\right) dG_x(b). \quad (8)$$

In what follows, we will occasionally abuse notation and write π_x instead of π_{G_x} . Lemma 5 implies that π_x is increasing in x , and so citizens who got a higher b_i become more optimistic about the size of protests, even though they are aware of a fixed threshold strategy that other citizens use. Intuitively, a high b_i serves as a signal of their individual (low) utility under the incumbent regime, and since it is also a signal of the aggregate, such citizen believes that many other people feel bad about the incumbent as well. Thus, more of them fall above the protest cutoff z_G , and therefore the share of protesters and the chance of success is higher.

For any individual i with $b_i = x$, the expected continuation utilities from protesting and staying

this paper, we make a departure from the standard approach. Technically, we assume that b_i is not merely a signal about the aggregate variable b ; it is also a parameter that enters the payoff of citizen i directly. There are two reasons for this approach. First, we believe that citizens have heterogeneous benefit from removing the dictator; their conflict of interests would not vanish if they met together and aggregated their signals, and thus it is realistic to think of b_i as a preference parameter which just happens to be informative of the whole distribution. Second, we are interested in a unique equilibrium even if there is no uncertainty about the underlying variable b , because in our model, the dictator has the ability to reveal b by organizing fair competitive elections.

¹³More precisely, G_x is the probability distribution of b conditional on $b + \delta_i = x$, given by $G_x(y) = \Pr(b \leq y \mid b + \delta_i = x) = \frac{\int_{-\infty}^y f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)}{\int_{-\infty}^{+\infty} f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)}$.

¹⁴An open interval $R_G = (z_G, +\infty)$ is also possible, but we can just assume that the indifferent individuals protest without any loss of generality.

at home are equal to

$$\mathbb{E}U_p(x) = \pi_x \alpha x + (1 - \pi_x) \gamma x - c + [\pi_x a_0 + (1 - \pi_x) x], \quad (9)$$

$$\mathbb{E}U_s(x) = [\pi_x a_0 + (1 - \pi_x) x], \quad (10)$$

respectively. The terms in brackets reflect the second-period utility and are the same in both cases, as no single individual may affect the chance of success. The threshold citizen with $b_i = z_G$ must be indifferent between protesting and not. Consequently, the cutoff z_G must satisfy

$$z_G = \frac{c}{(\alpha - \gamma) \pi_{z_G} + \gamma}. \quad (11)$$

Taking into account (8), which must hold for $x = z_G$, we conclude that the equilibrium threshold z_G is defined by the following equation:

$$z_G = \frac{c}{(\alpha - \gamma) \int_{-\infty}^{+\infty} F\left(\frac{b - z_G}{\sigma_\delta}\right) dG_{z_G}(b) + \gamma}. \quad (12)$$

In the Appendix, we prove that for any distribution G of b , the threshold z_G exists, is unique, and satisfies (12).¹⁵

Proposition 1 *For any posterior distribution G of beliefs about the difference between politicians' competences, $b = a_0 - a_D$, that is obtained by the time of protests using publicly available information, there exists a unique protest equilibrium. It is characterized by threshold $z = z_G$ given by (12) that determines which citizens (those with $b_i \geq z_G$) participate in the protest.*

While the detailed proof is relegated to the Appendix, it is instructive to see some detail in the special case when the difference in abilities b is public information, and thus the posterior distribution G is an atom at b . This corresponds to the case where citizens have learnt the truth about dictator's competence, for example, by aggregating the results of an election. Equation (12) then becomes

$$z_b = \frac{c}{(\alpha - \gamma) F\left(\frac{b - z_b}{\sigma_\delta}\right) + \gamma}, \quad (13)$$

where we abused notation and wrote z_b instead of z_G (likewise, we will use $\hat{\pi}_b$ instead of $\hat{\pi}_G$). Existence follows, since as left-hand side varies from $-\infty$ to $+\infty$, the right-hand side increases

¹⁵Notice that the threshold z_G is known to both politicians and citizens, since function G is common knowledge. The probability of success, however, is in the eye of the beholder. The dictator D knows the true value of b and thus the true distribution of $\{b_j\}$, whereas citizens have heterogenous beliefs, except for the case where G is degenerate and b is common knowledge.

from $\frac{c}{\alpha}$ to $\frac{c}{\gamma}$. Uniqueness is less obvious as the right-hand side is also increasing in z . Intuitively, as the protest threshold z becomes higher, the success of protests become less likely, and thus fewer citizens are willing to protest. As a result, a citizen must hate the dictator very much to be willing to protest, which also raises the threshold. Thus, there is a potential for multiple thresholds, due to the following strategic complementarity: more citizens protesting makes the success of a revolt more likely, and this encourages even more people to protest. However, uniqueness follows from Assumption 2; it ensures that the derivative of the right-hand side with respect to z is less than 1. The same Assumption 2 guarantees that there are no non-threshold equilibria.

If b is not known, then (12) exhibits an additional effect in the right-hand side as G_z becomes a different distribution as z changes. However, we show that this effect does not hurt uniqueness. Intuitively, a higher z implies that the threshold citizen updates on the distribution of b and becomes more confident in the success of the protests. This mitigates the effect that $F\left(\frac{b-z_G}{\sigma_\delta}\right)$ is decreasing in the threshold z , and thus the derivative of the right-hand side of (12) cannot exceed 1 in this case as well. In other words, a decrease of the threshold not only makes citizens more enthusiastic about the probability of success; it also has the opposite effect: the threshold citizen is more sceptical about the overall negative attitude towards the dictator as compared to the challenger, b . This decreases the strategic complementarity effect described earlier, and makes uniqueness of equilibrium easier to obtain. Notice that we have not made any assumptions about the distribution G .

Proposition 1, coupled with equation (12), allow us to obtain the following comparative statics result.

Proposition 2 *For any distribution of b conditional on public information, G , the participation threshold z_G is increasing in c , the cost of protests, and decreasing in α and γ , the utilities that a citizen receives from participating in successful and unsuccessful protests, respectively. Moreover, if distribution G_1 first-order stochastically dominates G_2 , then $z_{G_1} < z_{G_2}$. In particular, if the average attitude b is publicly known, then the participation threshold z_b is decreasing in b .*

The comparative statics is very intuitive. The threshold is lower, and thus the probability of success is higher, if protests are less costly (c is low), because for any fixed chance of success more people are willing to protest. Similarly, if a person who dislikes the dictator has a stronger incentive to protest (either α or γ is higher), more people will protest. Now, if for two distributions G_1 and G_2 , the former dominates the latter, then the chance of success if all citizens above a certain

threshold protest is higher under G_1 than under G_2 ; this, in turn, makes more people willing to protest in the former case. This last part has general implications: A dictator who is perceived to be incompetent or who faces an opponent believed to be competent will experience a lower threshold z_G and thus larger-scale protests.

To understand the decision to have a competitive election which would reveal the dictator's true popularity, consider the perception of the threshold citizen with $b_i = z_G$ about the size of the protests. He believes that the share of protesters and the probability of success equal π_{z_G} , which may be less than π_G , the objective probability of success, or greater than that. If $\pi_{z_G} < \hat{\pi}_G$, the dictator expects larger-scale protests than the threshold citizen. For such dictator, revealing the true value b to citizens would be dangerous: in this case, all citizens, including the threshold one, will update their beliefs and think that for the same protesting strategies (with threshold z_G), the share of protesters would be higher: $\hat{\pi}_G > \pi_{z_G}$. But in this case, the citizen who got $b_i = z_G$ would no longer be indifferent; he would strictly prefer to protest, as would citizens with slightly lower signals. This would make protests even bigger and overall, the threshold would decrease, endangering the dictator further. In case $\pi_{z_G} > \hat{\pi}_G$, the logic is the opposite. Here, the threshold citizen z_G is too optimistic about the chances to oust the dictator. Revealing true b would make him more skeptical, and he would then strictly prefer to stay at home. Thus, fewer people would protest, thus increasing the chance that the dictator survives. Consequently, we have the following result about the dictator's incentives to reveal the information he has on b in face of protests.

Lemma 1 *For any distribution G without atoms, there is a unique threshold b_G^* such that if the average attitude toward the dictator $b < b_G^*$, the protest threshold is lower than the protest threshold conditional on b being public: $z_G < z_b$, and thus the dictator is better off by revealing b . Moreover, $\pi_{z_G} > \hat{\pi}_G(b) > \hat{\pi}_b$, so the chance of success perceived by the threshold citizen with $b_i = z_G$ is higher than that perceived by dictator, which is in turn higher than the chance of success if b were revealed. For $b > b_G^*$, the situation is reversed: $z_G > z_b$, and the dictator is better off not revealing the average attitude to him (in this case, $\pi_{z_G} < \hat{\pi}_G(b) < \hat{\pi}_b$). This threshold b_G^* satisfies $G(b_G^*) \in (0, 1)$, i.e., there is a positive mass of b on both sides of b_G^* .*

Before studying equilibrium decision to have competitive elections, we first study voting strategies.

3.2 Voting

Consider citizens' voting behavior. Their preferences are simple: a citizen i with $b_i < 0$ wants the dictator to stay in office, while citizen with $b_i > 0$ wants to see him replaced. Thus, sincere voting strategies prescribe individuals to vote for the incumbent if and only if $b_i < 0$. This is indeed an equilibrium, for the simple reason that each citizen is infinitesimal. This also involves no dominated strategies; however, the standard reasoning for such voting behavior is not sufficient. In this game, not only the voting outcome matters, but also the protests that may follow, and the share of votes that the dictator gets will serve as a signal about b , which will affect z_b , the protest threshold.

Fortunately, voting and signaling incentives of citizens are aligned. Suppose, for the sake of the argument, that citizen i controls a small but positive mass ε of votes, and other citizens vote sincerely. Suppose that he wants the dictator to stay ($b_i < 0$). If he deviated and voted against the dictator, it would have two effects. First, the dictator would lose elections with at least as high probability. Second, for any voting profile of other citizens, the dictator's share of votes will decrease. Consequently, other citizens would believe that the share of those with $b_i > 0$ is higher, so b is higher than it actually is. This would decrease the protest threshold and increase the chance that the dictator loses the office, which is unambiguously bad for citizen i (regardless of whether he protests or not). Hence, such a citizen would not want to deviate. Similarly, a citizen with $b_i > 0$ would not deviate because a deviation would increase the chance of the dictator winning elections, and also lead to smaller-scale protests. Hence, we have the following result.

Lemma 2 *Sincere voting strategies, where citizens with $b_i \leq 0$ vote for the incumbent dictator D and those with $b_i > 0$ vote for the challenger C , constitute a voting equilibrium in undominated strategies.*

In this equilibrium, the share of votes obtained by the dictator is

$$\tau(b) = F\left(\frac{b}{\sigma_\delta}\right). \quad (14)$$

We cannot, however, claim that sincere voting is the only equilibrium in undominated strategies. Indeed, consider the opposite strategies: vote for the dictator if and only if $b_i > 0$, i.e., only if person i wants the dictator to lose elections. If such a person with $b_i > 0$ deviated and voted against the dictator (suppose again, for the sake of the argument, that he controls a small positive mass of votes), there would be two effects. First, the chance that the dictator loses elections would be higher. Second, in case he wins, he would get fewer votes. However, since the voting strategies in

this candidate equilibrium are reversed, the change in vote tally would be interpreted by Bayesian citizens as more support for the dictator, not less, and this would reduce the share of protesters (see Proposition 2). Citizen i wants the dictator to leave and thus wants protests to be more massive. Thus, if he believes that he is unlikely to be pivotal (which must be true if the dictator dared to have competitive elections), then deviation to sincere voting is not profitable because of signaling value that this vote carries. Such voting strategies will deliver the same results as sincere voting as everyone will update correctly, but they are not reasonable. For this reason, we focus on equilibria where voting is sincere.

3.3 Elections

Our next step is to analyze the dictator's incentives to allow competitive elections. Let us hold his decision to repress fixed, and suppose that conditional on not repressing, the distribution of b is G . Ideally, the dictator would have elections if and only if $b < b_G^*$ (this follows from Lemma 1). But he faces two problems. First, it is possible that the share of votes that the dictator receives, $\tau(b)$, satisfies $\tau(b_G^*) < \tilde{\tau}$, so there are dictators who would want to have fair elections because of their signaling value, but are afraid of losing. But even when this is not a constraint, there is a second problem. If the citizens expect the challenger to get on the ballot provided that the dictator did not forbid him ($\eta > 0$), then citizens would know that a dictator who does not run against the contender comes with b taken not from G , but from another distribution G' , given by

$$G'(x) = \begin{cases} \frac{(1-\eta)G(x)}{1-\eta G(b_G^*)} & \text{if } x < b_G^* \\ \frac{G(x)-\eta G(b_G^*)}{1-\eta G(b_G^*)} & \text{if } x \geq b_G^* \end{cases}$$

This distribution first-order stochastically dominates G , and thus the protest threshold under G' would be $z_{G'} < z_G = z_{b_G^*}$, and since the inequality is strict, $z_{G'} < z_b$ for some $b > b_G^*$. If so, the dictator would be better off revealing such value of b . This argument suggests unraveling: Since dictators with sufficiently low b have elections, those who do not are believed to know that b is high, and the borderline ones have to have elections to reveal that b is not too high.

However, there is a limit to this unraveling, even if the constraint $\tau(b) \geq \tilde{\tau}$ is not binding. To see why, notice first that for any belief H about the distribution of b for dictators who failed to have competitive elections, the dictator's best response must follow a threshold rule: have elections

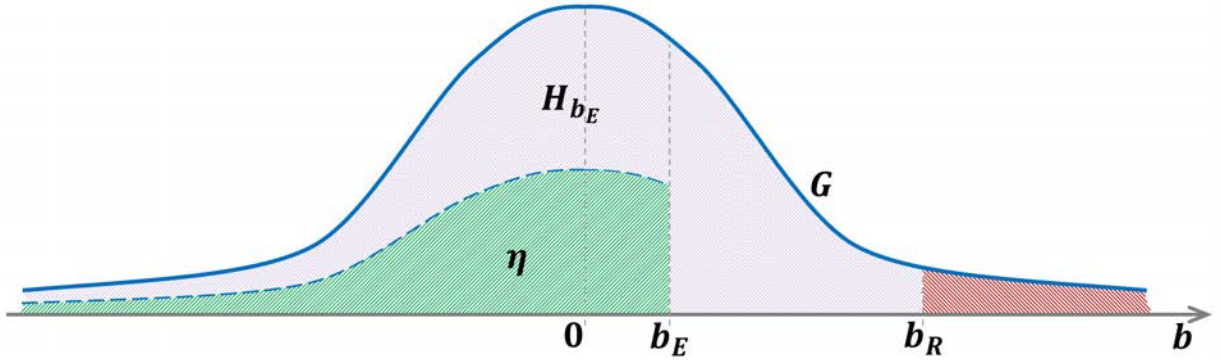


Figure 1: The distribution H_y for $y = b_E$. Conditional on not repressing ($b \leq b_R$), the dictator wants to participate in competitive election if $b < b_E$, and is able to do so with probability η .

if and only if $b \leq b_E$. Now, for any threshold y , define

$$H_y(x) = \begin{cases} \frac{(1-\eta)G(x)}{1-\eta G(y)} & \text{if } x < y \\ \frac{G(x)-\eta G(y)}{1-\eta G(y)} & \text{if } x \geq y \end{cases}. \quad (15)$$

Notice that this distribution converges to G in distribution both when $y \rightarrow -\infty$ and when $y \rightarrow +\infty$, but G first-order stochastically dominates H_y for any finite y , implying $z_{H_y} < z_G$. Thus, if y is very low, so elections are almost never held, then some types of dictators would be willing to reveal b . On the other hand, if y is sufficiently high, so almost every dictator allows competitive elections whenever he can, then failure to do so would not be held against the dictator, and in particular would not be a signal about b . In this case, sufficiently unpopular dictators would not want to run against the challenger. Ultimately, there is a threshold $y = b_E$, and it is unique; this threshold satisfies the condition $z_{H_y} = z_y$. The distribution H_y for $y = b_E$ is depicted on Figure 1.

Lemma 3 *There is a unique threshold b_E such that the dictator attempts to hold elections if and only if $b \leq b_E$. The threshold satisfies $\tau(b_E) \geq \tilde{\tau}$, but does not depend on $\tilde{\tau}$ otherwise. Moreover, $y = b_E$ is the unique minimand of z_{H_y} over $(-\infty, \tilde{b}) \cap (\text{support of } G)$ and it satisfies $0 < G(b_E) < 1$.*

In other words, there are positive probabilities that the dictator will have competitive elections and that he will choose not to. Quite interestingly, the equilibrium b_E satisfies the following property: of all possible thresholds for having elections, the equilibrium one makes the dictators who fail to run against the challenger worse off; they are going to face protests of the largest size. In the Appendix, we show that function z_{H_y} is strictly quasiconvex, and its lowest point corresponds precisely to the equilibrium.¹⁶

¹⁶Ostaszewski and Gietzmann (2008) prove a similar result in the context of a model of (non-)disclosure of

Let us now consider comparative statics. Suppose first that the dictator has almost full discretion, i.e., the challenger is sure to agree, and so η is high. In this case, the citizens know that absence of a challenger on the ballot is likely a signal that the dictator is unpopular, rather than that the challenger decided not to run. In this case, z_{H_y} is lower, so without competitive elections protests will be larger, which is more dangerous for the dictator. As a result, the threshold b_E increases in η . There only limit for this increase is $\tilde{b} = \tau^{-1}(\tilde{\tau})$, as b_E can never exceed it. For η high enough, $\tau(b_E) = \tilde{\tau}$, and the dictator tries to run in every competitive election that he can win.

One may wonder about the impact of the cost of protesting, c , as well as of parameters α and γ , on the choice to have competitive elections. Here, the effect is more subtle. Indeed, a higher cost of protests c raises both the threshold z_b for the case where citizens know b and the threshold z_{H_y} if citizens only know its distribution in case there are no competitive elections (this follows from Proposition 2). Similarly, a higher α or γ make citizens more likely to protest for any information they may know. Yet we get an unambiguous prediction.

Proposition 3 *The threshold b_E is (weakly) increasing in η , the probability that the dictator is able to organize competitive elections (strictly if $\tau(b_E) > \tilde{\tau}$). For η high enough, the condition $\tau(b_E) \geq \tilde{\tau}$ is binding; for η low enough, b_E converges to b_G^* . A higher cost of protests c makes elections more likely (threshold b_E increases). The effects of increases in α or γ are the opposite (b_E decreases).*

Despite its subtleness, the result follows naturally from the following intuition. ■ higher c discourages protesting both after fair elections and when there were no fair elections. However, there is an additional effect: lower level of participation also makes citizens more pessimistic about the success of an uprising, which further decreases participation. The first effect is similar in size in both cases, but the second effect is more pronounced if citizens are better informed about b . Intuitively, when citizens know b , the borderline citizen, who is indifferent between protesting and not protesting, interprets his increased scepticism (higher b_i) as a sign that fewer people will protest, not as a sign that b is actually high. In contrast, when citizens only know a distribution of b , the effect is mitigated by that this citizen also updates on b : the fact that he hates the dictator more makes him more optimistic about the success of the uprising. Consequently, a higher cost of protests c is more likely to deter citizens who know the true value of b , and this makes the dictator

information in Dye (1985), and Shadmehr and Bernhardt (2013) do so in a model of state censorship.

choose elections for a larger set of values of b . Overall, there is the following pattern: if it is more costly for citizens to protest, only the most convicted, and thus disproportionately optimistic protesters, remain, and the dictator will want to reveal information to moderate their optimism. The intuitions for changes in α and γ are very similar.

The comparative static results above were obtained for a fixed repression strategy. As we will see below, these results continue to hold if the dictator's repression strategy is endogenous and adjusts to the changes in parameter values.

3.4 Repression

Let us now analyze the decision to repress. Our question is when the dictator will use this option, and how this will affect his choice about having competitive elections if he decides not to repress.

We have already shown that for every repression strategy and thus for any distribution G if the dictator decided not to repress, there is a threshold b_E such that the dictator has competitive elections if and only if $b \leq b_E$. As a result, the effective (equilibrium) threshold of citizens' protests depends on b in the following way: it equals $\tilde{z}(b) = z_b$ for $b \leq b_E$ and it is $\tilde{z}(b) = z_{b_E}$ for $b > b_E$, i.e., $\tilde{z}(b) = z_{\min\{b, b_E\}}$. Our next goal is to show that the repression strategy also takes the form of a threshold, which depends on threshold b_E . We then analyze how these thresholds affect one another, and this allows us to establish both the existence and uniqueness of equilibrium, as well as all the comparative statics results.

For an arbitrary value of $b_E = y$, consider the dictator's decision to repress, if in case he does not repress, citizens protest if and only if $b_i \geq \tilde{z}^y(b)$, where

$$\tilde{z}^y(b) = z_{\min\{b, y\}}. \quad (16)$$

Then the expected utility of a dictator who does not repress is

$$\begin{aligned} V_D^y(b) &= A(1 - \hat{\pi}) = A \int_{-\infty}^{\tilde{z}^y(b)} f\left(\frac{x-b}{\sigma_\delta}\right) dx \\ &= A \int_{-\infty}^{\tilde{z}^y(b)-b} f\left(\frac{x}{\sigma_\delta}\right) dx. \end{aligned}$$

Since $\tilde{z}^y(b)$ is nonincreasing in b for any given y , we have that $V_D^y(b)$ is strictly decreasing in b . If the dictator represses, he gets $A - k$. Consequently, the dictator will never repress if $k \geq A$, whereas for $A < k$, the dictator's decision to repress takes the form of a threshold; he will repress if $b > R(y)$, where the function R is given by

$$\int_{\tilde{z}^y(R(y))-R(y)}^{+\infty} f\left(\frac{x}{\sigma_\delta}\right) dx = \frac{k}{A}. \quad (17)$$

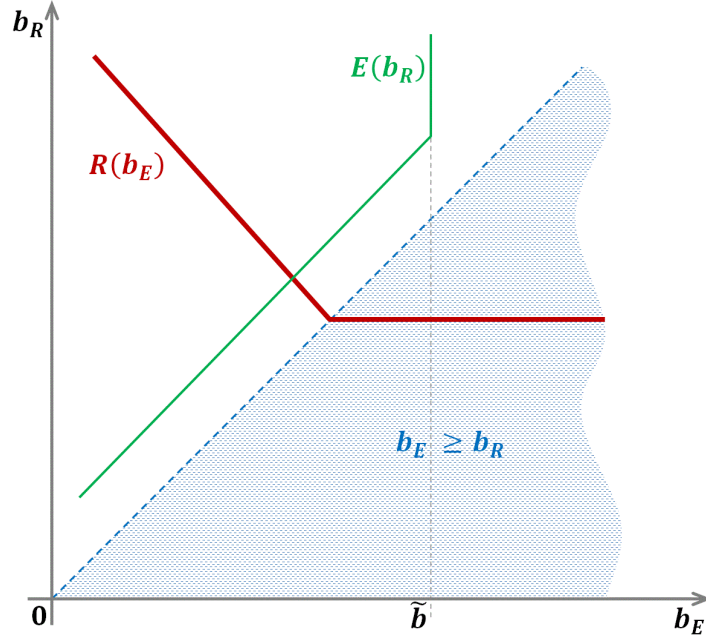


Figure 2: The reaction functions $E(\cdot)$ and $R(\cdot)$.

This function $R(y)$ is nonincreasing in y , because $\tilde{z}^y(b) - b$ is decreasing (strictly) in b and (weakly) in y . This admits a simple intuition: a higher y means that citizens protest in weakly higher numbers for any b , which is worse for the dictator, and he is more likely to repress. In equilibrium, the thresholds b_R (for repression) and b_E must satisfy $b_R = R(b_E)$.

At the same time, every repression threshold $b_R = r$ defines a posterior distribution of b if the dictator does not repress:

$$G^r(b) = \begin{cases} \left(\int_{-\infty}^b f\left(\frac{x}{\sigma_b}\right) dx \right) / \left(\int_{-\infty}^r f\left(\frac{x}{\sigma_b}\right) dx \right) & \text{if } b \leq r \\ 1 & \text{if } b > r \end{cases}.$$

This, in turn, defines a unique equilibrium threshold for elections $b_E = E(r)$, which is a nondecreasing function of r (it is increasing if $b_E < \tilde{b}$ and has a flat part above \tilde{b}). Indeed, the distribution $G^{r'}$ first order stochastically dominates G^r if $r' > r$, and this implies that $E(r') \geq E(r)$. Intuitively, a higher repression threshold implies that a dictator who did not use repression is, on average, less popular (has a higher b), so the citizens are more optimistic about the success of protests, and the dictator sees more internal pressure to reveal his true popularity. These two reaction functions are depicted in Figure 2.

In equilibrium, we must have $b_E = E(b_R)$ and $b_R = R(b_E)$, and these two equations simultaneously determine the two equilibrium thresholds b_E and b_R . We have the following result.

Lemma 4 *If the cost of repression k is high ($k \geq A$), there are no repressions ($b_R = +\infty$), and there is a unique equilibrium value of b_E . If $k < A$, there is a unique equilibrium given by a pair of thresholds b_E and b_R satisfying $b_E < b_R$: for $b > b_R$, the dictator represses at cost k ; for $b_E < b \leq b_R$, the dictator does not repress, but does not allow competitive elections either; for $b \leq b_E$ he does not repress and tries to organize a competitive election.*

The characterization of equilibrium in terms of reaction functions $E(\cdot)$ and $R(\cdot)$ facilitates obtaining the comparative statics results. For example, if either A increases or k decreases, i.e., the cost of repression as a fraction of utility from having office becomes lower, $R(y)$ decreases for any given y . The function $b_E = E(b_R)$ is preserved (the election threshold does not depend on A or k directly), and it is easy to see from Figure 2 that both equilibrium thresholds b_R and b_E decrease. The intuition for b_R is straightforward; as repression becomes cheaper, the dictator will use it more often. This, however, has the following spillover: a dictator who does not use repression when it is relatively cheap signals that he is sufficiently strong (has low b), and this makes citizens less willing to protest, absent further information. Not surprisingly, in this case, the dictator is less willing to provide this further information, so he prefers to have elections less often, and this results in a lower b_E .

A similar effect is present if we vary η . By Proposition 3, $E(r)$ increases for any r , because citizens interpret absence of competitive elections as incumbent's weakness, and this prompts him to elections for more parameter values. However, if this is the case, then protests in the case of absence of competitive elections become larger, the dictator's expected utility if he does not repress goes down, and he is more likely to repress. Thus, b_E increases and b_R decreases, while the response function $b_R = R(b_E)$ is preserved. These, and other results are summarized in the next proposition.

Proposition 4 *An increase in the benefits of being in office, A , or a decrease in the cost of repressions, k , decreases thresholds b_E (strictly if $b_E < \tilde{b}$) and b_R , making repression more likely and competitive elections less likely. An increase in η , the probability that the challenger runs if given a chance, increases b_E and decreases b_R (strictly in both cases if $b_E < \tilde{b}$), making both repression and competitive elections more likely at the expense of uncontested elections. In addition, an increase in c , the personal cost of protest, or a decrease in α or γ , the utilities that a citizen receives from participating in successful and unsuccessful protests, respectively, make competitive elections more likely (b_E increases), and repression less likely (b_R increases).*

The comparative statics results on the regime's cruelty c , or the warm glow parameters α and γ ,

is not that straightforward, because they affect both $R(y)$ and $E(r)$. As shown above in Proposition 3, an increase in c or a decrease in α or γ increase b_E and thus make competitive elections more likely. Thus, $E(r)$ increases for any given r . At the same time, these changes decrease the scale of protests for any given beliefs about b that citizens may have, and this makes the dictator less willing to repress, thereby increasing the threshold $b_R = R(y)$ for any election threshold y . Thus, the effect of these changes on b_E is unambiguous: it increases, making competitive elections more likely, and this confirms that Proposition 3 holds even if the decision to repress is endogenized. The effect on b_R is a priori ambiguous: while the direct effect leads to a higher b_R and fewer repressions, there is another, indirect effect: When b_E increases, then, as in the case of an increase in η , citizens believe that a dictator who failed to have competitive elections is unpopular, which makes them protest in large numbers, and which makes the dictator more willing to repress. In the Appendix we show that the direct effect always dominates, and b_R increases. All comparative statics results are summarized in the table below.

	Competitive election	Repression
$A \uparrow$ (more rents from office)	<i>Less likely</i>	<i>More likely</i>
$k \downarrow$ (repression is cheaper)		
$c \downarrow$ (protesting less costly)		
$\alpha, \gamma \uparrow$ (more warm glow)		
$\eta \uparrow$ (challenger is likely to run)	<i>More likely</i>	

4 Extensions

In this section, we consider extensions of the model that show robustness of our main results and highlight additional effects that we did not cover in the main body of the paper. First, we endogenize the opposition's willingness to participate in an election organized by the dictator. Then, we analyze what happens if the dictator is able to commit fraud and, in particular, win elections that he would lose without fraud.

4.1 Endogenous Decision to Run

So far, we assumed that the challenger runs with probability $\eta < 1$ if he is given this option. We now endogenize his decision to run. We will see that not running is not only possible (which would

justify our assumption $\eta < 1$); in fact, not participating in elections is a more attractive option for the challenger who wants to come to power, unless he has other reasons to be on the ballot.

Consider the following modification of our basic game. The dictator first decides whether or not to repress the opposition and, if not, whether or not to allow competitive elections. If he allows competitive elections, the challenger decides whether to participate or not. If he decides to participate, competitive elections take place, and if he does not, then citizens observe absence of competitive elections and of repression; still, citizens cannot observe whether it is the challenger who decided to pass or the dictator who put enough obstacles to keep the challenger off the ballot. (Notice that the decisions by the dictator and the challenger may be made in a different order or simultaneously: the dictator's decision to have a competitive election matters only if the challenger agrees to run, and vice versa.)

We assume the following payoffs for the challenger. If he comes to power, either through an election or mass protests, he gets a benefit A from being in power, similar to the incumbent dictator. Let us assume that participating in a competitive election gives the challenger utility w , which may be positive (e.g., nonpecuniary benefits of political participation) or negative (e.g., the opportunity cost of a campaign), and this utility is realized regardless of the outcome, in particular if he loses (the dictator will organize only such elections in our model). This utility is drawn from some distribution with c.d.f. $\Phi(w)$, which has full support on $(-\infty, +\infty)$. The realization of w is observed by the challenger at the time of making the decision whether or not to run. We do not introduce notation for the challenger's disutility from being repressed because he decides whether or not to run after he knows he is not repressed. The question we are asking is for what values of w the challenger will decide to participate in the election.

Naturally, the challenger will participate in the election if and only if $w \geq \tilde{w}$ for some threshold \tilde{w} . Naive intuition may suggest that $\tilde{w} < 0$, meaning the challenger runs in all elections which give him a nonnegative utility, plus perhaps some more, because of the chance of winning and getting A in the next period. However, because in this game elections are organized by the dictator who has superior information about his popularity, the result is going to be different. Indeed, suppose that $w = 0$, so the challenger does not get any immediate benefits nor bears costs from elections per se. The dictator organizes a competitive election if and only if $b \leq b_E$, consequently, if $b > b_E$, the challenger's decision to run does not have payoff consequences. If $b < b_E$, then deciding to run leads to the following: the election is competitive, the dictator gets share of votes $\tau(b)$ given by (14) and wins, the true value of b gets revealed, and protests follow, with threshold z_b depending

on b . If, however, the challenger does not run, the election is not competitive, b is not revealed, the citizens' posterior distribution of b is given by H_{b_E} given by (15), and the subsequent protests will have threshold $z_{H_{b_E}}$. However, we know that $z_{H_{b_E}} \leq z_b$ for all $b \leq b_E$, and for $b < b_E$ the inequality is strict. In other words, if the election is not competitive, the citizens will protest in larger numbers. Given that, not participating in the election is a best response for the challenger for any value of b . This means that he will not participate even if w is small but positive, in other words, we must have $\tilde{w} > 0$, contrary to the naive intuition above.

What determines the equilibrium value of \tilde{w} ? For any given \tilde{w} , from the perspective of the dictator, the challenger participates with probability η , where η is now endogenously given by

$$\eta = 1 - \Phi(\tilde{w}).$$

Now suppose that \tilde{w} decreases to $\tilde{w}' < \tilde{w}$ (η increases to $\eta' > \eta$), so the challenger is more likely to participate. Proposition 4 implies that in this case, competitive elections will be used more often, so b_E increases (at least up to the point where the constraint $b_E \leq \tilde{b}$ starts binding). In equilibrium, b_E is chosen such that the protest threshold in case of no elections is equal to the threshold z_{b_E} , which is the threshold if a competitive election is held for $b = b_E$. This implies the following: as η increases to η' , the threshold b_E increases to $b'_E > b_E$. The consequences for protest thresholds are the following. If the challenger runs, the threshold will remain at z_b for $b < b_E$, it will decrease from z_{b_E} to z_b for $b_E < b < b'_E$ (there were no elections before the change, but there are elections after), and it will decrease from z_{b_E} to $z_{b'_E}$ for $b > b'_E$. If the challenger does not run, the threshold will decrease from z_{b_E} to $z_{b'_E}$ for all b . This means that every value of b , the marginal decrease in the protest threshold is larger if the challenger does not run than if he runs, which makes him less willing to participate in the election.

We thus have shown that there is the following strategic interaction. The more likely the challenger is to run, the more willing the dictator is to allow competitive elections. But this, in turn, prompts the challenger to avoid running in elections, because uncontested elections will cast even more doubt on the popularity of the dictator. This suggests that even if the challenger makes his decision to run strategically, there is a unique equilibrium, and in this equilibrium, some challengers run and some do not.

Suppose that the benefit from participating in election increases, so the distribution Φ changes to Φ' , which first order stochastically dominates Φ . This may happen, for example, if the challenger is more compelled to run in elections, e.g., if there is a democratic tradition that gives him disutility if he does not participate in an election, or if his position as the leader of the opposition is not secure

and he will be sidelined in the future if he fails to participate now. This makes η higher for any fixed threshold \tilde{w} . As a result, the equilibrium will change in the following way: the dictator will be more likely to allow competitive elections and also more likely to repress, while the challenger will be more likely to run, even though the threshold \tilde{w} will increase as well. This suggests that if the country had a democratic tradition earlier, or if there is competition for the position of the opposition leader, then the country will see more competitive elections, but more repressions as well.

This subsection suggests that our results are robust to endogenizing the challenger's decision to participate in elections. More importantly, these results suggest that studying the behavior of the opposition and the dictator in an environment where the dictator possesses special incumbency advantages is tractable within our framework, and this may be an interesting avenue for future research.

4.2 Electoral Fraud

Our analysis so far assumed that in a competitive election there is no fraud, the vote tally is reported truthfully, and citizens are thus able to infer the dictator's true popularity, which is precisely the reason the dictator organized this election. The possibility of fraud could prevent the citizens from making this inference: consider, for instance, a dictator who always reports that he won with 99.9% (as was typical in the Soviet Union), and failure to do so is interpreted as extreme unpopularity and leads to major protests. We take the view that such elections with such unconstrained fraud may not be viewed as competitive, and in our classification this scenario would fall under 'absence of competitive elections'.

In many cases, however, technological or political constraints prevent reporting arbitrarily high numbers. For example, the presence of observers from opposition parties or using automated voting or counting systems may limit or even prevent fraud in some precincts. (Enikolopov et al., 2013, use a field experiment to show that the presence of trained observers in select precincts in Moscow lowered the vote share of the ruling party, United Russia, in the election to the Russian parliament in 2012, by at least 11 percentage points, from 47 to 36 percent on average.) In this Subsection, we argue that the possibility of fraud does not change our main results as long as the vote tally is informative about the dictator's true support. The possibility of fraud may make the dictator more likely to have competitive elections and also more likely to repress.

Consider a very simple model of electoral fraud. Suppose that if the incumbent gets share of

votes τ in the election, he is able to report any result up to $q(\tau)$, where $q(\tau)$ is a strictly increasing function and $q(\tau) \geq \tau$. For example, this function q may take a simple linear form: if share α of precincts report truthfully whereas the dictator has full control in the rest, the maximum he can report is

$$q(\tau) = \alpha\tau + (1 - \alpha).$$

If so, in equilibrium, the dictator will report $\hat{\tau} = q(\tau)$, the citizens will observe the value of $\hat{\tau}$ and infer $\tau = q^{-1}(\hat{\tau})$, and thus $b = \sigma_\delta F^{-1}(q^{-1}(\hat{\tau}))$. Thus, with such fraud, the voters are still able to make a perfect inference about the dictator's popularity. We see that even dictators who are sure to win may use fraud; the reason is that if they do not, citizens will believe that they are less popular and used fraud anyway (see also Simpson, 2013).

To understand the impact of fraud, suppose that we start with a scenario where fraud is impossible, and then introduce fraud. If in equilibrium the condition $b_E \leq \tilde{b}$ was not binding, then after fraud is introduced, all types of the dictator that have competitive elections will engage in fraud, but this will not change the voters' posterior beliefs about b , and thus it will not change the incentives to have elections, because every type of the dictator made an optimal choice to reveal b or not. In the opposite case, where the condition $b_E \leq \tilde{b}$ was binding, there are types of dictators with $b > \tilde{b}$ who wanted to reveal their type b in equilibrium, but could not do so through competitive elections because they would lose. Once fraud becomes possible, some of these dictators with b close to \tilde{b} , i.e., those who would lose by a narrow margin, are now able to use fraud to win elections. This implies that introducing fraud will enable more dictators to have competitive elections. But this also implies that those dictators who opt against having elections would be perceived as even weaker, and would face larger protests (this follows from the fact that the threshold z_{H_y} is quasiconvex in threshold y for any given level of repressions, as we show in the proof of Lemma 3 in the Appendix). Now, the repression threshold b_R will not remain constant; since dictators with b close to b_R now face larger protests if they fail to repress, their incentives to repress go up, and therefore b_R goes down in equilibrium.

We have thus shown that the possibility of fraud has no consequences for the incumbent's strategy to have elections or to repress in the case where all dictators who wanted to have competitive elections could do so. If, however, there are dictators who choose not to have competitive elections only because they are afraid of losing, then the possibility of fraud relaxes this constraint, thus enabling more competitive (though, needless to say, fraudulent) elections. On the flip side, a country where fraud is possible will also see more repressions.

5 Conclusion

In this paper, we considered a leader who wants to dissuade citizens from protesting by communicating that he has a sufficient number of supporters who prefer to keep him in power and that any protests are unlikely to succeed. We showed that competitive elections which are at least somewhat fair and thus informative about the dictator's true popularity may be used for this purpose. Similarly, refraining from costly repression is also a signal of confidence. We studied how the use of these strategies, to run in competitive elections and to repress opposition, interact with each other, and we obtained rich comparative statics results which provide testable implications for empirical studies of non-democratic regimes.

In particular, we show that competitive elections are more likely if the incumbent does not value office too much, if protesting is costly for citizens, or if opposition has incentives to challenge the dictator in elections. They are less likely to take place if the dictator values office significantly or repression is relatively cheap or easy to organize. For example, we expect countries rich in natural resources such as oil or totalitarian regimes to ban competitive elections and countries which are ruled by the military or which had a democratic tradition before falling into dictatorship to allow competitive elections more often. The countries where dictators do not allow competitive elections are also likely to be the countries where the opposition is repressed. One notable exception to this pattern are countries with a democratic tradition and/or competitive and well-organized opposition, so an opposition leader is compelled to run: in these countries, opposition is unlikely to be ignored by the incumbent, and we expect to see more competitive elections on the one hand and more repressions on the other.

Our model is simple and tractable but not without limitations. Most importantly, in our view, the interaction between the incumbent and the opposition is one-shot. It would be interesting to analyze how the incumbent's use of his strategies of staying in power changes throughout his tenure, how opposition leaders are selected, and what strategies they are likely to use after coming to power. Answering these questions requires studying a dynamic multiperiod model, and we believe this to be a promising direction for future research.

References

- Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin (2008) “Coalition Formation in Non-Democracies,” *Review of Economic Studies*, 75(4), 987-1010.
- Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin (2010) “Political Selection and Persistence of Bad Governments”, *Quarterly Journal of Economics*, 125(4): 1511-1575.
- Acemoglu, Daron, Mikhail Golosov, and Aleh Tsyvinski (2008) “Political Economy of Mechanisms” *Econometrica*, 76(3): 619-641.
- Acemoglu, Daron and Robinson, James (2001). “A Theory of Political Transitions,” *American Economic Review*, 91, 938-963.
- Acemoglu, Daron and Robinson, James (2006). *Political Origins of Dictatorship and Democracy*. MIT Press.
- Acemoglu, Daron, James Robinson, and Thierry Verdier, (2004) “Kleptocracy and Divide-and-Rule: A Model of Personal Rule”, *Journal of the European Economic Association Papers and Proceedings*, April–May 2004, Vol. 2, 162–192.
- Ashworth, Scott (2005) “Reputational Dynamics and Political Careers” *Journal of Law, Economics and Organization* 21 (2): 441–66.
- Ashworth, Scott and Ethan Bueno de Mesquita (2008) “Electoral Selection, Strategic Challenger Entry, and the Incumbency Advantage” *Journal of Politics* 70(4): 1006-1025.
- Besley, Timothy and Masayuki Kudamatsu, “Making Autocracy Work,” in *Institutions and Economic Performance*, ed. Elhanan Helpman (Cambridge, MA: Harvard University Press, 2008).
- Blaydes, Lisa (2006) *Who Votes in Authoritarian Elections and Why? Determinants of Voter Turnout in Contemporary Egypt*, mimeo.
- Blaydes, Lisa (2008) *Authoritarian Elections and Elite Management: Theory and Evidence from Egypt*, mimeo.
- Boix, Carles and Milan Svolik (2011), “The Foundations of Limited Authoritarian Government: Institutions and Power-sharing in Dictatorships,” mimeo.
- Brandt, Loren and Matthew Turner (2007) “The Usefulness of Imperfect Elections: The Case of Village Elections in Rural China,” *Economics and Politics*, 19(3): 453-480
- Bueno de Mesquita, Bruce, Smith, Alastair, Silverson, Randolph, and Morrow, James (2003) *The Logic of Political Survival*, MIT Press, Cambridge.
- Bueno de Mesquita, Ethan (2010) “Regime Change and Revolutionary Entrepreneurs,” *American Political Science Review* 104(3): 446-466.

- Carlsson, Hans, and Eric van Damme (1993) “Global Games and Equilibrium Selection,” *Econometrica* 61: 989-1018.
- Cornell, Svante E., and Fariz Ismailzade (2004) “Azerbaijan”, in Freedom House (eds) Nations in Transit 2004: Democratization in East Central Europe and Eurasia. Rowman & Littlefield Publishers, Inc: Oxford, UK.
- Cherepanov, Vadim, Timothy Feddersen, and Alvaro Sandroni (2013) Revealed Preferences and Aspirations in Warm Glow Theory, *Economic Theory* 54(3): 501-535.
- Debs, Alexandre (2010) Living by the Sword and Dying by the Sword? Leadership Transitions In and Out of Dictatorships, mimeo.
- Dye, Ronald A. (1985) “Disclosure of Nonproprietary Information,” *Journal of Accounting Research* 23(1): 123-145.
- 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.
- Egorov, Georgy, and Konstantin Sonin (2005) The Killing Game: A Theory of Non-Democratic Succession, mimeo.
- Egorov, Georgy, and Konstantin Sonin (2011) Dictators and Their Viziers: Loyalty-vs.-Competence Trade-off, *Journal of European Economic Association*, 9(5): 903-930.
- Enikolopov, Ruben, Vassily Korovkin, Maria Petrova, Konstantin Sonin, and Alexei Zakharov (2013) Field Experiment Evidence of Electoral Fraud in Russian Parliamentary Elections, *Proceedings of National Academy of Sciences*, 110(2):448-52.
- Fearon, James (2011) “Self-Enforcing Democracy,” *Quarterly Journal of Economics*, 126: 1661–1708.
- Feddersen, Timothy, Sean Gailmard, and Alvaro Sandroni (2009) Moral Bias in Large Elections: Theory and Experimental Evidence, *American Political Science Review* 103(2): 175-192.
- Gandhi, Jennifer and Ellen Lust-Okar (2009) “Elections under Authoritarianism,” *Annual Review of Political Science*, 12: 403–22.
- Gandhi, Jennifer, and Adam Przeworski (2007) “Authoritarian Institutions and the Survival of Autocrat”, *Comparative Political Studies*.
- Geddes, Barbara (2006) Why Parties and Elections in Authoritarian Regimes?, mimeo
- Geddes, Barbara (2009) How Autocrats Defend Themselves against Armed Rivals, mimeo.
- Gehlbach, Scott and Paul Keefer (2011) “Investment Without Democracy: Ruling-Party Institu-

- tionalization and Credible Commitment in Autocracies,” *Journal of Comparative Economics* 39 123–139.
- Gehlbach, Scott and Alberto Simpser (2014) “Electoral Manipulation as a Bureaucratic Control”, *American Journal of Political Science*, forthcoming.
- Howard, Marc and Philip Roessler (2006) “Liberalizing Electoral Outcomes in Competitive Authoritarian Regimes.” *American Journal of Political Science* 50.2 : 362-378.
- Kricheli, Ruth, Beatriz Magaloni, and Yair Livne (2011) “Taking to the Streets: Theory and Evidence on Protests under Authoritarianism,” mimeo.
- Kuhn, Patrick M. (2010) “Fraud, Coordination Problems, and Successful Post-Electoral Protests in Competitive Electoral Autocracies,” MPSA 2010 Annual Meeting Paper.
- Little, Andrew (2012) Fraud and Monitoring in Noncompetitive Elections, mimeo.
- Little, Andrew (2013) Rational Expectations and Electoral Fraud, *Quarterly Journal of Political Science*, 7: 249–283.
- Little, Andrew, Joshua Tucker, and Tom LaGatta (2013) Elections, Protest, and Alternation of Power, mimeo.
- Londregan, John and Andrea Vindigni (2006) “Voting as a Credible Threat,” mimeo.
- Lorentzen, Peter L. (2013) Regularizing Rioting: Permitting Public Protest in an Authoritarian Regime, *Quarterly Journal of Political Science*, 8: 127-158.
- Magaloni, Beatriz (2006) *Voting for Autocracy: Hegemonic Party Survival and Its Demise in Mexico* (2006) 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): 751-65.
- Martinez-Bravo, Monica, Gerard Padró i Miquel, Nancy Qian, and Yang Yao (2014) “Political Reform in China: Elections, Public Goods and Income Distribution,” mimeo.
- Miller, Michael (2010) Information Revelation and Policy Responsiveness in Electoral Authoritarian Regimes, mimeo.
- Myerson, Roger (2008) “The Autocrat’s Credibility Problem and the Foundations of the Constitutional State,” *American Political Science Review*, 102: 125–139.
- Ostaszewski, Adam J., and Miles B. Gietzmann (2008) “Value Creation with Dye’s Disclosure Option: Optimal Risk-shielding with an Upper Tailed Disclosure Strategy,” *Review of Quantitative Finance and Accounting*, 31(1): 1-27.
- Padró i Miquel, Gerard (2007) “The Control of Politicians in Divided Societies: The Politics of

- Fear,” *Review of Economic Studies* 74(4): 1259-1274.
- Padró i Miquel, Gerard and Pierre Yared (2012) The Political Economy of Indirect Control, *Quarterly Journal of Economics* 127(2): 947-1015, May 2012.
- Persson, Torsten, and Guido Tabellini (2000) *Political Economics: Explaining Economic Policy*. MIT Press: Cambridge.
- Persson, Torsten and Guido Tabellini (2009) Democratic Capital: The Nexus of Political and Economic Change, *American Economic Journal: Macroeconomics* 1(2):88-126.
- Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub, and Fernando Limongi (2000) *Democracy and development: Political institutions and well-being in the world*, Cambridge Studies in the Theory of Democracy, Cambridge Univ. Press.
- Przeworski, Adam (2009) Constraints and Choices: Electoral Participation in Historical Perspective, *Comparative Political Studies*, 42(4): 4-30.
- Rozenas, Arturas (2013) Forcing Consent: Information and Power in Non-Democratic Elections, mimeo.
- Shadmehr, Mehdi (2013) Repression Backfire, mimeo.
- 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 (2013) State Censorship, *American Economic Journal: Microeconomics*, forthcoming.
- Shadmehr, Mehdi, and Dan Bernhardt (2014) Vanguardism in Revolution: Sacrifice, Radicalism, and Coercion, mimeo.
- Simpser, Alberto (2013) *More than Winning: Why Parties and Governments Manipulate Elections*, Cambridge Univ. Press.
- Svolik, Milan (2008) Authoritarian Reversals and Democratic Consolidation, *American Political Science Review*, 102(2): 153-168.
- Svolik, Milan (2009) Power-sharing and Leadership Dynamics in Authoritarian Regimes, *American Journal of Political Science*, 53(2): 477-494.
- Wintrobe, Ronald (1990) The Tinpot and the Totalitarian: An Economic Theory of Dictatorship. *American Political Science Review*, Vol. 84, No. 3, September 1990, 849–872.
- Wintrobe, Ronald (1998) *The Political Economy of Dictatorship*, Cambridge Univ. Press.

Appendix

Lemma 5 Suppose agent i got $b_i = x$ and agent j got $b_j = y$, and $x > y$. Then G_x first-order stochastically dominates G_y : For any $\xi \in \mathbb{R}$ such that $0 < G(\xi) < 1$, we have $G_x(\xi) < G_y(\xi)$.

Proof. Let us prove that for two values of b_i , x and y such that $x > y$, G_x first-order stochastically dominates G_y (wherever $G(z) \in (0, 1)$).

We need to prove that $G_x(z)$ is decreasing in x for any fixed $z \in \mathbb{R}$ such that $G(z) \in (0, 1)$.

We have

$$G_x(z) = \frac{\int_{-\infty}^z \frac{1}{\sigma_\delta} f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)}{\int_{-\infty}^{+\infty} \frac{1}{\sigma_\delta} f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)} = \frac{1}{1 + \frac{\int_z^{+\infty} f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)}{\int_{-\infty}^z f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)}}.$$

This is decreasing in x if and only if

$$\ln\left(\frac{1}{G_x(z)} - 1\right) = \ln \int_z^{+\infty} f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi) - \ln \int_{-\infty}^z f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)$$

is increasing in x (for $0 < G_x(z) < 1$ the left-hand side is well-defined). Differentiating with respect to x , we get

$$\frac{\partial}{\partial x} \left[\ln\left(\frac{1}{G_x(z)} - 1\right) \right] = \frac{\int_z^{+\infty} \frac{1}{\sigma_\delta} f'\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)}{\int_z^{+\infty} f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)} - \frac{\int_{-\infty}^z \frac{1}{\sigma_\delta} f'\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)}{\int_{-\infty}^z f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)}.$$

But for normal distribution, $\frac{f'(a)}{f(a)}$ increasing in a , thus $\frac{f'\left(\frac{x-\xi}{\sigma_\delta}\right)}{f\left(\frac{x-\xi}{\sigma_\delta}\right)} > \frac{f'\left(\frac{x-z}{\sigma_\delta}\right)}{f\left(\frac{x-z}{\sigma_\delta}\right)}$ if $\xi > z$ and $\frac{f'\left(\frac{x-\xi}{\sigma_\delta}\right)}{f\left(\frac{x-\xi}{\sigma_\delta}\right)} < \frac{f'\left(\frac{x-z}{\sigma_\delta}\right)}{f\left(\frac{x-z}{\sigma_\delta}\right)}$ if $\xi < z$, and therefore

$$\begin{aligned} \frac{\int_z^{+\infty} f'\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)}{\int_z^{+\infty} f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)} &= \frac{\int_z^{+\infty} f\left(\frac{x-\xi}{\sigma_\delta}\right) \frac{f'\left(\frac{x-\xi}{\sigma_\delta}\right)}{f\left(\frac{x-\xi}{\sigma_\delta}\right)} dG(\xi)}{\int_z^{+\infty} f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)} > \frac{\int_z^{+\infty} f\left(\frac{x-\xi}{\sigma_\delta}\right) \frac{f'\left(\frac{x-z}{\sigma_\delta}\right)}{f\left(\frac{x-z}{\sigma_\delta}\right)} dG(\xi)}{\int_z^{+\infty} f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)} = \frac{f'\left(\frac{x-z}{\sigma_\delta}\right)}{f\left(\frac{x-z}{\sigma_\delta}\right)}, \\ \frac{\int_{-\infty}^z f'\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)}{\int_{-\infty}^z f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)} &= \frac{\int_{-\infty}^z f\left(\frac{x-\xi}{\sigma_\delta}\right) \frac{f'\left(\frac{x-\xi}{\sigma_\delta}\right)}{f\left(\frac{x-\xi}{\sigma_\delta}\right)} dG(\xi)}{\int_{-\infty}^z f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)} < \frac{\int_{-\infty}^z f\left(\frac{x-\xi}{\sigma_\delta}\right) \frac{f'\left(\frac{x-z}{\sigma_\delta}\right)}{f\left(\frac{x-z}{\sigma_\delta}\right)} dG(\xi)}{\int_{-\infty}^z f\left(\frac{x-\xi}{\sigma_\delta}\right) dG(\xi)} = \frac{f'\left(\frac{x-z}{\sigma_\delta}\right)}{f\left(\frac{x-z}{\sigma_\delta}\right)}. \end{aligned}$$

This proves that $\frac{\partial}{\partial x} \left[\ln\left(\frac{1}{G_x(z)} - 1\right) \right] > 0$, and thus $G_x(z)$ is decreasing in x for any z . ■

Lemma 6 Suppose that it is public information that $b \sim G$, and citizens i protests if and only if $b_i \in R$, where R satisfies the following: If $x > \frac{c}{\gamma}$, then $x \in R$; if $x < \frac{c}{\alpha}$, then $x \notin R$. Then the probability of success as perceived by citizen i with $b_i = x$,

$$\pi_{G_x} = \Pr(b + \delta_j \in R \mid b + \delta_i = x), \quad (18)$$

is increasing in x (strictly if G is not degenerate).

Proof. Consider the probability of success for a fixed and known value of b :

$$\hat{\pi}_R(b) = \Pr(b + \delta_j \in R) = \int_{x \in R} \frac{1}{\sigma_\delta} f\left(\frac{x-b}{\sigma_\delta}\right) dx. \quad (19)$$

Take two citizens with values b_i equal to x and y with $x > y$; we have

$$\pi_{G_x} = \int_{-\infty}^{+\infty} \hat{\pi}_G(b) dG_x(b)$$

and, similarly, for y . By Lemma 5, G_x first-order stochastically dominates G_y . Therefore, to prove that $\pi_{G_x} \geq \pi_{G_y}$, with strict inequality if G is not degenerate, it suffices to prove that $\hat{\pi}_G(b)$ is increasing in b .

To do this, consider the following cases. Suppose first $b < \frac{c}{\alpha}$. We can rewrite (19) as

$$\hat{\pi}_R(b) = 1 - F\left(\frac{\frac{c}{\alpha} - b}{\sigma_\delta}\right) + \frac{1}{\sigma_\delta} \int_{x \in R, \frac{c}{\alpha} < x < \frac{c}{\gamma}} f\left(\frac{x-b}{\sigma_\delta}\right) dx.$$

Since $b < \frac{c}{\alpha}$, then $x > b$ in the integral, thus $f\left(\frac{x-b}{\sigma_\delta}\right)$ is decreasing in its argument and thus increasing in b , and so $\hat{\pi}_R(b)$ is increasing in b .

Second, consider the case $b > \frac{c}{\gamma}$. Let us rewrite (19) as

$$\begin{aligned} \hat{\pi}_R(b) &= 1 - \int_{x \notin R} \frac{1}{\sigma_\delta} f\left(\frac{x-b}{\sigma_\delta}\right) dx \\ &= 1 - F\left(\frac{\frac{c}{\alpha} - b}{\sigma_\delta}\right) - \frac{1}{\sigma_\delta} \int_{x \notin R, \frac{c}{\alpha} < x < \frac{c}{\gamma}} f\left(\frac{x-b}{\sigma_\delta}\right) dx. \end{aligned}$$

Here, $x < b$ in the integral, so $f\left(\frac{x-b}{\sigma_\delta}\right)$ is increasing in its argument, and thus decreasing in b ; consequently, $\hat{\pi}_G$ is increasing in b in this case as well.

Finally, consider the case $\frac{c}{\alpha} < b < \frac{c}{\gamma}$. In this case, differentiating with respect to b under the integral (this is a valid operation here) yields

$$\begin{aligned} \frac{d\hat{\pi}_R(b)}{db} &= \frac{d}{db} \left(1 - F\left(\frac{\frac{c}{\alpha} - b}{\sigma_\delta}\right) + \frac{1}{\sigma_\delta} \int_{x \in R, \frac{c}{\alpha} < x < \frac{c}{\gamma}} f\left(\frac{x-b}{\sigma_\delta}\right) dx \right) \\ &= \frac{1}{\sigma_\delta} \left(f\left(\frac{\frac{c}{\alpha} - b}{\sigma_\delta}\right) - \int_{x \in R, \frac{c}{\alpha} < x < \frac{c}{\gamma}} \frac{1}{\sigma_\delta} f'\left(\frac{x-b}{\sigma_\delta}\right) dx \right) \\ &= \frac{1}{\sigma_\delta} \left(f\left(\frac{\frac{c}{\alpha} - b}{\sigma_\delta}\right) - \int_{x \in R, \frac{c}{\alpha} < x < b} \frac{1}{\sigma_\delta} f'\left(\frac{x-b}{\sigma_\delta}\right) dx - \int_{x \in R, b < x < \frac{c}{\gamma}} \frac{1}{\sigma_\delta} f'\left(\frac{x-b}{\sigma_\delta}\right) dx \right) \\ &> \frac{1}{\sigma_\delta} \left(f\left(\frac{\frac{c}{\alpha} - b}{\sigma_\delta}\right) - \int_{\frac{c}{\alpha}}^b \frac{1}{\sigma_\delta} f'\left(\frac{x-b}{\sigma_\delta}\right) dx \right) \\ &= \frac{1}{\sigma_\delta} \left(f\left(\frac{\frac{c}{\alpha} - b}{\sigma_\delta}\right) + f\left(\frac{\frac{c}{\alpha} - b}{\sigma_\delta}\right) - f(0) \right). \end{aligned}$$

The last term is obviously positive for $b = \frac{c}{\alpha}$ and $b = \frac{c}{\gamma}$. It is also positive for $b = \frac{1}{2} \left(\frac{c}{\alpha} + \frac{c}{\gamma} \right)$ by Assumption 2. Indeed, for such b , the last term is positive if and only if $2f \left(\frac{\frac{c}{\alpha} - \frac{c}{\gamma}}{2\sigma_\delta} \right) > f(0)$, which is equivalent to $\sigma_\delta > \frac{1}{2\sqrt{2\ln 2}} c \left(\frac{1}{\gamma} - \frac{1}{\alpha} \right) = \frac{1}{2\sqrt{2\ln 2}} \frac{c(\alpha - \gamma)}{\alpha\gamma}$. Since $\alpha > \gamma$, this follows from Assumption 2.

It remains to show that the last term is positive for $b \in \left(\frac{1}{2} \left(\frac{c}{\alpha} + \frac{c}{\gamma} \right), \frac{c}{\gamma} \right)$ (the case $b \in \left(\frac{c}{\alpha}, \frac{1}{2} \left(\frac{c}{\alpha} + \frac{c}{\gamma} \right) \right)$ is symmetric). Let $w = \frac{1}{2\sigma_\delta} \left(\frac{c}{\gamma} - \frac{c}{\alpha} \right)$ and $x = \frac{1}{\sigma_\delta} \left(\frac{1}{2} \left(\frac{c}{\alpha} + \frac{c}{\gamma} \right) - b \right)$. For a fixed w , consider the function $h(x) = f(x - w) + f(x + w)$. In terms of function h , we know that $h(0) > f(0)$ and $h(w) > f(0)$, and we need to show that $h(x) > f(0)$ for $x \in (0, w)$.

To do this, it suffices to show that h is quasiconcave on $(0, +\infty)$ for any w . We have

$$\begin{aligned} \frac{d}{dx} h(x) &= \frac{1}{\sqrt{2\pi}} \left((w - x) e^{-\frac{(x-w)^2}{2}} - (w + x) e^{-\frac{(x+w)^2}{2}} \right); \\ \frac{d^2}{dx^2} h(x) &= \frac{1}{\sqrt{2\pi}} \left(\left((w - x)^2 - 1 \right) e^{-\frac{(x-w)^2}{2}} + \left((w + x)^2 - 1 \right) e^{-\frac{(x+w)^2}{2}} \right). \end{aligned}$$

The derivative $\frac{d}{dx} h(x)$ equals zero if and only if

$$\frac{1 - t}{1 + t} - e^{-2w^2 t} = 0, \quad (20)$$

where $t = \frac{x}{w}$. If $w \leq 1$, the equation (20) has a unique solution $t = 0$, so $x = 0$ is a global maximum and h is quasiconcave. If $w > 1$, (20) has two nonnegative solutions: $t = 0$ corresponds to a local minimum $x = 0$ and another root $t = t^*$, to a global maximum $x = x^* = wt^*$; this proves the quasiconcavity of h on $(0, +\infty)$ in this case as well.¹⁷

We have thus shown that $h(x) > f(0)$ for $x \in (0, w)$, since this is true for $x = 0$ and $x = w$. This finishes the proof that $\frac{d\hat{\pi}_R(b)}{db} > 0$ for $b \in \left(\frac{c}{\alpha}, \frac{c}{\gamma} \right)$. Therefore, $\hat{\pi}_R(b)$ is strictly increasing in b on $b \in (-\infty, +\infty)$, and this completes the proof. ■

Proof of Proposition 1. Let us first consider the case of a degenerate distribution G , so assume b is known. In this case, citizens know $\hat{\pi}_G$, which equals π_{G_x} for any x . Thus, $\mathbb{E}U_p(b_i) \geq \mathbb{E}U_s(b_i)$

¹⁷Indeed (20) becomes 0 for $t = 0$ for any w . Differentiating the left-hand side yields

$$\frac{d}{dt} \left(\frac{1 - t}{1 + t} - e^{-2w^2 t} \right) = -\frac{2}{(1 + t)^2} + 2w^2 e^{-2w^2 t},$$

which equals zero if and only if $1 + t = \frac{1}{w} e^{w^2 t}$; thus, the derivative has no positive roots if $w \leq 1$ and a unique positive root if $w > 1$. This proves that for $w \leq 1$, (20) is a monotone function, and it is decreasing rather than increasing, because it is negative if t is large. This also proves that for $w > 1$, (20) has a unique positive root. Indeed, it has root because the left-hand side is positive for small t (if $w > 1$) and is negative for large t . If it had two roots $0 < t_1 < t_2$, then the derivative would have to equal zero at some points $t_3 \in (0, t_1)$ and $t_4 \in (t_1, t_2)$, but we just showed that it has only one root. Thus, the root t^* is unique, and the function changes its sign from positive to negative, i.e., it is a global maximum on $(0, +\infty)$.

if and only if $b_i \geq \frac{c}{(\alpha-\gamma)\hat{\pi}_G+\gamma}$. Therefore, there must exist a threshold $z = z_b \in (-\infty, +\infty)$ such that citizens with $b_i \geq z_b$ protest and those with $b_i < z_b$ do not.

Since $\hat{\pi}_G$ is given by (7), this threshold $z = z_b$ constitutes an equilibrium if and only if $Q(z) = 0$, where

$$Q(z) = z - \frac{c}{(\alpha - \gamma) \left(1 - F\left(\frac{z-b}{\sigma_\delta}\right)\right) + \gamma}. \quad (21)$$

By the Assumptions 1 and 2, $\frac{dQ(z)}{dz} > 0$, and thus $Q(z)$ is an increasing function of z . Indeed,

$$\frac{dQ(z)}{dz} = 1 - \frac{1}{\sigma_\delta} \frac{c(\alpha - \gamma) f\left(\frac{z-b}{\sigma_\delta}\right)}{\left((\alpha - \gamma) \left(1 - F\left(\frac{z-b}{\sigma_\delta}\right)\right) + \gamma\right)^2} > 1 - \frac{1}{\sigma_\delta} \frac{c(\alpha - \gamma)}{\gamma^2} \frac{1}{\sqrt{2\pi}} > 0,$$

because $\sqrt{2\pi} > 2\sqrt{2\ln 2}$. Moreover, $\lim_{z \rightarrow -\infty} Q(z) = -\infty$ and $\lim_{z \rightarrow +\infty} Q(z) = +\infty$. This means that there is exactly one value of $z = z_b$ such that $Q(z) = 0$. This proves that there exists a unique equilibrium.

Now consider the case where b is not an atom and is distributed with c.d.f. G . Let us show that the equilibrium must take the form of a threshold. Suppose that the set of values b_i such that citizens with these realizations protest in equilibrium is R_G . Citizen i protests if and only if

$$x \geq \frac{c}{(\alpha - \gamma) \pi_{G_x} + \gamma}. \quad (22)$$

Since $\pi_{G_x} \in [0, 1]$, citizens with $b_i > \frac{c}{\gamma}$ must protest and citizens with $b_i < \frac{c}{\alpha}$ must not (these types have a dominant strategy), thus $\left\{x : x > \frac{c}{\gamma}\right\} \subset R_G$ and $\left\{x : x > \frac{c}{\gamma}\right\} \cap R_G = \emptyset$. Therefore, Lemma 6 is applicable, which implies that π_{G_x} is increasing in x . Since the left-hand side of (22) is increasing in x and its right-hand side is decreasing in x , it must be that a citizen i protests if and only if $b_i \geq z_G$ for some z_G .

It remains to show that the equilibrium threshold exists and is unique. The threshold $z = z_G$ must satisfy $\tilde{Q}(z) = 0$, where

$$\tilde{Q}(z) = z - \frac{c}{(\alpha - \gamma) \left(1 - \int_{-\infty}^{+\infty} F\left(\frac{z-b}{\sigma_\delta}\right) dG_z(b)\right) + \gamma}. \quad (23)$$

Let us prove that

$$\frac{d}{dz} \int_{-\infty}^{+\infty} F\left(\frac{z-b}{\sigma_\delta}\right) dG_z(b) < \frac{1}{\sqrt{2\pi}\sigma_\delta}. \quad (24)$$

Notice that the following identity holds, due to integration by parts:

$$\begin{aligned} \int_{-\infty}^{+\infty} F\left(\frac{z-b}{\sigma_\delta}\right) dG_z(b) &= F\left(\frac{z-b}{\sigma_\delta}\right) G_z(b) \Big|_{b=-\infty}^{b=+\infty} + \int_{-\infty}^{+\infty} G_z(b) \frac{1}{\sigma_\delta} f\left(\frac{z-b}{\sigma_\delta}\right) db \\ &= \int_{-\infty}^{+\infty} G_z(b) \frac{1}{\sigma_\delta} f\left(\frac{z-b}{\sigma_\delta}\right) db. \end{aligned}$$

Using this last formula to differentiate with respect to the second inclusion of z (in $G_z(b)$), we have

$$\begin{aligned} \frac{d}{dz} \int_{-\infty}^{+\infty} F\left(\frac{z-b}{\sigma_\delta}\right) dG_z(b) &= \int_{-\infty}^{+\infty} \frac{1}{\sigma_\delta} f\left(\frac{z-b}{\sigma_\delta}\right) dG_z(b) + \int_{-\infty}^{+\infty} \left(\frac{\partial}{\partial z} G_z(b)\right) \frac{1}{\sigma_\delta} f\left(\frac{z-b}{\sigma_\delta}\right) db \\ &< \frac{1}{\sigma_\delta} \frac{1}{\sqrt{2\pi}} + 0 = \frac{1}{\sqrt{2\pi}\sigma_\delta}, \end{aligned}$$

where we used the fact that $G_z(b)$ is decreasing in b , as proved in Lemma 5. This proves (24),

which we now use to substitute the numeraire in

$$\frac{d\tilde{Q}(z)}{dz} = 1 - \frac{c(\alpha - \gamma) \frac{d}{dz} \int_{-\infty}^{+\infty} F\left(\frac{z-b}{\sigma_\delta}\right) dG_z(b)}{\left((\alpha - \gamma) \left(1 - \int_{-\infty}^{+\infty} F\left(\frac{z-b}{\sigma_\delta}\right) dG_z(b)\right) + \gamma\right)^2} > 1 - \frac{1}{\sigma_\delta} \frac{c(\alpha - \gamma)}{\gamma^2} \frac{1}{\sqrt{2\pi}} > 0.$$

This shows that $\tilde{Q}(z)$ is strictly increasing in z and the equilibrium threshold $z = z_G$ is unique. Its existence follows, as before, from that $\lim_{z \rightarrow -\infty} \tilde{Q}(z) = -\infty$ and $\lim_{z \rightarrow +\infty} \tilde{Q}(z) = +\infty$. Consequently, there is a unique equilibrium threshold z_G for any distribution G . ■

Proof of Proposition 2. If b is fixed, then treating Q (from (21)) as a function of z, c, γ, α, b , we get

$$\begin{aligned} \frac{\partial Q}{\partial c} &= -\frac{1}{(\alpha - \gamma) \left(1 - F\left(\frac{z-b}{\sigma_\delta}\right)\right) + \gamma} < 0, \\ \frac{\partial Q}{\partial \alpha} &= \frac{c \left(1 - F\left(\frac{z-b}{\sigma_\delta}\right)\right)}{\left((\alpha - \gamma) \left(1 - F\left(\frac{z-b}{\sigma_\delta}\right)\right) + \gamma\right)^2} > 0, \\ \frac{\partial Q}{\partial \gamma} &= \frac{cF\left(\frac{z-b}{\sigma_\delta}\right)}{\left((\alpha - \gamma) \left(1 - F\left(\frac{z-b}{\sigma_\delta}\right)\right) + \gamma\right)^2} > 0, \\ \frac{\partial Q}{\partial b} &= \frac{1}{\sigma_\delta} \frac{cf\left(\frac{z-b}{\sigma_\delta}\right)}{\left((\alpha - \gamma) \left(1 - F\left(\frac{z-b}{\sigma_\delta}\right)\right) + \gamma\right)^2} > 0. \end{aligned}$$

Moreover, we know from the proof of Proposition 1 that $\frac{\partial Q}{\partial z} > 0$. Consequently, $\frac{\partial z_b}{\partial c} > 0$, $\frac{\partial z_b}{\partial \gamma} < 0$, $\frac{\partial z_b}{\partial \alpha} > 0$, $\frac{\partial z_b}{\partial b} < 0$. If b is not known but is distributed as G , the same comparative statics with respect to c, α, γ follows by differentiating \tilde{Q} (from (23)) with respect to these variables (this is analogous) and using $\frac{\partial \tilde{Q}}{\partial z} > 0$, also established in the proof of Proposition 1.

Finally, consider two distributions of b , G_1 and G_2 , such that G_1 first-order stochastically dominates G_2 . Then we have

$$\begin{aligned} 1 - \int_{-\infty}^{+\infty} F\left(\frac{z-b}{\sigma_\delta}\right) d(G_1)_z(b) &= \int_{-\infty}^{+\infty} \left(1 - F\left(\frac{z-b}{\sigma_\delta}\right)\right) d(G_1)_z(b) \\ &> \int_{-\infty}^{+\infty} \left(1 - F\left(\frac{z-b}{\sigma_\delta}\right)\right) d(G_2)_z(b) = 1 - \int_{-\infty}^{+\infty} F\left(\frac{z-b}{\sigma_\delta}\right) d(G_2)_z(b), \end{aligned}$$

because $1 - F\left(\frac{z-b}{\sigma_\delta}\right)$ is a monotonically increasing function of b . We thus have $\tilde{Q}(z_1; G_2) < \tilde{Q}(z_1; G_1) = 0$. But $\tilde{Q}(z_2; G_2) = 0 > \tilde{Q}(z_1; G_2)$, and this implies $z_2 > z_1$. ■

Proof of Lemma 1. As shown in the proof of Proposition 1, the distribution G uniquely determines the threshold z_G and the probability of success $\hat{\pi}_G$, as well as perceived probabilities of success for all citizens, π_{G_x} for $b_i = x$. Let us show that there is a unique value b such that $z_b = z_G$. We know that z_G solves $\tilde{Q}(z_G) = 0$ and z_b solves $Q(z_b) = 0$; so z_G does not depend on b whereas z_b is decreasing in b by Proposition 2. Therefore, there is at most one value b such that $z_b = z_G$. Moreover, $\frac{c}{\alpha} < z_G < \frac{c}{\gamma}$: indeed, we have $\frac{c}{\alpha} \leq z_G \leq \frac{c}{\gamma}$ because citizens with $b_i < \frac{c}{\alpha}$ never protest and those with $b_i > \frac{c}{\gamma}$ always protest, since both parts contain a positive mass of citizens, it must be that $0 < \hat{\pi}_{G_b} < 1$ for any b , but this means that citizens with $b_i = \frac{c}{\alpha}$ and $b_i = \frac{c}{\gamma}$ are no longer indifferent and the inequalities are strict. From (21) it is easy to see that the function mapping b to solution z_b maps $(-\infty, +\infty)$ onto the entire interval $\left(\frac{c}{\alpha}, \frac{c}{\gamma}\right)$, and thus there exists a unique b such that $z_b = z_G$. Denote this value b_G^* . In what follows, we let $Q(z; b)$ be the value of function $Q(z)$ for a given value of b ; by definition of b_G^* , $Q(z_G; b_G^*) = 0$.

For $b = b_G^*$, we have $\hat{\pi}_G(b_G^*) = \hat{\pi}_{b_G^*}(b_G^*) \equiv \hat{\pi}_{b_G^*}$; this follows immediately from (7). Furthermore, $Q(z_G; b_G^*) = 0 = \tilde{Q}(z_G)$, and from (21) and (23) it follows that

$$\int_{-\infty}^{+\infty} F\left(\frac{z_G - \xi}{\sigma_\delta}\right) dG_{z_G}(\xi) = F\left(\frac{z_G - b_G^*}{\sigma_\delta}\right). \quad (25)$$

Therefore, from (7) and (8), we have

$$\hat{\pi}_G(b_G^*) = \hat{\pi}_{b_G^*} = 1 - F\left(\frac{z_G - b_G^*}{\sigma_\delta}\right) = 1 - \int_{-\infty}^{+\infty} F\left(\frac{z_G - \xi}{\sigma_\delta}\right) dG_{z_G}(\xi) = \pi_{z_G},$$

so $\pi_{z_G} = \hat{\pi}_G(b_G^*)$; in other words, the objective probabilities of success $\hat{\pi}_G(b_G^*) = \hat{\pi}_{b_G^*}$ indeed coincide with the belief of a citizen with $b_i = z_G$ if b were not revealed.

Take $b < b_G^*$. For such values of b , $z_b > z_G$, because z_b is decreasing in b . We then have

$$\hat{\pi}_G(b) = 1 - F\left(\frac{z_G - b}{\sigma_\delta}\right) > 1 - F\left(\frac{z_b - b}{\sigma_\delta}\right) = \hat{\pi}_b.$$

This means that for such b , the dictator would be better off if b is revealed. It remains to prove that $\pi_{z_G} > \hat{\pi}_G(b)$. Notice that $Q(z_G; b_G^*) = 0$ and $b < b_G^*$ imply $Q(z_G; b) < 0$ (this follows from the proof of Proposition 2); since $\tilde{Q}(z_G) = 0$, we have

$$\int_{-\infty}^{+\infty} F\left(\frac{z_G - \xi}{\sigma_\delta}\right) dG_{z_G}(\xi) > F\left(\frac{z_G - b}{\sigma_\delta}\right)$$

and thus

$$\hat{\pi}_G(b) = 1 - F\left(\frac{z_G - b}{\sigma_\delta}\right) < 1 - \int_{-\infty}^{+\infty} F\left(\frac{z_G - \xi}{\sigma_\delta}\right) dG_{z_G}(\xi) = \pi_{z_G},$$

so $\pi_{z_G} > \hat{\pi}_G(b)$.

If $b > b_G^*$, then, analogously, we get that $z_b < z_G$ and $\hat{\pi}_G(b) < \hat{\pi}_b$, so the dictator is better off concealing b , and $\pi_{z_G} < \hat{\pi}_G(b)$.

It remains to prove that $G(b_G^*) \in (0, 1)$. Suppose not; consider the case $G(b_G^*) = 0$ (the case $G(b_G^*) = 1$ is analogous). This means that $b \geq b_G^*$ in the support of the distribution G and thus in the support of the conditional distribution G_{z_G} , and consequently,

$$\int_{-\infty}^{+\infty} F\left(\frac{z_G - \xi}{\sigma_\delta}\right) dG_{z_G}(\xi) < \int_{-\infty}^{+\infty} F\left(\frac{z_G - b_G^*}{\sigma_\delta}\right) dG_{z_G}(\xi) = F\left(\frac{z_G - b_G^*}{\sigma_\delta}\right)$$

(the inequality is strict, because G is assumed to have no atoms and is therefore nondegenerate). But this contradicts (25), and the contradiction completes the proof. ■

Proof of Lemma 2. If such strategies are followed, then the share of votes that the dictator gets is given by (14). Consider a citizen with $b_i < 0$ who in equilibrium votes for the dictator, and suppose that he deviates to voting for C . For a citizen with an infinitesimal share of votes ε , this deviation will result in the dictator getting $\tau' = \tau(b) - \varepsilon$ votes, and other citizens observing τ' and concluding that the value of b is $b' = \sigma_\delta F^{-1}(\tau') > b$. As a result, the dictator gets fewer votes and this weakly decreases his chance of winning elections (weakly because he could, in principle, only have elections where he would win by a wide margin, and if citizens knew that this is his strategy, then a deviation by an infinitesimal citizen had zero chance to prevent him from winning). At the same time, all citizens except for the one who deviated choose strategies based on the cutoff $z_{b'}$ rather than z_b . Since $b' > b$, $z_b > z_{b'}$, and hence strictly more people participate in protests as a result of this deviation. Consequently, such a deviation by a citizen with $b_i < 0$ increases the chance that the dictator will leave office. It also does not affect this citizen's payoff from protesting, because he would not protest in any case. Hence, such deviation is not profitable.

If we consider a citizen with $b_i \geq 0$, we can similarly show that his deviation to voting for the dictator may only help the dictator win, and if the dictator wins, it makes citizen believe that b equals to $b' < b$ rather than the true value. Thus, fewer citizens protest, and this also reduces the likelihood that the dictator is removed from office. The deviating citizen may only switch from protesting to staying home, but not the other way around. In any case, this deviation is not profitable. Hence, this is an equilibrium.

Since in all cases except $b_i = 0$ a deviation made the citizen strictly worse off, and for $b_i = 0$, the citizen is indifferent, this equilibrium is in undominated strategies. ■

Proof of Lemma 3. Suppose that without competitive elections, b is distributed according to

some distribution H^* . Then there is a protest threshold $z_{H^*} \in \left(\frac{c}{\alpha}, \frac{c}{\gamma}\right)$ and, by Lemma 1, the dictator would prefer to have elections if and only if b satisfies $z_b \geq z_{H^*}$, i.e., when $b \leq y$ for some y . Consequently, the equilibrium decision to have elections must take the form of a threshold. Moreover, this threshold y must satisfy $\tau(y) \geq \tilde{\tau}$, because the opposite would imply that some dictators with b satisfying $\tau(b) < \tilde{\tau}$ have competitive elections and lose; this cannot happen in equilibrium because canceling elections yields strictly higher utility.

Consider the distribution $H_y(x)$ given by (15) for different y . Clearly, as $y \rightarrow -\infty$ or $y \rightarrow +\infty$, $H_y(x)$ pointwisely converges to the same distribution $G(x)$. Consider the function $s_y = z_{H_y}$; this function maps $[-\infty, +\infty]$ to $\left(\frac{c}{\alpha}, \frac{c}{\gamma}\right)$ and is continuous, therefore, its image is compact. In what follows, we show that it is strictly quasiconvex on the support of G and has a unique minimum which is interior.

It is straightforward to see that y such that $\tau(y) \geq \tilde{\tau}$ is an equilibrium threshold if and only if $z_y = z_{H_y}$: sufficiency follows from Lemma 1 and necessity follows immediately from continuity of all functions involved. Since the function $y \mapsto z_y$ maps $(-\infty, +\infty)$ onto $\left(\frac{c}{\alpha}, \frac{c}{\gamma}\right)$, we have $z_y < s_y$ for y high enough and $z_y > s_y$ for y low enough. Therefore, there exists y for which $z_y = s_y = z_{H_y}$, therefore, there is an equilibrium (provided that there is such y satisfying $\tau(y) \geq \tilde{\tau}$). If for all such y , $\tau(y) < \tilde{\tau}$, then \tilde{b} satisfying $\tau(\tilde{b}) = \tilde{\tau}$ is an equilibrium, because for all $b \leq \tilde{b}$, $z_y > s_y = z_{H_y}$, and thus the dictator prefers to have elections. Therefore, an equilibrium exists, and moreover, in the latter case, it is unique.

Take some value y for which $z_y = s_y$, and let us prove that s_y is quasiconvex with minimum achieved at y . First, take $y' > y$, and consider the distribution H' given by

$$H'(x) = \begin{cases} 0 & \text{if } x \leq y \\ \frac{G(x)-G(y)}{G(y')-G(y)} & \text{if } y < x \leq y' \\ 1 & \text{if } x > y' \end{cases} .$$

It is straightforward to verify that $H_y \equiv pH' + (1-p)H_{y'}$, where $p = (G(y') - G(y)) / (1 - G(y))$, and since $y' > y$, $p \in (0, 1)$. Now, we know that $z_{H_y} = z_y$. Now, the distribution H' first-order stochastically dominates the degenerate distribution concentrated in y , and by Proposition 2, $z_{H'} < z_y$. From this it follows that $z_{H_{y'}} > z_y$. Indeed, suppose, to obtain a contradiction, that $z_{H_{y'}} \leq z_y$. Then using the function $\tilde{Q}(z)$ defined by 23, we have $\tilde{Q}(z_y; H_y) = 0$, and also $\tilde{Q}(z_{H_{y'}}; H_{y'}) = 0$ and $\tilde{Q}(z_{H'}; H') = 0$, and thus $\tilde{Q}(z_y; H_{y'}) \geq 0$ and $\tilde{Q}(z_y; H') > 0$. This implies that from the standpoint of person with signal z_y , $\pi_{(H_{y'})_{z_y}} \geq \pi_{(H_y)_{z_y}}$ and $\pi_{(H')_{z_y}} > \pi_{(H_y)_{z_y}}$. At the

same time, for a given threshold z_y , $\pi_{G_{z_y}} = 1 - \int_{-\infty}^{+\infty} F\left(\frac{z_y - \xi}{\sigma_\delta}\right) dG_{z_y}(\xi)$ is linear in the distribution function, and thus satisfies $\pi_{(H_y)_{z_y}} = p\pi_{(H')_{z_y}} + (1-p)\pi_{(H_{y'})_{z_y}}$, a contradiction. Thus, $z_{H_{y'}} > z_y$.

In the case $y' < y$, let H'' be given by

$$H''(x) = \begin{cases} 0 & \text{if } x \leq y' \\ \frac{G(x) - G(y')}{G(y) - G(y')} & \text{if } y' < x \leq y \\ 1 & \text{if } x > y \end{cases} .$$

It is straightforward to verify that $H_{y'} \equiv pH'' + (1-p)H_y$, where $p = (G(y) - G(y')) / (1 - G(y'))$, and since $y > y'$, $p \in (0, 1)$. As before $z_{H_y} = z_y$. The degenerate distribution with an atom in y first-order stochastically dominates H'' , and by Proposition 2, $z_{H''} > z_y$. From this it follows that $z_{H_{y'}} > z_y$. Indeed, suppose, to obtain a contradiction, that $z_{H_{y'}} \leq z_y$. From Proposition 1, we have $\tilde{Q}(z_y; H_y) = 0$, and also $\tilde{Q}(z_{H_{y'}}; H_{y'}) = 0$ and $\tilde{Q}(z_{H''}; H'') = 0$, and thus $\tilde{Q}(z_y; H_{y'}) \geq 0$ and $\tilde{Q}(z_y; H'') < 0$. This implies that from the standpoint of person with signal z_y , $\pi_{(H_{y'})_{z_y}} \geq \pi_{(H_y)_{z_y}}$ and $\pi_{(H'')_{z_y}} < \pi_{(H_y)_{z_y}}$. But since $\pi_{G_{z_y}}$ is linear in G , we have $\pi_{(H_{y'})_{z_y}} = p\pi_{(H'')_{z_y}} + (1-p)\pi_{(H_y)_{z_y}}$, a contradiction. Thus, $z_{H_{y'}} > z_y$ in this case as well.

We have proven that any y such that $z_y = s_y$ is a unique global minimum of s_y , which proves uniqueness of such y . It is straightforward to see that $G(y) \in (0, 1)$; indeed, if $G(y) = 0$, then H_y first-order stochastically dominates the atom in y , and thus $s_y = z_{H_y} < z_y$, and if $G(y) = 1$, then, similarly, $s_y > z_y$; in either case $s_y \neq z_y$, a contradiction.

Let us prove that s_y is indeed quasiconvex; this would prove the result that s_{b_e} minimizes s_y over $(-\infty, \tilde{b}) \cap (\text{support of } G)$ even if $\tau(y) \geq \tilde{\tau}$ constraint is binding. Take $y' > \tilde{y} > y$ and let us show that $s_{y'} > s_{\tilde{y}}$. Since the equation $s_y = z_y$ has exactly one solution, we must have $z_{\tilde{y}} < s_{\tilde{y}}$. Consequently, $s_{\tilde{y}} = z_{\tilde{y}'}$ for some $\tilde{y}' < \tilde{y}$. Thus, in some vicinity of \tilde{y} , we have $y' > \tilde{y}'$. We then can use the same argument as before: for example, if $y' > \tilde{y}$, take H''' given by

$$H'''(x) = \begin{cases} 0 & \text{if } x \leq \tilde{y} \\ \frac{G(x) - G(\tilde{y})}{G(y') - G(\tilde{y})} & \text{if } \tilde{y} < x \leq y' \\ 1 & \text{if } x > y' \end{cases} .$$

As before, $H_{\tilde{y}} \equiv pH''' + (1-p)H_{y'}$, where $p = (G(y') - G(\tilde{y})) / (1 - G(\tilde{y}))$, and $p \in (0, 1)$. Now, we know that $z_{H_{\tilde{y}}} = s_{\tilde{y}'} = z_{\tilde{y}'}$. Now, the distribution H''' first-order stochastically dominates the degenerate distribution concentrated in \tilde{y}' , and by Proposition 2, $z_{H'''} < z_{\tilde{y}'} = z_{H_{\tilde{y}}}$. Suppose, to obtain a contradiction, that $z_{H_{y'}} \leq z_{\tilde{y}'} = z_{H_{\tilde{y}}}$. Consequently, $\tilde{Q}(z_{\tilde{y}'}; H_{\tilde{y}}) = 0$, and also $\tilde{Q}(z_{H_{y'}}; H_{y'}) = 0$ and $\tilde{Q}(z_{H'''}; H''') = 0$, and thus $\tilde{Q}(z_{\tilde{y}'}; H_{y'}) \geq 0$ and $\tilde{Q}(z_{\tilde{y}'}; H') > 0$. Then, as

before, from the standpoint of person with signal $z_{\tilde{y}'}$, $\pi_{(H_{y'})_{z_{\tilde{y}'}}} \geq \pi_{(H_{\tilde{y}})_{z_{\tilde{y}'}}}$ and $\pi_{(H''')_{z_{\tilde{y}'}}} > \pi_{(H_y)_{z_{\tilde{y}'}}}$. At the same time, we can again show that $\pi_{(H_{\tilde{y}})_{z_{\tilde{y}'}}} = p\pi_{(H''')_{z_{\tilde{y}'}}} + (1-p)\pi_{(H_{y'})_{z_{\tilde{y}'}}}$, a contradiction. Thus, $z_{H_{y'}} > z_{\tilde{y}'} = z_{H_{\tilde{y}}}$, and so $s_{y'} > s_{\tilde{y}}$. The case $y' < \tilde{y} < y$ is considered similarly, and this proves quasiconvexity. ■

Proof of Proposition 3. First, notice that for any y , z_y does not depend on η (η does not enter equation (21)). Let us show that for any y , z_{H_y} is decreasing in η . It suffices to prove that $H_{y;\eta}$ first-order stochastically dominates $H_{y;\eta'}$ if $\eta > \eta'$, i.e., that for any x , $H_{y;\eta}(x)$ is decreasing in η . Differentiating, we get

$$\begin{aligned} \frac{dH_{y;\eta}(x)}{d\eta} &= \frac{-G(x)(1-\eta G(y)) + (1-\eta)G(x)G(y)}{(1-\eta G(y))^2} \\ &= \frac{G(x)(G(y)-1)}{(1-\eta G(y))^2} < 0. \end{aligned}$$

if $x < y$, and, similarly,

$$\begin{aligned} \frac{dH_{y;\rho}(x)}{dt} &= \frac{-G(y)(1-\eta G(y)) + (G(x)-\eta G(y))G(y)}{(1-\eta G(y))^2} \\ &= \frac{G(y)(G(x)-1)}{(1-\eta G(y))^2} < 0 \end{aligned}$$

if $x \geq y$. This shows that if η increases, then z_{H_y} decreases. This means that, since b_E satisfied $z_{b_E} = s_{b_E}$, then after η increases to η' , we have $z'_{b_E} = z_{b_E} = s_{b_E} > s'_{b_E}$. This means that under η' , the equilibrium threshold that satisfies $z'_{b'_E} = s'_{b'_E}$, must also satisfy $b'_E > b_E$. Consequently, b_E is increasing in η .

To show that b_E satisfies $\tau(b_E) = \tilde{\tau}$ if η is close to 1, suppose not; then there is a limit point $\bar{b} = \lim_{\eta \rightarrow 1} b_E$. Then the distributions $H_{\bar{b};\eta}$ converge, in distribution, to a distribution with support on $(-\infty, \bar{b}]$. But at the same time, the degenerate distributions with atoms in b_E converge, in distribution, to one with atom in \bar{b} , which first-order stochastically dominates the former limit. This means that in the limit, $z_{\bar{b}} < z_{H_{\bar{b};\eta=1}}$, and this contradicts that $z_{b_E} = z_{H_{b_E(\eta);\eta}}$ for all η . This contradiction proves that $\tau(b_E) = \tilde{\tau}$ (i.e., $b_E = \tilde{b}$) for η sufficiently close to 1.

Conversely, if η approaches 0, then $b_E(\eta)$ is decreasing, and converges to some point b' . In this case, distributions $H_{y;\eta}$ converge G for any fixed y . This means that distributions $H_{b_E(\eta);\eta}$ converge to G , and thus $z_{H_{b_E(\eta);\eta}}$ converges to z_G . Similarly, $z_{b_E(\eta)}$ converge to $z_{b'}$. But $z_G = z_{b'_G}$ by definition of b'_G ; thus, $z_{b'_G} = z_{b'}$. This implies that $\lim_{\eta \rightarrow 0} b_E(\eta) = b' = b'_G$.

To demonstrate the comparative statics result with respect to c , take any c , and suppose that b_E is the threshold. At this threshold, $z_{b_E;c} = z_{H_{b_E};c}$. Now suppose that c increases, say, to $c' > c$.

Then, for a given b_E , both z_{b_E} and $z_{H_{b_E}}$ increase (see Proposition 2). However, $z_{H_{b_E}}$ will increase by a lower margin than z_{b_E} . Indeed, at b_E , $\frac{\partial Q(z)}{\partial c} = \frac{\partial \tilde{Q}(z)}{\partial c}$, but $\frac{\partial Q}{\partial z} < \frac{\partial \tilde{Q}}{\partial z}$, as follows from the proof of Proposition 1. Therefore, $\frac{\partial z_{b_E}}{\partial c} > \frac{\partial z_{H_{b_E}}}{\partial c}$, and this implies that $z_{b_E;c'} > z_{H_{b_E};c'}$ for $c' > c$. Thus, the equilibrium threshold under c' , b'_E which satisfies $z_{b'_E;c'} > z_{H_{b'_E};c'}$, must satisfy $b'_E > b_E$.

The arguments in case of α and γ are completely identical and are omitted. ■

Proof of Lemma 4. Let us first prove that the function $E(r)$ is nondecreasing in r for $r \in (-\infty, +\infty]$. Take two values, r and $r' > r$ (where r' may take the value of $+\infty$). Let $b_E = E(r)$ and $b'_E = E(r')$; it suffices to consider the case $b'_E > \tilde{b}$ (otherwise $b'_E \geq b_E$ holds automatically). Consider the function $H_y(x)$ given by (15) for $G = G^r$ and the function $H'_y(x)$ given by (15) for $G = G^{r'}$. We then have that for any y , the function H'_y first order stochastically dominates H_y ; indeed,

$$H_y(x) = \begin{cases} \frac{H'_y(x)}{H'_y(r)} & \text{if } x \leq r \\ 1 & \text{if } x > r \end{cases}.$$

This means that, using the notation from the proof of Lemma 3, that $s'_y < s_y$ for any y . We also know that $z_{b_E} = s_{b_E}$; in addition, $z_{b_E} = z'_{b_E}$, because if the value of b is known, the protest threshold does not depend on the distribution. Consequently, $s'_{b_E} < z'_{b_E}$, and this means that the equilibrium value b'_E under repression threshold r' satisfies $b'_E > b_E$. Thus, $E(r') > E(r)$, and this proves that $E(r)$ is a nondecreasing function; moreover, it is strictly increasing in the neighborhood of (b_E, b_R) whenever $b_E > \tilde{b}$. In addition, $E(r) < r$ for any $r \in (-\infty, +\infty]$, since Lemma 3 implies that $0 < G^r(E(r)) < 1$.

Now, if $k \geq A$, then (17) may not hold for any finite $R(y)$, and there is no repression for any election threshold b_E . In this case, $(b_E = E(+\infty), b_R = +\infty)$ is the unique pair of thresholds that constitute an equilibrium. If $k < A$, then the function R is given implicitly by (17), where $\tilde{z}^y(b) = z_{\min\{b,y\}}$ from (16), and since $\tilde{z}^y(x)$ is bounded and decreasing in x , $R(y)$ is uniquely defined for any y . Moreover, \tilde{z} is decreasing in y for $y < b$ and is constant for $y \geq b$, therefore, $R(y)$ is nonincreasing in y . Therefore, there is at most one equilibrium pair (b_E, b_R) that satisfies $b_E = E(b_R)$ and $b_R = R(b_E)$. In addition, we get that the function $R(E(r))$ is nonincreasing, which implies that there exists a unique equilibrium where b_R is the fixed point of $R(E(r))$ and $b_E = E(b_R)$.

It remains to note, for comparative statics results, that $R(y)$ is decreasing in the neighborhood of the equilibrium (b_E, b_R) . Indeed, (17) is satisfied for $y = b_E$ and $R(y) = b_R$, and then $b_E < b_R$ implies $y < R(y)$. Hence, $\tilde{z}^y(R(y)) = z_{\min\{R(y),y\}} = z_y$ in the neighborhood of (b_E, b_R) , and this

is strictly decreasing in y . Hence, (17) becomes

$$\int_{z_y - R(y)}^{+\infty} f\left(\frac{x}{\sigma\delta}\right) dx = \frac{k}{A}, \quad (26)$$

which implies that $R(y)$ is strictly decreasing in y . This completes the proof. ■

Proof of Proposition 4. If A increases or k decreases, then the right-hand side of (17) decreases; since the left-hand side is decreasing in $R(y)$, this implies that $R'(y) < R(y)$ for any y (where R' is the function after the change in A or k). At the same time, $E(r)$ does not change. Since R is strictly decreasing and E is nondecreasing, this leads to a decrease in b_R and also an decrease in b_E , unless $E(r)$ is constant in the neighborhood of b_R , i.e., unless $b_E = \tilde{b}$. The comparative statics result with respect to η is established similarly.

Suppose that c increases to $c' > c$, and suppose that this increase is small enough, so that $b'_R > b_E$ (considering a small deviation is sufficient). In this case, $E'(r) \geq E(r)$ for all r as follows from Proposition 3 (the inequality is strict if $b_E < \tilde{b}$). At the same time, $R'(y) > R(y)$, because z_y becomes higher for any given y by Proposition 2, and thus $R(y)$ must increase for any y , as follows from 26. This already implies that b_E decreases, strictly if $b_E < \tilde{b}$, and it remains to prove that b_R increases.

Suppose, to obtain a contradiction, that b_R does not increase, i.e., $b'_R \leq b_R$. In this case, consider the functions $s_y = z_{H_y}$ introduced in the proof of Lemma 3 and, similarly, s'_y for b'_R instead of b_R and c' instead of c . Since $b'_R \leq b_R$, then the distribution H_y weakly first-order stochastically dominates H'_y for any y ; in addition, for any distribution, the protest threshold increases because $c' > c$. This implies that for any y , $s'_y > s_y$. This, in turn, means that the same relation holds for minima as well, i.e., $\min_y s'_y > \min_y s_y$. However, by Lemma 3, these minima are achieved at the equilibrium values of b_E ; therefore, $s'_{b'_E} > s_{b_E}$. Since $b'_E < b'_R \leq b_R$ and $b_E < b'_R \leq b_R$, we have that the equilibrium protest thresholds are the same at b_R and b'_R for both c and c' (as the dictator does not have competitive elections in either case), and are equal to $s_{b_E} = z_{b_E}$ for c and to $s'_{b'_E} = z'_{b'_E}$ for c' . Now, $s'_{b'_E} > s_{b_E}$ implies that the equilibrium utility of a dictator with any $b > \max(b_E, b'_E)$ is higher under c' than under c , in case he decides not to repress. In particular, if under c , the dictator with $b = b_R$ was indifferent between repressing and not, he now strictly prefers not to repress, which implies $b'_R > b_R$. But this contradicts the assertion that $b'_R \leq b_R$. This contradiction proves that b_R must increase as c increases.

The proof in the case of α and γ is completely analogous and is omitted. ■