WORKING PAPER · NO. 2023-52

Why Did Putin Invade Ukraine? A Theory of Degenerate Autocracy

Georgy Egorov and Konstantin Sonin

APRIL 2023

Becker Friedman Institute | Political Economics Initiative

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

Why Did Putin Invade Ukraine? A Theory of Degenerate Autocracy*

Georgy Egorov
Northwestern University

Konstantin Sonin
University of Chicago

April 8, 2023

Abstract

Many, if not most, personalistic dictatorships end up with a disastrous decision such as Hitler's attack on the Soviet Union, Hirohito's government launching a war against the United States, or Putin's invasion of Ukraine in February 2022. Even if the decision is not ultimately fatal for the regime, such as Mao's Big Leap Forward or the Pol Pot's collectivization drive, they typically involve both a monumental miscalculation and an institutional environment in which better-informed subordinates have no chance to prevent the decision from being implemented. We offer a dynamic model of non-democratic politics, in which repression and bad decision-making are self-reinforcing. Repressions reduce the threat, yet raise the stakes for the incumbent; with higher stakes, the incumbent puts more emphasis on loyalty than competence. Our theory sheds light on the mechanism of disastrous individual decisions in highly institutionalized authoritarian regimes.

Keywords: nondemocratic politics, authoritarianism, dictatorship.

JEL Classification: P16, C73, D72, D83.

*The authors are grateful to Roger Myerson, Adam Przeworski, Barkley Rosser, and Jorgen Weibull for their helpful comments, and Sam Liberatore for excellent research assistance.

Introduction

The war that Russia launched against Ukraine in February 2022 must be an intellectual puzzle for theorists of authoritarian regimes. A naive rationalist analysis before the war almost inevitably concluded with an outcome that did not involve any actual warfare. Regardless of the estimates of parties' relative power, an actual war is associated with such costs for both sides that avoiding them seems to provide strong incentives to compromise (Fearon, 1995; Powell, 2000; Blattman, 2022). Not surprisingly, most public commentators who based their analysis on rationalist models were discussing Putin's "bluff" on the brink of the invasion. Yet the Russia-Ukraine war has fast become one of the bloodiest and costliest inter-state conflict post-World War II, involving the use of modern war tools at a scale unheard of in decades. Even more importantly, while it might take years to end the conflict and decades to fully evaluate the consequences, it is already clear that the decision to invade Ukraine caused a lot of harm to Putin and functionaries of his regime; with a high probability, it will be the regime's undoing (Gomza, 2022; Stoner, 2022).[1]

Putin's decision to invade Ukraine, a large European country with a recently modernized military, resilient political system, and well-developed sense of national identity, is just one of many dictators' decisions, which look unbelievably misguided in hindsight. Emperor Nicholas II's decision to enter the World War in 1914, in the absence of any threat to Russia's core security interests led to a revolution, the demise of the empire, the destruction of the elite of the ancient regime, and the death of the entire emperor's family (Lieven, 2015). Hitler's decision to invade USSR after failing to defeat Britain and then declaring war against USA, the world's largest industrial power, before defeating Soviets made the war unwinnable for Germany as early as in 1941 (Harrison,

[1] See also "Putin's Colossal Intelligence Failure", *The Russia File*, Kennan Center, March 12, 2022; "Hubris and isolation led Vladimir Putin to misjudge Ukraine", *The Washington Post*, April 12, 2022; "How Putin's War in Ukraine Became a Catastrophe for Russia", *The New York Times,* December 16, 2022; "Putin, Isolated and Distrustful, Leans on Handful of Hard-Line Advisers" , *The Wall Street Journal*, December 23, 2022; "What Russia Got Wrong", *Foreign Affairs*, February 8, 2023; "Putin Should Have Known His Invasion Would Fail", *Foreign Policy*, February 23, 2023.

1998; O'Brien, 2015). Saddam Hussein's decision to invade Kuwait in 1991 led to a highly predictable military defeat at the hands of an international coalition, reparation payments that stretch well over two decades, and, ultimately, Saddam's fall and execution (Karsh and Rautsi, 2007). The 1982 decision by General Leopoldo Galtieri, the leader of Argentina's military junta, to invade the Falkland Islands, led to a highly predictable military defeat, Galtieri's ouster, and the soon-to-follow demise of the regime (Lewis, 2002).

The disastrous decisions by top leaders of authoritarian regimes do not necessarily involve launching a war. For our theory, these are simply episodes that exemplify decisions that should have not been made by the respective leaders – on the modern explanations of war, we refer to excellent surveys in Fearon (1995), Jackson and Morelli (2011), and Herrera, Morelli and Nunnari (2019) and recent work on "democratic peace" (Debs and Goemans, 2010; Bueno de Mesquita and Smith, 2012; Weeks, 2012).[2] Economic campaigns might be equally ill-advised and self-defeating. In 1958-62, Mao's Great Leap Forward, a combination of economic reforms with a political campaign aimed to jump-start industrial development in then-predominantly agrarian China, led to mass famine, an economic disaster, and nearly cost Mao his political pre-eminence (Meng, Qian and Yared, 2015; Shih, 2022). What is surprising that the dictators that make these decisions are not ancient emperors whom their subjects cannot approach. Rather, they operate in institutionalized environments, with councils and advisors with presumably specialized expertise; their decisions are carried through by career professionals within structured hierarchies.

In this paper, we offer a model of non-democratic regimes which accounts for this apparent contradiction. This regime is both institutionalized and personalized at the same time. The leader's decisions are based on the input from his subordinates, yet the

[2]As it is well-documented in studies of "democratic peace", nearly almost wars since 1945 were initiated by dictatorships. Still, there is a prominent example of a modern war initiated by a functioning democracy, the U.S. invasion of Iraq in 2003. The decision to start the war relied, in part, on a dramatic underestimation of the future cost of the civil war and the U.S.-born cost of post-war reconstruction. However, from the pure military standpoint, the operation went according to the pre-war analysis. Despite the massive humanitarian cost for Iraqis, the war in no way threatened the US domestic stability in the way their own countries were affected by Galtieri's, Hussein's, or Putin's wars.

leader chooses the quality of advice that he receives himself – with the sole interest of keeping power. Specifically, the leader appoints an agent (or a council) that has better expertise and relies on the agent's advice while making the policy decision. Choosing the advisor, the leader could choose an agent of any information-processing quality: from an agent who knows whether or not the regime is vulnerable for sure to an agent who cannot make any distinction at all. This is an institutional choice that the leader makes.

The leader might lose power only if the regime is vulnerable and the policy choice, made basing on the agent's advice, is wrong. In such an environment, it seems to be a no-brainer to appoint an agent of the highest information-processing quality and to make, based on the informed advice, the correct policy choice. The problem is that an agent has another characteristic unobserved by the leader, the affinity with the opposition. Deciding what advice to provide to the incumbent leader, the agent weighs two factors: first, the vulnerability of the incumbent, and, second, the agent's own prospects under the new regime if the incumbent fails. Since the leader himself is not as well-informed as the advisor, the advisor has leverage: the advisor might tell the leader that he is not vulnerable, when he actually is. The wrong policy choice by a vulnerable leader will result in a change at the top.

In our model, the incumbent's vulnerability in the current period is an external shock. Yet other factors in the leader's decision-making are intertwined. The dictator's decision to repress opposition reduces the chances of a challenge, yet, simultaneously, raises the stakes in the future power struggles for him. Specifically, if the incumbent has repressed opposition in the past, then, once overthrown, he represents a more serious threat for the new leader – if not repressed, he might return to power in the future and, with his hard reputation, repress those who he just overthrown. A dethroned leader with a history of repression is more likely to be repressed than a leader with no repression in his past. This makes the leader who has already repressed in the past to be more likely to be repressed in the future. Thus, history has a bearing on the leader's choice of the level of repression and the quality of the advisor and, ultimately, on his survival prospects.

Worsening survival prospects generate a vicious cycle. Once the leader set on the repression path, the stakes become higher. The raising stakes – the fear to be tried and executed if dethroned – result in choosing the advisors with a lower information-processing capacity. Such advisors are more loyal in equilibrium – even with a high affinity with the opposition, they, having low ability to process information, are uncertain about the opposition's chances to oust the incumbent. Thus, they stay loyal. However, the quality of policy making with such advisors becomes worse. As a result, a fully rational, strategic dictator who has chosen to repress opposition to reduce the probability of a strong challenge ends up surrounded by low-quality subordinates and making low-quality policy choices.

What does our theory predict further concerning the dynamics of individual dictatorship?[3] A new leader might be surrounded by brilliant people when he comes to power, since he is (a) an outcome of equilibrium selection of contenders, and (b) conditional on becoming the dictator, he was able to overthrow the previous one. His tenure might follow either a bloody or peaceful path, depending on many circumstances, including both luck and rational decisions. Those dictators that stay in power long enough to witness their power fading away and become fearful to step down (e.g, fearing that some old scores could be settled) sacrifice achievements of the early years by replacing competent subordinates with the loyal ones and, ultimately, making bad policy.[4]

The one-period interaction in our model aligns with the now-standard theoretical intuition on the loyalty vs. competence trade-off in authoritarian regimes (Besley and Kudamatsu, 2009; Egorov and Sonin, 2011; McMahon and Slantchev, 2015; Zakharov, 2016; Kosterina, 2017; Hollyer, Rosendorff and Vreeland, 2018; Tyson and Smith, 2018). The presence of this trade-off has been confirmed empirically in Jia, Kudamatsu and Seim

[3]There is a recent surge in literature that focuses on formal models of authoritarian dynamics: see, e.g., Paine (2021, 2022); Meng and Paine (2022); Gratton and Lee (2023); see also Subsection 4.3.

[4]There are examples of autocrats appointing competent reformers closer to the end of their tenure: Francisco Franco technocratic reforms in 1957-59 laid foundation for the "Spanish economic miracle". Still, the systematic evidence points out to the relationship that our model predicts: towards the end of their tenure, authoritarians increasingly pursue bad policy. Jones and Olken (2005) used unexpected deaths of leaders as a source of exogenous variation to demonstrate that negative effects of individual leaders are strongest for unconstrained autocrats. Easterly and Pennings (2017) replicated, using an expanded data set, the Jones and Olken's results with respect to low growth episodes.

(2015); Bai and Zhou (2019); Shih (2022); Mattingly (2022).

In this paper, we do not directly discuss authoritarian power-sharing (Gandhi and Przeworski, 2006; Svolik, 2009, 2012; Powell, 2013; Francois, Trebbi and Xiao, 2023; Paine, 2021; Meng, Paine and Powell, 2022). In Svolik (2012) dichotomy of "authoritarian power-sharing" vs. "authoritarian control", our model is a theory of authoritarian control. Still, our model contributes to understanding of authoritarian power sharing as well. Specifically, relying on someone's privately obtained information is, effectively, sharing power. If a dictator chooses to allow media freedom as in Egorov, Guriev and Sonin (2009); Lorentzen (2014); Tyson and Smith (2018), or Shadmehr and Bernhardt (2011), then the autocrat's power is shared with whoever controls or influences the media. If the dictator relies, as he relies in our model, on his subordinates for the information, then the power is shared with these subordinates.

Finally, let us make two general points about the methodology. First, we want to point out that our game-theoretic approach to study of authoritarian regimes is not an antithesis to the quantitative and qualitative studies in political science, economics, sociology, history and other disciplines. (See, e.g., Przeworski, 2022 for the recent scathing critique.) Game-theoretic models of politics, by construction, are bound to be a simplistic representation of the reality. First and foremost, choosing a narrow set of actors and their possible actions, which is necessary to make a game-theoretic model tractable, invariably involves depriving all other potential actors of agency. For researchers that are unwilling to Occam-razor down details like public perception of the dictator, hierarchical relations binding citizens, religious, ideological, philosophical, or cultural history of the polity, etc., this results in grave doubts about the validity of the whole study. Still, there is a benefit in analyzing models whose main advantage is tractability. This particular paper demonstrates the critical complementarity between the emphasis on loyalty and repression of opposition, which results, dynamically, in a "degenerate autocracy". This could have been challenging to do, in a logically consistent way, without using game theory.

Second, we seek to expand the methodology of dynamic games beyond the now-standard Markovian approach to authoritarian dynamics (Bueno de Mesquita et al., 2003; Acemoglu, 2003; Acemoglu and Robinson, 2005; Paine, 2021). Concepts of path-dependence of economic processes and multiple equilibria have become inter-related since the pioneering work of Douglass North (North, 1981). Acemoglu and Robinson (2001) and Acemoglu (2003) have developed a workable framework for dynamic analysis of political processes (see also Lagunoff, 2009; Bai and Lagunoff, 2011). However, the reliance on Markov-type dynamic models limits the ability of these theories to explicitly focus on mechanisms of path-dependence. Indeed, any formal theory of this kind models path-dependence as multiple stable equilibria. Our focus on reputational concerns allows us to go beyond the existing models by explicitly demonstrating the workings of such a mechanism.

The rest of the paper is organized as follows. Section 2 introduces our theoretical model. Section 3 analyses decisions that the incumbent makes in one period. Section 4 studies the joint dynamics of repression and policy choice. Section 5 concludes.

2 Setup

We assume time to be discrete, $t = 1, 2, \ldots \infty$. The game is an infinite sequence of interactions between an incumbent leader, their subordinates, and the opposition. Each incumbent leader decides whether or not to repress the opposition, determines the quality of information to base his policy choice on, and makes the policy choice. Depending on whether or not the opposition is repressed, the policy choice, and an external shock, the incumbent might remain in power or lose to a challenger. The winner is determined as a result of a lottery with the odds determined endogenously: the odds depend on decisions made by the incumbent earlier. If the challenger has overthrown the incumbent, the former leader becomes the opposition.

The first thing that the winner of the power struggle – either the former incumbent or the former challenger – decides is whether or not to repress the opposition. If the opposition is not repressed, then the incumbent faces a challenge in the next period. If the opposition is repressed, then, with probability $1-\mu$, $\mu < 1$, there is no challenge. With the remaining probability, there is a new challenger from a set of potential challengers, despite the repressions.

If there is a challenge, the odds that the incumbent faces in the struggle for power are determined as follows. In each period, the incumbent is vulnerable with probability q, $q < 1$. In this situation, the challenger has a chance to take over – yet only if the leader chose a weak policy, $d = L$. A strong policy, $d = H$, guarantees the leader's survival. At the same time, the strong policy costs $C > 0$, while a weak policy is cost-less.

The incumbent dictator himself does not know whether or not he is vulnerable in this period, yet can gather information by appointing a lieutenant of competence $\theta \in [0, 1]$. Let v denote the state of the world, in which the incumbent is vulnerable; otherwise, the state of the world is n, $\Pr(v) = q$. If a lieutenant of competence θ is appointed, the lieutenant gets informative signal $S \in \{n, v\}$, $\Pr(S = v|v) = 1$ and $\Pr(S = n|n) = \theta$. That is, the lieutenant knows for sure that the dictator is vulnerable if the signal is $S = v$, yet the leader can be both vulnerable and safe if the lieutenant's signal is that the leader is not vulnerable, $S = n$.

If the dictator would have had full information about his vulnerability in a given period, he would prefer to choose the strong policy $d = H$ if and only if the state of the world is v. The lieutenant, however, may choose to betray the incumbent – that is, to misinform the leader about the signal he received, which changes the odds with which the incumbent defeats the challenger. (In any equilibrium, the incumbent would always follow the lieutenant's advice, so misinforming does change the odds.) The signal that the lieutenant received and the action, $d \in \{L, H\}$ that he recommended to the incumbent, becomes observable to the winner of the power struggle *ex post*.

If the dictator wins, the lieutenant gets wage w if he did not betray and suffers punishment $-\pi$ if he did. If the challenger wins, the lieutenant gets reward R. Thus, R is parameterizes the lieutenant's affinity with the current challenger. Values w and π are fixed and known to everyone; R is a random variable, which becomes known to the lieutenant before he makes decision. Assume that R is distributed on $(0, \infty)$ with c.d.f. $G(x)$, p.d.f. $g(x)$ such that $g(x) > 0$ and $g'(x) < 0$ for $x > 0$. For simplicity and without loss of generality, we assume that if the lieutenant is indifferent whether to obey or betray, he obeys the dictator. Lieutenants live for one period; it is possible to make lieutenants long-lived, yet this makes the model overburdened without bringing significant insights to the issues we focus on.

The incumbent i maximizes his life-time utility

$$U^i = \sum_{\tau=1}^{\infty} \beta^\tau U_\tau^i,$$

where U_t^i is the instantaneous utility player i receives in period t, $\beta < 1$ is the discount factor, common for all dictators. The winner of the power struggle gets one-period utility Y; the loser gets $-D < 0$ when repressed. When a player dies peacefully, which happens with probability δ, $K > 0$ is subtracted from their utility. We assume $K < D$, $\delta K < Y$. In all other circumstances, their one-period utility equals 0.

In each period t, the timing of the stage game is as follows.

1. The incumbent leader appoints a lieutenant of competence θ_t.

2. The incumbent faces a challenger with probability $\mu \in (0,1)$ if the opposition was repressed, and with probability 1 otherwise. If there is no challenger, the current incumbent remains in power and the game moves to stage 6.

3. If there is a challenger, the incumbent's lieutenant learns the realization of the signal about the incumbent's vulnerability, $S_t \in \{n, v\}$ and of his affinity with the challenger $R_t \in (0, +\infty)$. After that, the lieutenant chooses recommendation $d_t \in \{L, H\}$.

4. The outcome of the power struggle is determined, depending on the state of the world and the incumbent's decision based on the lieutenant's advice.

5. The winner of the power struggle decides whether or not to repress the opposition.

6. The opposition leader, if not repressed, and the incumbent die independently with probability $\delta \in (0,1)$. If only the winner dies, the loser (if she is alive) becomes the next dictator automatically, and faces a challenger next period. If only the loser

dies, a new challenger is chosen for the dictator. If both die, a new dictator and a new challenger are chosen. If none die, the winner is the incumbent of the next period, and the loser (if she exists) is the next challenger.

In a generic dynamic game, strategies may depend on the whole history, and this may produce a large number of equilibria. We restrict ourselves to symmetric equilibria where leaders' strategies, i.e., choices of the lieutenant's competence and the repression regime, may depend on two binary variables: whether or not the decision-maker has ever resorted to repressions before and whether or not the opposition leader, whose fate the winner is deciding, has ever resorted to repression.

Admittedly, our assumption that the players' strategies might depend on "reputation" state variables, the reputation of the winner and the reputation of the loser, is a short-cut. It is possible to do the same model with newcoming leaders having, with some probability, a commitment type that always represses the opposition, and other players' having uncertainty about whether or not the leader has this commitment type as in the canonical models of reputation (Kreps and Wilson, 1982; Milgrom and Roberts, 1982). Our approach allows to carry out the basic intuition and economize on notation and standard algebra.

Finally, we allow lieutenants' strategies to depend on whole histories; however, in equilibrium they will depend on values of S_t and R_t only. (In what follows, we skip the period index whenever it leads to no ambiguity.)

We relegate a full formal definition of strategies, as well as a full formal definition of equilibria we are interested in, to the Appendix. Here, we focus on essential details. Let us say that a leader has good reputation (G) if he has never resorted to repression before, and has bad reputation (B) otherwise. Thus, all possible combinations of the winner's and the loser's types belong to the set $\Lambda = \{(B,G),(G,B),(G,G)\}$.

Definition 1. *A tuple $(\alpha^*, \theta^*, d^*)$, where α^* and θ^* are mappings from Λ to $[0,1]$ and d^* is a mapping from $[0,1] \times \{s, \bar{s}\} \times [0,\infty)$ to $\{H, L\}$ is called an equilibrium if and only if*

(a) for any $(X,Y) \in \Lambda$, choosing repression with probability $\alpha^*((X,Y))$ is weakly optimal

for a winner with reputation X if loser's reputation is Y;

(b) *for any $(X,Y) \in \Lambda$, choosing a lieutenant with competence $\theta^*((X,Y))$ is weakly optimal for an incumbent with reputation X if challenger's reputation is Y;*

(c) *for a lieutenant of competence θ who received signal S and has learned the reward R, it is optimal to recommend the strong costly policy if and only if $d^*(\theta,S,r)=H$;*

(d) *for an incumbent that appointed a lieutenant of competence θ^*, it is optimal to follow the lieutenant's recommendation about the policy.*

3 Static Regime Formation

To make our analysis tractable, we split it into several steps. We start with analyzing the choice of the regime by the incumbent leader. First, we study the lieutenant's behavior, since it is least connected with past and future decisions of players. Then, treating lieutenants' behavior as given, we find dictator's optimal choice of lieutenant's competence; these will depend on dictator's expected continuation utilities, which we will for a moment treat as given. After analyzing the static institutional choice, we will characterize dictaror's utilities in the case of no repression and repression, again treating future behavior of all players, including himself, as given. This will allow us to find out in when players choose the repression regime. The next important step is to find dictators' best responses if they correctly predict future winners' decisions on repression, but consider competences of future lieutenants to be fixed arbitrarily. We will call the corresponding strategy profiles protoequilibria. Finally, to find equilibria, we check in which protoequilibria future dictators do hire lieutenants of the quality expected by the current incumbent.

3.1 The Information-Gathering Trade off

We begin by studying the behavior of a lieutenant of a fixed type θ who has received a noisy signal S about the incumbent's vulnerability and the value of potential reward R

from the challenger if the incumbent is overthrown. The agent betrays the dictator as far as he knows that his expected utility from betrayal exceeds that in the case of no betrayal. Both expectations are conditional on the agent's signal S and thus are functions of the agent's competence θ. To calculate them, the agent uses the Bayes formula:

$$\begin{aligned} \Pr(v|S=v) &= \frac{\Pr(S=v|v)\Pr(v)}{\Pr(S=v|v)\Pr(v)+\Pr(S=v|n)\Pr(n)} \\ &= \frac{q}{q+(1-\theta)(1-q)}. \end{aligned}$$

If the agent gets signal $S = n$, then the agent does not betray: as $\Pr(S = n|v) = 0$, it follows that $\Pr(v|S = n) = 0$. So, betrayal yields $-\pi < 0$ instead of $w > 0$, which the lieutenant gets if he did not betray.

If the agent gets signal $S = v$, then betrayal yields

$$R\frac{q}{q+(1-\theta)(1-q)} - \pi\frac{(1-\theta)(1-q)}{q+(1-\theta)(1-q)},$$

which is greater than w, the agent's utility if he does not betray the leader, if and only if

$$R > w + (1-\theta)\frac{1-q}{q}(w+\pi).$$

This gives us the following formal result.

Proposition 1. *If the lieutenant's signal is that the leader is not vulnerable, $S = n$, the lieutenant stays loyal regardless of the affinity with the opposition. If the signal is that the leader is vulnerable, $S = v$, the lieutenant stays loyal if and only if*

$$R < R^*(\theta) \equiv w + (1-\theta)\frac{1-q}{q}(w+\pi).$$

The threshold level $R^*(\theta)$ of reward that is required by the lieutenant of a fixed type θ, having received signal $S = v$, to "betray" the dictator, increases with the lieutenant's wage w, the level of punishment for treason π, and decreases with the ex ante probability of the dictator being vulnerable q. In particular, a more competent lieutenant, one with a higher θ, betrays the dictator for lower values of reward.

The intuition behind the comparative statics in Proposition 1 is as follows. When the lieutenant of competence θ receives the signal that the incumbent is vulnerable, he faces a trade-off between the probability of the reward from a victorious challenger and the probability of a punishment from a surviving incumbent. A lieutenant of high competence has a very precise signal: such a lieutenant knows with near-certainty that the incumbent is vulnerable. Thus, a competent lieutenant might accept a lower reward for misinforming the incumbent about the signal, which results in the poor policy choice, and, by doing so, the lieutenant guarantees the incumbent's loss.

Both a higher reward for remaining loyal and a higher punishment for the opposite increase the lieutenant's incentives to be loyal. An increase in q leads to an increase of the probability that conditions are favorable for the enemy, as perceived by the lieutenant. This, in turn, decreases lieutenant's fear of being punished, and makes him more likely to betray. Finally, though a smarter lieutenant receives a signal that the enemy is likely to win less frequently than a less competent one does, once he does, he is more sure that the enemy will win if he betrays, which also decreases his fear of punishment.

3.2 Equilibrium Choice of Loyalty vs. Competence

The dictator does not observe the affinity between the lieutenant and the challenger (the value of the reward for betrayal R), but knows its distribution. From the leader's standpoint, the probability of betrayal (i.e., recommending $d = L$ when $S = v$) conditional on the fact that the agent gets signal $S = v$ as

$$\Pr(d = L|S = v) = 1 - G\left(R^*\left(\theta\right)\right).$$

The probability of losing the struggle is therefore

$$\begin{aligned} p\left(\theta\right) &= \Pr(d = L, v) \\ &= \Pr(d = L, S = v, v) \\ &= \Pr(v)\Pr(d = L|S = v) \\ &= q\left(1 - G\left(R^*\left(\theta\right)\right)\right). \end{aligned}$$

Another issue that the dictator is concerned about is the policy cost. The lieutenant recommends costly policy $(d = H)$ with probability

$$\begin{aligned} r(\theta) &= \Pr(d = H) = \Pr(S = v)\Pr(d = H|S = v) + \Pr(S = n)\Pr(d = H|S = n) \\ &= (q + (1-\theta)(1-q))\, G(R^*(\theta)). \end{aligned}$$

It is easy to see that $p(\theta)$ is increasing, and $r(\theta)$ is decreasing with respect to θ.

Being aware of the constraints imposed by the lieutenant's possible disloyalty, the dictator faces the following maximization problem. Let his expected utilities of winning and losing in the current struggle be denoted by U and V, respectively. Let $p(\theta)$ be the probability of losing and $r(\theta)$ the probability of facing high costs as functions of lieutenant's competence θ. Then the dictator's optimization problem is

$$\max_{\theta}\{(1-p(\theta))\,U + p(\theta)\,V - r(\theta)\,C\}.$$

The dictator's solution of this maximization problem is given by the following proposition.

Proposition 2. *There is a unique equilibrium in the one-period-choice-of-a-lieutenant game. The dictator chooses a lieutenant characterized by θ^*, who is more competent (θ^* is high) when (a) the dictator is unlikely to be vulnerable (q is low); (b) the stakes are low for the dictator ($U - V$ is low); and (c) the measures that have to be taken are more costly (C is high).*

Basically, this proposition says that an insecure dictator, e.g., the one that fears that he will be executed upon removal from power, is bound to select less competent lieutenants. Indeed, as we know from Proposition 1, a more competent lieutenant is more likely to betray the incumbent. With higher stakes, loyalty, the flip side, in equilibrium, of competence, becomes relatively more important for the incumbent.

Proposition 2 is an important building block of our dynamic story: it shows that when the stakes for the incumbent leader are high, the leader chooses to select a less competent lieutenant, thus increasing the probability of a bad policy choice. When are the stakes high for the incumbent? When his reputation has deteriorated as a result of past repression decisions. So, repressing the opposition does reduce the probability of a challenge in the current period, yet has an endogenous opportunity cost – a leader with a bad reputation is bound to select less competent subordinates, increasing, ultimately, the probability of a fatal policy mistake.

4 Joint Dynamics of Repression and Bad Policy

In the previous section, we analyzed the leaders' choice of his information-gathering institution. Now, we are going to analyze how this choice and the policy choice evolve over time, responding to the leader's choice of the repressiveness of his regime.

To study formally the dynamics of an authoritarian regime, we write down the continuation values that correspond to different choices of the winner of a power struggle. As before, $(X,Y) \in \{(B,G),(G,B),(G,G)\}$ describes the history of the winner-loser pair at the point when the winner makes the repression decision. Then U^E_{XY} is the continuation value of the winner when the choice is to repress the opposition; U^S_{XY} – when the decision is not to repress the opposition, and $U_{XY} = \max\left(U^E_{XY}, U^S_{XY}\right)$ is the optimal choice. V_{XY} is the continuation value of the loser of the power struggle, and W_{XY} is the result of the optimal choice of the lieutenant's competence. It is straightforward to demonstrate, and this is done formally in the Appendix, that any of the values $U_{XY}, U^E_{XY}, U^S_{XY}, V_{XY}, W_{XY}$ is not greater than $-D$, and is smaller than $\frac{Y-\delta K}{1-(1-\delta)\beta}$, because a player may lose at most $-D$, and only once in his life, and likewise, he may not expect to get more than $Y - \delta K$ each period.

Now let us write down the equations that link these expected utilities to each other. Suppose for a moment that α_{XY}, the probability of repressions when the leader-opposition type is (X,Y), and p_{XY}, the probability that the power struggle is won, which is a function of θ, are given. Then utilities must satisfy the following conditions.

$$U_{XY} = \max\left(U^E_{XY}, U^S_{XY}\right) \tag{1}$$

$$U^E_{XY} = Y + (1-\delta)\,\beta\,((1-\mu)\,(Y + (1-\delta)\,\beta W_{BG} - \delta K) + \mu W_{BG}) - \delta K \tag{2}$$

$$U^S_{XY} = Y + (1-\delta)\,\beta\,((1-\delta)\,W_{XY} + \delta W_{XG}) - \delta K \tag{3}$$

$$\begin{aligned} V_{XY} &= (1-\alpha_{XY})\,(1-\delta)\,\beta\,((1-\delta)\,((1-p_{YX})\,V_{XY} + p_{YX}U_{YX}) + \delta W_{YG}) \\ &\quad - (1-\alpha_{XY})\,\delta K - \alpha_{XY} D \end{aligned} \tag{4}$$

$$W_{XY} = \max_\theta W_{XY}(\theta) = \max_\theta \{(1-p(\theta))\,U_{XY} + p(\theta)\,V_{YX} - r(\theta)\,C\} \tag{5}$$

The first equation simply says that the winner of the power struggle maximizes his expected utility when he decides whether or not to repress the opposition. If the opposition is repressed, he earns bad reputation, but his next opponent will necessarily have good reputation – as this would be the opponent's first entry. Following repression, this opponent will appear in the next period with probability μ, and after one period with probability $1-\mu$. If the opposition is not repressed, the reputation of the winner (the new does not change, and he will face a challenger (if he dies not die), who will be the same, unless the challenger dies (in this case, a new opponent with a good reputation emerges). The loser, in his turn, expects to be repressed with probability α_{XY}, and even if he is not, he may die with a certain probability. However, if he survives, he has a chance to regain power – either through a struggle, or simply because the winner peacefully dies himself. Finally, W_{XY} is just a weighted average of utilities from winning and from losing (accounting for costs of a policy choice, of course).

In the Appendix, we use real analysis to prove that if we fix any α_{XY} and p_{XY} in $[0,1]$, equations (1) – (5) have a unique solution $\left(U_{XY}, U^E_{XY}, U^S_{XY}, V_{XY}, W_{XY}\right)$ – 15 variables in total – which continuously depend on all parameters: α_{XY}, p_{XY}, β, δ, Y, C, K, D. This makes analysis of utilities if α_{XY} and p_{XY} very simple, and that is what we will start with. In the full game, however, values α_{XY} and p_{XY} are determined endogenously rather than fixed.

To form an equilibrium, they must satisfy the following conditions. First,

$$\alpha_{XY} \in \begin{cases} \{0\}, & \text{if } U_{XY}^{E} < U_{XY}^{S}; \\ [0,1], & \text{if } U_{XY}^{E} = U_{XY}^{S}; \\ \{1\}, & \text{if } U_{XY}^{E} > U_{XY}^{S}. \end{cases} \tag{6}$$

Second,

$$p_{XY} = p\left(\theta_{XY}\right), \tag{7}$$

where

$$\theta_{XY} = \arg\max_{\theta} W_{XY}\left(\theta\right) = \arg\max_{\theta}\left(\left(1 - p\left(\theta\right)\right) U_{XY} + p\left(\theta\right) V_{YX} - r\left(\theta\right) C\right),$$

(this complies with the definition of θ_{XY} above).

To proceed, it appears worthwhile to consider these conditions independently. If a set of utilities and α_{XY}'s satisfies (1) – (5) and (6) for a certain given set of p_{XY}'s, we call it a protoequilibrium. More precisely, we get the following definition.

Definition 2. *A vector of state-dependent probabilities of repression and odds of winning the power struggle* $\left(\alpha_{XY}, p_{XY}\right) \in [0,1]^6$ *is said to form a protoequilibrium, if utilities* $\left(U_{XY}, U_{XY}^{E}, U_{XY}^{S}, V_{XY}, W_{XY}\right)$*, uniquely identified by equations* (1) – (5)*, satisfy conditions on* α_{XY}*'s* (6).

With this definition, an equilibrium is a protoequilibrium for which (7) holds, because rationality of lieutenants is already incorporated in equations (1) – (5). Analytically, it is convenient to focus, for the time being, on protoequilibria, which allows us to take the probabilities of power transition as given. We proceed by characterizing the behavior of those who kept or won power in different situations that may be the case in (proto)equilibria, starting with the case, in which either the leader or the opposition leader already has a bad reputation.

Proposition 3. *Suppose that in a power struggle there is a politician with a good reputation and a politician with reputation of repression:*

(i) A politician with past experience of repression (B) *always represses opposition with no experience of repression* (G)*.*

(ii) A politician with no past experience of repression (G) represses an opposition leader with a bad reputation (B), provided that δ, the discount factor, is sufficiently small.

(iii) In each of these cases, lieutenants hired are of equal competence. More precisely, in any protoequilibrium, $\theta_{BG} = \theta_{GB}$, and these values do not depend on p_{XY}'s. Furthermore, in any equilibrium, $p_{BG} = p_{GB}$, and these values may be found as $p(\theta_{BG})$.

The dynamic mechanism behind this proposition is intuitive. A leader who has repressed in the past cannot undo the bad reputation. Therefore, he does not suffer any negative effects from repression, while repression results in a safe period with a positive probabilities. In terms of α's, this means $\alpha_{BG} = 1$, and therefore $V_{BG} = -D$. Now, if the loser is committed to repress the winner if spared and then returned to power at some point, he needs to be repressed. Naturally, the condition that the probability of exogenous death, δ, is sufficiently small ensures that it does not make sense not to repress the loser with a bad reputation in the hope that he will die on his own. As a consequence, the lieutenants chosen in both cases will have equal competence. The reason for this is the incumbent's understanding (in either of these cases) that he will not survive if he loses. For that reason, he is not interested in their chances of coming back to power, and therefore chooses his lieutenant without taking any of p_{XY}'s into account. Furthermore, it is easy to find from (7) that $p_{BG} = p_{GB}$ in any equilibrium. Denote them, for simplicity, by $\hat{p}$. It is also evident that W_{BG} are the same in all (proto)equilibria, and thus $U^E_{BG} = U^E_{GB} = U^E_{GG}$ are also the same. Denote the latter by $\hat{U}^E$.

4.1 Characterizing Repression

Our next goal is to find, whether or not it is possible that a winner with a good reputation represses the loser with a good reputation. This is critical for the stability of the good (no repression) equilibrium. Technically, the question is what values α_{GG} may take in an equilibrium. First, note that U^E_{GG} is known, since it equals U^E_{BG} (see(2)). It is easy to see that $\alpha_{GG} = 1$ always forms a protoequilibrium (and if p_{GG} equals p_{BG}, it is a real equilibrium). In other words, if one expects to be repressed regardless of his past actions, he opts to repression when he has a chance himself. Therefore, there always exists a 'repressive' equilibrium, which is uniquely defined by condition $U^E_{GG} > U^S_{GG}$. However, there may exist other equilibria. Consider the case $U^E_{GG} \leq U^S_{GG}$, or equivalently, $U_{GG} = U^S_{GG}$. Denote U^S_{GG} as a function of α_{GG} (holding all p_{XY}'s and $\alpha_{BG} = \alpha_{GB} = 1$ constant, and replacing $U_{GG} = U^S_{GG}$) with $U^S_{GG}(\alpha_{GG})$ in (1). In the Appendix, we prove that $U^S_{GG}(\alpha_{GG})$ is a continuous strictly decreasing function of α_{GG}, and $U^S_{GG}(1) < \hat{U}^E$. Now we formulate the following proposition.

Proposition 4. *The following is true about any (proto)equilibria:*

(a) For any p_{XY}'s there always exists a unique protoequilibrium in which $\alpha_{BG} = \alpha_{GB} = \alpha_{GG} = 1$. If p_{XY}'s are set to be equal to p^*, then it is an equilibrium, and moreover, it is the only equilibrium where all α_{XY}'s are equal to 1.

(b) If $U^S_{GG}(0) \geq \hat{U}^E$ then there is a protoequilibrium with $\alpha_{GG} = 0$ (the 'good' protoequilibrium), and if $U^S_{GG}(0) > \hat{U}^E$, then there also exists a mixed strategy protoequilibrium where $0 < \alpha_{GG} < 1$, and $U^S_{GG} = U^E_{GG} = \hat{U}^E$. There may exist at most one equilibrium such that $0 < \alpha_{GG} < 1$.

(c) For any given power transition probabilities p_{XY}, good (no repression) protoequilibria and good equilibria exist for a wide range of parameters when D is high and μ is high.

While Proposition 4 looks exceedingly technical, it is both important and intuitive. Part (a) simply states that it is always an equilibrium when every leader chooses repression after winning the power struggle. Indeed, if the current decision-maker expects to be repressed once out of power regardless of his reputation, choosing repression today does not have negative implications. Part (b) describes conditions, under which there exists a no-repression equilibrium; if such an equilibrium exists, there exist "intermediate equilibria", in which a string of no-repression periods might end with repression.

The intuition behind the comparative statics, part (c), is as follows. First, a higher μ (the lower effectiveness of repression) means that it is less profitable to repress, and high D means that gaining bad reputation is more dangerous. Both effects cause no-repression (proto)equilibria to exist for a wider range of other parameters.

In general, there are two basic channels through which the pay-off of the loser affects the incentives the decision-maker has. First, if the disutility of being removed from power decreases, then for the leaders that makes the decision at the moment, the costs associated with repressing the opposition and gaining bad reputation as a result decreases as well. Second, the current decision maker takes into account thoughts of the next decision maker (the next successful challenger), the one who will be deciding his fate once he loses. Since the next leader also faces lower costs of repressing the opposition, the reputation becomes less valuable for the current one. The impact of the discount factor is straightforward. A decrease in β makes the absence of challengers, which is not achievable if the opposition is not repressed, more valuable.

4.2 Equilibrium Paths and Competence

Though equilibria of the game may lead to a variety of different paths, as there are random shocks that influence the paths, we may delineate three substantially different paths, which correspond to different equilibria. In this paper, we specifically focus on the "degenerate autocracy", which corresponds to a set of equilibria that generates a specific path. In what follows, we show that these paths that feature frequent, even if not necessarily every-period, repression are robust. Then, we demonstrate that these paths are the worst in terms of the quality of information-gathering and decision-making.

The first group of paths are 'stable autocracies'. When a leader with no-repression reputation wins the power struggle, he does not repress the opposition. When he eventually loses power, he is not repressed and has an opportunity return to power. Thus, he faces relatively low stakes in the power struggles, which allows him to appoint competent subordinates, make good policy decisions with a higher probability, and have, as a result, a high probability of surviving the next power struggle.[5] New leaders that ap-

[5] Authoritarian regimes, in which the losers of political struggles we able to stage a comeback – that is, regimes that follow our "good path", are not exceptional. Of 54 leaders of Mexico in the 19th century,

pear because we allow for an exogenous death with a positive probability, play the same strategies. On this path, if we add a possibility of democratization, it does not result in behavioral change of the dictators prior to democratization.

Another group of stable paths are "consecutive degenerate autocracies". Every time a power struggle occurs, it is followed by repressions. Though some people might die peacefully, this does not lead to the escape from repression trap, in which the new-coming leader organizes repressions against the former regime as he fears their return to power. For our purposes, the most interesting dynamics is the complementarity between the repressiveness of the regime and the low quality of decision-making along this group of stable paths. A vulnerable dictator represses opposition to reduce the probability of a challenge in the next period. This raises the stakes for the leader as, after the repression, the probability that he will be repressed if (when) he loses power, increases. In response to the increasing stakes, the leader has to appoint more loyal, that is, less competent subordinates which results in a higher probability of a bad policy choice.

Our dynamic game has more equilibria. The rest of stable paths are "mixed": the incumbent and the opposition may swap their positions several times, or be replaced by newcomers, if they die by chance, but eventually the winner of the power struggle chooses repression rather than no repression. This is followed by a sequence of "repression only" periods: a winner with a history of repression represses a loser with no history of repression, and vice versa. Still, if either incumbent or challenger with a reputation of repression dies, the chain of repression may end, as now both the incumbent and opposition do not have history of repressions. Then, the story repeats.

On the "mixed" equilibrium paths the autocracy does not "stay degenerate". It is degenerate – repressive and incompetent – from time to time, probabilistically. Very much in line with the existing anecdotal evidence, the dictatorships become repressive and incompetent at later, rather than earlier, stages of the dictator's political life.

17 have held this positions more than one time, and 7 came back to power at least two times. General de Santa Anna, "the Napoleon of Mexico", came back at least 5 times; most of power changes were military coups. In Chile, General Ramon Freire came back 5 times. In Venezuela, among 56 changes in leadership, elected, military, and provisional in 1830-1910, there were 14 comebacks by 10 leaders who had been leaders before.

Our next Proposition 5 states formally the core result about the quality of governance on the equilibrium paths. It is straightforward to observe that on any equilibrium path where dictators always repress the opposition, the competence of lieutenants is the same $\theta_{BG} = \theta_{GB} = \theta_{GG}^{E}$. The interesting case, the one that illustrates our "degenerate autocracy" theory, requires comparing θ_{GG}^{E} and θ_{GG}^{S} in the case where both equilibria exist. As we shall see, the quality of governance – the competence of subordinates and, consequently, the probability of a wrong policy choice – is lower on those equilibrium paths that feature more repression.

Proposition 5. *In any equilibrium, the competence of subordinates satisfies the following comparative statics results.*[6]

(a) If the power struggle features at least one politician with history of repressions, the incumbent's choice of lieutenant's competence is the same regardless of the equilibrium played.

(b) If the power struggle features two politicians with no history of repressions, the incumbent's choice of lieutenant's competence in the good equilibrium is (weakly) higher than the choice of the lieutenant's competence in any mixed equilbrium, which is in turn (weakly) higher than the lieutenant's competence in any "consecutive degenerate autocracy" equilibrium.

(c) The probability that repression is effective, μ, has no effect on the competence of lieutenant in the good equilibrium. In any other equilibrium, the higher effectiveness of repression leads to less competent (and endogenously more loyal) lieutenants.

The intuition behind the comparison of government competence and, correspondingly, the probability of a wrong policy choice along different equilibrium paths was discussed above. The higher is the probability of repression, the higher are stakes for the dictator in the struggle. This forces him to look for more security in at the expense of

[6]The same is true for protoequilibria. In Appendix, we will prove this (and use this fact to prove the Proposition 5 itself).

spending additional resources, and the subordinates they hire are less competent and more loyal. Therefore, cruel and insecure dictators are more likely to choose poor policymakers in equilibrium.

4.3 Robustness

There are several modeling choices that require an additional discussion. First, while reputation plays an important role in the mechanism underlying the degenerate autocracy, our model of reputation is, effectively, a short-cut. Second, trying to go beyond the Markovian approach in modeling authoritarian dynamics, we restrict possible strategies to those that rely on the binary reputation variable. Finally, it might seem that the static model is limited by the fact that the leader chooses a single subordinate; more realistically, there should have been more than one.

Our choice to make "reputation" a binary variable instead of using a game of imperfect information with a commitment type following the pioneer work of Kreps and Wilson (1982) and Milgrom and Roberts (1982) has one goal – to simplify the algebra to be able to focus on the substantive issues. With the new leader having a type that is "bloodthirsty" (prefers to repress opposition regardless of circumstances) with some ex ante probability, each episode of repression increases the posterior probability. At some point, a new opponent, after overthrowing the incumbent who is highly likely to be the committed "bloodthirsty" type, will have to execute him, thus increasing own posterior probability to be "bloodthirsty". The resulting dynamics will be very similar to the one that we study, albeit at a high cost of additional notation and cumbersome algebra.

The Markovian approach has become a staple of studies in authoritarian dynamics, thanks to the combination of tractability and richness of potential strategic interactions (Tornell and Lane, 1999; Acemoglu and Robinson, 2001; Bueno de Mesquita et al., 2003; Acemoglu, 2003; Gallego and Pitchik, 2004; Acemoglu and Robinson, 2005; Jack and Lagunoff, 2006; McGillivray and Smith, 2006; Padró i Miquel, 2007; Robinson and Torvik, 2009; Svolik, 2009; Herrera and Martinelli, 2013; Leventoğlu, 2014). Still, for the subjects like ours assuming the Markov property creates a substantive problem: we do want to study the endogenous dynamic, in which changes in the previous periods lead to different choices in the subsequent ones. As a result, we opted for a limited extension: players' strategies depend not on whole theories but on binary variables that summarize the acquired reputation of the winner (the new incumbent) and the loser (the new opposition) of a power struggle.[7]

In Egorov and Sonin (2015), we allowed the winner's strategy to depend on the total number of periods in which he, the winner, and the loser have chosen repression in the past. As a result, there are equilibria, in which the probability that the winner of the power struggle who has already repressed in the past is growing monotonically for a finite number of periods. In fact, for a range of parameters, one might construct an equilibrium in which this probability strictly increases for any fixed number of periods. With power struggle lottery probabilities fixed, this is an extension of Proposition 4. However, the complexity of that model makes it impossible to study the core issue of the current paper, the dynamic relationship between repressiveness and quality of government. Egorov and Sonin (2015) is a model in which the power struggle probabilities are exogenous. Still, its results a robustness check for the inter-period part of our story.

In our basic one-period model, the leader appoints a single subordinate that gathers information for the leader to use in the power struggle. In Egorov and Sonin (2011), we considered a static model, in which the probability of the incumbent losing power depended, stochastically, on both the type and actions of the challenger. The current model does not have this complication. This allows, instead, to make the competence-loyalty model a construction block in a dynamic model, without which it is not possible to have a theory of degenerate autocracy. It is also possible to have the leader appointing a council of advisors, with advisors of, potentially, heterogeneous information-processing capac-

[7] It needs to be noted that, formally, nearly any game can be made Markov as defined in Fudenberg and Tirole, 1991, with a proper labeling of histories as states. In our case, if the reputation pairs are considered states, our equilibrium is Markov perfect. (A natural additional requirement would be "symmetry" that requires a different new winner from the sequence of infinitely many new winners to do the same in similar circumstances.) Still, we prefer to make a distinction between models in which the possible state include, e.g., the state of the economy as in Acemoglu and Robinson, 2001 – the incumbent's "vulnerability" is such a state in our model, and models like ours in which we allow strategies to rely on a pair of binary variables summarizing reputation.

ity. This would not change the qualitative results – higher-competence subordinates will still have stronger incentives to betray the incumbent. These results are available from authors upon request.

5 Conclusion

Many modern dictatorships end up with a disastrous, suicidal decisions such as Galtieri's attack on the Falklands in 1982 or Putin's invasion of Ukraine in 2022. Even if the disastrous decision is not ultimately fatal for the regime such as Mao's Big Leap Forward in China or Pol Pot's collectivization drive in Cambodia, they typically involve both a monumental miscalculation and an institutional environment in which subordinates have no chance to prevent the decision from being implemented. In this paper, we develop a theory of degenerate autocracy, a stage in an authoritarian regime life-cycle which is characterized by increased repressiveness and, simultaneously, deteriorating quality of decision-making. We show that these two tendencies reinforce each other. Repressions against political opponents increase stakes for the incumbent dictator, which in turn shifts his priorities from competence to loyalty. Our theory sheds light on governing mechanisms in repressive, inefficient authoritarian regimes.

References

Acemoglu, Daron. 2003. "Why Not a Political Coase Theorem? Social Conflict, Commitment, and Politics." *Journal of Comparative Economics* 31(4):620–652.

Acemoglu, Daron and James A Robinson. 2001. "A Theory of Political Transitions." *American Economic Review* 91(4):938–963.

Acemoglu, Daron and James A Robinson. 2005. *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.

Bai, Jinhui H and Roger Lagunoff. 2011. "On the Faustian Dynamics of Policy and Political Power." *Review of Economic Studies* 78(1):17–48.

Bai, Ying and Titi Zhou. 2019. ""Mao's Last Revolution": a Dictator's Loyalty — Competence Tradeoff." *Public Choice* 180(3):469–500.

Besley, Timothy and Masayuki Kudamatsu. 2009. *Making Autocracy Work*. Harvard University Press pp. 452–510.

Blattman, Chris. 2022. *Why We Fight: The Roots of War and the Paths to Peace*. Penguin Publishing Group.

Bueno de Mesquita, Bruce and Alastair Smith. 2012. "Domestic explanations of international relations." *Annual Review of Political Science* 15:161–181.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph Siverson and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.

Debs, Alexandre and Hein E Goemans. 2010. "Regime type, the fate of leaders, and war." *American Political Science Review* 104(3):430–445.

Easterly, William and Steven Pennings. 2017. "Shrinking dictators: how much economic growth can we attribute to national leaders?" *mimeo* .

Egorov, Georgy and Konstantin Sonin. 2011. "Dictators and Their Viziers: Endogenizing the Loyalty–Competence Trade-off." *Journal of the European Economic Association* 9(5):903–930.

Egorov, Georgy and Konstantin Sonin. 2015. "The Killing Game: A Theory of Non-democratic Succession." *Research in Economics* 69(3):398–411.

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* pp. 645–668.

Fearon, James. 1995. "Rationalist explanations for war." *International Organization* 49(3):379–414.

Francois, Patrick, Francesco Trebbi and Kairong Xiao. 2023. "Factions in nondemocracies: Theory and evidence from the Chinese Communist Party." *Econometrica* 91(2):565–603.

Fudenberg, Drew and Jean Tirole. 1991. *Game theory*. MIT press.

Gallego, Maria and Carolyn Pitchik. 2004. "An Economic Theory of Leadership Turnover." *Journal of Public Economics* 88(12):2361–2382.

Gandhi, Jennifer and Adam Przeworski. 2006. "Cooperation, Cooptation, and Rebellion under Dictatorships." *Economics and Politics* 18(1):1–26.

Gomza, Ivan. 2022. "The War in Ukraine: Putin's Inevitable Invasion." *Journal of Democracy* 33(3):23–30.

Gratton, Gabriele and Barton E Lee. 2023. "Liberty, security, and accountability: The rise and fall of illiberal democracies." *Review of Economic Studies* p. forthcoming.

Harrison, Mark. 1998. "The economics of World War II: an overview." *The economics of World War II: Six great powers in international comparison* pp. 1–42.

Herrera, Helios and César Martinelli. 2013. "Oligarchy, democracy, and state capacity." *Economic Theory* 52(1):165–186.

Herrera, Helios, Massimo Morelli and Salvatore Nunnari. 2019. "A theory of power wars.".

Hollyer, James R, B Peter Rosendorff and James Raymond Vreeland. 2018. *Transparency, Democracy, and Autocracy: Economic Transparency and Political (In) Stability*. Cambridge University Press.

Jack, William and Roger Lagunoff. 2006. "Dynamic Enfranchisement." *Journal of Public Economics* 90(4-5):551–572.

Jackson, Matthew O and Massimo Morelli. 2011. The reasons for wars: an updated survey. In *The handbook on the political economy of war*. Edward Elgar Publishing.

Jia, Ruixue, Masayuki Kudamatsu and David Seim. 2015. "Political Selection in China: The Complementary Roles of Connections and Performance." *Journal of the European Economic Association* 13(4):631–668.

Jones, Benjamin F. and Benjamin A. Olken. 2005. "Do Leaders Matter? National Leadership and Growth Since World War II." *Quarterly Journal of Economics* 120(3):835–864.

Karsh, Efraim and Inari Rautsi. 2007. *Saddam Hussein: A Political Biography*. Grove Atlantic.

Kosterina, Svetlana. 2017. "Ambition, personalist regimes, and control of authoritarian leaders." *Journal of Theoretical Politics* 29(2):167–190.

Kreps, David M and Robert Wilson. 1982. "Reputation and imperfect information." *Journal of Economic Theory* 27(2):253–279.

Lagunoff, Roger. 2009. "Dynamic stability and reform of political institutions." *Games and Economic Behavior* 67(2):569–583.

Leventoğlu, Bahar. 2014. "Social Mobility, Middle Class, and Political Transitions." *Journal of Conflict Resolution* 58(5):825–864.

Lewis, Paul H. 2002. *Guerrillas and generals: the" Dirty War" in Argentina.* Greenwood Publishing Group.

Lieven, Dominic. 2015. *The End of Tsarist Russia: The March to World War I and Revolution.* Penguin Publishing Group.

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

Mattingly, Daniel C. 2022. "How the Party Commands the Gun: The Foreign–Domestic Threat Dilemma in China." *American Journal of Political Science* forthcoming.

McGillivray, Fiona and Alastair Smith. 2006. "Credibility in compliance and punishment: Leader specific punishments and credibility." *The Journal of Politics* 68(2):248–258.

McMahon, R Blake and Branislav L Slantchev. 2015. "The guardianship dilemma: Regime security through and from the armed forces." *American Political Science Review* 109(2):297–313.

Meng, Anne and Jack Paine. 2022. "Power sharing and authoritarian stability: How rebel regimes solve the guardianship dilemma." *American Political Science Review* 116(4):1208–1225.

Meng, Anne, Jack Paine and Robert Powell. 2022. "Authoritarian Power Sharing: Concepts, Mechanisms, and Strategies." *Annual Review of Political Science* 26.

Meng, Xin, Nancy Qian and Pierre Yared. 2015. "The Institutional Causes of China's Great Famine, 1959–1961." *Review of Economic Studies* 82(4):1568–1611.

Milgrom, Paul and John Roberts. 1982. "Predation, reputation, and entry deterrence." *Journal of Economic Theory* 27(2):280–312.

North, Douglass C. 1981. *Structure and Change in Economic History.* New York: Norton.

O'Brien, Phillips Payson. 2015. *How the War Was Won: Air-Sea Power and Allied Victory in World War II.* Cambridge Military Histories Cambridge University Press.

Padró i Miquel, Gerard. 2007. "The Control of Politicians in Divided Societies: The Politics of Fear." *Review of Economic Studies* 74(4):1259–1274.

Paine, Jack. 2021. “The Dictator’s Power-Sharing Dilemma: Countering Dual Outsider Threats.” *American Journal of Political Science* 65(2):510–527.

Paine, Jack. 2022. “Reframing the Guardianship Dilemma: How the Military’s Dual Disloyalty Options Imperil Dictators.” *American Political Science Review* 116(4):1425–1442.

Powell, G. B. 2000. *Elections as instruments of democracy: Majoritarian and proportional visions.* Yale University Press.

Powell, Robert. 2013. “Monopolizing Violence and Consolidating Power.” *Quarterly Journal of Economics* 128(2):807–859.

Przeworski, Adam. 2022. “Formal Models of Authoritarian Regimes: A Critique.” *Perspectives on Politics* pp. 1–10.

Robinson, James A and Ragnar Torvik. 2009. “A Political Economy Theory of the Soft Budget Constraint.” *European Economic Review* 53(7):786–798.

Shadmehr, Mehdi and Dan Bernhardt. 2011. “Collective Action with Uncertain Payoffs: Coordination, Public Signals, and Punishment Dilemmas.” *American Political Science Review* pp. 829–851.

Shih, Victor. 2022. *Coalitions of the Weak.* Cambridge Studies in Comparative Politics Cambridge University Press.

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

Svolik, Milan W. 2009. “Power Sharing and Leadership Dynamics in Authoritarian Regimes.” *American Journal of Political Science* 53(2):477–494.

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

Tornell, Aaron and Philip R Lane. 1999. “The Voracity Effect.” *American Economic Review* 89(1):22–46.

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

Weeks, Jessica L. 2012. “Strongmen and straw men: Authoritarian regimes and the initiation of international conflict.” *American Political Science Review* 106(2):326–347.

Zakharov, Alexei V. 2016. “The Loyalty-Competence Trade-off in Dictatorships and outside Options for Subordinates.” *Journal of Politics* 78(2):457–466.

Appendix

Formal Setup. There is an infinite sequence of identical players $i = 1, 2, \ldots \infty$ who join the active part of the game in this sequence. In each period t, one player, denoted by $D_t \in \mathbb{N}$, is the incumbent dictator in this period. There may also be or not be a contender. We write $C_t \in \mathbb{N}$ to denote the contender's number in the sequence if there is one, and $C_t = 0$ if there is none. Initially, $D_1 = 1$ and $C_1 = 2$.

For each period t, denote the least of the identities of players who have not joined the active part of the game yet by N_t (for instance, $N_1 = 3$). Let W_t and L_t denote the winner and the loser, respectively.

Denote the instantaneous utility player i receives in period t by $U_t(i)$. We assume that if $i \neq D_t$ and $i \neq C_t$, then $U_t(i) = 0$. In other words, only actively participating players can get a non-trivial utility in the current period. At each period t, agent i (actually, only agent W_t) maximizes his utility $U(i) = \sum_{\tau=1}^{\infty} \beta^\tau U_\tau(i)$, where $\beta < 1$ is the discount factor, common for all agents. In each period t, the sequence of actions and events is as follows.

1. If $C_t \neq 0$, then the contender attempts to become the dictator. If $C_t = 0$, then $W_t = D_t$, $L_t = C_t = 0$, $C_{t+1} = N_t$, and $N_{t+1} = N_t + 1$, and in this case steps 2 – 4 are skipped.

2. The struggle breaks out, and the contender wins with probability $0 < p_t < 1$, which is determined endogenously. In other words, $\Pr(L_t = D_t) = \Pr(W_t = C_t) = p_t$, and $\Pr(W_t = D_t) = \Pr(L_t = C_t) = 1 - p_t$.

3. W_t decides on his action A_t, whether to execute ($A_t = E$) or spare ($A_t = S$) the loser L_t.

4. If $A_t = E$, then $U_t(L_t) = -D$, and with probability $\mu < 1$ there is still a successor in the next period ($C_{t+1} = A_t$ and $A_{t+1} = A_t + 1$), and with probability $1 - \mu$ there is no successor ($C_{t+1} = 0$). If $A_t = S$, then $U_t(L_t) = 0$, and $C_{t+1} = L_t$.

5. The winner gets $U_t(W_t) = Y$, and becomes the next dictator, i.e. $D_{t+1} = W_t$.

6. D_{t+1} and C_{t+1} (if there is one) die independently with probability δ. If only the contender dies, he is replaced by the next one in the line. If the dictator dies, then the contender (if any) becomes the dictator, and the next player becomes the contender. If both die, or the dictator with no contender dies, then a pair of players become the dictator and the contender for the next period.

Players get one-period utility Y if they are the winners of the period and the chosen policy was cost-less (see below), and $Y - C > 0$ if they are the winners, but the policy was strong. They get $-D < 0$ if they are repressed at this period. When they die peacefully, $K < D$ is subtracted from their utility. In all other circumstances, their one-period utility equals 0.

The lieutenants' strategies may depend on his signal S and value R only. We call a player good (G) if he has never killed before, and bad (B) if he ever has killed. Now we formally define the equilibrium concept of the game. We restrict to Markovian strategies (depending on the state of the world only), though the game is not truly Markovian, since the identity of the decision-maker (player) may change over time without being included in the state variables. Moreover, in the definition, we confine ourselves to one-shot deviations only (this is without loss of generality, Fudenberg and Tirole, 1991).

Definition 3. *A strategy profile consists of*

(a) winner's loser's-fate function $\alpha^* : \{G, B\} \times \{G, B\} \longrightarrow [0, 1]$, *where* $\alpha^*_t(X, Y)$ *is the probability with which the winner of reputation* X *represses the loser with reputation* Y *in period* t

(b) incumbent's choice-of-lieutenant function $\theta^* : \{G, B\} \times \{G, B\} \longrightarrow [0, 1]$, *where* $\theta^*(X, Y)$ *is the competence of the lieutenant chosen by the winner of reputation* X *that expects to meet a contender with reputation* Y

(c) lieutenant's decision function $d^* : [0, 1] \times \{s, \overline{s}\} \times [\overline{r}, +\infty) \longrightarrow \{H, L\}$ *such that for any* $\theta \in [0, 1]$, $d^*(\theta, \cdot, \cdot)$ *is lieutenant's decision on a recommended policy, if his competence is* θ.

We define equilibrium in 'Markovian' strategies in the following way.

Definition 4. *A strategy profile* $(\alpha^*, \theta^*, d^*)$ *is called an equilibrium if and only if*

(a) In any decision node where a winner of reputation X *determines the fate of a loser with reputation, executing with probability* $\alpha^*(X, Y)$ *weakly dominates other options;*

(b) In any decision node where an incumbent of reputation X *chooses a lieutenant to counter a contender with reputation* Y*, choosing lieutenants with competence* $\theta^*(X, Y)$ *is weakly dominant;*

(c) In any decision node where a lieutenant of competence θ *suggests the policy choice, having received signal* S *and expecting reward* R*, he finds choosing* $d^*(\theta, S, R)$ *weakly optimal, and if he is indifferent,* $d^*(\theta, S, R) = H$*;*

(d) For an incumbent that appointed a lieutenant of competence θ^**, it is optimal to follow the lieutenant's recommendation about the policy.*

Definition 5. *The tuple* $\left(\alpha^*_{BG}, \alpha^*_{GB}, \alpha^*_{GG}, \theta^*_{BG}, \theta^*_{GB}, \theta^*_{GG}, d^*(\theta, S, R)\right)$*, where* α^*_{XY} *and* θ^*_{XY} *are numbers on* $[0,1]$ *for* $(X, Y) \in \{(B,G),(G,B),(G,G)\}$*, and* d^* *is a mapping from tuple* (θ, S, R) *of lieutenant's competence* $\theta \in [0,1]$*, signal* $S \in \{s, \overline{s}\}$ *and stochastic part of lieutenant's reward* R *into the recommended policy* $d \in \{H, L\}$*, is called* equilibrium *if and only if the following conditions are satisfied:*

(i) When making his decision on the recommended policy, a lieutenant of competence θ *chooses* $d = d^*(\theta, S, R)$ *after having received signal* S *and learned the stochastic part of reward* R[8]*;*

(ii) Given the strategies of all other people, when dictator with reputation X *faces contender with reputation* Y*, he opts to hire an lieutenant of competence* θ^*_{XY} *(for all* $XY \in \{BG, GB, GG\}$*);*

(iii) Given the strategies of all other people, when winner with reputation X *determines the fate of the loser with reputation* Y*, he represses him with probability* α^*_{XY} *(for all* $XY \in \{BG, GB, GG\}$*), i.e. no other strategy yields higher expected utility.*

[8]Since every agent lives for one period only, once he becomes the decision-maker, he does not care about the strategies of players and other agents. For this reason, we do not require that he acts being aware of the equilibrium strategy of other people.

Proposition 6. *(i) $U_{GG}^{S}(\alpha_{GG})$ is a continuous strictly decreasing function of α_{GG}, and $U_{GG}^{S}(1) < \hat{U}^{E}$. If $U_{GG}^{S}(0) \geq \hat{U}^{E}$ then there is a protoequilibrium with $\alpha_{GG} = 0$ ('good' protoequilibrium), and if $U_{GG}^{S}(0) > \hat{U}^{E}$, then there also exists a mixed strategy protoequilibrium where $0 < \alpha_{GG} < 1$, and $U_{GG}^{S} = U_{GG}^{E} = \hat{U}^{E}$. These protoequilibria exist for a wider range for other parameters when D is higher and μ is higher. Moreover, competence θ_{GG} in the good protoequilibrium is at least as high as in the mixed one, which is in turn at least as high as in the bloody one.*
(ii) Good equilibria exist for a wider range of parameters if the same conditions as in the case of protoequilibria (i.e. D is high and μ is higher) hold. There may exist at most one equilibrium such that $0 < \alpha_{GG} < 1$. The equilibrium competence in the mixed equilibrium is at least as high as in the bloody one, and the competence in a good equilibrium is at least as high as in mixed and bloody ones.

Proof of Proposition 6.

Proofs for (i) and (ii) are similar and we will not repeat the algebra separately for protoequilibria and equilibria. As demonstrated in Section 3, if the dictator expects to get U after winning and V after losing, then his maximization problem may be written as

$$Z(\theta) = (1 - p(\theta))U + p(\theta)V - r(\theta)C \to \max_{\theta}.$$

Let us prove that it has a unique solution $\theta^{*}(U, V)$.

$$Z(\theta) = (1 - q)U + qV + (q(U - V) - (q + (1 - \theta)(1 - q))C)G(R^{*}(\theta)).$$

For brevity, denote $M \equiv q(U - V) - (q + (1 - \theta)(1 - q))C$ (remember that X depends on θ) and $N = \frac{1-q}{q}(w + \pi)$ (this is a positive constant). Obviously,

$$\frac{\partial Z}{\partial \theta} = (1 - q)CG - MNG'(R^{*}(\theta)), \quad (8)$$

$$\frac{\partial^{2} Z}{\partial \theta^{2}} = -2(1 - q)CNG'(R^{*}(\theta)) + MN^{2}G''(R^{*}(\theta)). \quad (9)$$

First, note that $Z(\theta)$ has no interior minima, for in such point, $\frac{\partial Z}{\partial \theta} = 0$ (implying $M > 0$), and therefore $\frac{\partial^{2} Z}{\partial \theta^{2}} < 0$ (recall that $G'' < 0$), which violates second-order necessary condition. Since it has no interior minima, it cannot have two distinct local maxima, and hence it is single-peaked. In particular, the dictator's optimal choice θ^*, is either 0 or 1) or satisfies first-order condition $\frac{\partial Z}{\partial \theta}\big|_{\theta=\theta^*} = 0$. It is worth noting that points with $\frac{\partial Z}{\partial \theta} < 0$ lie to the right of θ^*, while those with $\frac{\partial Z}{\partial \theta} > 0$ lie to the left of it.

To determine, how optimal choice depends on variations of other parameters, we compute the signs of cross-derivatives of $Z(\theta)$ at optimal points; evidently, we may restrict our attention to those cases where maximum satisfies first-order condition (otherwise a slight change in any of the parameters does not affect optimal choice θ^*). Of course, at such points $M > 0$. So, at these points (we omit the argument $R^*(\theta)$ at the derivatives of G),

$$\begin{aligned}
\frac{\partial^2 Z}{\partial \theta \partial q} &= -\frac{\frac{\partial Z}{\partial \theta}}{1-q} - \frac{N(1-\theta)}{q}\left(2CG' - MNG''\right) \leq 0, \\
\frac{\partial^2 Z}{\partial \theta \partial U} &= -\frac{\partial^2 Z}{\partial \theta \partial V} = -(1-q)(w+\pi)G' < 0, \\
\frac{\partial^2 Z}{\partial \theta \partial \overline{R}} &= -(1-q)CG' + MNG'' < 0, \\
\frac{\partial^2 Z}{\partial \theta \partial C} &= (1-q)G + (q + (1-\theta)(1-q))NG' > 0.
\end{aligned}$$

This gives us the necessary comparative statics (θ^* (weakly) increases with those parameters that yield a positive cross-derivative at θ^*, and decreases with those that yield a negative one. ∎

Lemma 7. *Let X be any of the values $U_{XY}, U^E_{XY}, U^S_{XY}, V_{XY}, W_{XY}$. Then $-D \leq X < \frac{Y-\delta K}{1-(1-\delta)\beta}$.*

Proof of Lemma 7. At each period (including the current one) a player may not expect to get more than $Y - \delta K$. The probability that a player will survive till the next period is not greater than $1-\delta$, and the discount factor is β. Therefore, the sum of discounted expected utilities does not exceed $\frac{Y-\delta K}{1-(1-\delta)\beta}$, and since there is a non-trivial chance that the player will eventually be away from office, the inequality is strict.

Observe that utility in each period, utility may not be less than $-D$, and it may be negative only if the player dies or is repressed (another source of negative utility is expenditures on policy C, but they are always compensated by utility from winning Y enjoyed in that period). Therefore, a player may get negative utility in only one period (probably current one), and so expected utility cannot be less than $-D$. ■

Lemma 8. *For any α_{XY} and p_{XY} in $[0,1]$, equations (1) – (5) have a unique solution $\left(U_{XY}, U^E_{XY}, U^S_{XY}, V_{XY}, W_{XY}\right)$. Moreover, these values continuously depend on parameters $\alpha_{XY}, p_{XY}, \beta, \delta, Y, C, K, D$.*

Proof of Lemma 8. The mapping from 15-dimensional space of $\left(U_{XY}, U^E_{XY}, U^S_{XY}, V_{XY}, W_{XY}\right)$ to itself, set by equations (1) – (5), is contracting in metrics induced by norm $\left\|U_{XY}, U^E_{XY}, U^S_{XY}, V_{XY}, W_{XY}\right\| = \max_{XY\in\{BG,GB,GG\}} \max\left\{U_{XY}, \beta^{-\frac{1}{2}}U^E_{XY}, \beta^{-\frac{1}{2}}U^S_{XY}, V_{XY}, W_{XY}\right\}$, and the proof is straightforward. This proves the first part of the statement. To prove the latter part, we cannot use the implicit function theorem directly, for equations are not differentiable (because of max in the first equation). We proceed as follows. Denote any set of parameters $(\alpha_{XY}, p_{XY}, \beta, \delta, Y, C, K, D)$ by λ and consider a sequence $\lambda_1, \lambda_2, \ldots$ converging to λ_0 (in Euclidean metric). For any i, let U_i be the stable point for the set of parameters λ_i. Let us prove that sequence $U_1, U_2 \ldots$ converges to U_0. If it is not the case, there exists $\varepsilon > 0$ and subsequence $U_{k_1}, U_{k_2} \ldots$, the elements of which do not lie in ε-proximity of U_0. By Proposition 7, all utilities U_k (starting from a certain k) lie between $-D$ and $\frac{Y}{1-\beta_0/2}$, and therefore there is a subsequence $U_{m_1}, U_{m_2} \ldots$ converging to some finite point $U \neq U_0$. Since the right-hand side of (1) – (5) is continuous with respect to all variables, its value at (λ_0, U) equals the double limit of values at $m_1, m_2 \ldots$ and $U_{m_1}, U_{m_2} \ldots$, and in particular equals the limit of values at $\left(m_i, U_{m_i}\right)$. These values equal U_{m_i}, respectively, and tend to U, which means that U is a stable point of the mapping, and this contradicts $U \neq U_0$. This contradiction completes the proof.■

Bad Always Executes Good. Here we establish that B always represses G. In particular, $\alpha_{GB} = 1$.

Assume the contrary. Then $U^E_{BG} \leq U^S_{BG}$. From (2) and (3) it follows that

$$(1-\mu)\left(Y + (1-\delta)\,\beta W_{BG} - \delta K\right) + \mu W_{BG} \leq W_{BG}.$$

Rearranging and dividing by $1-\mu$, we get

$$Y + (1-\delta)\,\beta W_{BG} - \delta K \le W_{BG},$$

which implies

$$W_{BG} \ge \frac{Y-\delta K}{1-(1-\delta)\,\beta}.$$

This contradiction completes the proof.■

Good Executes Bad If Delta Is Low. By Proposition 3, $\alpha_{BG} = 1$, and thus $V_{BG} = -D$. By Proposition 7, $V_{BG} \ge -D$. We prove that for sufficiently small δ, $U^E_{GB} > U^S_{GB}$. To do that, we prove this strict inequality for $\delta = 0$, and then use continuity stated in Proposition 8.

Now assume the contrary, i.e. $U^E_{GB} \le U^S_{GB}$. Then $U_{GB} = U^S_{GB}$, while $U_{BG} \ge U^S_{BG}$. Evidently, $V_{BG} = -D$ and $V_{GB} \ge -D$.

Let $\tilde{\theta}$ be such that

$$W_{GB} = \left(1-p\left(\tilde{\theta}\right)\right)U_{GB} + p\left(\tilde{\theta}\right)V_{BG} - r\left(\tilde{\theta}\right)C.$$

Then

$$W_{BG} \ge \left(1-p\left(\tilde{\theta}\right)\right)U_{BG} + p\left(\tilde{\theta}\right)V_{GB} - r\left(\tilde{\theta}\right)C,$$

because this holds for any $\tilde{\theta}$. Now, notice that

$$\begin{aligned} U^S_{GB} &= Y + \beta W_{GB} = Y + \beta\left(\left(1-p\left(\tilde{\theta}\right)\right)U_{GB} + p\left(\tilde{\theta}\right)V_{BG} - r\left(\tilde{\theta}\right)C\right) \\ &= Y + \beta\left(\left(1-p\left(\tilde{\theta}\right)\right)U^S_{GB} - p\left(\tilde{\theta}\right)D - r\left(\tilde{\theta}\right)C\right) \end{aligned}$$

and

$$\begin{aligned} U^S_{BG} &= Y + \beta W_{BG} \ge Y + \beta\left(\left(1-p\left(\tilde{\theta}\right)\right)U_{BG} + p\left(\tilde{\theta}\right)V_{GB} - r\left(\tilde{\theta}\right)C\right) \\ &\ge Y + \beta\left(\left(1-p\left(\tilde{\theta}\right)\right)U^S_{BG} - p\left(\tilde{\theta}\right)D - r\left(\tilde{\theta}\right)C\right). \end{aligned}$$

Therefore,

$$U_{BG} \ge U^S_{BG} \ge \frac{Y-\beta\left(p\left(\tilde{\theta}\right)D + r\left(\tilde{\theta}\right)C\right)}{1-\beta\left(1-p\left(\tilde{\theta}\right)\right)} \ge U^S_{GB} = U_{GB},$$

which, combined with $V_{BG} \ge V_{GB}$, implies that

$$(1-p(\theta))\,U_{BG} + p(\theta)\,V_{BG} - r(\theta)\,C \ge (1-p(\theta))\,U_{GB} + p(\theta)\,V_{GB} - r(\theta)\,C$$

for any given θ. Therefore, $W_{BG} \geq W_{GB}$. Hence, from the known inequality for W_{BG} we find that

$$(1-\mu)(Y+\beta W_{BG}) + \mu W_{BG} < W_{BG} \leq W_{GB}.$$

Therefore, for $\delta = 0$, $U_{GB}^E > U_{GB}^S$ holds. Therefore, the same holds for δ small enough. ■

Same Competence If One Is Bad. Propositions 3 and 3 imply that $V_{BG} = V_{GB} = -D$ and $U_{BG} = U_{BG}^E = U_{GB}^E = U_{GB}$, since U_{BG}^E and U_{GB}^E are equal, as can be seen from (2). This, in turn, implies $W_{BG} = W_{GB}$, since problems in (5) become identical. By Proposition 2, the solution to maximization problem (5) is unique, and hence in a protoequilibrium, $\theta_{BG} = \theta_{GB}$. To prove that these values do not depend on p_{XY}'s, note that they are determined uniquely from a subset of equations (1) – (5) for $XY \in \{BG, GB\}$. These equations do not include p_{GG}, and if $\alpha_{BG} = \alpha_{GB} = 1$, they do not include p_{BG} and p_{GB} either. Hence, θ_{BG} and θ_{GB} do not depend on p_{XY}'s. Since these values are equal to each other and the same for any protoequilibrium, they are the same in any equilibrium. In an equilibrium, (7) must hold, therefore, $p_{BG} = p_{GB}$, and these values are the same for any equilibrium. ■

Existence Of Bloody Equilibrium. It is evident that $\alpha_{GG} = 1$, $V_{GG} = -D$, $U_{GG} = U_{GG}^E = U_{GB}$, $U_{GG}^S = U_{GB}^S$, $W_{GG} = W_{GB}$ satisfy all equations (1) – (5) and (6), because equations for $XY = GG$ become the same as for $XY = GB$. Therefore, this forms a protoequilibrium. If we set $p_{XY} = p(\theta_{XY})$ for all XY, we will get an equilibrium, since protoequilibrium does not depend on p_{XY}'s (follows Proposition 3 and formulae for variables with index GG). Conditions $\alpha_{BG} = \alpha_{GB} = \alpha_{GG} = 1$ detemine all utilities, and thus θ_{XY}'s, uniquely, because in this case p_{XY}'s are not included in the system of equations. This implies that p_{XY}'s are also uniquely determined, and thus only one equilibrium with all α_{XY}'s equal to 1 exists, which completes the proof.■

Existence And Competence Of Good And Mixed Protoequilibria. If $U_{GG}^E > U_{GG}^S$, then in an equilibrium, α_{GG} should equal 1. It is straightforward to check that it is indeed an equilibrium. Otherwise, $U_{GG}^E \leq U_{GG}^S$, or equivalently, $U_{GG} = U_{GG}^S$, holds. Now, our goal is to find U_{GG}^S as a function of α_{GG} and to compare it with U_{GG}^E which is known. Obviously, $\alpha_{GG} = 0$ would form an equilibrium if $U_{GG}^S \geq U_{GG}^E$, while $\alpha > 0$ would if $U_{GG}^S = U_{GG}^E$.

Let us prove that U_{GG}^S is a strictly decreasing function of α_{GG}. Since for $\alpha_{GG} < 1$ we have $U_{GG} = U_{GG}^S$, then by (3),

$$U_{GG}\left(W_{GG}\right) = Y + (1-\delta)\,\beta W_{GG} - \delta K. \tag{10}$$

Then, from (4) we get

$$V_{GG}\left(W_{GG},\alpha_{GG}\right) = \frac{\left(1-\alpha_{GG}\right)\left(\left(1-\delta\right)\beta\left(\left(1-\delta\right)p_{GG}\left(Y+\left(1-\delta\right)\beta W_{GG}-\delta K\right)+\delta W_{GG}\right)-\delta K\right)-\alpha_{GG}D}{1-\left(1-\alpha_{GG}\right)\left(1-\delta\right)\beta\left(1-\delta\right)\left(1-p_{GG}\right)}. \tag{11}$$

Therefore, W_{GG} satisfies equation

$$L\left(W_{GG},\alpha_{GG}\right) \equiv W_{GG} - \max_{\theta}\left(\left(1-p\left(\theta\right)\right)U_{GG}\left(W_{GG}\right)+p\left(\theta\right)V_{GG}\left(W_{GG},\alpha_{GG}\right)-r\left(\theta\right)C\right)=0, \tag{12}$$

where $U_{GG}\left(W_{GG}\right)$ and $V_{GG}\left(W_{GG},\alpha_{GG}\right)$ are taken from (10) and (11). $L\left(W_{GG},\alpha_{GG}\right)$ is strictly increasing with respect to both W_{GG} and α_{GG}. To prove that, we use envelope's theorem, we get

$$\frac{\partial L}{\partial \alpha_{GG}} = -p\left(\theta\right)\frac{\partial V_{GG}}{\partial \alpha_{GG}}, \tag{13}$$

while (omitting GG subscript for brevity)

$$\frac{\partial V_{GG}}{\partial \alpha_{GG}} = \frac{K\delta - W\beta\delta\left(1-\delta\right) - U\beta p\left(1-\delta\right)^2 - D\left(1-\beta\left(1-\delta\right)^2\left(1-p\right)\right)}{\left(1-\left(1-\alpha\right)\left(1-\delta\right)\beta\left(1-\delta\right)\left(1-p\right)\right)^2}.$$

Since $K < D$, $U \geq -D$, $w \geq -D$, the numerator is less than $-D\left(1-\beta\right)\left(1-\delta\right) < 0$. Therefore, $\frac{\partial Z}{\partial \alpha_{GG}} > 0$. Now,

$$\frac{\partial L}{\partial W_{GG}} = 1-\left(1-p\left(\theta\right)\right)\frac{dU_{GG}}{dW_{GG}} - p\left(\theta\right)\frac{\partial V_{GG}}{\partial W_{GG}}. \tag{14}$$

Evidently, $\frac{dU_{GG}}{dW_{GG}} = \left(1-\delta\right)\beta < 1$. It is sufficient to demonstrate that $\frac{\partial V_{GG}}{\partial W_{GG}} < 1$, but we would later need $\frac{\partial V_{GG}}{\partial W_{GG}} < \left(1-\delta\right)\beta$, which we establish right away.

$$\frac{\partial V_{GG}}{\partial W_{GG}} = \frac{\left(1-\alpha\right)\left(1-\delta\right)\beta\left(\left(1-\delta\right)p\left(\left(1-\delta\right)\beta+\delta\right)\right)}{1-\left(1-\alpha\right)\left(1-\delta\right)\beta\left(1-\delta\right)\left(1-p\right)}.$$

The denominator is clearly greater than 0. Thus, we need to check that

$$\left(1-\delta\right)\beta\left(1-\left(1-\alpha\right)\left(1-\delta\right)\beta\left(1-\delta\right)\left(1-p\right)\right)-\left(1-\alpha\right)\left(1-\delta\right)\beta\left(\left(1-\delta\right)p\left(\left(1-\delta\right)\beta+\delta\right)\right)>0.$$

This is equivalent to

$$\beta(1-\delta)(1-(1-\delta)(1-\alpha)(\beta(1-\delta)+p\delta))>0.$$

Clearly, this expression lies between 0 and 1. Therefore, $\frac{\partial V_{GG}}{\partial W_{GG}} < (1-\delta)\beta$, which implies $(1-p(\theta))\frac{dU_{GG}}{dW_{GG}} + p(\theta)\frac{\partial V_{GG}}{\partial W_{GG}} < (1-\delta)\beta < 1$. Hence, $\frac{\partial L}{\partial W_{GG}} > 0$, and therefore, function $W_{GG}(\alpha_{GG})$, defined as implicit function by (12), satisfies

$$\frac{dW_{GG}}{d\alpha_{GG}} = -\frac{\frac{\partial L}{\partial \alpha_{GG}}}{\frac{\partial L}{\partial W_{GG}}} < 0. \tag{15}$$

This proves that $W_{GG}(\alpha_{GG})$ and, by (10), $U_{GG}(W_{GG}(\alpha_{GG}))$ are strictly decreasing functions of α_{GG}.

One can immediately check that $U_{GG}(W_{GG}(1)) < \hat{U}^E$. If it were not the case, there would be a protoequilibrium with $\alpha_{GG} = 1$ and $U_{GG} = U^S_{GG}$, while Proposition **??** states that there is only one protoequilibrium with $\alpha_{GG} = 1$ in which, by Proposition 3, $U_{GG} = U_{BG} > U^S_{BG} = U^S_{GG}$. This leads us to a contradiction, and we conclude that $U_{GG}(W_{GG}(1)) < \hat{U}^E$.

It is obvious that $\alpha_{GG} < 1$ leads to a protoequilibrium either if $\alpha_{GG} = 0$ and $U_{GG}(W_{GG}(0)) \geq \hat{U}^E$ or if $U_{GG}(W_{GG}(\alpha_{GG})) = \hat{U}^E$. Hence, continuity and monotonicity of $U_{GG}(W_{GG}(\alpha_{GG}))$ implies that if $U_{GG}(W_{GG}(0)) = \hat{U}^E$, there is a protoequilibrium with $\alpha_{GG} = 0$, and if $U_{GG}(W_{GG}(0)) > \hat{U}^E$, there is one protoequilibrium with $\alpha_{GG} = 0$ and one with $0 < \alpha_{GG} < 1$.

To proceed, denote utilities that correspond to bloody, mixed, and good protoequilibria with upper indices (B), (M), and (G), respectively. Our goal is to demonstrate that $U^{(B)}_{GG} - V^{(B)}_{GG} > U^{(M)}_{GG} - V^{(M)}_{GG} > U^{(G)}_{GG} - V^{(G)}_{GG}$, since then the necessary result would immediately follow from Proposition 2. To prove the first inequality, notice that $U^{(B)}_{GG} = U^{(M)}_{GG} = \hat{U}^E$, $V^{(B)}_{GG} = -D$, while $V^{(M)}_{GG} > -D$ (the latter follows from (4), noting that if $\alpha_{GG} < 1$, there is a non-trivial chance of getting $-\delta K$ which is greater than $-D$). The proof of the latter inequality is a bit trickier. We will prove that $U_{GG}(W_{GG}(\alpha_{GG})) - V_{GG}(W_{GG}(\alpha_{GG}), \alpha_{GG})$ is an increasing function of α_{GG}. For brevity, we will omit index GG in formulae. First,

$$\frac{d(U-V)}{d\alpha} = -\frac{\partial V}{\partial \alpha} + \left(\frac{dU}{dW} - \frac{\partial V}{\partial W}\right)\frac{dW}{d\alpha}.$$

We have already proved that $\frac{\partial V}{\partial w} < (1-\delta)\,\beta$. Using (15), (13), and (14), we find

$$\frac{dW}{d\alpha} = -\frac{\frac{\partial L}{\partial \alpha}}{\frac{\partial L}{\partial W}} = -\frac{-p(\theta)\frac{\partial V}{\partial \alpha}}{1-\frac{dU}{dW}+p(\theta)\left(\frac{dU}{dW}-\frac{\partial V}{\partial W}\right)}.$$

Consequently,

$$\frac{d(U-V)}{d\alpha} = -\frac{\partial V}{\partial \alpha}\left(1-\frac{p(\theta)\left(\frac{dU}{dW}-\frac{\partial V}{\partial W}\right)}{1-\frac{dU}{dW}+p(\theta)\left(\frac{dU}{dW}-\frac{\partial V}{\partial W}\right)}\right) = -\frac{\partial V}{\partial \alpha}\frac{1-\frac{dU}{dW}}{1-\frac{dU}{dW}+p(\theta)\left(\frac{dU}{dW}-\frac{\partial V}{\partial W}\right)} > 0,$$

because, as we demonstrated before, $\frac{\partial V}{\partial \alpha} < 0$, $\frac{dU}{dW} = (1-\delta)\,\beta < 1$ and $\frac{\partial V}{\partial W} < (1-\delta)\,\beta$, thus making $\frac{dU}{dW} - \frac{\partial V}{\partial W} > 0$. From this, we conclude that $\theta^{(B)}_{GG} \le \theta^{(M)}_{GG} \le \theta^{(G)}_{GG}$.

To finish the proof, we need to show that $U_{GG}(W_{GG}(0)) - \hat{U}^E$ is increasing with respect to both D and μ, and decreasing with respect to Y. The first thing to observe is that $U_{GG}(W_{GG}(0))$ does not depend on D and μ. Now consider derivatives of $\hat{U}^E$.

$$\frac{d\hat{U}^E}{dD} = (1-\delta)\,\beta\,((1-\mu)(1-\delta)\,\beta+\mu)\left((1-p_{BG})\frac{d\hat{U}^E}{dD} - p_{BG}\right), \tag{16}$$

because $V_{BG} = -D$. Similarly,

$$\frac{d\hat{U}^E}{d\mu} = (1-\delta)\,\beta\,((1-(1-\delta)\,\beta)\,W_{BG} - (Y-\delta K)) + (1-\delta)\,\beta\,((1-\mu)(1-\delta)\,\beta+\mu)(1-p_{BG})\frac{d\hat{U}^E}{d\mu}. \tag{17}$$

Rearranging (16) and (17), we get $\frac{d\hat{U}^E}{dD} < 0$ and $\frac{d\hat{U}^E}{d\mu} < 0$ (recalling that $W_{BG} < \frac{Y-\delta K}{1-(1-\delta)\beta}$). Therefore, good equilibrium is more likely to exist when D or μ are higher. ■

Existence And Competence Of Good And Mixed Equilibria. When studying equilibria, we restrict ourselves to the case where $p_{BG} = p_{GB} = p(\theta_{BG})$ (recall that the latter does not depend on p_{GG}). First, we check that U_{GG} is increasing with respect to p_{GG}. By (3), it is sufficient to check that $W_{GG}(0)$ is increasing with respect to p_{GG}. We obtain

$$\frac{dL}{dp_{GG}} = -p(\theta_{GG})\frac{\beta(1-\delta)^2\,\delta K + \left(1-\beta(1-\delta)^2\right)Y + \beta(1-\delta)^2(1-\beta(1-\delta))\,W}{(1-(1-\delta)\,\beta(1-\delta)(1-p))^2}.$$

Since α_{GG}, all utilities with GG subscript are not less than $-\delta K$, which means that the numeretor is not less than $\left(1-\beta(1-\delta)^2\right)Y + \beta^2(1-\delta)^3\,\delta K$, i.e. it is positive (in other words, $\frac{\partial V}{\partial \alpha} > 0$). Since $\frac{\partial L}{\partial W_{GG}} > 0$, $\frac{dW_{GG}(0)}{dp_{GG}} > 0$.

Now let us demonstrate that $U_{GG} - V_{GG}$ is decreasing with respect to p_{GG}. Evidently (omitting GG subscript),

$$\frac{d(U-V)}{dp} = -\frac{\partial V}{\partial p} + \left(\frac{dU}{dW} - \frac{\partial V}{\partial W}\right)\frac{dW}{dp}.$$

Since

$$\frac{dW}{dp} = -\frac{\frac{\partial L}{\partial p}}{\frac{\partial L}{\partial W}} = -\frac{-p(\theta)\frac{\partial V}{\partial p}}{1 - \frac{dU}{dW} + p(\theta)\left(\frac{dU}{dW} - \frac{\partial V}{\partial W}\right)},$$

we conclude that

$$\frac{d(U-V)}{dp} = -\frac{\partial V}{\partial p}\frac{1 - \frac{dU}{dW}}{1 - \frac{dU}{dW} + p(\theta)\left(\frac{dU}{dW} - \frac{\partial V}{\partial W}\right)} < 0,$$

because $\frac{\partial V}{\partial \alpha} > 0$. This means that θ_{GG} is (weakly) increasing with respect to p_{GG}, and so does $p(\theta_{GG})$. We will need this later; for now, notice that when p_{GG} grows from 0 to 1, $p(\theta_{GG}(p_{GG}))$ lies strictly between 0 and 1, and it does not depend on D and μ. Therefore, there is a fixed set of values p_{GG} such that $p_{GG} = p(\theta_{GG}(p_{GG}))$, and the question is only whether for this value $\hat{U}^E \leq U_{GG}(W_{GG}(0))$. This is more likely if D is high or μ is high.

In a mixed protoequilibrium, $U_{GG} = U^S_{GG} = \hat{U}^E$, and $\hat{U}^E$ does not depend on p_{GG}. Therefore, (3) implies that W_{GG} is fixed, and (4) in its turn implies that V_{GG} is fixed (otherwise W_{GG} would be greater for larger V_{GG}. Hence, $U_{GG} - V_{GG}$ is the same, and therefore θ_{GG} and p_{GG} are fixed. This means that there may exist only one mixed equilibrium, and for this equilibrium (if it exists) p_{GG} is found as described. Mixed equilibrium exists if and only if for this p_{GG} there is a good protoequilibrium.

Let us prove that competence in bloody equilibrium is at least as low as in any other one. Denote competence in the latter equilibrium by θ^*, and let $p^* = p(\theta^*)$. For $p_{GG} = p^*$, there is a bloody protoequilibrium as well, and in this protoequilibrium, $\theta^{(B)}_{GG} \leq \theta^*$ by (i) in Proposition 6. However, $\theta^{(B)}_{GG}$ is fixed and equal to θ_{BG} by Proposition 3. In bloody equilibrium, competence is also equal to θ_{BG}, and therefore it is not greater than θ^*.

Finally, let us demonstrate that competence in the mixed equilibrium (if there exists one) is not greater than competence in a good equilibrium. Denote competence in the latter one by θ^*, and let $p^* = p(\theta^*)$. Consider the mixed protoequilibrium for $p_{GG} = p^*$.

In this protoequilibrium, competence $\theta_{GG}^{(M)}$ is less than θ^*. However, in all mixed protoequilibria (and thus in the mixed equilibrium) competence is the same, which means that in the mixed equilibrium, competence does not exceed that in a good equilibrium. ■

Competence On Different Paths. Competence on the path where dictators execute their enemies coincides with competence in GG case in the bloody equilibrium. Therefore, according to (ii) in Proposition 6, competence of agents is better (not worse) on the good path. Utilities, and thus competence, of the good equilibrium does not depend on μ, as it was shown above. On the other hand, $\frac{d\hat{U}^E}{d\mu} < 0$ and $\frac{dV_{XY}}{d\mu} = 0$ on the bloody path, since $V_{XY} = -D$. Consequently, higher μ leads to lower difference in utilities and thus to higher competence. ■