

A Relational Theory of Autocracy

Scott Gehlbach

University of Chicago

gehlbach@uchicago.edu

Zhaotian Luo

University of Chicago

luozhaotian@uchicago.edu

Spencer Pantoja

Stanford University

spantoja@stanford.edu

This version: September 23, 2026

Abstract

We develop a relational theory of autocracy that sheds light on both the concentration of power and elite competence in authoritarian regimes. Our theoretical framework is a dynamic agency model in which the agent in one period can attempt to become the ruler in the next, capturing the ever-present threat of military and palace coups in autocracies. Under this assumption of “contestable control rights,” the ruler’s decision to recruit an agent of observed competence is endogenous to the anticipated cost of deterring a coup. Three key principles emerge from our analysis: 1) Incompetent rulers earning high “natural rents” from office rule alone. 2) Autocrats generally choose right-hand men who are neither too competent nor too incompetent. 3) More competent rulers are willing to recruit more competent agents.

1 Introduction

Authoritarian regimes vary widely in the extent to which power is concentrated in a single individual or small clique. They also differ in the competence of the governing elite responsible for turning dictates into policy. Although typically studied separately, these two dimensions of autocratic governance emerge jointly from bargaining between a ruler and his agents in the hothouse environment that characterizes authoritarian politics. With no third party to enforce contracts, any agreement between ruler and agent is necessarily relational, that is, supported by the value of the ongoing relationship rather than the threat of formal sanction. Every agent of the ruler, in turn, has his eye on the ultimate prize—control of the state and its associated rewards. A dictator deciding whether to recruit an agent at all, and if so of what competence, therefore weighs the benefit of competent governance against the cost not only of incentivizing effort, but also of deterring the agent from attempting a coup.

In this paper, we develop these ideas in a relational theory of autocracies. Our theoretical framework is a dynamic agency model in which the agent in one period can attempt to become the ruler in the next. It is precisely this friction—what we call *contestable control rights*—that distinguishes our setting from the typical environment in the literature on relational contracts. Crucially, we do not take as given the presence of an agent who can try to seize power. Rather, any ruler decides whether to recruit an agent, of randomly drawn and observed competence, as well as the size of the transfer paid to that agent. Any agent, in turn, decides whether to exert effort and whether to attempt a coup. We assume that effort is complementary to competence, so that a more competent agent generates more surplus, but also that competence increases an agent’s ability to win a contest for power against the ruler. Crucially, none of these actions is contractible, which implies that any agreement between a ruler and an agent must be self-enforcing—that is, supported by the value of the relationship over an infinite horizon.

The assumption of contestable control rights complicates the analysis of this dynamic game

of perfect information. It proves useful to define a *regime*, which is that portion of the game that begins after a ruler acquires or retains power and ends with a contest for power, regardless of whether the ruler wins or loses. We restrict attention to equilibria in which each ruler maximizes the value of his regime, under the assumption that all subsequent rulers do the same—a generalization of the idea of “principal-optimal” equilibrium to our setting, in which there are multiple possible principals. The rationale for this selection is that the aura that accompanies victory in a contest for power positions the ruler to set the terms of his relationship with any new agent.¹

In any ruler-optimal regime-symmetric equilibrium, a new ruler chooses to recruit an agent only if the subsequent relationship is *stable*: the agent exerts effort and does not attempt a coup, while the ruler provides a transfer sufficient to incentivize this behavior, with deviations punished by the breakdown of the relationship.² This introduces a constraint on the set of types a ruler would recruit. We establish that recruitment is always from a “competence interval”: for any ruler, there are actors who are either too competent or too incompetent to serve as agents. For some parameters of the model, this interval is empty, implying that the ruler chooses not to recruit any agent.

Three principles of autocratic governance emerge from this analysis, which we illustrate with reference to three paradigmatic cases of authoritarian rule. First, some rulers—bumbling or favorably endowed with large resource rents—choose to rule without any key agent who could threaten their hold on power. We call this the *Idi Amin Principle*. Second, the typical ruler who does recruit a right-hand man does so among potential agents who are neither mediocrities nor brilliant strategists. We term this the *Molotov Principle*. Finally,

¹Our notion of a “regime” is narrower than a leadership spell—in contrast, say, to [Geddes, Wright and Frantz \(2018\)](#), who conceptualize a regime as a system for selecting rulers that may span more than one leader. Intuitively, we have in mind periods such as “early Stalin,” “mid-Stalin,” and so forth, each ending with a contest for power that terminates the existing “regime” and inaugurates a new one.

²Empirically, most coups are off the equilibrium path. Figure [A1](#), which merges leadership data from the Archigos project ([Goemans, Gleditsch and Chiozza, 2009](#)) with Boix, Miller, and Rosato’s (2013) data on regime type, demonstrates that while coups are more common in autocracies than in democracies, they are far less common than regular transfers of power. The vast majority of authoritarian leaders survive from one year to the next.

more competent rulers are willing to recruit more competent agents, a form of assortative matching. We call this the *Atatürk Principle*.

Our paper builds on several literatures. A first immediate connection is to growing theoretical work on agency concerns in authoritarian regimes (e.g., [Myerson, 2008](#); [Egorov, Guriev and Sonin, 2009](#); [Gehlbach and Simpser, 2015](#); [Rundlett and Svolik, 2016](#); [Tyson, 2018](#)). Our focus is closest to models of the loyalty-competence tradeoff in autocracies ([Egorov and Sonin, 2009](#); [Zakharov, 2016](#); [Egorov and Sonin, 2025](#)).³ As in that work, we assume that competent subordinates are dangerous subordinates, but the mechanism is different. In our model, the threat to the ruler comes not from outside the regime, but from the agent himself. Rulers recruit agents to help them extract rents, but agents can exploit that position to try to replace the ruler. From any ruler’s perspective, the fundamental tradeoff is that a more competent agent is better able to extract rents but also more tempted to contest power, and perhaps more likely to win that contest. The three principles of autocratic governance that emerge from our analysis of this tradeoff are new to the scholarship on autocratic politics.

Our work is also related to a vast literature on institutional variation in authoritarian regimes. At a high level, this work distinguishes between “personalist” autocracies, where power is concentrated in a single individual or narrow clique, and “institutionalized” regimes, where power is more dispersed (e.g., [Svolik, 2012](#); [Geddes, Wright and Frantz, 2014, 2018](#); [Gandhi and Sumner, 2020](#); [Meng, 2020](#)). For the most part, this literature has not focused on the consequences of power concentration for the competence of ruling elites, though there are a few contributions tying institutional variation in autocracies to economic growth and investment ([Gandhi, 2008](#); [Wright, 2008](#); [Gehlbach and Keefer, 2011, 2012](#); [Blattman, Gehlbach and Yu, 2026](#)). One exception is [Pan \(2025\)](#), who documents a decline in bureaucratic competence in contemporary China as power concentration increases. Our theoretical framework explicitly ties the competence of agents in authoritarian regimes to the willingness of a ruler

³The idea of loyalty also figures centrally in the selectorate theory of [Bueno de Mesquita et al. \(2003\)](#) and [Bueno de Mesquita and Smith \(2025\)](#), though without a corresponding focus on competence.

to govern alone.

A final connection is to theoretical work on relational contracts in organizations (e.g., MacLeod and Malcomson, 1989; Baker, Gibbons and Murphy, 1994, 1999, 2002; Levin, 2003), including in states (Dixit, 2004; Tyson, 2020; Callander, Foarta and Sugaya, 2022; Acharya, Haber and Lee, 2024).⁴ Most such models feature a fixed principal and agent. We model recruitment of agents of heterogeneous types, as in Board and Meyer-Ter-Vehn (2015), though with the novelty that a newly recruited agent can attempt to replace the principal.⁵

2 Agency Relationships in Authoritarian Regimes

When we think of dictators, we often—but not always—think of the agents who served them. Some authoritarian leaders have well-known right-hand men. Others occupy center stage alone. Our theoretical framework helps to make sense of these relationships, or lack thereof. To preview the key principles that emerge from our analysis, we illustrate by way of example, showing how the dictator’s own competence and the broader political-economic environment shape the recruitment of agents, given the possibility that any recruited agent can subsequently attempt to replace the ruler.

Idi Amin

The poster child for personalist authoritarianism, Idi Amin Dada Oumee ruled Uganda from 1971 to 1979. A career soldier, and later officer, Amin served as army commander under President Milton Obote, whom he overthrew in a military coup in 1971. His eight years of rule are remembered for their erraticism and brutality. Amin’s foreign policy fluctuated wildly, from an alliance with Israel to support for the hijackers of an Air France flight from Tel Aviv. He expelled the country’s “Asian” (Indian-Ugandan) population, at enormous

⁴More generally, our paper builds on efforts to extend the domain of organizational economics to political institutions (e.g., Moe, 1984; Weingast and Marshall, 1988). For a synthesis that connects to organization theory more broadly, see Ahlquist and Grzymala-Busse (2025). Other recent work includes Snowberg and Ting (2022), Foarta and Ting (2023), and Angelucci, Meraglia and Voigtländer (2025).

⁵In related work, Debs (2007, Chapter 1) analyzes an infinite-horizon model in which rulers can “shuffle” agents of unobservable type to discourage popular insurrection and the ruler’s replacement.

cost to the economy. His euphemistically named State Research Bureau (SRB) tortured and murdered hundreds of thousands of Ugandan citizens. In 1978, he invaded neighboring Tanzania. He lost the resulting war, fled the country, and lived out his remaining years in exile.

Not even a dictator with a reputation for cannibalism can murder thousands without assistance. Throughout his period of rule, however, Idi Amin had no consistent right-hand man. The extent to which he concentrated power in his own person is reflected in the titles he bestowed upon himself: “His Excellency, President for Life, Field Marshal Al Hadji Doctor Idi Amin Dada, VC, DSO, MC, CBE, Lord of All the Beasts of the Earth and Fishes of the Seas and Conqueror of the British Empire in Africa in General and Uganda in Particular.” Amin himself served as minister of defense on two occasions—a typical arrangement in “big man” autocracies with little power-sharing ([Meng and Paine, 2022](#); [Clarke, Meng and Paine, 2025](#)). For most of his rule, he had no vice president.

Uganda under Idi Amin was poorly governed, but it was rich in natural and human resources. Gold and ivory smuggling were central sources of revenue for the Amin regime. So, in a sense, was the Asian population that Amin expelled: their accumulated assets, separated from any legitimate property claim, were available for redistribution to regime supporters. Amin, meanwhile, may not have been the incompetent buffoon commonly portrayed by foreign media, but he was no Lee Kuan Yew. He showed little talent for economic policy in particular, leaving him vulnerable to removal when the end came—notably, at the hands of foreign military forces rather than through any palace or military coup. These characteristics of Amin and the country he governed illustrate a first key principle of authoritarian governance: *rulers who are incompetent or who lead regimes capable of producing high “natural” rents are more likely to rule alone.*

Joseph Stalin and Vyacheslav Molotov

As he consolidated power in the late 1920s, Joseph Stalin systematically removed other members of the Soviet Union's collective leadership: Leon Trotsky, Grigory Zinoviev, Lev Kamenev, Nikolai Bukharin, Alexei Rykov, Mikhail Tomsy. During the Great Terror, he murdered countless members of the party and military elite. Yet a few key allies survived these events, remaining at Stalin's side to his death in 1953. Among these was Vyacheslav Molotov, Stalin's predecessor as party secretary and eventual head of government and foreign minister.

In exile and in power, the Bolshevik Party had many brilliant theorists and charismatic leaders. Molotov was neither. Trotsky, who combined both qualities, viewed Molotov as "mediocrity personified." Lenin, damning with faint praise, called him the "best filing clerk in Russia." Once a student at St. Petersburg Polytechnic University, Molotov lacked the intellectual gravitas not only of Bukharin but also the autodidacticism of Stalin and Trotsky. He played a minor role in the Russian Civil War—especially compared to Trotsky, who served as People's Commissar for military affairs before, during, and after the conflict. By temperament, Molotov was a bureaucrat.

Trotsky's view notwithstanding, however, Molotov was no mediocrity. Rather, he was an efficient executor of Stalin's vision. Best known outside the Soviet Union for his negotiation of the Hitler-Stalin Pact—otherwise known as the Molotov-Ribbentrop Pact—that launched the Second World War in Europe, Molotov played an important role in numerous aspects of domestic and foreign policy. He was an agent of Stalinist collectivization, including the campaign against the so-called "kulaks." Following the German invasion in 1941, Molotov served as the inaugural deputy chairman, under Stalin, of the State Defense Committee. (Less successful was Molotov's leadership of the Soviet atomic-bomb project, which languished under his direction.) After the war, the Soviet bloc countered the Marshall Plan with what came to be known as the Molotov Plan.

Vyacheslav Molotov exemplifies a second key principle of authoritarian governance: *autocrats generally choose right-hand men who are neither too competent nor too incompetent*. Stalin's vision could only be implemented with the assistance of capable functionaries. At the same time, too capable an agent—a Trotsky rather than a Molotov—posed a threat to Stalin's rule. Molotov was in the sweet spot that ensured survival in a brutal regime. His limited room for maneuver can be seen in his acquiescence even to the arrest in 1948 of his wife, Polina Zhemchuzhina. Molotov figured in the palace intrigue that followed Stalin's death in 1953, but his star dimmed as Nikita Khrushchev's rose. In a series of humiliations, he was expelled from the Presidium, posted to Mongolia, and finally relieved of his party membership.

Kemal Atatürk and İsmet İnönü

Among twentieth-century autocrats, few possessed the self-assurance of Mustafa Kemal Atatürk. His confidence was earned. A gifted student, Atatürk graduated from the Ottoman Military Academy and soon thereafter joined the Ottoman Army. He served in all of the Empire's wars that followed, distinguishing himself especially during the First World War. The Treaty of Sèvres that followed provoked a Turkish nationalist uprising, with Atatürk at its head. A national hero with a reputation for efficiency, Atatürk led the Turkish Republic and its sole political party, the Republican People's Party, from the country's founding until his death in 1938.

Atatürk's competence afforded him the liberty to select a right-hand man of similar ability. İsmet İnönü was this man. In many respects, İnönü's profile was similar to that of Leon Trotsky: an autodidact who spoke multiple languages, a military leader who reorganized elements of the imperial army into an effective fighting force, an effective diplomat who secured peace with the country's enemies, a revolutionary who was unafraid to suppress opposition (though not with Trotsky's zeal for terror). Only three years Atatürk's junior, it is not too difficult to imagine İnönü as the Turkish Republic's founding father had Atatürk been more seriously injured during the Italo-Turkish War.

Unlike Trotsky, however, İnönü was not expelled from the regime. Instead, as prime minister for nearly all of Atatürk’s rule, he implemented many of the policies that characterize this period: secular education, Western dress, industrialization, gender equality, the capital’s move to Ankara, and more. Not least, he led the establishment of a one-party state, banning all rivals to the Republican People’s Party. Only in 1937, a year before Atatürk’s death, was İnönü removed as prime minister, as Atatürk pivoted to a less statist economic policy than that championed by İnönü. Nonetheless, when Atatürk died in 1938, İnönü succeeded him as president and party leader.

This is a final principle of authoritarian governance captured by our theoretical framework: *more competent rulers are willing to recruit more competent agents*. Yet motive must coincide with opportunity. Atatürk was lucky that İnönü was available when he was. Secure in his power, Atatürk benefited from the presence of a competent right-hand man through a period of immense social, political, and economic change.

3 Model and solution concept

3.1 Setup

Consider a polity comprising a continuum of actors. Actors are indexed by their *competence*, $\theta \in \Theta := [0, 1]$, distributed according to F with full support on Θ . At each $t \in \{0, 1, 2, \dots\}$, one actor, $R_t \in \Theta$, is the *ruler* of the polity; at $t = 0$, this actor is randomly drawn from Θ . The ruler may have an *agent*, $A_t \in \Theta$, either inherited from the previous period or recruited at t , where $a_t \in \{0, 1\}$ indicates the presence of an agent at t . If there is an agent at t , there may be a *contest for power* between the ruler and agent, $c_t \in \{0, 1\}$, in which case the (possibly new) ruler at $t + 1$ does not inherit an agent from t . All actors other than the ruler and agent at t are *inactive*.

Recruitment. If there was an agent at $t - 1$ but no contest for power, then $A_t = A_{t-1}$; there is no recruitment. In contrast, if the ruler at t inherits no agent from the previous period, which follows from either $t = 0$ or $a_{t-1}(1 - c_{t-1}) = 0$, then an actor $\theta_t \in \Theta$ is randomly

drawn from the polity. After observing θ_t , the ruler chooses whether to recruit that actor as his agent, $m_t \in \{0, 1\}$. If $m_t = 1$, then $a_t = 1$ and $A_t = \theta_t$, whereas if $m_t = 0$, $a_t = 0$. Hence, whether the ruler has an agent at t is determined by

$$\begin{aligned} a_0 &= m_0 \\ a_t &= a_{t-1}(1 - c_{t-1}) + (1 - a_{t-1}(1 - c_{t-1}))m_t, \quad t \geq 1. \end{aligned}$$

Rent extraction and transfers. If the ruler has an agent, $a_t = 1$, the agent chooses whether to exert effort, $e_t \in \{0, 1\}$. The ruler's total rent extraction at t is

$$y_t := \begin{cases} \rho + \alpha R_t + e_t A_t, & a_t = 1 \\ \rho + \alpha R_t, & a_t = 0, \end{cases} \quad (1)$$

where $\rho > 0$ is the ruler's "natural rent" and $\alpha > 0$ measures the contribution (if any) of the ruler's competence to rent extraction. When an agent is present and exerts effort, the marginal rent extraction equals the agent's competence, A_t . After collecting rents, the ruler chooses a transfer to the agent, if present, $x_t \geq 0$.

Contest for power. If an agent is present, $a_t = 1$, he chooses whether to challenge the ruler's position at the end of t , $c_t \in \{0, 1\}$. A challenge triggers a *contest for power*, where $W_t \in \{R_t, A_t\}$ indicates the winner of the contest. Hence, period t concludes in one of three ways. First, if the agent refrains from challenging the ruler, $c_t = 0$, then both the ruler and the agent keep their positions at $t + 1$, $R_{t+1} = R_t$ and $A_{t+1} = A_t$. Second, if the agent challenges the ruler but loses the contest, $c_t = 1$ but $W_t = R_t$, then the ruler retains power, $R_{t+1} = W_t = R_t$, while the defeated agent becomes forever inactive, $A_t \notin \{R_\tau, A_\tau\}$ for all $\tau > t$. Third, if the agent wins the contest, $c_t = 1$ and $W_t = A_t$, then the agent usurps power and becomes the ruler at $t + 1$, $R_{t+1} = W_t = A_t$, while the defeated ruler becomes forever inactive, $R_t \notin \{R_\tau, A_\tau\}$ for all $\tau > t$.

We assume that more competent agents are more threatening to the ruler, whereas more competent rulers are better able to defend their position. In particular, let $g(A_t, R_t)$ denote the probability that the agent wins the contest for power, where $g: \Theta^2 \rightarrow (0, 1)$ is weakly increasing and weakly convex in its first argument and weakly decreasing and weakly concave in its second argument.⁶

Timing. The timing of the stage game at each t is as follows. First, if $t = 0$ or $a_{t-1}(1 - c_{t-1}) = 0$, an actor $\theta_t \in \Theta$ is randomly drawn. The ruler observes the actor's type and chooses whether to recruit θ_t as his agent, $m_t \in \{0, 1\}$. Second, if an agent is present, $a_t = 1$, the agent chooses whether to exert effort, $e_t \in \{0, 1\}$. Third, the ruler collects rents and chooses a transfer $x_t \geq 0$ to the agent, if present. Finally, if present, the agent decides whether to challenge the ruler, $c_t \in \{0, 1\}$.

Preferences. Each actor $\theta \in \Theta$ receives at t the flow payoff

$$u_{\theta,t} := \begin{cases} y_t - x_t, & \theta = R_t \\ x_t - e_t k, & \theta = A_t \\ 0, & \theta \notin \{R_t, A_t\}, \end{cases}$$

where the parameter $k > 0$ is any agent's cost of effort. Actors share a common discount factor $\delta \in (0, 1)$, which throughout we assume satisfies the following condition:

$$\delta > \underline{\delta} := \frac{1 + g(1, 0)(1 + \rho + \alpha)}{1 + g(1, 0)(1 + \rho + \alpha) + g(0, 1)\rho}. \quad (2)$$

In what follows, we normalize payoffs as per-period averages.

⁶As an example, the following function satisfies these properties:

$$g(A_t, R_t) = \eta + (1 - \epsilon) \mathbf{A}_t (\mathbf{1} - \mathbf{R}_t),$$

where $0 < \eta < \epsilon < 1$. The expression in bold is a common contest function, whereas the remaining terms represent residual uncertainty to ensure that g is bounded by 0 and 1.

3.2 Discussion

A few aspects of our modeling strategy deserve discussion beyond the brief overview in the introduction. The “agent” in our model, for example, is a stylized representation of a ruler’s coalition, which may be narrow or broad and of generally greater or lesser competence. The question of collective-action concerns within that coalition is beyond the scope of this paper but potentially an important topic for future research.⁷

At first glance, it may seem odd that the ruler has no explicit opportunity to replace one agent with another. In fact, a ruler can attempt to “purge” an agent by refusing to provide the agreed transfer. In equilibrium, as we will show, this sets up a contest for power, which the ruler may or may not win. As in [Montagnes and Wolton \(2019\)](#), following a successful purge, the ruler draws a fresh recruit from the replacement pool—in our setting, this is the full support Θ , given that a purged agent has measure zero. We leave to future work the (non-trivial) extension in which the ruler observes a potential replacement’s type before deciding whether purge the current agent.

In our telling, the “output” to which an agent contributes is rent extraction—not, say, economic growth. This is without loss of generality, as the variable y_t can be interpreted as any output valued by the ruler that is a function of the ruler’s and agent’s competence and the agent’s effort. Although the welfare implications of equilibrium behavior depend on the particular interpretation, the model itself is sufficiently stylized to encompass both positive and negative interpretations.

In this context, the “natural rent” measured by ρ in Equation 1 represents not only mineral wealth, but also the contribution to output (whatever that may be) of unmodeled but politically non-threatening agents. In autocracies, technical competence is valued, whereas political competence is a double-edged sword ([Lee and Schuler, 2020](#)). The expression $\rho + \alpha R_t$

⁷Alternatively, our theory is precisely a model of the Sith in the Star Wars universe, where Darth Bane’s Rule of Two dictated that there could be only one master and one apprentice at any point in time. We are grateful to David Siegel for the analogy.

represents what is possible for a ruler of given competence to achieve before recruiting an agent who aspires to be in charge himself.

The “competence” of rulers and agents, in turn, reflects a general ability to diagnose problems, translate policy into programs, mobilize resources, coordinate behavior, and otherwise generate “output” (extract rents) and successfully contest power. In what follows, we initially adopt only the weak assumptions above regarding the mapping g from competence to contest success: weak monotonicity and weak convexity/concavity. We then show how the progressive introduction of stronger assumptions generates additional results that correspond to the three principles of autocratic rule outlined in the introduction.

3.3 Solution concept

Our baseline solution concept is *subgame perfect Nash equilibrium* (SPNE). In the literature on relational contracts, it is conventional to restrict attention to SPNEs that are “ruler-optimal,” that is, that maximize the expected payoff of the ruler at the beginning of the game. The motivation for this equilibrium selection is that, in proposing a contract to the agent, the ruler’s preferred equilibrium becomes focal. In generalizing this idea to our setting, we focus on SPNEs that are “ruler-optimal,” that is, that maximize the expected payoff of the ruler at the moment the ruler acquires or retains power. This selection follows from the observation that any ruler who has newly acquired or retained power has won a contest for power. The aura that accompanies this victory positions the “new” ruler to focus expectations on mutually beneficial behavior that maximally profits the ruler.

The difficulty in applying this principle is that, in our model, the ruler’s position is contestable and subject to change. Because a ruler who has newly acquired or retained power is able to “reselect” an equilibrium after winning a contest for power, any ruler’s ability to select an equilibrium cannot extend beyond the termination of his regime. Our equilibrium selection, therefore, necessarily allows each ruler to optimally select only the “part” of an equilibrium corresponding to his own regime, while anticipating that other rulers do the

same with their “parts.” Such a selection should rule out, for example, equilibria in which the fall of one ruler implies the complete breakdown of society, with subsequent rulers and agents trapped in mutually destructive behavior.

To formalize our equilibrium selection, we first define for each t the *period history* h^t that records the actions of all players and realizations of all random variables prior to t :

$$h^0 := R_0$$

$$h^t := (h^{t-1}, \theta_{t-1}, q_{t-1}, a_{t-1}, A_{t-1}, e_{t-1}, x_{t-1}, c_{t-1}, W_{t-1}, R_t).$$

The set of all possible period histories is denoted $\mathcal{H}$. A period history h^t is a *regime history* if $t = 0$ or $c_{t-1} = 1$, that is, if it is the initial history or immediately follows the resolution of a contest for power. The set of all possible regime histories is denoted $\mathcal{B} \subseteq \mathcal{H}$.

Following each regime history, a new “regime” begins, which persists as long as no contest for power occurs. Formally, the *regime* that begins following $h^t \in \mathcal{B}$ is a subset of period histories, $\mathcal{D}(h^t) \subseteq \mathcal{H}$, such that $h^t \in \mathcal{D}(h^t)$ and, for each $\tau > t$, $h^\tau \in \mathcal{D}(h^t)$ if and only if $h^{\tau-1} \in \mathcal{D}(h^t)$ and $c_{\tau-1} = 0$. A regime has the following useful properties. First, $\mathcal{D}(h^t) \cap \mathcal{B} = \{h^t\}$, so that each regime can be uniquely identified by the regime history from which it begins. Hence, in what follows we refer to $\mathcal{D}(h^t)$ as regime h^t . Second, for each $h^\tau \in \mathcal{D}(h^t)$, $R_\tau = R_t$, so that each regime is ruled by one and only one ruler. Hence, in what follows we use $\mathcal{R}(h^t) \in \Theta$ to denote the ruler of regime h^t . Finally, for $h^t, h^{t'} \in \mathcal{B}$ such that $h^t \neq h^{t'}$, $\mathcal{D}(h^t) \cap \mathcal{D}(h^{t'}) = \emptyset$ and $\bigcup_{h^t \in \mathcal{B}} \mathcal{D}(h^t) = \mathcal{H}$, so that the collection of regimes constitutes a partition of the set of all period histories.

Our equilibrium selection has two components. The first is a bounded function $V: \Theta \rightarrow \mathbb{R}_+$, the *value of power*. This function maps the competence of each actor θ to a (discounted average) expected payoff $V(\theta)$ received from winning a contest for power. The value of power allows us to analyze each regime separately from all other regimes. By definition,

a regime ends following a contest for power and its resolution. Hence, a value of power V induces for each regime $h^t \in \mathcal{B}$ a *regime game* $G(\mathcal{R}(h^t), V)$ that begins with its ruler $\mathcal{R}(h^t)$ in power, has the same stage-game timing as the original game, and *terminates* with the resolution of a contest for power, following which the winner W_t receives the “terminal payoff” $\delta V(W_t)$ and the loser receives 0. Given that the terminal payoffs are pinned down by V , the regime game $G(\mathcal{R}(h^t), V)$ is independent of all other regimes and can be analyzed in isolation. This game admits a set of SPNEs, which we denote as $\Sigma(\mathcal{R}(h^t), V)$. Each $\sigma \in \Sigma(\mathcal{R}(h^t), V)$ provides the ruler of regime h^t the (discounted average) expected payoff

$$U(\sigma \mid \mathcal{R}(h^t), V) := (1 - \delta) \mathbb{E}^\sigma \left(\sum_{\tau=t}^{T(\sigma)} \delta^{\tau-t} (y_\tau - x_\tau) + g(A_{T(\sigma)}, \mathcal{R}(h^t)) \delta^{T(\sigma)-t+1} V(\mathcal{R}(h^t)) \right),$$

where $T(\sigma) := \inf \{ \tau \geq t : c_\tau = 1 \mid \sigma \}$. Intuitively, the ruler receives $y_t - x_t$, as specified by σ , for the duration of his regime, following which he receives his value of power if and only if he wins the contest for power that terminates regime h^t .

The second component of our equilibrium selection is a selection rule

$$\xi : \mathcal{B} \rightarrow \bigcup_{h^t \in \mathcal{B}} \Sigma(\mathcal{R}(h^t), V)$$

such that for each regime $h^t \in \mathcal{B}$, $\xi(h^t) \in \Sigma(\mathcal{R}(h^t), V)$ is the SPNE selected by its ruler $\mathcal{R}(h^t)$.

Given these two components, a *regime perfect equilibrium* (RPE) is a pair (V, ξ) , where V is a value of power and ξ is a selection rule such that the following two conditions hold:

1. (*Ruler optimality*) $U(\xi(h^t) \mid \mathcal{R}(h^t), V) \geq U(\sigma \mid \mathcal{R}(h^t), V)$ for all $h^t \in \mathcal{B}$ and all $\sigma \in \Sigma(\mathcal{R}(h^t), V)$;
2. (*Consistency*) $V(\theta) = U(\xi(h^t) \mid \theta, V)$ for all $\theta \in \Theta$ and all $h^t \in \mathcal{B}$ such that $\mathcal{R}(h^t) = \theta$.

Ruler-optimality is a “within-regime” condition. It requires that, given the “terminal pay-offs” following a contest for power specified by V , the selection rule ξ ensures the ruler of each regime an SPNE that maximizes his expected payoff at the beginning of the regime. Consistency, in turn, is a “cross-regime” condition. It stipulates that for each actor θ , the value of power $V(\theta)$ is derived from his expected payoff of optimally ruling a regime. These two conditions together imply that, in any RPE, every ruler optimally rules his regime under the assumption that all subsequent rulers do the same. Intuitively, each ruler selects only the “part” of an equilibrium that corresponds to his own regime, while the equilibrium selections of all possible rulers constitute an “equilibrium” in which no ruler has a unilateral profitable deviation.

Finally, from any RPE (V, ξ) , we can derive an SPNE of the original game. In particular, consider the strategy profile in the original game that prescribes for each $h^t \in \mathcal{H}$ the same actions as $\xi(h)$, where $h \in \mathcal{B}$ is the unique regime history such that $h^t \in \mathcal{D}(h)$. In the remainder of the paper, we focus on characterizing regime perfect equilibria. For convenience, we omit the script “ t ” whenever it causes no confusion.

4 Analysis

We begin by establishing a lower bound on the value of power for any ruler of competence θ .

Lemma 1. *In any regime perfect equilibrium (V, ξ) ,*

$$V(\theta) \geq \rho + \alpha\theta$$

holds for all $\theta \in \Theta$.

Proof. Omitted. □

Once becoming the ruler, an actor of competence θ can ensure a payoff each period of $\rho + \alpha\theta$

by recruiting no agent and ruling alone. If $V(\theta) > \rho + \alpha\theta$, the gap must consist of what θ gains from recruiting and managing an agent.

4.1 Stable ruling coalition

When a ruler recruits an agent, they form a *ruling coalition*. If the agent's competence is greater than the cost of effort, $A > k$, the surplus available to such a coalition is maximized when the agent exerts effort. A ruling coalition, however, is shadowed by two frictions. First, after the agent has helped the ruler to extract rents, the ruler has an incentive to keep all the rents for himself. Second, even after the ruler has shared rents with the agent, the agent has an incentive to contest the ruler's hold on power to gain control of future rent extraction. A *stable ruling coalition* is one in which the ruler and the agent overcome these two frictions and maintain a reciprocal relationship: in any period after the agent is recruited, the agent exerts effort and refrains from challenging the ruler, while the ruler shares rents sufficient to incentivize this behavior.

To explore this idea, fix a value of power V , consider a regime ruled by R , and suppose R recruits an agent A at t . Recalling that we normalize payoffs as per-period averages, the total surplus shared by the ruler and agent in a stable ruling coalition is $\rho + \alpha R + A - k$. The agent's share of this surplus is

$$(1 - \delta) \sum_{\tau=t}^{\infty} \delta^{\tau-t} (x_{\tau} - k),$$

whereas the ruler's share is

$$(1 - \delta) \sum_{\tau=t}^{\infty} \delta^{\tau-t} (\rho + \alpha R + A - x_{\tau}).$$

We establish the feasibility of a stable ruling coalition as follows. First, by challenging the

ruler at any $\tau \geq t$, the agent receives an expected payoff of

$$g(A, R) \cdot \delta V(A) \geq g(A, R) \cdot \delta(\rho + \alpha A) > 0,$$

where $g(A, R)$ is the agent's probability of winning the contest for power and $\delta V(A)$ is his continuation payoff from usurping power. With complementary probability, the agent loses the contest for power and receives 0 thereafter. To form a stable ruling coalition, the agent must find it profitable to refrain from challenging the ruler, so that

$$(1 - \delta) \sum_{\tau'=\tau+1}^{\infty} \delta^{\tau'-\tau} (x_{\tau'} - k) \geq g(A, R) \cdot \delta V(A) \quad (3)$$

must hold for all $\tau \geq t$, where the left-hand side is the agent's continuation value in a stable ruling coalition.

Second, at any $\tau \geq t$, by not sharing rents with the agent after he has exerted effort, the worst possible payoff the ruler can expect to receive is

$$(1 - \delta)(\rho + \alpha R + A) + (1 - g(A, R)) \cdot \delta V(R),$$

where the first term is total rents in the current period and the second term is the ruler's expected payoff in a contest for power with the agent. A stable ruling coalition requires the ruler to be willing to share rents, which necessitates that

$$\begin{aligned} (1 - \delta) \sum_{\tau'=\tau}^{\infty} \delta^{\tau'-\tau} (\rho + \alpha R + A - x_{\tau'}) \\ \geq (1 - \delta)(\rho + \alpha R + A) + (1 - g(A, R)) \cdot \delta V(R). \end{aligned} \quad (4)$$

holds for all $\tau \geq t$.

Third, at any $\tau \geq t$, by not exerting effort, the worst possible payoff the agent can expect

to receive is

$$g(A, R) \cdot \delta V(A),$$

which accrues when the ruler subsequently withdraws from sharing rents and the agent then chooses to challenge the ruler. A premise for a stable ruling coalition is that the agent is incentivized to exert effort and help the ruler to extract rents, so that

$$(1 - \delta) \sum_{\tau'=\tau}^{\infty} \delta^{\tau'-\tau} (x_{\tau'} - k) \geq g(A, R) \cdot \delta V(A) \quad (5)$$

must hold for all $\tau \geq t$.

The following lemma establishes that the same three conditions for a stable ruling coalition are sufficient for a SPNE of the regime game in which cooperation is supported, off the equilibrium path, by strategies that prescribe no effort, no transfer, and a contest for power.

Lemma 2. *Given any value of power V and any ruler $R \in \Theta$, a stable ruling coalition between R and an agent $A \in \Theta$ at $t \geq 0$ can be supported in at least one SPNE in $\Sigma(R, V)$ if and only if there exists a sequence $(x_\tau)_{\tau=t}^\infty \in \mathbb{R}_+^\infty$ for which Conditions 3, 4, and 5 hold for all $\tau \geq t$. Moreover, R 's largest SPNE payoff in a stable ruling coalition with A is*

$$\rho + \alpha R + A - [k + g(A, R) \cdot \delta V(A)]. \quad (6)$$

Proof. See online appendix. □

Unlike in some other models of relational contracts, the credibility of punishment in our setting holds trivially, given the analysis of each regime separately from all other regimes: off the equilibrium path, any deviation results in regime termination at most one period after punishment prescribes.

4.2 Optimal recruitment

Now fix an RPE (V, ξ) , consider a regime ruled by R , and suppose R inherits no agent at t from the previous period. By the definition of RPE, R 's equilibrium expected payoff at the beginning of t must be $V(R)$.

Consider the ruler's decision after some actor θ is drawn and thus available to be recruited as an agent. If a stable ruling coalition cannot be formed with that actor, so that recruitment would be followed by shirking and a contest for power, then the ruler unambiguously prefers not to recruit. Suppose instead that a stable ruling coalition can be formed. Following Lemma 2, by recruiting θ , the ruler receives at most

$$\rho + \alpha R + \theta - k - g(\theta, R) \cdot \delta V(\theta).$$

In contrast, by not recruiting θ and proceeding to the next period without an agent, the ruler receives

$$(1 - \delta)(\rho + \alpha R) + \delta V(R),$$

that is, whatever rent the ruler is able to extract on his own plus the discounted continuation value of starting the next period with no agent. The ruler thus prefers to recruit and establish a ruling coalition with θ only if

$$\rho + \alpha R + \theta - k - g(\theta, R) \cdot \delta V(\theta) \geq (1 - \delta)(\rho + \alpha R) + \delta V(R). \quad (7)$$

Lemma 3. *In any regime perfect equilibrium (V, ξ) , if a ruler $R \in \Theta$ at $t \geq 0$ inherits no agent from the previous period, he recruits a randomly drawn actor $\theta \in \Theta$ as his agent if and only if a stable ruling coalition between R and θ can be supported by at least one SPNE in $\Sigma(R, V)$ and Condition 7 holds. Moreover, if Condition 7 holds, a stable ruling coalition between R and θ is feasible.*

Proof. See appendix. □

The first sentence of Lemma 3 summarizes the discussion above. The second sentence is perhaps more surprising. The intuition behind this result is that, if the ruler is sufficiently patient, $\delta > \underline{\delta}$ (Condition 2), he would never find it optimal to recruit an agent, promise compensation sufficient to induce effort, and then renege on that promise and face a contest for power. Such a contest necessarily entails a loss in expected utility from the possibility of losing power, which must outweigh the short-term gain from reneging if $\delta > \underline{\delta}$. Formally, for any $\tau > t$, where t is the period of recruitment, Conditions 3 and 5 impose a minimal set of concessions the ruler must grant to incentivize effort and avoid the agent's challenge:

$$(1 - \delta) \sum_{\tau'=\tau}^{\infty} \delta^{\tau'-\tau} x_{\tau'} \geq k + g(A, R) \cdot V(A).$$

Given $\delta > \underline{\delta}$, Condition 7 implies Condition 4 for all $\tau > t$, assuming the ruler makes the minimal necessary concessions.⁸

4.3 Equilibrium value of power and the loyalty-competence tradeoff

The following theorem establishes the foundation for our main results.

Theorem 1. *In any regime perfect equilibrium (V, ξ) :*

1. *The value of power V is the unique fixed point of a contraction mapping Φ , where, for each $R \in \Theta$,*

$$\begin{aligned} & \Phi[V](R) \\ &:= \int_{\Theta} \max \{ \rho + \alpha R + \theta - [k + g(\theta, R) \cdot \delta V(\theta)], (1 - \delta)(\rho + \alpha R) + \delta V(R) \} dF(\theta). \end{aligned}$$

2. *For each $h \in \mathcal{B}$, $\xi(h)$ prescribes for $\mathcal{R}(h)$ a recruitment strategy such that actor θ is*

⁸Given the timing assumption that compensation precedes the decision to contest power, the optimal transfer is less in period t than in any period $\tau > t$, implying a slacker condition.

recruited as an agent if and only if $\theta \in Q(\mathcal{R}(h))$, where, for each $R \in \Theta$,

$$Q(R) := \{\theta \in \Theta : \rho + \alpha R + \theta - [k + g(\theta, R) \cdot \delta V(\theta)] \geq (1 - \delta)(\rho + \alpha R) + \delta V(R)\}.$$

3. The value of power $V(R)$ is strictly increasing and weakly convex in R .
4. The ruler's option value from recruiting an agent, $V(R) - (\rho + \alpha R)$, is weakly increasing in R .

Proof. See appendix. □

Theorem 1 establishes that, for each possible ruler, every RPE prescribes the same value of power V and same recruitment strategy Q . Hence, the RPE is essentially unique: any two equilibria share the same V and same Q . This result holds because the value of power is a fixed point of Φ , which is a contraction mapping on the space of bounded functions on Θ .⁹

Theorem 1 also states that the value of power is strictly increasing (and weakly convex) in a ruler's competence. To understand this result, observe that a more competent ruler extracts more rents on his own, while also sharing fewer rents with an agent of given competence (because the ruler has less to fear from a contest for power). The latter of these two direct effects, in turn, implies an indirect effect: more competent rulers can afford to recruit more competent agents, who help the ruler to extract more rents. The final statement in Theorem 1 immediately follows.

As noted in the introduction, a central idea in the literature on agency relationships in autocracies is the presence of a “loyalty-competence tradeoff.” This tradeoff emerges endogenously in our setting. The definition of Q implies that a ruler of competence R would recruit an

⁹The functional Φ is not monotone, so Blackwell's sufficient conditions do not apply, and the standard contraction argument fails. In particular, if V' majorizes V , it is not guaranteed that $\Phi[V']$ majorizes $\Phi[V]$, as

$$\rho + \alpha R + \theta - [k + g(\theta, R) \cdot \delta V(\theta)] \geq \rho + \alpha R + \theta - [k + g(\theta, R) \cdot \delta V'(\theta)].$$

actor of competence θ if and only if

$$\theta - [k + g(\theta, R) \cdot \delta V(\theta)] \geq \delta [V(R) - \rho - \alpha R]. \quad (8)$$

The right-hand side of this inequality is independent of θ , whereas the left-hand side represents the net profit R receives by recruiting θ and forming a stable ruling coalition with him. In this net profit, the first term is the marginal rent from recruitment of an agent of type θ , whereas the second is the transfer that induces effort by that agent and disincentivizes a contest for power. Even under the baseline assumption that $g(\theta, R)$ is only weakly increasing in θ , the transfer is increasing in θ , as the value of power is increasing in the agent's competence. Intuitively, a more competent agent anticipates that he can extract more rents as ruler, even if he recruits no agent. This increases the incentive to contest power, implying that a larger transfer is necessary to discourage that behavior.

Although the result is similar, our loyalty-competence tradeoff differs in mechanism from that in prior work on authoritarian politics. In [Egorov and Sonin \(2009\)](#), competent agents provide better protection to the ruler but have lower costs of betrayal, as they are better at discriminating among potential plotters. In [Zakharov \(2016\)](#), in turn, competent agents provide better protection to the ruler but have more value to a successor regime. In each case, the threat to the ruler comes from outside the regime. In our setting, in contrast, the agent himself is the threat, with more “disloyalty” when the agent anticipates being able to extract more rents if he successfully contests power.

5 Main results

We proceed to characterize the equilibrium of this model, considering in turn each of the three principles of autocratic governance discussed above.

5.1 The Idi Amin Principle

Our first principle states when a ruler would choose to recruit no agent, whatever a potential agent's competence. For completeness, we write the recruitment set Q as a function of the parameters ρ and α .

Proposition 1. *If $Q(R; \rho, \alpha) = \emptyset$, then:*

1. $Q(R'; \rho, \alpha) = \emptyset$ for all $R' < R$;
2. $Q(R; \rho, \alpha) = \emptyset$ for all $\rho' > \rho$.
3. $Q(R; \rho, \alpha) = \emptyset$ for all $\alpha' > \alpha$;

Moreover, there exists a unique $\rho^* > 0$ such that $Q(0; \rho, \alpha) = \emptyset$ if and only if $\rho > \rho^*$.

Proof. See online appendix. □

To understand Proposition 1, recall that the equilibrium transfer paid to any agent is

$$k + g(A, R) \cdot \delta V(A).$$

More competent rulers are more likely to win a contest for power (that is, g returns a smaller value), implying smaller equilibrium transfers and correspondingly greater incentive to recruit an agent. Equilibrium transfers, in turn, are increasing in an agent's value of power. That value is increasing in the natural rent, ρ , and in the degree to which any (future) ruler's competence matters for rent extraction, α . The compensation that must be paid to any agent is thus increasing in each of these parameters, again reducing the attractiveness of recruitment. (The current ruler's value of power is also increasing in each of these parameters, but this is true regardless of whether he recruits an agent.) Incompetent rulers who preside over countries endowed with large “natural rents” are less likely to have right-hand men: precisely the situation in Idi Amin's Uganda that we describe above.

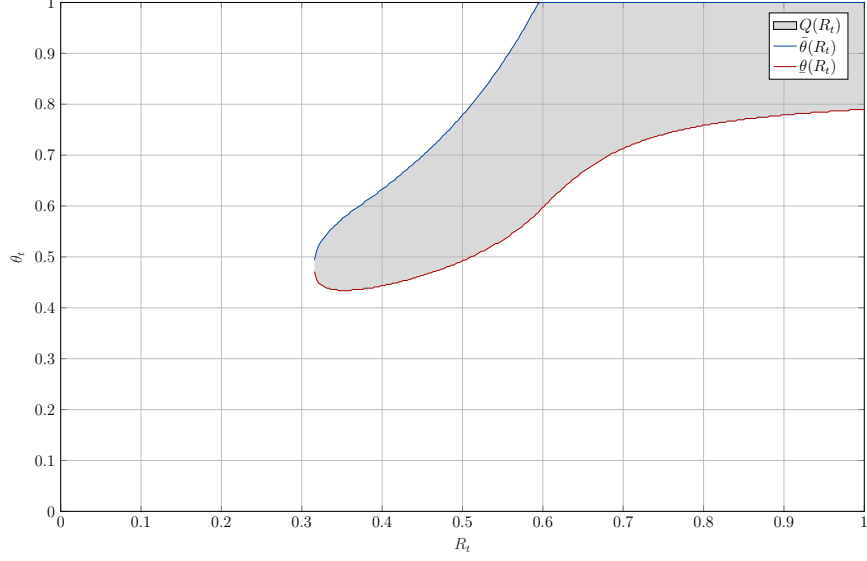


Figure 1: Equilibrium Retention Rule

5.2 The Molotov Principle

In the baseline model, we assume that the probability any ruler wins a contest for power is *weakly* increasing in the agent's (and ruler's) competence. Our second principle of autocratic governance requires that this relationship be strict.

Assumption 1. *The probability any agent wins a contest for power is strictly increasing in that agent's competence:*

$$g(A, R) < g(A', R),$$

for all $A < A'$ and all R .

Proposition 2. *Suppose Assumption 1 holds. Then each ruler recruits from an interval: there exist $\underline{\theta}, \bar{\theta}: \Theta \rightarrow \Theta$ such that $Q(R) = [\underline{\theta}(R), \bar{\theta}(R)]$.*

Proof. See online appendix. □

Theorem 1 gives the set of agents a ruler would recruit:

$$Q(R) = \{\theta \in \Theta : \theta - g(\theta, R) \cdot \delta V(\theta) \geq k + \delta [V(R) - \rho - \alpha R]\}.$$

Proposition 2 thus amounts to showing that $\theta - g(\theta, R) \cdot \delta V(\theta)$ is strictly concave in θ , implying that this set is convex—that is, an interval. This is true by Lemma 1, which establishes that the value of power is positive; by Theorem 1, which further states that the value of power is strictly increasing and weakly convex; and by Assumption 1, which imposes that the contest function g is strictly increasing in an agent’s competence. As seen in Figure 1, this “interval recruiting” often means enlisting agents who are neither especially competent nor particularly incompetent—agents who, in our telling, are like Vyacheslav Molotov.

5.3 The Atatürk Principle

Our final principle of autocratic governance rests on the additional assumption—a single-crossing condition—that a ruler’s competence is more important for contest success when an agent is more competent.

Assumption 2. *A ruler’s competence matters weakly more for contest success when the agent is more competent:*

$$g(A, R) - g(A, R') \leq g(A', R) - g(A', R'),$$

for all $A < A'$ and $R < R'$.

Proposition 3. *Suppose Assumption 2 holds. Then more competent rulers are willing to recruit more competent agents: $\bar{\theta}(R)$ is weakly increasing in R .*

Proof. See online appendix. □

Under Assumption 2, the marginal increase in compensation that a ruler must provide a more competent agent is smaller when the ruler is more competent. To see this, observe that the difference in equilibrium transfers between a more (R') and less (R) competent ruler is

$$[k + g(A, R') \cdot \delta V(A)] - [k + g(A, R) \cdot \delta V(A)] = -[g(A, R) - g(A, R')] \delta V(A).$$

Assumption 2 guarantees that the right-hand expression in brackets is weakly increasing in A , while Theorem 1 establishes that the value of power is strictly increasing, so the difference is weakly negative. A more competent ruler is more inclined to recruit a more competent agent, as he need not share as many rents with him.

There is, however, a force that pushes in the opposition direction. The option value of deferring recruitment and drawing another potential agent the next period,

$$V(R) - (\rho + \alpha R),$$

is also larger for more competent rulers, as established by Theorem 1. For the most competent type that a ruler might recruit, this effect is second-order, as there are not many agents more competent than the current potential recruit. In contrast, as Figure 1 illustrates, either effect may dominate at the lower bound of the recruitment interval.

The proper reading of Proposition 3 is that more competent rulers are willing to recruit more competent agents, not that they always do so. Kemal Atatürk was lucky to get a good draw in İsmet İnönü. It was Atatürk’s manifest competence that made it attractive to keep him at his side.

6 Conclusion

In this paper, we develop a relational theory of autocracy. Relative to prior work on relational contracts, our key departure is to assume that relational contracts are contestable—that is, that the agent can attempt to become the principal. For an autocracy, this is the natural assumption, but control rights are also contestable in some other organizations. Workplace democracies (such as the Mondragon Corporation), start-ups (including, famously, OpenAI during Sam Altman’s removal), and family firms all exhibit some ability of employees to replace the principal. Leadership challenges in parliamentary systems also have this flavor; the Blair and Johnson governments are analogous to regimes in our setting. Not least,

the leaders of terrorist groups and bosses of criminal gangs serve at their pleasure of their operatives.

There are many natural extensions of the theoretical framework that we develop here. These include: assuming that an agent may be forced upon a ruler (a measure of institutionalization of the autocratic regime); allowing the ruler to punish an agent beyond merely returning him to inactive status; and relaxing the assumption that an agent, once recruited, can only be replaced following a contest for power (though a contest for power may be the consequence of a purge). These and other extensions can lead to a more fully developed relational theory of autocracy.

Appendix

Proof of Lemma 3

The first part of the lemma is established in the text. Here, we show that if Condition 7 holds, then a stable ruling coalition between R and θ is feasible. According to Lemma 2, it is sufficient to prove that there exists a sequence $(x_\tau)_{\tau=t}^\infty$ that satisfies Conditions 3, 4, and 5 given $A = \theta$.

Consider $(x_\tau)_{\tau=t}^\infty$ such that $x_t = k$ and $x_\tau = k + g(\theta, R)V(\theta)$ for all $\tau > t$. Clearly, $(x_\tau)_{\tau=t}^\infty$ satisfies Conditions 3 and 5. For Condition 4, note that $\Theta = [0, 1]$ implies $V(\theta) \leq 1 + \rho + \alpha$ (effort, maximal competence of both agent and ruler) and $V(R) \geq \rho$ (per Lemma 1, assuming minimal competence). Hence, the assumption $\delta > \underline{\delta}$ leads to

$$\delta > \frac{\theta + g(\theta, R)V(\theta)}{\theta + g(\theta, R)V(\theta) + g(\theta, R)V(R)},$$

which is equivalent to

$$\begin{aligned} & (1 - \delta)(\rho + \alpha R) + \delta V(R) \\ & > (1 - \delta)(\rho + \alpha R + \theta) + (1 - g(\theta, R))\delta V(R) + g(\theta, R)(1 - \delta)V(\theta). \end{aligned}$$

Thus, Condition 7 implies

$$\begin{aligned} & \rho + \alpha R + \theta - k - g(\theta, R)\delta V(\theta) \\ & > (1 - \delta)(\rho + \alpha R + \theta) + (1 - g(\theta, R))\delta V(R) + g(\theta, R)(1 - \delta)V(\theta), \end{aligned}$$

which in turn implies

$$\begin{aligned} (1 - \delta) \sum_{\tau'=\tau}^{\infty} \delta^{\tau'-\tau} (\rho + \alpha R + \theta - x_{\tau'}) &= \rho + \alpha R + \theta - k - g(\theta, R)V(\theta) \\ &> (1 - \delta)(\rho + \alpha R + \theta) + (1 - g(\theta, R))\delta V(R) \end{aligned}$$

holds for all $\tau > t$. Finally, for $\tau = t$,

$$\begin{aligned} (1 - \delta) \sum_{\tau'=t}^{\infty} \delta^{\tau'-\tau} (\rho + \alpha R + \theta - x_{\tau'}) &= \rho + \alpha R + \theta - k - \delta g(\theta, R)V(\theta) \\ &> \rho + \alpha R + \theta - k - g(\theta, R)V(\theta) \\ &> (1 - \delta)(\rho + \alpha R + \theta) + (1 - g(\theta, R))\delta V(R). \end{aligned}$$

Therefore, $(x_{\tau})_{\tau=t}^{\infty}$ also satisfies Condition 4. □

Proof of Theorem 1

Step 1

We first show that Φ is a contraction mapping in the space of nonnegative bounded functions on Θ with the sup norm “ $\|\cdot\|$.” Note that for each bounded function s on Θ and $R \in \Theta$,

$$\begin{aligned} & \Phi[s](R) \\ &= \rho + \alpha R + \int_{\Theta} \max \{ \theta - k - g(\theta, R)\delta s(\theta), \delta(s(R) - \rho - \alpha R) \} dF(\theta) \\ &= \rho + \alpha R + \delta(s(R) - \rho - \alpha R) + \int_{Q_s(R)} (\theta - k - g(\theta, R)\delta s(\theta) - \delta(s(R) - \rho - \alpha R)) dF(\theta), \end{aligned}$$

where

$$Q_s(R) := \{\theta \in \Theta : \theta - k - g(\theta, R)\delta s(\theta) \geq \delta(s(R) - \rho - \alpha R)\}.$$

For reference, note that by definition of $Q_s(R)$,

$$\begin{aligned} & \int_{Q_s(R)} (\theta - k - g(\theta, R)\delta s(\theta) - \delta(s(R) - \rho - \alpha R)) dF(\theta) \\ & \geq \int_X (\theta - k - g(\theta, R)\delta s(\theta) - \delta(s(R) - \rho - \alpha R)) dF(\theta) \end{aligned} \quad (9)$$

holds for any subset $X \subseteq \Theta$.

Consider two nonnegative bounded functions on Θ , s and s' . For $R \in \Theta$ such that $\Phi[s](R) \geq \Phi[s'](R)$,

$$\begin{aligned} & |\Phi[s](R) - \Phi[s'](R)| \\ & = \Phi[s](R) - \Phi[s'](R) \\ & \leq \delta(s(R) - s'(R)) - \int_{Q_s(R)} (g(\theta, R)\delta(s(\theta) - s'(\theta)) + \delta(s(R) - s'(R))) dF(\theta) \\ & = \left(1 - \int_{Q_s(R)} dF(\theta)\right) \delta(s(R) - s'(R)) + \int_{Q_s(R)} g(\theta, R)\delta(s'(\theta) - s(\theta)) dF(\theta) \\ & \leq \left(1 - \int_{Q_s(R)} dF(\theta)\right) \cdot \delta\|s - s'\| + \left(\int_{Q_s(R)} dF(\theta)\right) \cdot \left(\sup_{\theta' \in \Theta} g(\theta', R)\right) \delta\|s - s'\| \leq \delta\|s - s'\|, \end{aligned}$$

where the first weak inequality holds, per Condition 9, because

$$\begin{aligned} & \int_{Q_{s'}(R)} (\theta - k - g(\theta, R)\delta s'(\theta) - \delta(s'(R) - \rho - \alpha R)) dF(\theta) \\ & \geq \int_{Q_s(R)} (\theta - k - g(\theta, R)\delta s'(\theta) - \delta(s'(R) - \rho - \alpha R)) dF(\theta). \end{aligned}$$

Similarly, for $R \in \Theta$ such that $\Phi[s](R) < \Phi[s'](R)$,

$$\begin{aligned}
& |\Phi[s](R) - \Phi[s'](R)| \\
&= \Phi[s'](R) - \Phi[s](R) \\
&\leq \delta(s'(R) - s(R)) - \int_{Q_{s'}(R)} (g(\theta, R)\delta(s'(\theta) - s(\theta)) + \delta(s'(R) - s(R))) dF(\theta) \\
&= \left(1 - \int_{Q_{s'}(R)} dF(\theta)\right) \delta(s'(R) - s(R)) + \int_{Q_{s'}(R)} g(\theta, R)\delta(s(\theta) - s'(\theta)) dF(\theta) \\
&\leq \left(1 - \int_{Q_{s'}(R)} dF(\theta)\right) \cdot \delta\|s - s'\| + \left(\int_{Q_{s'}(R)} dF(\theta)\right) \cdot \left(\sup_{\theta' \in \Theta} g(\theta', R)\right) \delta\|s - s'\| \leq \delta\|s - s'\|,
\end{aligned}$$

where the first weak inequality holds, per Condition 9, because

$$\begin{aligned}
& \int_{Q_s(R)} (\theta - k - g(\theta, R)\delta s(\theta) - \delta(s(R) - \rho - \alpha R)) dF(\theta) \\
&\geq \int_{Q_{s'}(R)} (\theta - k - g(\theta, R)\delta s(\theta) - \delta(s(R) - \rho - \alpha R)) dF(\theta).
\end{aligned}$$

Therefore,

$$\|\Phi[s] - \Phi[s']\| = \sup_{R \in \Theta} |\Phi[s](R) - \Phi[s'](R)| \leq \delta\|s - s'\|,$$

so that Φ is a contraction mapping and admits a unique fixed point V .

Step 2

As the unique fixed point of Φ , V is the unique value of power. The second statement in Theorem 1 is therefore established by defining $Q(R) := Q_V(R)$ for each $R \in \Theta$.

Step 3

We now prove that $V(R) - \rho - \alpha R$ is weakly increasing in R (the fourth statement in Theorem 1), which in turn implies that V is strictly increasing (the first half of the third statement in Theorem 1).

Consider any nonnegative bounded function s on Θ such that $s(R) - (\rho + \alpha R)$ is weakly increasing in R , and $R, R' \in \Theta$ such that $R < R'$. Then,

$$\begin{aligned}
& \Phi[s](R') - \rho - \alpha R' \\
& \geq \delta(s(R') - \rho - \alpha R') + \int_{Q_s(R)} (\theta - k - g(\theta, R') \delta s(\theta) - \delta(s(R') - \rho - \alpha R')) dF(\theta) \\
& = \left(1 - \int_{Q_s(R)} dF(\theta)\right) \delta(s(R') - \rho - \alpha R') + \int_{Q_s(R)} (\theta - k - g(\theta, R') \delta s(\theta)) dF(\theta) \\
& \geq \left(1 - \int_{Q_s(R)} dF(\theta)\right) \delta(s(R) - \rho - \alpha R) + \int_{Q_s(R)} (\theta - k - g(\theta, R) \delta s(\theta)) dF(\theta) \\
& = \delta(s(R) - \rho - \alpha R) + \int_{Q_s(R)} (\theta - k - g(\theta, R) \delta s(\theta) - \delta(s(R) - \rho - \alpha R)) dF(\theta) \\
& = \Phi[s](R) - \rho - \alpha R,
\end{aligned}$$

where the first weak inequality is because the definition of $Q_s(R')$ implies that

$$\begin{aligned}
& \int_{Q_s(R')} (\theta - k - g(\theta, R') \delta s(\theta) - \delta(s(R') - \rho - \alpha R')) dF(\theta) \\
& \geq \int_{Q_s(R)} (\theta - k - g(\theta, R') \delta s(\theta) - \delta(s(R') - \rho - \alpha R')) dF(\theta),
\end{aligned}$$

while the second weak inequality is because $s(R) - \rho - \alpha R$ is weakly increasing in R and $g(\theta, R)$ is weakly decreasing in R . Hence, $\Phi[s](R) - \rho - \alpha R$ is weakly increasing in R .

Therefore, for any nonnegative bounded function s on Θ such that $s(R) - (\rho + \alpha R)$ is weakly increasing in R , $\Phi[s](R) - (\rho + \alpha R)$ is also weakly increasing in R . Consequently, V , as the unique fixed point of Φ , must satisfy that $V(R) - (\rho + \alpha R)$ is weakly increasing in R .

Step 4

Finally, we prove that V is weakly convex (the second half of the third statement in Theorem 1). It is sufficient to prove that for any nonnegative bounded function s such that $s(R) - (\rho + \alpha R)$ is weakly convex in R , $\Phi[s](R) - (\rho + \alpha R)$ is also weakly convex in R .

Consider $R', R'' \in \Theta$ and $\lambda \in (0, 1)$, and let $R := (1 - \lambda)R' + \lambda R''$. Note that for any weakly

convex function s on Θ , because

$$\begin{aligned}
& \Phi[s](R') - \rho - \alpha R' \\
& \geq \delta(s(R') - \rho - \alpha R') + \int_{Q_s(R)} (\theta - k - g(\theta, R') \delta s(\theta) - \delta(s(R') - \rho - \alpha R')) dF(\theta) \\
& = \left(1 - \int_{Q_s(R)} dF(\theta)\right) \delta(s(R') - \rho - \alpha R') + \int_{Q_s(R)} (\theta - k - g(\theta, R') \delta s(\theta)) dF(\theta)
\end{aligned}$$

and

$$\begin{aligned}
& \Phi[s](R'') - \rho - \alpha R'' \\
& \geq \delta(s(R'') - \rho - \alpha R'') + \int_{Q_s(R)} (\theta - k - g(\theta, R'') \delta s(\theta) - \delta(s(R'') - \rho - \alpha R'')) dF(\theta) \\
& = \left(1 - \int_{Q_s(R)} dF(\theta)\right) \delta(s(R'') - \rho - \alpha R'') + \int_{Q_s(R)} (\theta - k - g(\theta, R'') \delta s(\theta)) dF(\theta)
\end{aligned}$$

there holds

$$\begin{aligned}
& (1 - \lambda) (\Phi[s](R') - \rho - \alpha R') + \lambda (\Phi[s](R'') - \rho - \alpha R'') \\
& \geq \left(1 - \int_{Q_s(R)} dF(\theta)\right) \delta(s(R) - \rho - \alpha R) + \int_{Q_s(R)} (\theta - k - g(\theta, R) \delta s(\theta)) dF(\theta) \\
& = \delta(s(R) - \rho - \alpha R) + \int_{Q_s(R)} (\theta - k - g(\theta, R) \delta s(\theta) - \delta(s(R) - \rho - \alpha R)) dF(\theta) \\
& = \Phi[s](R) - \rho - \alpha R,
\end{aligned}$$

where the first weak inequality holds because s is weakly convex and $g(\theta, \cdot)$ is weakly concave.

Hence, $\Phi[s](R) - (\rho + \alpha R)$ is indeed weakly convex in R . $\square$

References

Acharya, Avidit, Stephen H. Haber and Alexander Lee. 2024. “The Political Economy of the Middle Income Trap.”.

Ahlquist, John S. and Anna Grzymala-Busse. 2025. “Political Organizations & Organiza-

- tional Politics.” Unpublished manuscript.
- Angelucci, Charles, Simone Meraglia and Nico Voigtländer. 2025. “Organizing a Kingdom.” NBER Working Paper 32542.
- Baker, George, Robert Gibbons and Kevin J. Murphy. 1994. “Subjective Performance Measures in Optimal Incentive Contracts.” *Quarterly Journal of Economics* 109(4):1125–1156.
- Baker, George, Robert Gibbons and Kevin J. Murphy. 1999. “Informal Authority in Organizations.” *Journal of Law, Economics, and Organization* 15(1):56–73.
- Baker, George, Robert Gibbons and Kevin J. Murphy. 2002. “Relational Contracts and the Theory of the Firm.” *Quarterly Journal of Economics* 117(1):39–84.
- Blattman, Christopher, Scott Gehlbach and Zeyang Yu. 2026. “The Personalist Penalty: Varieties of Autocracy and Economic Growth.” NBER Working Paper 34093.
- Board, Simon and Moritz Meyer-Ter-Vehn. 2015. “Relational Contracts in Competitive Labour Markets.” *Review of Economic Studies* 82(2):490–534.
- Boix, Carles, Michael Miller and Sebastian Rosato. 2013. “A Complete Data Set of Political Regimes, 1800–2007.” *Comparative Political Studies* 46(12):1523–1554.
- Bueno de Mesquita, Bruce and Alastair Smith. 2025. *Anticipating Instability: Assessing and Removing Risks before They Happen*. Cambridge, MA: MIT Press. Unpublished book manuscript.
- Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.
- Callander, Steven, Dana Foarta and Takuo Sugaya. 2022. “Market Competition and Political Influence: An Integrated Approach.” *Econometrica* 90(6):2723–2753.

- Clarke, Killian, Anne Meng and Jack Paine. 2025. “Violent Origins and Authoritarian Order: Divergent Trajectories after Successful Rebellions.” Unpublished manuscript.
- Debs, Alexandre. 2007. “On Dictatorships.” PhD thesis, MIT.
- Dixit, Avinash. 2004. *Lawlessness and Economics: Alternative Modes of Governance*. Princeton, NJ: Princeton University Press.
- Egorov, Georgy and Konstantin Sonin. 2009. “Dictators and Their Viziers: Endogenizing the Loyalty-Competence Tradeoff.” *Journal of the European Economic Association* 9(5):903–930.
- Egorov, Georgy and Konstantin Sonin. 2025. “Why Did Putin Invade Ukraine? A Theory of Degenerate Autocracy.” *American Journal of Political Science* Forthcoming. Unpublished manuscript.
- Egorov, Georgy, Sergei Guriev and Konstantin Sonin. 2009. “Why Resource-Poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data.” *American Political Science Review* 103(4):645–668.
- Foarta, Dana and Michael Ting. 2023. “Organizational Capacity and Project Dynamics.” Stigler Center New Working Paper Series 339.
- Gandhi, Jennifer. 2008. *Political Institutions under Dictatorship*. New York: Cambridge University Press.
- Gandhi, Jennifer and Jane Lawrence Sumner. 2020. “Measuring the Consolidation of Power in Nondemocracies.” *Journal of Politics* 82(4):1545–1558.
- Geddes, Barbara, Joseph Wright and Erica Frantz. 2014. “Autocratic Breakdown and Regime Transitions: A New Data Set.” *Perspectives on Politics* 12(2):313–331.
- Geddes, Barbara, Joseph Wright and Erica Frantz. 2018. *How Dictatorships Work: Power, Personalization, and Collapse*. New York: Cambridge University Press.

- Gehlbach, Scott and Alberto Simpser. 2015. "Electoral Manipulation as Bureaucratic Control." *American Journal of Political Science* 59(1):212–224.
- Gehlbach, Scott and Philip Keefer. 2011. "Investment without Democracy: Ruling-party Institutionalization and Credible Commitment in Autocracies." *Journal of Comparative Economics* 39(2):123–139.
- Gehlbach, Scott and Philip Keefer. 2012. "Private Investment and the Institutionalization of Collective Action in Autocracies: Ruling Parties and Legislatures." *Journal of Politics* 74(2):621–635.
- Goemans, Henk E., Kristian Skrede Gleditsch and Giacomo Chiozza. 2009. "Introducing Archigos: A Dataset of Political Leaders." *Journal of Peace Research* 46(2):269–283.
- Lee, Don S. and Paul Schuler. 2020. "Testing the "China Model" of Meritocratic Promotions: Do Democracies Reward Less Competent Ministers than Autocracies?" *Comparative Political Studies* 53(3-4):531–566.
- Levin, Jonathan. 2003. "Relational Incentive Contracts." *American Economic Review* 93(3):835–857.
- MacLeod, W. Bentley and James M. Malcomson. 1989. "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica* 57(2):447–480.
- Meng, Anne. 2020. *Constraining Dictatorship: From Personalized Rule to Institutionalized Regimes*. New York: Cambridge University Press.
- 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.
- Moe, Terry M. 1984. "The New Economics of Organization." *American Journal of Political Science* 28(4):739–777.

- Montagnes, B. Pablo and Stéphane Wolton. 2019. “Mass Purges: Top-down Accountability in Autocracy.” *American Political Science Review* 113(4):1045–1059.
- Myerson, Roger B. 2008. “The Autocrat’s Credibility Problem and Foundations of the Constitutional State.” *American Political Science Review* 102(1):125–139.
- Pan, Mingcong. 2025. “Personalization and Its Governance Consequences in Authoritarian Regimes: Theory and Evidence from China.” Unpublished manuscript.
- Rundlett, Ashlea and Milan W. Svolik. 2016. “Deliver the Vote! Micromotives and Macrobehavior in Electoral Fraud.” *American Political Science Review* 110(1):180–197.
- Snowberg, Erik and Michael M Ting. 2022. “An Organizational Theory of State Capacity.” Unpublished manuscript.
- Svolik, Milan W. 2012. *The Politics of Authoritarian Rule*. New York: Cambridge University Press.
- Tyson, Scott A. 2018. “The Agency Problem Underlying Repression.” *Journal of Politics* 80(4):1297–1310.
- Tyson, Scott A. 2020. “The Strategic Foundations of Political Sovereignty.” *Journal of Politics* 82(2):657–670.
- Weingast, Barry R. and William Marshall. 1988. “The Industrial Organization of Congress; or Why Legislatures, like Firms, Are Not Organized as Markets.” *Journal of Political Economy* 96(1):132–163.
- Wright, Joseph. 2008. “Do Authoritarian Institutions Constrain? How Legislatures Affect Economic Growth and Investment.” *American Journal of Political Science* 52(2):322–343.
- Zakharov, Alexei V. 2016. “The Loyalty-Competence Trade-off in Dictatorships and Outside Options for Subordinates.” *Journal of Politics* 78(2):457–466.

A Relational Theory of Autocracy: Online Appendix

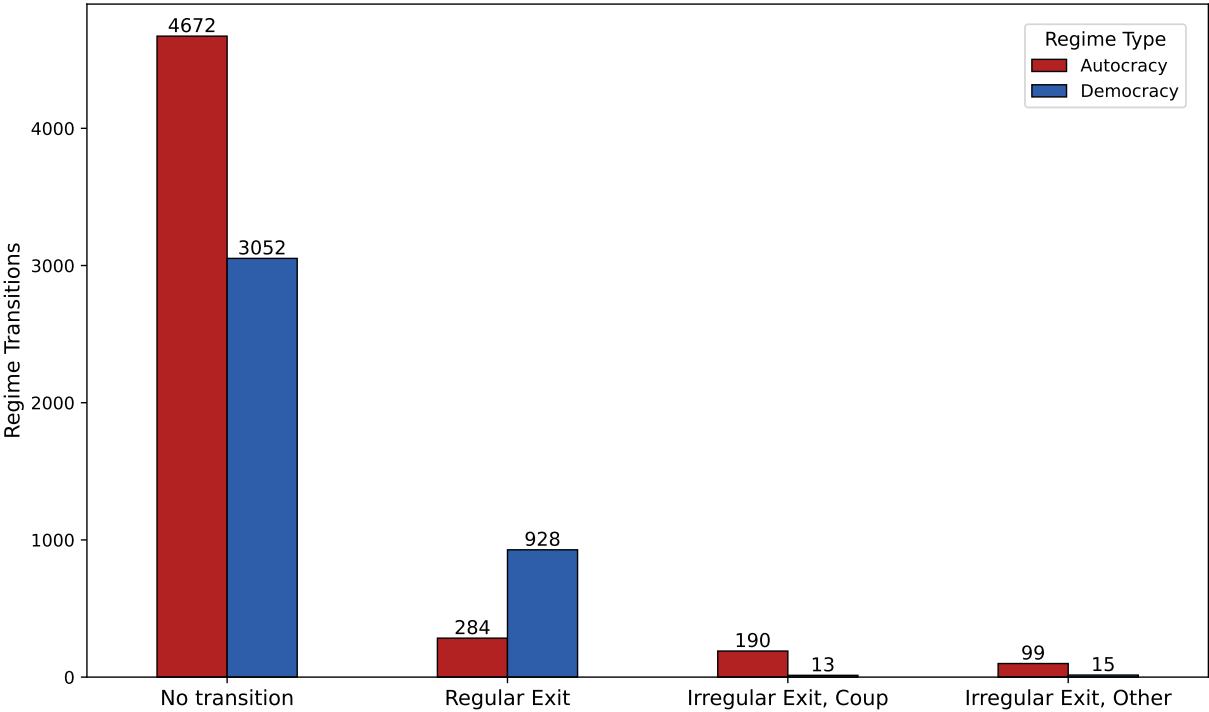


Figure A1: Exits from power in autocracies and democracies.

A.1 Proof of Lemma 2

Necessity is established in the main text. For sufficiency, consider the strategy profile in $\Sigma(R, V)$ that prescribes, on the equilibrium path, for each $\tau \geq t$, $e_\tau = 1$, x_τ as specified by $(x_\tau)_{\tau=t}^\infty$, and $c_\tau = 0$; and that prescribes, in response to a deviation by either player, no effort, zero transfer, and a challenge. On the equilibrium path, Conditions 3, 4, and 5 ensure that there is no profitable one-shot deviation. Off the equilibrium path, sequential rationality is satisfied trivially, given that the regime terminates with the resolution of a contest for power. Together, these imply that the strategy profile is an SPNE.

The upper bound of the ruler's payoff in Expression 6 is derived from Condition 5, which can be rewritten as

$$(1 - \delta) \sum_{\tau=t}^{\infty} \delta^{\tau-t} x_\tau \geq k + g(A, R) \cdot \delta V(A),$$

where the right-hand side represents the smallest concession the ruler need offer the agent in a stable ruling coalition.

□

A.2 Proof of Proposition 1

In this proof, we write $V(R; \alpha, \rho)$, $\Phi[V](R; \alpha, \rho)$, and $Q(R; \alpha, \rho)$ to clarify the explicit dependence on parameters. The same arguments in Step 1 in the proof of Theorem 1 can prove that Φ is a contraction mapping in the space of nonnegative bounded functions on $\Theta \times \mathbb{R}_+^2$.

Step 1

We first show that for any $R \in \Theta$, $Q(R; \alpha, \rho) = \emptyset$ holds if and only if

$$\sup_{\theta \in \Theta} (\theta - k - g(\theta, R) \delta V(\theta; \alpha, \rho)) < 0. \quad (\text{A1})$$

Consider any $R \in \Theta$. First, suppose Condition A1 holds. As shown in Step 1 in the proof

of Theorem 1,

$$Q(R; \alpha, \rho) = \{\theta \in \Theta : \theta - k - g(\theta, R)\delta V(\theta; \alpha, \rho) \geq \delta(V(R; \alpha, \rho) - \rho - \alpha R)\}.$$

By Lemma 1, $V(R; \alpha, \rho) - \rho - \alpha R \geq 0$, so that Condition **A1** implies that for all $\theta \in \Theta$,

$$\theta - k - g(\theta, R)\delta V(\theta; \alpha, \rho) < 0 \leq V(R; \alpha, \rho) - \rho - \alpha R,$$

which in turn implies $Q(R; \alpha, \rho) = \emptyset$. Conversely, suppose $Q(R; \alpha, \rho) = \emptyset$, then as shown in Step 1 in the proof of Theorem 1,

$$V(R; \alpha, \rho) = \rho + \alpha R + \delta(V(R; \alpha, \rho) - \rho - \alpha R) = \rho + \alpha R,$$

so that

$$Q(R; \alpha, \rho) = \{\theta \in \Theta : \theta - k - g(\theta, R)\delta V(\theta; \alpha, \rho) \geq 0\}.$$

Then, $Q(R; \alpha, \rho) = \emptyset$ implies that $\theta - k - g(\theta, R)\delta V(\theta; \alpha, \rho) < 0$ holds for all $\theta \in \Theta$, so that Condition **A1** holds.

Step 2

Here we prove the first result in the proposition. For any $\theta \in \Theta$, because $g(\theta, \cdot)$ is weakly decreasing, $\theta - k - g(\theta, R)\delta V(\theta; \alpha, \rho) < 0$ implies $\theta - k - g(\theta, R')\delta V(\theta; \alpha, \rho) < 0$ holds for all $R' < R$. Hence, if Condition **A1** holds for R , it holds as well for all $R' > R$.

Step 3

Now we prove the second and third results in the proposition.

Consider any nonnegative bounded function s on $\Theta \times \mathbb{R}_+^2$ such that $\frac{s(R; \alpha, \rho') - s(R; \alpha, \rho)}{\rho' - \rho} \in [0, 1]$ holds for all $R \in \Theta$, $\alpha \in \mathbb{R}_+$, and $\rho, \rho' \in \mathbb{R}_+$ such that $\rho < \rho'$.

Note that for any $R \in \Theta$, $\alpha \in \mathbb{R}_+$, and $\rho \in \mathbb{R}_+$,

$$\begin{aligned}
& \Phi[s](R; \alpha, \rho) - \rho - \alpha R \\
&= \delta(s(R; \alpha, \rho) - \rho - \alpha R) \\
&\quad + \int_{Q_s(R; \alpha, \rho)} (\theta - k - g(\theta, R) \delta s(\theta; \alpha, \rho) - \delta(s(R; \alpha, \rho) - \rho - \alpha R)) dF(\theta) \\
&= \left(1 - \int_{Q_s(R; \alpha, \rho)} dF(\theta)\right) \delta(s(R; \alpha, \rho) - \rho - \alpha R) + \int_{Q_s(R; \alpha, \rho)} (\theta - k - g(\theta, R) \delta s(\theta; \alpha, \rho)) dF(\theta)
\end{aligned}$$

and, for any subset $X \subseteq \Theta$, because

$$\begin{aligned}
& \int_{Q_s(R; \alpha, \rho)} (\theta - k - g(\theta, R) \delta s(\theta; \alpha, \rho) - \delta(s(R; \alpha, \rho) - \rho - \alpha R)) dF(\theta) \\
&\geq \int_X (\theta - k - g(\theta, R) \delta s(\theta; \alpha, \rho) - \delta(s(R; \alpha, \rho) - \rho - \alpha R)) dF(\theta),
\end{aligned}$$

there holds

$$\begin{aligned}
& \Phi[s](R; \alpha, \rho) - \rho - \alpha R \\
&\geq \delta(s(R; \alpha, \rho) - \rho - \alpha R) + \int_X (\theta - k - g(\theta, R) \delta s(\theta; \alpha, \rho) - \delta(s(R; \alpha, \rho) - \rho - \alpha R)) dF(\theta) \\
&= \left(1 - \int_X dF(\theta)\right) \delta(s(R; \alpha, \rho) - \rho - \alpha R) + \int_X (\theta - k - g(\theta, R) \delta s(\theta; \alpha, \rho)) dF(\theta).
\end{aligned}$$

Now fixing $R \in \Theta$ and $\alpha \in \mathbb{R}_+$ and consider $\rho, \rho' \in \mathbb{R}_+$ such that $\rho < \rho'$. First, because

$$\begin{aligned}
& \Phi[s](R; \alpha, \rho') - \rho' - \alpha R \\
&\geq \left(1 - \int_{Q_s(R; \alpha, \rho)} dF(\theta)\right) \delta(s(R; \alpha, \rho') - \rho' - \alpha R) + \int_{Q_s(R; \alpha, \rho)} (\theta - k - g(\theta, R) \delta s(\theta; \alpha, \rho')) dF(\theta),
\end{aligned}$$

there is

$$\begin{aligned}
& \Phi[s](R; \alpha, \rho') - \Phi[s](R; \alpha, \rho) \\
& \geq \rho' - \rho + \left(1 - \int_{Q_s(R; \alpha, \rho)} dF(\theta)\right) \delta(s(R; \alpha, \rho') - s(R; \alpha, \rho) - (\rho' - \rho)) \\
& \quad - \int_{Q_s(R; \alpha, \rho)} g(\theta, R) \delta(s(\theta; \alpha, \rho') - s(\theta; \alpha, \rho)) dF(\theta) \\
& = (1 - \delta)(\rho' - \rho) + \left(1 - \int_{Q_s(R; \alpha, \rho)} dF(\theta)\right) \delta(s(R; \alpha, \rho') - s(R; \alpha, \rho)) \\
& \quad + \delta(\rho' - \rho) \int_{Q_s(R; \alpha, \rho)} \left(1 - g(\theta, R) \frac{s(\theta; \alpha, \rho') - s(\theta; \alpha, \rho)}{\rho' - \rho}\right) dF(\theta) \\
& \geq (1 - \delta)(\rho' - \rho) > 0,
\end{aligned}$$

where the second weak inequality holds because $s(R; \alpha, \rho') - s(R; \alpha, \rho) \geq 0$ and $\frac{s(\theta; \alpha, \rho') - s(\theta; \alpha, \rho)}{\rho' - \rho} \leq 1$ for all $\theta \in \Theta$. Second, because

$$\begin{aligned}
& \Phi[s](R; \alpha, \rho) - \rho - \alpha R \\
& \geq \left(1 - \int_{Q_s(R; \alpha, \rho')} dF(\theta)\right) \delta(s(R; \alpha, \rho) - \rho - \alpha R) + \int_{Q_s(R; \alpha, \rho')} (\theta - k - g(\theta, R) \delta s(\theta; \alpha, \rho)) dF(\theta),
\end{aligned}$$

there is

$$\begin{aligned}
& \frac{\Phi[s](R; \alpha, \rho') - \Phi[s](R; \alpha, \rho)}{\rho' - \rho} \\
& \leq 1 - \delta + \delta \left(1 - \int_{Q_s(R; \alpha, \rho')} dF(\theta)\right) \frac{s(R; \alpha, \rho') - s(R; \alpha, \rho)}{\rho' - \rho} \\
& \quad + \delta \int_{Q_s(R; \alpha, \rho')} \left(1 - g(\theta, R) \frac{s(\theta; \alpha, \rho') - s(\theta; \alpha, \rho)}{\rho' - \rho}\right) dF(\theta) \\
& \leq 1 - \delta + \delta \left(1 - \int_{Q_s(R; \alpha, \rho')} dF(\theta)\right) + \delta \int_{Q_s(R; \alpha, \rho')} dF(\theta) = 1,
\end{aligned}$$

where the second weak inequality holds because $\frac{s(R; \alpha, \rho') - s(R; \alpha, \rho)}{\rho' - \rho} \geq 0$ and $\frac{s(\theta; \alpha, \rho') - s(\theta; \alpha, \rho)}{\rho' - \rho} \leq 1$ for all $\theta \in \Theta$. Hence, $\frac{\Phi[s](R; \alpha, \rho') - \Phi[s](R; \alpha, \rho)}{\rho' - \rho} \in (0, 1]$.

Therefore, as the unique fixed point of Φ , V must satisfy that $\frac{V(R; \alpha, \rho') - V(R; \alpha, \rho)}{\rho' - \rho} \in (0, 1]$ for

all $R \in \Theta$, $\alpha \in \mathbb{R}_+$, and $\rho, \rho' \in \mathbb{R}_+$ such that $\rho < \rho'$. The same arguments can prove that $\frac{V(R; \alpha', \rho) - V(R; \alpha, \rho)}{\alpha' R - \alpha R} \in (0, 1]$ holds for all $R \in \Theta$ such that $R > 0$, $\rho \in \mathbb{R}_+$, and $\alpha, \alpha' \in \mathbb{R}_+$ such that $\alpha < \alpha'$.

Consequently, if Condition **A1** holds for ρ , it holds for all $\rho' > \rho$; and if Condition **A1** holds for α , it holds for all $\alpha' > \alpha$.

Step 4

Now we prove the last result in the proposition. It is sufficient to prove that there exists some $\bar{\rho} \geq 0$ such that Condition **A1** holds for $R = 0$ as long as $\rho > \bar{\rho}$. Note that for all $\theta \in \Theta$,

$$\theta - k - g(\theta, 0)\delta V(\theta; \alpha, \rho) \leq \theta - k - g(\theta, 0)\delta(\rho + \alpha\theta).$$

Hence, the result can be proved by defining

$$\bar{\rho} := \max_{\theta \in \Theta} \left(\frac{\theta - k}{g(\theta, 0)\delta} - \alpha\theta \right) \geq 0.$$

For all $\rho > \bar{\rho}$,

$$\theta - k - g(\theta, 0)\delta V(\theta; \alpha, \rho) \leq \theta - k - g(\theta, 0)\delta(\rho + \alpha\theta) < 0.$$

□

A.3 Proof of Proposition 2

Consider any $R \in \Theta$. Recall that $\theta \in Q(R)$ if and only if

$$\theta - k - g(\theta, R)\delta V(\theta) \geq \delta(V(R) - \rho - \alpha R).$$

Consider any $\theta', \theta'' \in Q(R)$ such that $\theta' < \theta''$ and $\lambda \in (0, 1)$. Let $\theta := (1 - \lambda)\theta' + \lambda\theta''$. Note

that

$$\begin{aligned}
& (1 - \lambda)g(\theta', R)V(\theta') + \lambda g(\theta'', R)V(\theta'') \\
&= g(\theta, R) ((1 - \lambda)V(\theta') + \lambda V(\theta'')) \\
&+ (1 - \lambda) (g(\theta', R) - g(\theta, R)) V(\theta') + \lambda (g(\theta'', R) - g(\theta, R)) V(\theta'').
\end{aligned}$$

First, because V is weakly convex, $g(\theta, R) ((1 - \lambda)V(\theta') + \lambda V(\theta'')) \geq g(\theta, R)V(\theta)$. Second, because $g(\cdot, R)$ is weakly convex,

$$\frac{g(\theta'', R) - g(\theta, R)}{\theta'' - \theta} \geq \frac{g(\theta, R) - g(\theta', R)}{\theta - \theta'}.$$

Then, because $\frac{1-\lambda}{\lambda} = \frac{\theta''-\theta}{\theta-\theta'}$,

$$g(\theta'', R) - g(\theta, R) \geq \frac{1 - \lambda}{\lambda} (g(\theta, R) - g(\theta', R)).$$

Hence,

$$\begin{aligned}
& (1 - \lambda) (g(\theta', R) - g(\theta, R)) V(\theta') + \lambda (g(\theta'', R) - g(\theta, R)) V(\theta'') \\
& \geq (1 - \lambda) (g(\theta, R) - g(\theta', R)) (V(\theta'') - V(\theta')) > 0,
\end{aligned}$$

where the last strict inequality holds because $g(\cdot, R)$ and V are strictly increasing. Consequently,

$$(1 - \lambda)g(\theta', R)V(\theta') + \lambda g(\theta'', R)V(\theta'') > g(\theta, R)V(\theta),$$

so that

$$\begin{aligned}
& \theta - k - g(\theta, R)\delta V(\theta) \\
& > \theta - k - \delta((1 - \lambda)g(\theta', R)V(\theta') + \lambda g(\theta'', R)V(\theta'')) \\
& = (1 - \lambda)(\theta' - k - g(\theta', R)\delta V(\theta')) + \lambda(\theta'' - k - g(\theta'', R)\delta V(\theta'')) \\
& \geq \delta(V(R) - \rho - \alpha R),
\end{aligned}$$

where the last weak inequality holds because $\theta', \theta'' \in Q(R)$. It follows that $\theta \in Q(R)$. Therefore, $Q(R)$ must be a closed interval. $\square$

A.4 Proof of Proposition 3

Consider any $R, R' \in \Theta$ such that $R < R'$ and assume en route of a contradiction that $\bar{\theta}(R) > \bar{\theta}(R')$.

Note that

$$\begin{aligned}
& V(R) - \rho - \alpha R \\
& = \delta(V(R) - \rho - \alpha R) + \int_{Q(R)} (\theta - k - g(\theta, R)\delta V(\theta) - \delta(V(R) - \rho - \alpha R)) dF(\theta) \\
& \geq \delta(V(R) - \rho - \alpha R) + \int_{Q(R')} (\theta - k - g(\theta, R)\delta V(\theta) - \delta(V(R) - \rho - \alpha R)) dF(\theta) \\
& = \left(1 - \int_{Q(R')} dF(\theta)\right) \delta(V(R) - \rho - \alpha R) + \int_{Q(R')} (\theta - k - g(\theta, R)\delta V(\theta)) dF(\theta) \\
& \geq \frac{\int_{Q(R')} (\theta - k - g(\theta, R)\delta V(\theta)) dF(\theta)}{1 - \delta + \delta \int_{Q(R')} dF(\theta)}
\end{aligned}$$

and, similarly,

$$V(R') - \rho - \alpha R' = \frac{\int_{Q(R')} (\theta - k - g(\theta, R')\delta V(\theta)) dF(\theta)}{1 - \delta + \delta \int_{Q(R')} dF(\theta)}.$$

Hence,

$$\begin{aligned}
& V(R') - \rho - \alpha R' - (V(R) - \rho - \alpha R) \\
& \leq \frac{\delta \int_{Q(R')} (g(\theta, R) - g(\theta, R')) V(\theta) dF(\theta)}{1 - \delta + \delta \int_{Q(R')} dF(\theta)} \\
& \leq \frac{\delta \int_{Q(R')} dF(\theta)}{1 - \delta + \delta \int_{Q(R')} dF(\theta)} (g(\bar{\theta}(R), R) - g(\bar{\theta}(R), R')) V(\bar{\theta}(R)) \\
& < (g(\bar{\theta}(R), R) - g(\bar{\theta}(R), R')) V(\bar{\theta}(R)),
\end{aligned}$$

where the third strict inequality holds because $g(\cdot, R) - g(\cdot, R')$ is weakly increasing and V is strictly increasing and because $\theta \leq \bar{\theta}(R') < \bar{\theta}(R)$ holds for all $\theta \in Q(R')$. It follows that

$$\begin{aligned}
& \bar{\theta}(R) - k - g(\bar{\theta}(R), R') \delta V(\bar{\theta}(R)) - \delta (V(R') - \rho - \alpha R') \\
& > \bar{\theta}(R) - k - g(\bar{\theta}(R), R) \delta V(\bar{\theta}(R)) - \delta (V(R) - \rho - \alpha R) \geq 0,
\end{aligned}$$

where the last weak inequality holds because $\bar{\theta}(R) \in Q(R)$. Consequently,

$$\bar{\theta}(R) - k - g(\bar{\theta}(R), R') \delta V(\bar{\theta}(R)) > \delta (V(R') - \rho - \alpha R'),$$

so that $\bar{\theta}(R) \in Q(R')$, leading to a contradiction that $\bar{\theta}(R) \leq \bar{\theta}(R')$. □