The Political Economics of Non-democracy

Georgy Egorov and Konstantin Sonin

JUNE 2022
The Political Economics of Non-democracy*

Georgy Egorov  Konstantin Sonin
Northwestern University  University of Chicago

May 19, 2022

Abstract

We survey recent theoretical and empirical literature on political economics of non-democracies. Nondemocratic regimes face multiple challenges to their rule, both internal, such as palace coups or breakdown of their support coalition, and external, such as mass protests or revolutions. We analyze strategic decisions made by dictators from the standpoint of maximizing the chances of regime survival in the light of these challenges, and show how it explains multiple patterns, from hiring political loyalists to positions that require competence, to restricting media freedom at the cost of sacrificing bureaucratic efficiency, to running propaganda campaigns, organizing election fraud, purging opponents and associates, and repressing citizens.

Keywords: nondemocratic politics, authoritarianism, dictatorship, bureaucracy, electoral fraud, protests, revolutions, coup d’etat, media freedom, propaganda, censorship, repressions, institutionalized ruling party.

JEL Classification: P16, D74, D72, D82, C73, D83.

*The authors are grateful to Scott Gehlbach, Sergei Guriev, Ruixue Jia, Natalia Lamberova, Arturas Rozenas, Anton Sobolev, Michael Sonin, Dan Treisman, and Ekaterina Zhuravskaya for their helpful comments.
## Contents

1 Introduction ....................................... 1

2 Direct Information Control ...................... 6
   2.1 Information Manipulation ........................ 7
   2.2 The Costs of Information Control ................. 10

3 Control over Collective Action .................. 15
   3.1 Protests and Revolutions .......................... 16
   3.2 Authoritarian Elections and Electoral Fraud ....... 19

4 Nondemocratic Coalition Formation .............. 21
   4.1 Support Coalitions and Ruling Parties ............ 21
   4.2 Disenfranchisement and Repressions ............... 25
   4.3 Divided Autocratic Government .................. 30

5 Dynamics of Nondemocratic Power .............. 32
   5.1 The Challenge of Democratization ................. 33
   5.2 Path Dependence .................................. 37
   5.3 Succession ....................................... 39

6 The Modernization Debate and the Future of Authoritarianism 41

7 Conclusion ....................................... 45
“In war you can only be killed once, but in politics many times.”

Winston Churchill

“Mr. Churchill just cannot get it right.”

Joseph Stalin

1

1 Introduction

The 20th century ended with a collapse of most communist dictatorships amid predictions of “the end of history”, the final victory for liberal democracy as a political regime (Fukuyama, 1992). Yet, after thirty years of transition, many former socialist countries have transformed from aspiring democracies into imperfect democracies or full-blown dictatorships. Even countries that seemed to have established mature democratic institutions, such as Hungary and Poland, have been recently balancing on the brink of sliding back into authoritarian rule.

The phenomenon is hardly limited to post-socialist democracies: Turkey’s and Venezuela’s periods of competitive, democratic elections spanned decades, yet this did not prevent them from turning back. In China, the hopes that the pressure of the thriving economy and trade and the improving quality of life will lead to democratization have recently subsided as the government increased its control over media and eliminated restrictions on the paramount leader’s power. Around the world, the share of non-democracies has stabilized; their share of the world’s GDP has been increasing, largely due to the Chinese enormous contribution (see Figure 1).

Perhaps not surprisingly, nondemocratic politics has recently become an active area of research in economics, even if lagging behind studies of democratic processes and politics. In this essay, we attempt to synthesize the growing literature on political economics of non-democracies.

Traditionally, a non-democracy is a regime defined by negation. In a democracy, the country leadership is accountable to the population; that is, people have a regular opportunity to replace the government by voting. In a non-democracy, this mechanism is absent -

1 Churchill: in a conversation with Harold Begbie, as cited in Master Workers (Begbie, 1906); Stalin: an interview in Pravda, March 8, 1946.

even if modern dictatorships learned to imitate, superficially, democratic institutions (Guriev and Treisman, 2019). Authoritarian regimes are defined by means that their leaders employ to deny citizens the opportunity to replace them. Dictators use state funds to pay their supporters, purge their would-be opponents, falsify election results, repress the populace, censor information, etc. In response, citizens participate in protests and revolutions. Vying for power themselves or just fearing that the dictator would consider them a threat, dictator’s subordinates organize coups d’etat.

Even the most powerful of dictators have to weigh expected costs and benefits of their decisions, resolving numerous trade-offs. Executing predecessors or purging challengers raises stakes if the regime is eventually overthrown; censoring information and restricting media freedom result in suboptimal policy decisions and inefficient implementation; expropriating, in the absence of rule of law, businesses disincentivizes production; sharing rents with loyalists decrease the amount of resources that can be spent on police and secret service; hiring competent subordinates increases the risk of betrayal, and so on.

In Section 2, we start by focusing on the economics of informational control that autocrats exert over their subjects. In Subsection 2.1, we discuss research on censorship and propaganda, the mechanisms that dictators use to manipulate people’s beliefs about the quality of the government and its alignment with their interests. Although propaganda, and ideology more generally, has always been at the forefront of analysis of nondemocratic regimes starting with the classic work of Arendt (1951) and Friedrich and Brzezinski (1965), recent advances in econometric inference and field experiments have elevated the discussion
to a new level.

Yet propaganda is only one part of the informational trade-offs that every authoritarian leader has to resolve. A major prerequisite for efficient governance and, ultimately, for the dictator’s survival is his ability to gather and process information. To keep power, even the most sultanistic of dictators need to know the ever-changing needs of their subjects. This makes it even more surprising that, almost as a rule, dictators end up in an informational vacuum, surrounded by loyal yet incompetent subordinates. Subsection 2.2 deals with this puzzle.

Though all the trade-offs that we discussed above are, essentially, trade-offs between “informational openness” and “security,” they deal with different types of threats to the incumbent leader. The “loyalty vs. competence” model explains government structure as a survival strategy of a dictator who fights coups d’etat, an internal threat to the regime. The “media freedom” models deal with protests and revolutions, an external threat to the incumbent leader. (Svolik, 2012, classifies all problems that a dictator faces into problems of authoritarian power-sharing and problems of authoritarian control.)

The main external threat to any dictatorship is a “revolution” (Subsection 3.1). Such a revolution might be largely peaceful as the Velvet Revolutions in Hungary, Czechoslovakia, Poland in late 1980s, may involve limited violence as in Iran in 1979 or during the Arab Spring in 2011, or may escalate into a full–blown civil war as in France in 1793, Mexico in 1910, or in Russia in 1917. Even if a revolution does not actually happen, a dictatorship might end by a voluntary devolution of power by the dictator under the shadow of a revolution (Acemoglu and Robinson, 2005; Aidt and Franck, 2015). To prevent mass protests and revolutions,dictators try to “project strength” by organizing faux elections and reporting vote shares unheard of in democracies (Subsection 3.2). Importantly, protests and revolutions might matter “off the equilibrium path”, i.e., as threats that restrict the dictator’s options. For example, Fearon (2011) treats the threat of protests as the only means for the society to enforce regular elections, which are in turn critical for accountability and public goods provision.

Understanding the critical role of information in authoritarian politics was made possible by the development of modern contract theory and principal-agent models. Although informational control has arguably played as great a role for 20th-century caudillos as for modern dictators, they were not conceptualized as such. For early biographers of Adolf Hitler
or Juan Peron, attacks on critical journalists were a civil rights violation, not information management. For a 20th-century biographer of Stalin, propaganda was a mobilization tool of the regime, not a strategic instrument in power struggle. Histories and biographies that would analyze information management as a strategic tool that the political economists now consider one of the main mechanism of maintaining power are yet to be written. At the same time, modern economists approach repression, be it elite purges (Tyson, 2018), strategic mass killings (Esteban, Morelli and Rohner, 2015), deportations (Gregory, Schröder and Sonin, 2011), or ethnic cleansing (Rozenas, 2020) as an informational problem. We hope that our exposure of the informational approach will inform the broader historical discussion.

Information is no less relevant for regimes dealing with internal threats. Svolik (2009) estimates that between 1945 and 2002, out of 303, authoritarian rulers, 205 (more than two-thirds) were deposed by a coup (see updated data in Figure 2). The key element of survival in power is building a support coalition, which we discuss in Section 4. One form of organizing a machine of patronage and repression that keeps leaders in power is via an institutionalized ruling party such as the Communist Party of the Soviet Union, the Chinese Communist Party, or the Institutional Revolutionary Party of Mexico, each of which was the primary governing mechanism in their country for the large part of the 20th century. In Example 4.1 we sketch an informational model of a ruling party, the idea pioneered in Gehlbach and Keefer (2011).

Though information lies at the core of many important trade-offs that any dictator faces, there are many other trade-offs, no less critical to the functioning of an authoritarian regime. For example, while signaling and screening play an important role in both mass and elite
repression, the direct effect of repression – that the fear of being executed or sent to prison deters people from protesting – might be even more important. The three major groups of non-information models that feature in studies of non-democracies are as follows. First, there are models, which rely on commitment or absence thereof to explain the presence or absence of certain institutions or policy choices. In non-democratic politics, such models were used to explain why an extension of a franchise might help when a policy commitment is impossible. They are also indispensable in explaining why peaceful, non-revolutionary transitions from dictatorship to democracy have been historically so rare. Second, many models of bargaining and coalition formation do not rely on asymmetric information; the driving force in each of these models is the same with perfect information. Finally, there are important models of collective action, in which information plays a significant yet secondary role.

Subsection 4.2 deals with two related instruments of authoritarian control — repressions and disenfranchisement. Both serve the same goal, reducing the government accountability, and allowing the leader to pursue his preferred, unpopular (or, at least, lacking a majority support) policy. If the dictator cannot repress the opposition, he would have to either distribute resources or make policy concessions. In Example 4.3, we demonstrate the complementarity between information manipulation and repression, using models of propaganda (Example 2.1) and repression (Example 4.2) as building blocks. They might be further combined with models of revolution (Example 3.1) and internal coups (Example 2.2).

Newson and Trebbi (2018) argue that the “winner-take-all” metaphor, often applied to brutal leadership battles in authoritarian regimes, is misleading. In many nondemocratic regimes, the typical state of affairs is not an unconstrained rule of a single individual, but a balance between different factions that are unable to defeat each other. In Example 4.4, we show that such an equilibrium balance might include many warring factions. In Example 5.2, we show that not only those who are in power, but those who are just a part of a broader oligarchy might prefer keeping their rents intact rather than challenging the system.

One distinct feature of political dynamics in non-democracies is that new authoritarian leaders more often than not come to power as a result of overthrowing the incumbent leader. This by itself creates path-dependent dynamics, whereby events that fail to occur and unfulfilled threats play a role as important as those that actually materialized (North, 1981; Acemoglu, Egorov and Sonin, 2020). In Section 5 we discuss both the now-standard (Markov)
approach to model political dynamics (Acemoglu and Robinson, 2001, 2005) and alternative approaches that allow for path dependence. By its nature, Markov games cannot account for any kind of path dependence, a very important feature when we deal with life-long, rather than term-limited, tenure and possibilities of violent comebacks. This contributes to the “succession problem”, an acute issue for any personalized authoritarian regime. In Example 5.4, we combine the static model of divided autocratic government with a model of path-dependent dynamics to analyze how institutionalized ruling parties manage to guarantee regular replacement of top leadership.

survey of formal models in nondemocratic politics, Gehlbach, Sonin and Svolik (2016), emphasizes the critical role that institutions of nondemocratic government play. For a survey on long-term institutional change, which necessarily deals with long-term institutions of non-democracies, we refer to Acemoglu, Egorov and Sonin (2020). In this essay, we focus on institutions that can be changed by strategic individuals — either by insiders such as the incumbent leader or members of the ruling oligarchy (Subsection 5.1), or by outsiders such as protesters (Subsection 3.1).

The rest of the paper is organized as follows. Section 2 deals with direct informational control such as censorship and propaganda. Section 3 focuses on the role of collective action such as revolutions and protests. In Section 4, we consider the structure of support coalitions of authoritarian rule. Section 5 discusses the issues of nondemocratic dynamics. Section 6 surveys recent literature on causes and consequences of democratization, and Section 7 concludes.

2 Direct Information Control

Any authoritarian leader faces at least two major problems related to information flows. First, he has to design the optimal information structure that would determine how much and what kind of information should be available to citizens. Second, he needs to organize the information flow in the other direction so that he can gather information, be it by allowing a certain degree of media freedom or relying on secret surveys and opinion polls. We start with discussing the trade-offs an autocrat faces when manipulating people’s beliefs about his strength in Subsection 2.1. We will later use this propaganda model to augment other mechanisms of authoritarian control. In Subsection 2.2, we focus on mechanisms of
gathering information.

2.1 Information Manipulation

We start with the simplest possible model of information manipulation. An authoritarian leader chooses an information structure, under which information is released to citizens whose interests are not necessarily that same as that of the autocrat. We will see that manipulating information in an optimal, from the dictator’s standpoint, way requires providing truthful information with some positive probability. In other words, optimal propaganda is necessarily a mix of some truthful information with some information favorable to the dictator. An information manipulation model of this kind can be added, as a building block, to many models of nondemocratic politics that we discuss below.

Example 2.1 (A Model of Censorship) Consider an autocrat who is strong with probability $\theta$ and weak with the remaining probability. The people prefer to keep him in power if he is strong, and prefer to remove him if he is weak. (Until Section 3 we will assume that there is no collective action problem in a revolution.) There is an experiment that tests the leader’s strength, yet the leader is able to design, in advance, the signal that will be reporting the result of the experiment to the public. The optimal persuasion mechanism for the leader is to commit to the following signal: to report “keep” if the outcome of the experiment is “strong”, and to report “keep” with some probability $\beta > 0$ if the outcome is “weak”.

Suppose further that the people get the utility of 1 if they keep the strong leader or remove the weak leader, and the utility of 0 otherwise. If $\theta < \frac{1}{2}$, then without a signal, people would remove the leader. With a signal of slant $\beta$, people would keep the leader if the signal is “keep” whenever it is incentive-compatible to do so, $\frac{\theta}{\theta + (1 - \theta)\beta} \geq \frac{(1 - \theta)\beta}{\theta + (1 - \theta)\beta}$, or, equivalently, whenever $\beta \leq \frac{\theta}{1 - \theta}$. As the leader is interested in as high $\beta$ as possible provided that the incentive constraint is fulfilled, the optimal slant is given by

$$\beta^* = \frac{\theta}{1 - \theta}.$$ 

If the slant $\beta$ is too high, i.e., $\beta > \beta^*$, then following the signal is incentive-incompatible for citizens. They ignore the signal, and, if $\theta < \frac{1}{2}$, remove the leader. Conversely, if $\beta < \beta^*$, then the probability of survival is not maximized. With optimal slant, the leader survives with probability $\theta + (1 - \theta)\beta^* = 2\theta$, which means that the propaganda is working. If $\theta < \frac{1}{2}$, then the propaganda gives the leader a chance to survive.
Example 2.1 is a particular case of the general model of Bayesian persuasion (Bergemann and Morris, 2019). For authoritarian regimes, such a model was developed as a model of government censorship in Gehlbach and Sonin (2014). It is possible to base models of information manipulation in autocracies on other celebrated communication protocols such as cheap talk in Crawford and Sobel (1982), verifiable messaging in Milgrom (1981), and signaling in Spence (1973). In all these models, the leader holds certain information and sends a message to citizens, who act based on their updated priors. Relative to other communication protocols, a Bayesian persuasion mechanism such as the one used in Example 2.1, has two advantages. First, it is much simpler to analyze, while the basic intuition is still conveyed. Second, as demonstrated Kamenica and Gentzkow (2011), this mechanism results in larger effect of information manipulation than any communication protocol. In other words, this approach provides the upper limit on the effect of any propaganda, censorship, or any other form of information manipulation.\(^3\)

At the same time, using Bayesian persuasion to model information manipulation by the regime comes at cost. The mechanism requires to assume a significant commitment power on the sender’s part. In some circumstances, this assumption is too strong to be realistic: a dictator’s ability to commit to a certain signal structure, required by the mechanism, might be limited compared to that of a democratic leader. In other circumstances, it might be without loss of any realism. Manipulating information by restricting media freedom is an example. Dictators do not edit news in real time: instead, they appoint biased editors or establish institutions of censorship to generate the slanted signal.

In Shadmehr and Bernhardt (2015), the state does not censor moderately bad news to prevent citizens from making inferences from the absence of news that the news could have been far worse. Hollyer, Rosendorff and Vreeland (2018) model the autocrats’ preference for transparency in a two-player bargaining model: the autocrat might want to increase transparency to reduce the risk of an inside challenge (see also Kosterina, 2017). In Boleslavsky, Shadmehr and Sonin (2021), the mechanism is similar: the autocrat might prefer more transparency as it might help mobilize citizens to protect him (in turn reducing incentives for a potential inside challenge). Bueno de Mesquita and Smith (2017) make the broad point that

\(^3\text{In a theoretical model with rational actors, there is no difference between propaganda and censorship as any form of information manipulation might be modelled as a truncation of some true signal. Of course, assuming that dictator’s subjects are fully rational in that they are fully aware that they are being manipulated and know the mechanism of the manipulation – the same way citizens in Example 2.1 know the game and the dictator’s chosen strategy, }\beta\text{ – does not always result in a realistic model.}\)
an increase in the risk of removal of the incumbent dictator via one means results in increased risks of removal by other means. Guriev and Treisman (2020) suggest that violence is much less common in modern dictatorships than in those of the past, and analyze the impact of cooptation of elites and propaganda on a dictator’s popularity.

Relying on extensive archival data and modern methods of econometric inference, Adena et al. (2015) studied the impact of Nazi propaganda. In particular, they found a significant effect of radio propaganda after consolidation of the Hitler’s regime: it incited anti-Semitic acts and denunciations of Jews to authorities by ordinary citizens. Remarkably, the propaganda can backfire: in areas with positive attitudes towards Jews, the anti-Semitic propaganda strengthened the positive attitude. Such propaganda might be used for both targeting of potential opponents (for example, preceding the disenfranchisement of a certain group, see Subsection 4.2), or as an instrument of totalitarian mobilization (Arendt, 1951). Yanagizawa-Drott (2014) used the radio propagation in an uneven terrain to establish the causal impact of anti-Tutsi propaganda on violence in Uganda. (Zhuravskaya, Petrova and Enikolopov, 2020 survey empirical literature on political effects of internet and social media.)

Qin, Strömberg and Wu (2018), using the data on government-owned newspapers in China from 1981 to 2011, demonstrated that market competition has reduced the amount of political propaganda, confirming the theoretical prediction in Gehlbach and Sonin (2014). In earlier empirical analysis, King, Pan and Roberts (2013, 2014) unpacked the Chinese government’s strategic use of censorship. In particular, they found that the government did not censor government criticism, yet blocked information that would allow citizens to coordinate actions related to their grievances. Using daily news reports from Russia’s largest state-owned television network, Rozenas and Stukal (2019) found that instead of simply censoring economic facts, the media attributes positive news to competence of government officials, while blaming bad news on external factors.

In Example 2.1, the receivers of propaganda do not make a strategic choice whether or not to get the signal. If there is a cost of being a receiver — even as low as the opportunity cost of watching TV — this cost puts a limit on the amount of slant that the sender can use. Suppose that there is a cost of receiving information for citizens, \( c > 0 \). In this case, the optimal slant is given by

\[
\beta^* = \frac{\theta - c}{1 - \theta}.
\]

Naturally, the optimal slant decreases with the opportunity cost of getting information: if
citizens do not pay attention to what the government says, the latter survives for a smaller range of parameters. Knight and Tribin (2018) demonstrate, in the context of Venezuelan dictatorship, that availability of propaganda-free channels reduced the impact of state propaganda. Glaessel and Paula (2020) confirm that access to alternative sources of information limits propaganda with 1989 data on German Democratic Republic television: recipients disapproved of censorship if they were able to use Western television to detect misinformation.

Still, access to information is not sufficient unless citizens have incentives to consume information. Chen and Yang (2019) conducted a field experiment in China, giving citizens access to an uncensored internet. In a nutshell, the finding is that propaganda works: once citizens acquire new information, their knowledge, beliefs, attitudes, and intended behaviors change. At the same time, the demand for uncensored information is low; on their own, citizens do not look for additional information even if they have access.

2.2 The Costs of Information Control

In December 1989, the Romanian leader Nicolae Ceausescu called a mass meeting to demonstrate the strength of his support; the video recording reveals that he was genuinely surprised to see the anger and frustration of ordinary people. Apparently, he also overestimated the level of his support in the Romanian military and security services. In February 2011, Hosni Mubarak, in his 22nd year as the leader of Egypt, appeared unaware of his unpopularity — both among ordinary citizens and the elite — the day before he was ousted from power and put under arrest. In February 2022, in his 23rd year at the helm in Russia, Vladimir Putin ordered an invasion of Ukraine apparently relying on inadequate intelligence about Ukraine’s military and state capabilities and the readiness of his own forces. In each of these cases, the long-term dictators failed to be well-informed about political developments of the first-order importance to them.

There are two very distinct broad types of information gathering mechanisms that autocrats use. First, the government can use public sources of information, relying on competitive media, both domestic and foreign, and social networks. Second, the government can rely primarily on information gathered by various secret services or new-era methods such as digital surveillance (Xu, 2021). With the former, the downside for the autocrat is that free media provides the same information to citizens, which reduces the impact of propaganda.
and might help to facilitate anti-government protest. With the latter, the problem is that relying on information from secret services is, effectively, sharing power and rents with them.

As Example 2.1 demonstrates, censoring media increases the chances of an autocrat’s survival. However, there might be efficiency costs associated with restrictions on media freedom. In Egorov, Guriev and Sonin (2009), a resource-poor dictator allows media freedom as he is concerned with providing his bureaucrats with proper incentives. The paper confirms empirically the relationship between oil wealth and media freedom: in dictatorships, more oil means less media freedom, whereas in democracies the effect disappears. An oil-rich dictator can afford to stay out of touch with reality, censoring media, and yet stay in power; an oil-poor dictator does not have this luxury. Similar efficiency vs. propaganda trade-offs appear in the model of strategic protest restrictions (Lorentzen, 2013) and censorship (Lorentzen, 2014). As a result, media freedom varies a lot across nondemocratic regimes, from levels comparable to mature democracies to that of totalitarian regimes (see Figure 3).

The most straightforward way to present the dictator’s informational dilemma is to consider the choice and promotion of subordinates (Besley and Kudamatsu, 2009; Egorov and Sonin, 2011; Jia, Kudamatsu and Seim, 2015; Zakharov, 2016). A dictator needs a competent agent that controls the information flow. A more competent subordinate is more likely to differentiate a real threat to the dictator’s power such as a change in public mood, a need for an economic reform, an emergence of a new leader, or even a looming foreign invasion. At the same time, a more competent subordinate is more likely to side with the dictator’s enemies when the dictator is vulnerable, i.e., exactly when the subordinate’s loyalty is criti-
cal. An insecure or cautious dictator will therefore choose incompetent loyalists as ministers because he fears that a competent minister will betray him more easily than an incapable one, and this cripples dictator’s control even further.\(^4\)

Consider the following principal-agent relationship between an autocratic leader and a subordinate adopted from Egorov and Sonin (2011). While very simple, it highlights both the informational constraints on a dictator’s power and the critical difference between contracting with a third-party enforcement, a natural possibility in the presence of democratic institutions, and contracting in the absence of such enforcement.

**Example 2.2 (An Informational Model of Loyalty)** Consider a dictator who faces a potential challenge, an internal coup, or mass protests. With probability \(\theta\), the challenge is weak and will not succeed, so fighting this coup would be wasteful. With probability \(1 - \theta\), however, the challenge is strong, but it will still fail if the key lieutenant of the dictator remains loyal, doing what is optimal for the leader. Only if the challenge is strong and the lieutenant betrays the dictator, the dictator is removed from power. In contrast to the corporate world, the contract between the leader and the subordinate cannot be conditional on all possible outcomes, even if the outcomes are fully observable: the agent is not punished for betrayal if the dictator is overthrown.

The dictator needs this lieutenant to judge the seriousness of the threat; however, the agent himself is imperfectly informed. Denote the challenge’s type, which might be either strong or weak, by \(t \in S, W\) and the agent’s signal by \(s \in S, W\) and assume that an agent of type \(\lambda\) is characterized by

\[
\begin{align*}
P(s = S|t = S) &= \lambda, \\
P(s = W|t = W) &= 1.
\end{align*}
\]

Thus, \(\lambda\) characterizes the agent’s competence: an agent with a higher \(\lambda\) has a higher ability to distinguish a strong enemy from a weak one. For simplicity, we assumed that a weak signal is a definitive signal of a weak enemy, whereas a strong signal is possible when the enemy is either strong or weak. This means that the agent who received a weak signal will never betray.

\(^4\)The “loyalty vs. competence” dilemma is present beyond the political world. In the corporate world, it may have been be the fate of Jon Corzine, Goldman’s CEO, ousted in a “palace” coup by the firm’s board members in 1999, that made Richard Fuld, the CEO of Lehman Brothers, surround himself with incompetent loyalists in the wake of the financial crisis (Sorkin, 2010).
Figure 4: A model of endogenous loyalty: in equilibrium, a more competent lieutenant saves the dictator’s money, yet is more likely to betray.

The advantage of having an agent with high competence is that the dictator does not need to spend resources on countering weak threats. For instance, an agent with low $\lambda$ would always advise to spend resources on protective measures as he is unable to tell a serious threat from a weak one.

Now suppose that the agent received a strong signal. If he remains loyal, the dictator stays in power, and the agent collects some wage, which we normalize to zero. If he betrays for some bribe $b$, then the enemy wins with probability $P(t = S \mid s = S) = \frac{1-\theta}{1-\theta(1-\lambda)}$, in which case the agent collects rewards $b$ and the dictator wins with the complementary probability, in which case the agent is punished with utility $-c$. Thus, the agent betrays if the rewards exceed

$$b > \bar{b}(\lambda) = \frac{\theta}{1-\theta} c(1-\lambda) \quad (1)$$

As the threshold $\bar{b}(\lambda)$ is decreasing in $\lambda$, a more competent agent betrays for a lower reward. If the agent’s rewards conditional on betrayal are a random variable from the leader’s standpoint, a more competent agent is more likely to betray, so he is endogenously less loyal.

The model in Example 2.2 might be combined with the propaganda mechanism of Example 2.1. Indeed, if $\theta$ is everyone’s prior about the strength of the leader, then the leader could use the Bayesian persuasion mechanism of Example 2.1 to have, effectively, the probability of $2\theta$ of being strong. This will raise the bribe bar in equation (1); that is, propaganda will make betrayal by a subordinate less likely and thus increase the competence of the subordinate that the dictator hires.

Hiring a more competent agent saves resources on fighting weak or nonexistent enemies,
not to mention other helpful advice a competent person may give. Thus, the dictator faces a
trade-off between competence of his subordinates and their loyalty, and the way he resolves
this trade-off depends on his utility when he survives. Figure 4 illustrates the dictator’s
trade-off. Interestingly, the higher is the dictator’s loss when overthrown and the higher is the
probability that the enemy is strong, the higher are incentives to choose a less incompetent
lieutenant. Thus, there is an important endogenous constraint on the dictator’s power and
ability to choose and implement policies of his choice.

The informational model of loyalty and competence of Example 2.2 provides a par-
tial answer to the following puzzle. It is only natural that countries that have frequent
coups and recurrent revolutions exhibit poor economic performance. Yet why do countries
that have the same leaders for decades provide a disproportional number of growth failures
(Bueno de Mesquita et al., 2003; Jones and Olken, 2005)? Mancur Olson has offered a
powerful metaphor of “roving vs. stationary bandit” (Olson, 1991), further developed in the
concept of “encompassing interest” (McGuire and Olson, 1996). The power of this metaphor
is in its consistency with a basic premise of economics: incentives matter. The higher the
leader’s stake in the country’s welfare and the longer his horizon, the more interested he
should be in the country’s prosperity.

A major problem with the Olson metaphor is that it contradicts the accumulated em-
pirical evidence on modern dictatorship. The dictators who have had the most power –
Germany’s Hitler, Russia’s Stalin, China’s Mao – have led their countries to massive human-
itarian disasters and destruction of social welfare on a historic scale. The longer a dictator’s
tenure, the lower were economic and societal benefits of his rule (Bueno de Mesquita et al.,
2003). The loyalty vs. competence model demonstrates, theoretically, that economic stag-
nation might be a likely flip-side of the prolonged political stability under dictatorship, and
that loyalty at the expense of competence might be responsible for massive disasters such
as the Great Famine in China (Meng, Qian and Yared, 2015) or Holodomor in the Soviet
Union (Naumenko, 2021). Inefficient government is the flip side of the strong authoritarian
control.

Bai and Zhou (2019) confirmed the existence of the loyalty vs. competence trade-off:
during the Cultural Revolution in China (1966—1976), it was the most competent elite
members who were purged and replaced by mediocre substitutes. Not coincidentally, the
Mao Zedong’s years in power saw stagnation of the Chinese economy. Studying the post-
Mao period, Jia, Kudamatsu and Seim (2015) demonstrate empirically that the Chinese Communist Party has avoided the “loyalty vs. competence” trap through a system of job rotation and promotion within the party. In contrast, Shih, Adolph and Liu (2012) provide evidence that nepotism plays a major role in promotions within the highest tier of the CCP hierarchy.

Using a data set on over 1,400 of world leaders from 1848 to 2004, Besley and Reynal-Querol (2011) found that autocracies select leaders with 20% less education than democracies. At the same time, François, Panel and Weill (2020) established, employing a sample of 100 authoritarian regimes from 1973 to 2008, that more educated autocrats are better in, e.g., attracting foreign direct investment, a critical growth factor in the developing world. With autocrats having less education than democratically elected leaders, the performance, unsurprisingly, suffers. Jones and Olken (2005) demonstrated, using unexpected deaths of leaders in office as a source of exogenous variation in leadership, that leaders matter for growth, and negative effects of individual leaders are strongest for unconstrained autocrats.\(^5\)

### 3 Control over Collective Action

Since the classic work of Mancur Olson (Olson, 1965), the problem of collective action is well-understood to be an important mechanism that protects authoritarian leaders. Even if most citizens want the dictator removed, in the absence of free elections, they might find it difficult to organize and coordinate protests and revolutions. One critical element of collective action is information aggregation. For a citizen who wants the leader removed, it is crucial to know how many others have the same preferences, what information others possess, and what actions they are planning to undertake.

As there is an important informational component, all forms of information manipulation discussed in Section 2 are relevant here as well. However, the problem of collective action is difficult to resolve even if the dictator does not directly manipulate people’s beliefs. In protests and revolutions (Subsection 3.1), citizens need to know what other citizens plan to do: the payoff of an individual depends, critically, on actions of others. In authoritarian elections (Subsection 3.2), citizens use the official results to update their beliefs about other

---

\(^5\)In a recent working paper, Easterly and Pennings (2017) replicated, using an expanded data set, the Jones and Olken’s results with respect to very high or very low growth episodes; they also confirmed that autocracies produce higher growth volatility than democracies as suggested by Rodrik (2000) and Besley and Kudamatsu (2009). See also Section 6.
citizens’ attitudes. In these situations, information control over collective action is about manipulating the process of information exchange, rather than about providing manipulated information or censoring media. Furthermore, even in the absence of any uncertainty, the collective action problem is not necessarily resolved: when others stay at home, staying in home might be a dominant strategy even for a committed enemy of the regime.

3.1 Protests and Revolutions

The critical element of many modern models of a revolution is the mechanism that translates information dispersed among multiple agents into collective action. Individually, participating in a revolution is costly. However, when a mass of citizens participate, the costs are substantially lower and the chances of success are higher.

The early models of protests and revolutions assumed away the collective action problem. In Acemoglu and Robinson (2001, 2005), potential dissidents under the elite-controlled regime (“the poor”) are able to overcome, from time to time, the collective action problem and coordinate on protests. In other circumstances, “the rich” would overcome the collective action problem and launch a coup. As a next step, Shadmehr and Bernhardt (2011) model protests as a two-person coordination game and show that limiting public information available to citizens might increase the likelihood of protests as each individual citizen is forced to rely on others’ information to a larger extent.

Persson and Tabellini (2009) departed from the no-collective-action-problem assumption by using the global game approach to refine equilibria in a coordination model. In Bueno de Mesquita (2010), protests are modeled as a coordination game with multiple equilibria, and the vanguard of revolution moves first, thus altering the focal point for mass protesters. The vanguard does not have any informational advantage over the mass followers, and as such has no information revelation or signaling motive. Much like Vladimir Lenin’s “revolutionary vanguard,” the vanguard in Bueno de Mesquita (2010) or Apolte (2012) is essentially a device to choose the focal point, the critical element in any model of collective action. The following model of revolution is adopted from Bueno de Mesquita (2010) and Shadmehr (2018) to illustrate the main trade-off potential participants face.

Example 3.1 (A Model of Revolution) There is a continuum of citizens that decide

---

6Tullock (1971) argued that the coups d’etat are more common in non-democracies than revolutions because the collective action problem is much less acute in the case of a coup.
whether or not to participate in a revolution, \( r_i \in \{0, 1\} \), where \( r_i = 0 \) corresponds to non-participation. Suppose that the benefit from revolting is 1 if the revolution succeeds, 0 if the revolution is unsuccessful, and not participating in the revolution provides the payoff of \( a > 0 \).

The citizen \( i \)'s problem is \( \max_{r_i \in \{0, 1\}} pr_i + (1 - r_i) a \), where the probability that the revolution succeeds, \( p \), depends on the fraction of participants, \( r = \int_0^1 r_i di \), and the strength of the regime \( \theta \in (0, 1) \). Assume that

\[
p(r, \theta) = \begin{cases} 
1 & r \geq \theta, \\
0 & r < \theta.
\end{cases}
\]

In the case of complete information when the strength of the regime \( \theta \) is known, the outcome is deterministic. If \( \theta \leq 0 \), then there is a unique equilibrium, in which every citizen revolts, if \( \theta \geq 1 \), then there is a unique equilibrium, in which no citizen revolts, and if \( \theta \in (0, 1) \), then both equilibria are possible.

Now, suppose that each citizen \( i \) receives a private signal \( s_i = \theta + \varepsilon_i \) with \( \varepsilon_i \)'s independently drawn from the same atomless distribution \( F \) with full support on \( \mathbb{R} \). Now the decision whether or not to revolt \( r_i \) is a mapping \( r_i : \mathbb{R} \to \{0, 1\} \).

It can be shown that there is a unique equilibrium in symmetric monotone strategies

\[
r_i(s_i) = \begin{cases} 
1 & s_i \leq s^*, \\
0 & s_i > s^*,
\end{cases}
\]

where \( s^* = 1 - a + F^{-1}(1 - a) \). Naturally, the participation threshold is monotone increasing in the opportunity cost of the revolution: the higher the \( a \), the smaller is the share of agents who take part in the revolution, and the lower is the revolution’s probability. Furthermore, \( p(r, \theta) = 1 \) if and only if \( \theta \leq \theta^* = 1 - a \).

Example 3.1 makes use of the “global game” approach (Carlsson and Van Damme, 1993; Morris and Shin, 2001), a tool to select an equilibrium in a coordination game that would typically feature multiple equilibria. The idea of the refinement is to introduce some correlated asymmetric information and then select the risk-dominant equilibrium of the resulting game. Barbera and Jackson (2020), Casper and Tyson (2014), and Tyson and Smith (2018) use this approach to model revolutions when citizens have private information on either the regime’s strength or the common benefits from changing the regime. In Edmond (2013), the dictator has a costly technology to jam the signal available to citizens who might want to protest.
The search for a proper equilibrium refinement in models of collective action in politics arises from the fact that multiple equilibria is an immanent feature of models of collective action. The real-life counterpart of this phenomenon is that the same fundamental conditions are consistent with multiple drastically different equilibrium outcomes. In principle, nothing prevents the actors to switch to another equilibrium overnight. To explain why people might express positive or neutral attitude towards the regime, and then participate in mass protests or a revolution next day, Kuran (1989) hypothesized the existence of “preference falsification”. While Kuran’s work is an early and important precursor of the modern models of protests and revolutions such as one presented in Example 3.1, the agents in these models need not to falsify their preferences: the change their behavior in response to their private signals corresponds to switching to a different equilibrium. In another important early work, Lohmann (1993, 1994) models partially informed citizens who update their beliefs watching other citizens protesting, which results in “informational cascades” in unfolding revolutions.

The model of revolution in Example 3.1 can be combined with the model of propaganda in Example 2.1. Indeed, suppose that prior to the revolutionary stage, the dictator has a chance to design an information experiment. As Example 2.1 demonstrated, this will increase, in expectation, the people’s estimate of the dictator’s strength, $\theta$, which will result in a lower probability of a revolution at the revolution stage.

Boleslavsky, Shadmehr and Sonin (2021) combine Bayesian persuasion with global games to model the possibility of protests both against and for the incumbent leader. The possibility of protests for the leader following his dismissal in an internal coup alters the pre-coup calculus, discouraging the plotters. In contrast, when the incumbent is unpopular, a coup d’état might be organized by those who fear that they would go down with the leader dismissed by a popular uprising. Dorsch and Maarek (2018) investigated the connection between protests and coups using panel data on sub-Saharan Africa, and found that popular unrest causes an increase in the probability of a coup.

Instrumenting for the social media access with former classmates of a network founder, Enikolopov, Makarin and Petrova (2020) demonstrated that social media does help anti-government protest. Confirming the logic of the models that focus on coordination aspects of protests, they found that social media impact was through reducing the costs of protest coordination, rather than via spreading information critical of the government. Using high-resolution data on the expansion of cell phone coverage and protests in 2007–2014, Chris-
tensen and Garfias (2018) demonstrate that cell phone coverage increases the probability of anti-dictator protests by over half the mean. In a field experiment in Hong Kong, Cantoni et al. (2019) focused on incentives for people to participate in the protests as a function of information about others’ plans to participate. Their results suggest that agents consider their own participation as a substitute, rather than complement to others’ protest participation. This result partially aligns with Enikolopov, Makarin and Petrova (2020): information is relevant for coordination, rather than for changing attitudes to the regime.

Even a threat of a protest might affect the policy. Using an online field experiment, Chen, Pan and Xu (2016) unpack the mechanism through which Chinese citizens can effectively pressure their local governments via threats of collective action.

3.2 Authoritarian Elections and Electoral Fraud

Another form of authoritarian control is electoral fraud. In a democracy, electoral fraud might tip the outcome of a close election. If citizens do not know about the fraud, or consider its extent negligible, the outcome is legitimate. So, if the fraud goes undetected, it influences the outcome. In non-democracies, the electoral fraud is often so massive and violations of the due procedure so visible that it begs the question: why does it make sense to have fraudulent elections that citizens recognize as such? Why does it make sense to even organize such elections?

There is a substantial literature in political science that strives to explain elections held by autocrats (see Gandhi and Lust-Okar, 2009 and Miller, 2015, for recent surveys). Przeworski (2009) describes “plebiscitary elections,” which the regime uses to demonstrate that it can “force everyone to appear in a particular place on a particular day and perform the act of throwing a piece of paper into a designated box”. Collecting data from all over the world, Simpser (2013) suggests that electoral fraud can be used to demonstrate strength by showing the capacity to organize fraud. Egorov and Sonin (2021) model dictators “projecting strength” by organizing authoritarian elections to prevent a revolution. Though citizens realize that the reported numbers are manipulated by the incumbent, they still interpret his overwhelming “victory” as a sign of genuine support.

Authoritarian elections might play another role – they may define and enforce power-sharing or rent-sharing agreements among the elites (Boix and Svolik, 2013; Gandhi and Przeworski, 2007; Magaloni, 2010). In some circumstances, such agreements might help the
authoritarian elite to survive a transition to competitive elections (Martinez-Bravo, Mukherjee and Stegmann, 2017). Another explanation deals with gathering information and learning about local issues through elections. Martinez-Bravo et al. (2017) study the case of local (village-level) democracy in China to support this theory. Using a database on world-leaders’ exits from 1975 to 2004, Cox (2009) argues that elections help autocratic regimes to gather information that is needed to optimize succession. Miller (2015) finds that a negative shock to the election results prompts autocracies to spend more on education and social welfare.

In a model in which both fraud and protests are decisions made by unitary actors, Kuhn (2010) argues that protests are only possible if the election is won by the incumbent by a narrow margin and there is evidence of fraud. In Little, Tucker and LaGatta (2015), the results of an election convey the same information to the dictator and the citizens, and the main question is whether or not the dictator agrees to step down voluntarily after losing. In Little (2012), electoral fraud is modeled as a jamming of the public signal about the outcome of elections. In Gehlbach and Simpser (2015), electoral fraud, modeled as Bayesian persuasion, is the instrument the dictator uses to manipulate bureaucrats’ incentives. This makes efforts of different bureaucrats complementary — they participate in fraud if they expect others to participate — and results in a high variance of the winning (or losing) margins. Finally, Wig and Rod (2016) document the risks associated with authoritarian elections: if the election outcome points to a possible revolution, it might trigger a preemptive coup.

Empirical literature on electoral fraud is growing fast, though most of the studies are focused on imperfect democracies rather than authoritarian regimes. Enikolopov et al. (2013) used a field experiment to estimate the extent of electoral fraud in the Russian parliamentary elections in 2011. Rundlett and Svolik (2016) used data from the same election to illustrate the mechanism of beliefs formation when citizens observe electoral fraud. Cantu (2019) described mechanisms of electoral fraud in a one-party dictatorship employing data on the Mexican presidential elections in 1988.

Broader empirical literature addresses other elements of dictators’ tactics in authoritarian elections. Voigtländer and Voth (2014) demonstrated that focus on building infrastructure helped Hitler and the National Socialists to get a higher vote share in the 1933 elections and the 1934 plebiscite (organized with heavily limited competition, yet not fully uninformative). Importantly, the main channel of influence was not the limited direct benefits such
as employment in highway construction, but rather the propaganda effect.

4 Nondemocratic Coalition Formation

Assuming that there is always a single incumbent leader who makes the critical decision is analytically convenient, yet no dictator rules alone. Even the most personalistic of dictators need to resolve the problem of power-sharing to deal with elites and the problem of control over masses (Svolik, 2012). In this section, we discuss political coalition-formation in a nondemocratic context, which is radically different from the democratic one. One important difference is that a dictator can share the country’s resources with his supporters in ways, in which democratic leaders cannot (Subsection 4.1). In particular, many authoritarian regimes rely on institutionalized ruling parties as an instrument of maintaining support by sharing resources or imposing discipline on its members. Another difference is that autocratic leaders have a vast arsenal of repression and disenfranchisement at their disposal (Subsection 4.2). The final part of this section, Subsection 4.3, deals with the phenomenon of a “divided authoritarian government,” when the government structure, while not at all democratic, consists of competing antagonistic factions.

4.1 Support Coalitions and Ruling Parties

For an authoritarian leader, the most obvious way to build a coalition of support is to buy allegiance by sharing rents or making policy concessions. (We discuss repressions, both a complement and a substitute for sharing rents and policy concessions, in Subsection 4.2.) Gandhi and Przeworski (2006) theorize that when a dictator faces a threat of rebellion, he makes larger policy concessions, but also shares more rents, and confirm this prediction for all dictatorships that existed between 1946 and 1996. Powell (2013) analyzes a model of repeated bargaining between an autocrat and opposition, during which the autocrat also invests in increasing his powers. In equilibrium, the autocrat pays off the opposition in bad times, and tries to defeat it by force when there are plenty of resources.

Empirical research supports, broadly, the idea that autocrats buy their support. Desai, Olofsgård and Yousef (2009) use a panel of 80 non-democracies from 1975 to 1999 to demonstrate the existence of “authoritarian bargain”: autocrats pay off their citizens to surrender their political freedoms. Their bargaining model explains why non-democracies, in contrast
with democracies, feature a negative correlation between welfare spending and political liberalization. Caselli and Tesei (2016) find that in moderately entrenched autocracies, windfalls significantly exacerbate the autocratic nature of the political system. Brückner and Ciccone (2011) used within-country variation in rainfall to confirm that transitory negative shocks can open a window of opportunity for democratic improvement. Leon (2014) demonstrates empirically that military coups are more likely in countries that spend a relatively low share of GDP on military.\footnote{We refer to Geddes, Frantz and Wright (2014) for an overview of modern political science research on military regimes.} Using unique archival data on the allocation of cars for mid-level bureaucrats, Lazarev and Gregory (2003) analyze the microlevel of the dictator’s distribution of rents. Their evidence strongly suggest that allocation of cars, a prized commodity in the Soviet Union, was primarily a purchase of the bureaucrat’s political loyalty.

Empirical research has tried to measure the value of political connections to autocratic leaders. In a pioneer contribution, Fisman (2001) demonstrated that the share price of firms connected to Indonesian dictator Suharto suffered more than that of less-connected firms in response to news about serious health issues of the dictator. González and Prem (2020) use firm-level data from Chile to document resource misallocation in favor of politically connected firms during the transition from dictatorship to democracy: firms linked to the Pinochet regime (1973–1990) were relatively unproductive and benefited from resource misallocation under dictatorship, but then, after learning that the dictatorship was going to end, firms in the dictator’s network increased their productive capacity, experienced higher profits, and obtained more loans from the main state-owned bank.

Bueno de Mesquita et al. (2003) organize their analysis of nondemocratic regimes using the “selectorate theory” (see also an alternative model in Besley and Kudamatsu, 2009 and a discussion in Gehlbach, Sonin and Svolik, 2016). The selectorate are those who participate in choosing the winning coalition, a subset of the selectorate that guarantees power control. It might consist of a single person in a personalized dictatorship or of all citizens eligible to vote in a perfect democracy. Members of the winning coalition stick to the current leader as they are not sure they will be included in the winning coalition of a challenger. Thus, any challenger who wants to upend the status quo has to offer a premium over what the members of the winning coalition receive from the incumbent.

Myerson (2008) offers a game-theoretic model, in which supporting the leader is a focal
point for his (would-be) lieutenants. Unless the leader agrees, *ex ante*, to limit his authority, the lieutenants cannot be sure that he would remunerate them for the past support. As a result, in any renegotiation-proof equilibrium, the leader offers some constitutional restrictions on his power and the lieutenants’ support is conditional on the leader observing these restrictions.

One important example of a government structure in a non-democracy is an institutionalized ruling party, a quasi-state body that is structured like a political party in a democracy, yet does not actually compete in elections. Instead, its main function is to maintain control over the rest of the society.

There have been relatively few attempts to build a model of an institutionalized ruling party. Gehlbach and Keefer (2011) suggest an informational approach toward understanding the size and scope of operation of a ruling party. Citizens are divided into two groups, the insiders and the outsiders. The difference between the former and the latter is that when the party leadership expropriates property from an insider, other insiders are informed. Therefore, they have incentives to protect each other against expropriation. In contrast, outsiders have their property expropriated with other agents unaware, thus creating a premium for belonging to the elite.

In Example 4.1, we provide a simple model, where the leader determines the size of the ruling party. Those who are “in” receive information that helps them to make the correct investment decision; those who are “out” have to rely on publicly available information (prices). As Vladimir Lenin, the founder of the Communist Party of the Soviet Union, put it in 1921, the fourth year of the Russian Revolution: “We need full and truthful information. And the truth should not depend upon whom it has to serve. We can accept only the division between the unofficial information (for the Comintern Executive Committee only) and official information (for everybody).” Of course, an “investment” should be understood broadly — as any commitment of resources or efforts by a strategic individual in a situation when one choice is winning and the other losing. A choice of occupation or living location are standard economic examples of such investments.

**Example 4.1 (An Informational Model of Ruling Party)** There is a country with a unit continuum of citizens, and a leader who chooses the share of population that should be made members of the ruling party, $\gamma$.

There is a choice that every citizen has to make, to invest either in project $A$ or $B$. One
of the projects succeeds, while the other fails. The *ex ante* probability that project $A$ will succeed is $\theta$. Without loss of generality, we assume that $\theta < \frac{1}{2}$. Every citizen invests a unit of wealth into either project $A$ or project $B$. Citizens that invested in the losing project lose their investments; the ones that invested in the winning project receive their investment back plus their share of the losing project investment.

The difference between party members and the rest of the population is that the leader knows which project is winning and informs the party members. This is the simplest possible model with agents belonging to two different groups, informed and uninformed. Non-members do not know which project is winning, and cannot directly observe the party members’ choices. Still, there is a market for investments in $A$ and $B$, so non-members can base their decision on the market prices.

So, suppose that share $x$ of $1 - \gamma$ non-members invested in $A$. If the outcome is $A$, then each winner gets $\frac{1}{x(1-\gamma)+\gamma}$. If the outcome is $B$, each winner gets $\frac{1}{(1-x)(1-\gamma)+\gamma}$. Given the priors, the expected return of investing in $A$ is $\theta \frac{1}{x(1-\gamma)+\gamma}$, and the expected return of investing in $B$ is $(1 - \theta) \frac{1}{(1-x)(1-\gamma)+\gamma}$. In equilibrium, there should be no arbitrage, so the expected returns of the two projects should be equal:

$$\frac{\theta}{x (1-\gamma) + \gamma} = (1 - \theta) \frac{1}{(1-x)(1-\gamma) + \gamma}.$$

For any prior $\theta$ and choice of party structure $\gamma$, there exists a unique solution:

$$x^*(\theta, \gamma) = \frac{1}{1-\gamma} (\theta - (1-\theta) \gamma).$$

Critically, share $x^*$ should satisfy $0 \leq x^* \leq 1$, which implies that

$$\gamma \leq \min \left\{ \frac{\theta}{1-\theta}, \frac{1-\theta}{\theta} \right\} = \frac{\theta}{1-\theta},$$

the last equality following from the assumption that $\theta \leq \frac{1}{2}$. If the above condition is not fulfilled, then the no-arbitrage condition cannot hold, which means that everyone, including non-members, knows which project is going to win. In this case, there is no advantage of being in the party, so the choice of $\gamma$ above $\frac{\theta}{1-\theta}$ cannot be the optimal choice.

The total amount of resources that can be extracted from non-party members is $\gamma \frac{1}{\theta(1+\gamma)}$. If the leader were to maximize this amount, the optimal choice would be $\gamma^* = \frac{\theta}{1-\theta}$. At the same time, each party member gets $\frac{1}{\theta(1+\gamma)}$, and therefore prefers to be in as small a party as possible. Let us make additional assumptions that the leader appropriate share $\beta$ of
the surplus and that the cost of maintaining the party is $c$ per member. Then the leader’s maximization problem is
\[ \arg \max_{\gamma} \left\{ \frac{\gamma}{\theta (1 + \gamma)} - c\gamma, \right\} \]
and the optimal size of the party is
\[ \gamma^* = \sqrt{\frac{\beta}{\theta c}} - 1. \]
Naturally, the party size increases with the leader’s bargaining power, $\beta$, and decreases with the informational advantage that the membership gives, $\theta$, and the cost of maintaining the party, $c$.

Both models of Gehlbach and Keefer (2011) and that of Example 4.1 are particular cases of the Coasean approach to modeling a political party. Ronald Coase’s celebrated theory of the firm puts emphasis on the distinction between in-house and outsourced production, the difference being attributed to “transaction costs” (the agency problems). Similarly, an optimal organization of a political party would allocate some tasks in-house (what is done by party members) and some to outside producers (what is done by party supporters). The agency problems within the party would define the hierarchy of authority within the party the same way they define it in a firm. The promise of the Coasean approach is that it potentially encompasses parties in both democratic and nondemocratic environments.

The model of Example 4.1 focuses on a single aspect of party formation: the informational advantage that the insiders have over the outsiders. There is much more work to be done on analyzing the party structure and mechanisms that it uses to maintain discipline and cohesiveness inside and maintain control over the populace. In one of rare contributions, Francois, Trebbi and Xiao (2016) construct a hierarchical model of warring factions within the Communist Party of China. In Subsection 5.2, we will use Example 5.3 to discuss another important role that an institutionalized ruling party plays — that of a mechanism that ensures regular leadership replacement.

### 4.2 Disenfranchisement and Repressions

By definition, a nondemocratic government assumes that there are many citizens who are excluded from having say in political decisions. Dictators of the 20th century – Hitler, Stalin, and Mao, among others – and many less infamous ones purged the ranks of their political
supporters and repressed millions of those who did not support their policies. In the extreme case of totalitarian dictatorship in the Stalin’s Soviet Union or Mao Zedong’s China, the entire population was effectively disenfranchised, having no say in the policy choice or the choice of leaders. Many more dictatorships disenfranchised or purged whole social groups of population based, e.g., on wealth status, religious affiliation, or ethnicity.\footnote{In this survey, we do not discuss the breakdown of democracy — it is an issue in the realm of democratic politics — yet it is worth noting that disenfranchisement or outright elimination of certain groups was a first step to autocratic power for many elected leaders. Acemoglu, Egorov and Sonin (2015) is one model of a democracy breakdown, with the ultimate winner not necessarily the one who started disenfranchisement.}

The simplest argument why disenfranchisement benefits an autocratic leader comes from the standard spatial (Downsian) model of elections.\footnote{We spend little time discussing political positioning of leaders and challengers in non-democracies, though these certainly play a huge role. One reason for this is that basic cleavages, motivations, and tactics involved have certain resemblance — or at least important theoretical parallels — with those employed in democracies and are studied elsewhere (e.g., Persson and Tabellini, 2002; Duggan and Martinelli, 2017).} In a unique equilibrium of the basic model, candidates, regardless of their own ideological preferences, have to commit to the median voter position as a policy platform; otherwise, they have no chance to win. Unlike a democratic politician, the dictator might want to “trim the electorate,” thus making the move toward the median less necessary / less expensive, and to kill or exile his opponents. If those who are most opposed to the dictator’s preferred policy are repressed, it is easier to implement the desired policy. Consider the following example with the concept of protest on a one-dimensional policy space introduced by Dagaev, Lamberova and Sobolev (2019) and the effect of repression from Gregory, Schröder and Sonin (2011).

**Example 4.2 (A Model of Endogenous Protest and Repression)** There is a one-dimensional policy space, and citizens that have heterogeneous preferences over policy. Citizen $i$ has an ideal policy $i \in [0, 1]$, and the utility function is $u_i(x) = -|x - i|$. Suppose that there is an autocratic leader who is expected to pursue policy $x_D \in [0, 1]$, $d > \frac{1}{2}$. Citizen $i$ participates in protest $P$ if her utility from the policy $x_C = x_C(P)$ that results from the successful removal of the dictator exceeds her utility from the status quo $x_D$ plus the cost of participating in the protest $c : u_i(x_C(P)) - u_i(x_D) \geq c$. Define protest $P$ as the set of all citizens who are willing to participate:

$$P = \{i \in [0, 1]|u_i(x_C(P)) - u_i(x_D) \geq c\}.$$
Figure 5: Spatial model of protest and repressions. (a): A protest against the dictator who has policy platform $x_D$: agent $x_1$ participates in the protest, $x_2$ is indifferent, while $x_3$ supports the incumbent; $x_C$ is the expected position of the new leader. (b): A change of the dictator’s platform, from $x_D$ to $x'_D$, which makes $x'_2$ the new indifferent agent, and repressions (the blue area) reduce the size of the protest and thus the threat to the incumbent.

Figure 5 illustrates the basic logic of the model in Example 4.2. Suppose that the size of the protest in panel (a) (the red shaded area) is such that the dictator is overthrown. In panel (b), a group of citizens is repressed (the blue shaded area), which makes the potential protest smaller. In fact, there are two effects on the size of the protest, both negative. First, repression removes some potential protesters. Second, because of this, the expected outcome (the median) of a successful protest moves closer to the dictator’s policy $x_D$, which in turn reduces the number of potential protesters as the difference between the two outcomes shrinks. Finally, Figure 5 demonstrates that the dictator can complement repression with a policy adjustment from $x_D$ to $x'_D$, which makes participation in the protest less attractive. This is where an authoritarian leader is different from a democratic one: for the latter, a policy adjustment is the only way to increase support.

The notion of “protest” in Example 4.2 allows to sidestep the problem of collective action. If the agents were asked to make a strategic decision on participation in the protest that brings them, at a cost, a leader with a more appealing policy position, there will be another equilibrium, in which everyone would prefer to free-ride and not participate in the protest. Still, it is possible to combine the model of Example 4.2 with a model of revolution (Example 3.1) that features a unique equilibrium in symmetric monotonic strategies. This combined
model will retain the basic features of the spatial model, including the complementarity between repression and the dictator’s policy adjustment.

The first fully-fledged models of enfranchisement were introduced by Acemoglu and Robinson (2001, 2005), featuring essentially two-agent (“rich” and “poor”) dynamic games. The inability of those in power, i.e., the rich, to commit to a certain policy necessitated extension of the franchise to prevent a revolution. Several papers analyze strategic disenfranchisement by pushing voters out of the district (the “Curley effect,” a feature of democratic politics, in Glaeser and Shleifer, 2005) or repressing them (Gregory, Schröder and Sonin, 2011; Bove, Platteau and Sekeris, 2017). In the former case, the median voter of the remainder of the population is closer to the ideal point of the incumbent, which improves re-election prospects and increases utility derived from the policy. In the latter – diminishes the prospects of a revolution against the dictator. In Figure 5(b), the disenfranchisement of a group of voters with ideals points far from that of the incumbent leader makes the new median closer to the incumbent.

The early economic theories of nondemocratic government (Wintrobe, 1990, 1998) focused on a simple trade-off: the dictator was deciding how to optimally allocate resources between “repression” and “benefits.” Modern theories of repressions assume strategic targeting and selection. Myerson (2015) shows that the best incentives for the autocrat’s supporters are provided via randomized purges, appropriately combined with rewards for service. Tyson (2018) and Dragu and Przeworski (2019) combine an agency model of a dictatorship with targeted repressions. Esteban, Morelli and Rohner (2015) consider the determinants of the extreme case of citizens’ disenfranchisement by an authoritarian government, the “strategic mass killings”. Their empirical results confirm the role that the relative abundance of natural rents plays both at the country level and the ethnic group level. Montagnes and Wolton (2019) and Rozenas (2020) use communist purges in Stalin’s Russia and Mao’s China to demonstrate the effect of violence on performance and selection of subordinates.

Guriev and Treisman (2019) consider propaganda as a substitute for repression. Example 4.3 below, which combines repression and propaganda modeled in Example 2.1, demonstrates that repression and informational control could complement each other. The main mechanism is that repressing those who are most skeptical of the regime allows to increase the volume of propaganda for the others. Initially, propaganda slant was limited by the incentive constraints of the “skeptics”. When the skeptics are repressed, the incentive constraint is
relaxed, and the rest of the population receives more pro-regime information.

**Example 4.3 (A Model of Repression and Propaganda)** As in Example 2.1, suppose that there is a dictator who might be weak or strong, a continuum of citizens who think that the dictator is strong with probability \( \theta < \frac{1}{2} \), and who are willing to replace him if he is weak and keep in power if he is strong. The dictator simply wants to stay in power. As in Example 2.1, the dictator designs an information mechanism, which, in equilibrium, informs citizens that the dictator is strong if he is indeed strong, but misinforms them with probability \( \beta \) if he is weak. The optimal slant is \( \beta^* = \frac{\theta}{1-\theta} \), and the expected action is \( 2\theta \).

Now, let us introduce heterogeneous priors. Suppose that share \( \alpha \) of people received a signal that changed their prior to \( \theta' < \theta \), so they are more skeptical about the regime. What is then the optimal propaganda strategy for the leader?

If the skeptics were alone, the optimal propaganda is

\[
\beta'^* = \frac{\theta'}{1-\theta'} < \beta^*.
\]

If \( \beta > \beta'^* \), then “skeptics” do not follow the signal. If “skeptics” are ignored, the optimal slant is \( \beta^* \), and the total expected action is \( 2 (1 - \alpha) \theta \).

If \( \beta \leq \beta'^* \), then people of both groups follow the signal. Then the optimal slant is \( \beta'^* \), and the total expected action, from the sender’s standpoint, is

\[
\theta + (1 - \theta) \frac{\theta'}{1-\theta'}.
\]

The optimal choice depends on the share of skeptics: it is optimal to choose \( \beta^* \) (to ignore skeptics) if and only if

\[
(1 - 2\alpha) \frac{\theta}{1 - \theta} \geq \frac{\theta'}{1 - \theta'},
\]

and the total expected amount of support (action 1) is

\[
\max \left\{ 2 (1 - \alpha) \theta, \theta + (1 - \theta) \frac{\theta'}{1-\theta'} \right\}.
\]

Suppose that it is optimal not to ignore skeptics, so the optimal slant is \( \frac{\theta'}{1-\sigma} \) and the total expected action is \( \theta + (1 - \theta) \frac{\theta'}{1-\sigma} \).

Let \( \gamma \) be the required share of support for the leader to survive. We will focus on the situation when \( \theta + (1 - \theta) \frac{\theta'}{1-\sigma} < \gamma \), so the optimal propaganda is not sufficient.
Now, suppose that the leader is able to purge the share of $\lambda$ (and the purges are efficient enough to focus exclusively on the skeptics. Now, the total expected amount of support is

$$\max \left\{ 2 (1 - \alpha) \theta, (\alpha (1 - \lambda) + (1 - \alpha)) \left( \theta + (1 - \theta) \frac{\theta'}{1 - \theta'} \right) \right\}.$$

If $\lambda$ satisfies

$$2 (1 - \alpha) \theta \geq \gamma (\alpha (1 - \lambda) + (1 - \alpha)),$$

which is equivalent to

$$\lambda \geq \lambda(\alpha, \theta, \gamma) = \frac{1}{\alpha} - 2 \frac{\theta}{\gamma} \left( \frac{1}{\alpha} - 1 \right),$$

then repression makes propaganda sufficient for the leader to survive. Naturally, the critical threshold $\lambda(\alpha, \theta, \gamma)$ is increasing with $\alpha$ (a higher share of skeptics requires more repression) as long as $2 \theta > \gamma$ and $\gamma$ (a higher level of support that is needed for survival requires more repression), and is decreasing with $\theta$, the ex ante level of support for the leader. The implication is intuitive: repressing the regime skeptics allows propaganda to switch to a higher slant, guaranteeing more support for the leader from non-skeptics. Thus, repression and propaganda are complements: a regime that is more capable of repression is also able to use more propaganda.

### 4.3 Divided Autocratic Government

A classic method of nondemocratic control is “divide-and-rule,” whereby the leader maintains control by playing different factions of the society against each other. Structurally, this is similar to the pork-barrel model of democratic politics, in which the agenda-setter might build a coalition that comprises a majority in the parliament, imposing a tax on the rest. 

Roemer (1985) is an early model of coalition-building aimed at maintaining nondemocratic power: the challenger proposes income redistribution, the incumbent offers a list of penalties for joining the challenger’s coalition and then failing.

In many circumstances, a dictator might use the threat of violence by one group over another to extract resources from both. In Konrad and Skaperdas (2007) and Acemoglu, Robinson and Verdier (2004), the leader uses his power to redistribute from those who would depose him to those who are supportive of the status quo. Padró i Miquel (2007) explores the politics of fear as a tool of dividing and ruling: the fear that it is the other group is favored helps to control supporters. The model’s implications help to explain experiences of bad governance, ethnic bias, wasteful policies, and kleptocracy in post-colonial Africa.
Still, the “divide-and-rule” framework misses another, perhaps more salient, form of authoritarian government. Newson and Trebbi (2018), analyzing authoritarian elites in such diverse polities as Sub-Saharan Africa and China, conclude that “the prevailing view of winner-take-all contests can be clearly rejected.” In an authoritative study of the last decade of Stalin’s rule, historians conclude that the most appropriate model would be that of balance-of-power, in which no individual politician, even Stalin himself, is able to move without building a temporary coalition of support (Gorlizki and Khlevniuk, 2005).

The essential difference between democracies and dictatorships is that models of the latter cannot rely on commitment ensured by democratic institutions, established procedures, independent courts, etc. In the most extreme case, no commitment is possible at all. Example 4.4, which is based on Acemoglu, Egorov and Sonin (2008, 2009, 2012), analyzes an environment with total absence of commitment in nondemocratic politics. In this extreme example, a coalition that has sufficient number of (weighted) votes can vote to eliminate the rest; there might be more than one round of eliminations. The process stops when the ultimate winning coalition is reached. This is a drastic departure from the theory of democratic coalition formation, where coalitions can be perpetually formed and dissolved.

Example 4.4 (A Model of Coalition Formation in the Absence of Commitment)

There are $N$ agents, each of which has “power” $x_i$, $\sum_{i \in N} x_i = 1$. We say that configuration $(x_1, ..., x_N)$ is stable if there exists no subset $M \subset N$ such that (i) configuration $(x_1, ..., x_{i_M})$ is stable, and (ii) $\sum_{i \in M} x_i > \sum_{i \in N \setminus M} x_i$. By definition, all one-player power configurations are stable, and then all stable coalitions can be described by a recursive procedure.\(^{10}\) This notion of stability is consistent with the following dynamic process. A group of agents, each of which controls a certain amount of “power”, decides to eliminate (or simply strip of any power), by majority, some of them. The elimination continues until a stable state is reached. Each time an agent is eliminated, his “power” is distributed proportionally among the remaining agents. When a stable configuration is reached, the remaining agents split the pie of 1 in proportion to their power shares. Now, any two-player configuration is unstable. Then $(\frac{3}{12}, \frac{4}{12}, \frac{5}{12})$ is stable, $(\frac{3}{22}, \frac{4}{22}, \frac{5}{22}, \frac{10}{22})$ is unstable, and $(\frac{3}{42}, \frac{4}{42}, \frac{5}{42}, \frac{10}{42}, \frac{20}{42})$ is stable again.

Example 4.4 demonstrates that nondemocratic stability might be an equilibrium outcome

\(^{10}\)The above definition does not allow for “ties” when two coalitions have equal power. Still, it is straightforward to extend the definition to account for ties as well: $(x_1, ..., x_N)$ is stable if $(x_1 + \varepsilon_i, ..., x_N + \varepsilon_N)$ is stable for all, save for a subset of the Lebesgue measure of zero, small shocks $(\varepsilon_i, ..., \varepsilon_N)$. A simple exercise is to demonstrate that $(\frac{1}{N}, ..., \frac{1}{N})$ is stable if and only if $N = 2^K - 1$ for some positive integer $K$. 

31
even if there is no single dominating force, e.g., a leader or an agenda-setter who punishes a deviator. In this equilibrium, the ruling coalition is not necessarily minimal, and the agent with the highest amount of individual power is not necessarily included in the ultimate winning coalition. Acemoglu, Egorov and Sonin (2008) use the \( \left( \frac{3}{12}, \frac{4}{12}, \frac{5}{12}, \frac{10}{12}, \frac{20}{12} \right) \) example to tell the story of Joseph Stalin’s succession fight. The coalition is stable, yet the demise of the strongest members with power \( \frac{20}{12} \) (Stalin) makes the rest unstable. Then, three weakest members, \( \frac{3}{12}, \frac{4}{12}, \frac{5}{12} \) (Khrushchev, Malenkov, and Bulganin), eliminate the strongest of the remaining members, \( \frac{10}{12} \) (Beria), and form stable coalition \( \left( \frac{3}{12}, \frac{4}{12}, \frac{5}{12} \right) \). In Subsection 5.3, we make this model dynamic to account for the possibility of regular leadership turnover within the institutionalized ruling party.

While the above model is conceptual, it squares well with evidence on some modern autocracies. Francois, Rainer and Trebbi (2015) show that African ruling coalitions are relatively large and ethnic groups are proportionally represented. Newson and Trebbi (2018) found similar results analyzing authoritarian elites in Sub-Saharan Africa.

5 Dynamics of Nondemocratic Power

In a democracy, elections are the means of political change. Elections stimulate building new coalitions, changing political platforms, bringing in new leaders, and, eventually, implementing reforms. The political dynamics in authoritarian regimes are no less interesting, featuring both periods of extreme stability, with the same leader or party staying in power for decades, and moments of extreme turmoil, with drastic institutional changes happening in a matter of months, if not weeks.

In this section, we start with democratizations, peaceful transitions of power that involve a change of the political regime (Subsection 5.1). As we will see, a major constraint for such a transition stems from “slippery slope” considerations: a transfer of political power might be impossible if the preferences of the current power-holder and the ultimate power-holders after transition diverge too much. In Subsection 5.2, we discuss the phenomenon of “path dependence”. In Subsection 5.3 we focus on succession, another major challenge in any nondemocratic regime. In particular, we combine the model of path-dependent dynamics with a static model of ruling party to offer a model of a regular change of leadership in a nondemocratic context.
5.1 The Challenge of Democratization

There are many reasons why an authoritarian regime might turn to democratization. Democratizations themselves do produce growth benefits in both short- (Rodrik and Wacziarg, 2005) and long-run (Papaioannou and Siourounis, 2008). Even if he does not care about broad prosperity, a dictator might want to democratize to avoid facing mass protests or internal coups, which are costly even when unsuccessful (Balima, 2020). Partial democratization might be desirable if it provides a commitment device to protect property rights and thus improve agents’ incentives (Acemoglu, 2003; Acemoglu and Robinson, 2005; Myerson, 2008). We discuss empirical evidence on causes and consequences of democratization in Section 6.

Acemoglu, Egorov and Sonin (2012, 2015) analyze a general model of strategic enfranchisement and disenfranchisement with forward-looking agents. In the general model, the current decision-maker, e.g., a unitary dictator or the median voter of the enfranchised coalition, makes two decisions. First, she chooses this-period policy. Second, she determines who is going to be the decision-maker in the next period. As it turns out, there is a number of reasons to pass the political power to someone else, a representative of another social group or class. One reason is that someone else’s rule in the future might be preferred by the current decision maker.\footnote{Lizzeri and Persico (2004) apply this logic in a model of franchise extension to discuss the evolution of public spending without any threat from the disenfranchised.} This is a major general rationale for enfranchisement: e.g., with a king in power, the middle class fears expropriation, thus providing low efforts and generating small surplus. When enfranchised, the middle class has its property rights protected better, and exerts high efforts; for the king, the result is beneficial as he gets a smaller share of a larger pie.

Example 5.1 demonstrates the same dynamic force that creates an obstacle to partial democratization in the simplest possible setting. The current decision maker, who considers enfranchisement, might be willing to abide by policy decision made by the median voter of the extended franchise. However, she might dislike the consequences of the political choices of this median voter. As a result, the country is stuck with an inefficient autocratic rule despite the fact that the autocrat herself would prefer partial democratization.

**Example 5.1 (A Model of Inefficient Stability)** Consider an autocratic leader, \( A \), considering a reform that gives power to \( M \), the middle class, a more democratic arrangement.
There are three possible states of the world: the status quo, autocracy $a$, in which $A$ rules, limited franchise $f$, in which $M$ has greater security and is willing to invest; and democracy (full franchise) $d$, where $M$ becomes more influential and privileges of $A$ disappear. Stage payoffs satisfy

$$w_A(d) < w_A(a) < w_A(f),$$
$$w_M(a) < w_M(f) < w_M(d),$$

that is, $A$ prefers limited franchise to autocracy as greater investments by $M$ increase tax revenues, and $M$ prefers democracy to autocracy; $M$ is least well-off under autocracy. Both parties discount the stage payoffs at rate $\beta \in (0, 1)$. States $a$, $f$, and $d$ do not only determine payoffs, but also specify decision rules. In autocracy, $A$ decides which regime will prevail tomorrow; in both $f$ and $d$, $M$ decides the next period’s regime.

There are two possible long-term equilibria in this model. First, $d$ is such a state: $d$ is optimal for the decision-maker in $d$. In contrast, $f$ cannot be a long-term equilibrium, because, if the society ends up in $f$, the decision-maker in $f$ will move to their most-preferred state, $d$. Therefore, if, starting in state $a$, $A$ chooses the political reform towards $f$, this will ultimately lead to $d$ in the following period. Thus, the reform gives $A$ a discounted payoff of

$$U_A(\text{reform}) = w_A(f) + \frac{\beta}{1-\beta}w_A(d).$$

If $A$ decides to stay in $a$ forever, its payoff is $U_A(\text{no reform}) = \frac{1}{1-\beta}w_A(a)$. If $\beta$ is sufficiently small, then $U_A(\text{no reform}) < U_A(\text{reform})$, and the reform takes place. However, when players are sufficiently forward looking ($\beta$ is large), then $U_A(\text{no reform}) > U_A(\text{reform})$.

So, the initial state $a$ is made stable by the instability of the limited franchise state, $f$, which is preferred by those who are powerful in $a$. Note that both $A$ and $M$ would be strictly better off in $f$ than in $a$, so the stable state starting from $a$ is Pareto inefficient. It also illustrates that the reform is less likely when players are forward-looking (when $\beta$ is small, only $d$ is stable; when $\beta$ is large, both $a$ and $d$ are stable).

Example 5.1 explains why a rational dictator would not want to partially relinquish his power, fearing that this will lead to his ouster through the “slippery slope” (Schwarz and Sonin, 2008; Acemoglu, Egorov and Sonin, 2015). Then why do democratizations happen at all? Treisman (2020a), examining all episodes of democratization since 1800, offers an ingenious answer: there are indeed cases of deliberate democratization, but they happen
by mistake more than two thirds of the time. These mistakes might be agreeing to run in an election, losing a military conflict of choice, ignoring civil or military context, making a wrong succession decision, or simply “choosing the wrong combination of carrots and sticks against potential opposition”.

Myerson (2010) and Ellis and Fender (2010) identify conditions under which a peaceful transition to democracy is possible; Riedl et al. (2020) provides a recent review of the literature. Giavazzi and Tabellini (2005) demonstrate, using difference-in-difference estimates, that countries that first liberalized their economies and then become democracies do much better than countries that pursue the opposite sequence. Using Hofstede data on individualism/collectivism and a panel covering 1980–2010, Gorodnichenko and Roland (2021) provide evidence that countries with collectivist cultures are more likely to experience autocratic breakdowns that do not result in a transition to democracy.

In the “selectorate model” (Bueno de Mesquita et al., 2003), members of the winning coalition are unwilling to support a challenger as they are uncertain about being included in the winning coalition of the challenger. As the next Example 5.2 demonstrates, members of a ruling oligarchy might be unwilling to challenge the nondemocratic regime, because competing in elections would result in dissipation of their rents.

**Example 5.2 (A Model of Political Oligarchy)** Consider a polity with a limited number of politicians who can get power, \( N \), and a continuum \([0, 1]\) of citizens. Each politician in the office gets the rent \( R > 0 \) and ego boost \( B > 0 \).

Before the contest, each politician decides, whether to compete in open elections, or within the elite. Open elections happen if a single politician decides to enter the process. Within-party competition is a lottery with equal chances. In open elections, politicians spend money to buy votes. Politicians do not have money of their own, so they buy votes by promising patronage out of \( R \). For simplicity, they cannot renege, once in office, on the patronage promises. The process is a Bertrand competition between politicians: each voter supports the contender who promises her the most.

In an equilibrium that we are interested in, every contender promises the whole rent \( R \) to \( 1/2 \) voters. (The focal equilibrium will be in mixed strategies.) The winner’s payoff is then \( B \).
No politician competes in the open when \( \frac{1}{N}(R + B) > B \) or, equivalently,

\[
\frac{1}{N-1}R > B, \tag{2}
\]
i.e., when the (material) rent \( R \) is large, the ego boost \( B \) is small, and the number of contenders, \( N \), is limited.

The model of Example 5.2, though very simple, produces some natural implications. For example, the condition \( \frac{1}{N-1}R > B \) explains how a party dictatorship might switch to democracy: when the amount of rents \( R \) falls (e.g., oil rents when the oil price falls), the condition (2) is violated, and oligarchs are no longer interested in protecting their rents. The Mexican transition from a one-party dictatorship to a competitive democracy in the 1990s amid the collapsing oil prices, is a good example.

**Example 5.2, continued.** Observe that the \( N \) politicians in the model have strong preferences in limiting the number of entrants to their “club”: for each politician, the lower \( N \), the better. Yet it does not mean that the club will necessarily be small. Suppose that we augment Example 5.2 with the possibility to exclude, by majority vote, some of the oligarchs. The clubs of size 1, 2, or 3 oligarchs are stable, but ones of 4 or 5 are not. Indeed, as a club of size 3 is stable, any three members of a club of size 5 could exclude the other two. In this example, clubs of size \( 2^k \) and \( 2^k + 1 \), \( k \in \mathbb{N} \) are stable, while clubs of all other sizes are not. In particular, there might be large “elites”, which are, nevertheless, stable. Of course, the incentive compatibility condition (2) must be fulfilled, so there are only finitely many stable oligarchies.

In Example 5.2, there is neither upside, nor downside in the political competition between the oligarchs. There might be an upside, if political competition results, e.g., in better selection of leaders (Besley and Reynal-Querol, 2011). There might be a downside in resource loss when politicians kill each other or repress each other’s supporters. Even a larger problem is a loss of future benefits, when political uncertainty results in under-investment and sub-par efforts because several parties are stuck in a “prisoners’ dilemma”-type equilibrium. Greif (1998) analyzes *podesteria*, the institute of delegation of power by a group of competing oligarchs in medieval Venice, as an instrument to resolve this problem. Yet Guriev and Sonin (2009) argue that even if oligarchs are interested in appointing a leader who can
contain destructive rent-seeking, they would more often opt for a weak dictator fearing that a strong one will end up expropriating their property.

5.2 Path Dependence

The extent to which current developments are predicated on the history is a subject of ongoing debate. How does nondemocratic past of a country affect the democracy prospects in the future? Does a history of military coups make new coups easier? Do past protests predict unrest under a new regime? Douglass North has pioneered the idea of institutional path-dependence; we review the recent literature in Acemoglu, Egorov and Sonin (2020). In this essay, we focus on strategic decisions that individual actors take: in nondemocratic politics, history might play a critical role.

Since Acemoglu and Robinson (2001, 2005), Markov games have become a major tool in modeling political dynamics. A standard model has at least two states of the world that alternate following a Markov process: the probability with which a state occurs in the next period depends on the current state and actions that agents undertake in the current period, but not on what happened before the current period. For example, economic shocks alter the payoffs of economic agents, and they have stronger incentives to revolt in a crisis. If they revolt under a dictatorship, the next state is a democracy.

While analytically convenient, the Markovian property is a significant restriction. In particular, it does not allow to model any path dependence: to be Markov, a strategy cannot rely on the game history. At the same time, allowing players’ actions to depend on full histories leads to another modeling problem: the Folk Theorem guarantees that if players are sufficiently forward-looking, any static outcome might be realized as a subgame-perfect Nash equilibrium. (See Acemoglu, 2003, for a discussion in the context of nondemocratic political dynamics.)

Example 5.3 features a non-Markov dynamic model that allows to illustrate path-dependence but preserves the basic insights of Markov dynamics as well.

Example 5.3 (A Strategic Model of Path-Dependence) Consider an infinite sequence of potential leaders \( i \in \mathbb{N} \); a leader receives a positive payoff for each period in power. In each period \( t \), there is an incumbent \( i_t \in \mathbb{N} \), and there might be a challenger \( c_t \in \mathbb{N}\setminus\{i_t\} \); If there is a challenger, the battle for the throne in period \( t \) is modeled as a lottery in which the probability that the incumbent defeats the challenger is \( \theta \). After the lottery, the winner has
to decide whether to execute the loser or to spare him. If the loser is executed (he receives a negative payoff), then there is no challenger in period $t + 1$, and $c_{t+2} = i_t + 1$. If the loser is spared, then the loser becomes the new challenger in $t + 1$. So, the winner has incentives to kill the loser: this allows to survive the next period with probability 1.

We allow strategies to depend on the “reputation”, the number of killings that have been ordered by the loser in question during his tenure in power. The welfare-maximizing equilibrium is such that each winner spares the loser. On the equilibrium path, the first two leaders replace each other. However, there is another, “killing” equilibrium in the game, where the winner always (or above a certain threshold of loser’s killings) executes the loser. The mechanism at work is as follows. If dictator $X$ executed his predecessor, then dictator $Y$ if and when he eventually takes over power from $X$, will have higher incentives to kill $X$ since he cares about the reputation of $X$. And the reason why $Y$ will care about $X$’s reputation is that if $X$ is spared by $Y$, he might come back and decide, in turn, $Y$’s fate. And it is the fear that $X$, who has a reputation for cruelty, will execute $Y$ in the future that makes $Y$, the current decision maker, be more inclined to kill $X$ rather than spare him. One step back, this affects $X$’s motivations in dealing with his unsuccessful challenger: if $X$ executed his predecessor and is now deciding $Y$’s fate, the fact that the marginal impact on reputation of the second executions is lower than that of the first one makes him more likely to execute again.

This basic logic of Example 5.3 provides an immediate path-dependence: the current winner values his options differently depending on the type of his fallen enemy. If somebody takes over from a bloody dictator, he is more likely to become a bloody dictator himself than if he comes to power after a natural death of the previous ruler. Thus, the new ruler cannot switch to another equilibrium path, even though he knows that he would be better off in a “peaceful” equilibrium path. In many circumstances, the loser might be willing to commit not to be a contender in the future as such commitment would spare his life. Such commitment might be impossible for a dynastic ruler, whose rights to contend the throne are “divine” and thus virtually indispensable. Not surprisingly, countries with a limited and clear-cut set of contenders such as dynastic monarchies are more likely to witness executions of predecessors than, e.g., military dictatorships.

The “killing game” in Egorov and Sonin (2015) is a complete information game; it is straightforward to extend the logic to a model of commitment-type-based reputation in an imperfect information game (Kreps et al., 1982).
The model of Example 5.3 can be combined with the informational control models of Examples 2.1 and 2.2 as well as with regime change models (e.g., Example 3.1). A model of information control becomes a stage in the dynamic game, and the regime change model determines the probability of losing power. Each period, the incumbent might face a challenge with some odds to survive. For example, allowing more media freedom increases the probability to lose as the opponents have a better chance to organize a revolution.

Though equilibria of this game may lead to a variety of different paths, it is possible to single out three substantially different paths that correspond to different equilibria in Example 5.3. The first is the “Pareto optimal” path: if a winner with a reputation for benevolence spares a loser who also has a reputation for benevolence in equilibrium, these two actors remain the dictator and the challenger forever, swapping from time to time until one of them dies. On this equilibrium path, the “information parameter” is chosen to be high: bureaucrats are provided with good incentives, electoral fraud is limited, and propaganda is contained. In a “bloody path” equilibrium, every time a fight occurs, the loser is executed. Every incumbent limits media freedom, sacrificing efficiency, and the social welfare is minimized along this path. Finally, “the mixed path” allows for situations, in which a lucky string of outcomes switches the bad path to the Pareto optimal one (Egorov and Sonin, 2015).

With path-dependent dynamics, the initial conditions do matter. Overland, Simons and Spagat (2005) argue, theoretically, that dictators with a low level of starting capital tend to plunder the economy, while those with abundant capital invest in growth.

5.3 Succession

The “succession problem” is something that each autocrat, unlike a democratically elected leader, has to face (Herz, 1952; Konrad and Mui, 2017; Bueno de Mesquita and Smith, 2017). One critical difference is the loyalty problem: the power that an appointed successor will possess over the dictator’s fate makes his loyalty most important. At the same time, not having a successor has always been considered a destabilizing factor, threatening the incumbent regime. Studying 961 monarchs who ruled 42 European countries between 1000 and 1800, Kokkonen and Sundell (2014) argue that primogeniture has a significant advantage as a tool of building a strong state. Using data on coups against Danish monarchs between 935 and 1849, Kurrild-Klitgaard (2000) demonstrated that establishing “automatic hereditary
succession” reduces the number of coups.

Not surprisingly, few dictators have truly solved the succession problem. Most recently, the aging leaders of Egypt, Tunisia, Yemen, Libya, Turkmenistan, Uzbekistan and other countries for years failed to delegate any power to designated successors; ultimately, their succession plans failed. In Egypt and Libya, rumors of possible succession by a son had long circulated, yet no real power was ever transferred.

Besley and Reynal-Querol (2017) assemble a data set on leaders between 1874 and 2004 in which the leaders were classified as hereditary or non-hereditary based on their family history. One finding is that economic growth is higher in polities with hereditary leaders but only if executive constraints are weak.

One governing mechanism that does solve the succession problem is an institutionalized ruling party. The model of path-dependence allows us to extend the discussion of such a party that we started in Example 4.1. In a static setting, the difference between a party member and a non-member is access to information that allows the former to accumulate rents at the expense of the latter. In a dynamic environment, an important function of an authoritarian party is to provide a mechanism of leadership replacement. As we discussed in Subsection 2.2, long tenures of authoritarian leaders result in deteriorating quality of governance and poor economic performance. Authoritarian parties that were able to ensure regular rotation at the top such as Mexico’s PRI in 1930-1994 or the Chinese Communist Party since Mao Zedong’s death in 1976 were able to avoid this trap.

The following example of successful ruling party dynamics combines the features of Example 4.4, in which the autocratic government is divided yet short-term stable, and Example 5.3, in which the dynamics is history dependent. Naturally, the model of leadership replacement, which had no return option in both PRI and CCP cases, requires history to play a role in decisions about the future.

Example 5.4 (A Dynamic Model of In-party Leadership Replacement) We have a ruling party that consists of $N$ factions of possibly different sizes that play repeatedly the following game. Each period starts with each agent having the power of $x_u$. First, factions decide whether or not they want to eliminate some agents by “majority voting” by sums of their power. They eliminate until they reach a stable configuration as in Example 4.4. If a faction is eliminated, its payoff is 0 for the rest of the game. Second, factions decide whether or not they want to keep the current leader, who represents one of the factions. If the
majority wants to replace the leader, every faction pays a cost of some $\varepsilon$, $0 < \varepsilon < 1$. Third, proceeds of the current period are distributed proportionally to the powers of the factions. Finally, the leader’s faction adds $\Delta$ to its power; that is, the new powers are $x_{i,t+1} = \frac{x_i}{1+\Delta}$, if the leader does not belong to $i$, and $x_{k,t+1} = \frac{x_k}{1+\Delta}$ if the leader belongs to faction $k$.

Consider the following stable three-faction configuration, $(\frac{3}{12}, \frac{4}{12}, \frac{5}{12})$, let faction with power $\frac{5}{12}$ be in power, and let $\Delta = 1$. Then in period $t = 2$, the new power configuration, $(\frac{3}{13}, \frac{4}{13}, \frac{6}{13})$ is stable. However, the configuration $(\frac{3}{14}, \frac{4}{14}, \frac{7}{14})$, which will occur in $t = 3$, is not stable, as $\frac{7}{14}$ can defeat $(\frac{3}{14}, \frac{4}{14})$ (or vice versa — it does not matter here as the latter coalition is also unstable). Therefore, in period $t = 2$, the two smaller factions will vote to replace the leader.

In general, if the game starts with $(x_1, x_2, x_3)$ with $x_i < \sum_{j \in -i} x_j$, then the (welfare-maximizing) equilibrium strategy is to make changes a period before leader $k$ has $x_k = \sum_{j \in -k} x_j$ and to make $\min_i \{x_i\}$ the new leader. Why make the player with the minimum power the new leader? The rationale is that it minimizes the cost of replacements over the life-time.

The simple model of dynamics of inside-party succession in Example 5.4 ignores a number of important elements of the full model that was discussed before. It does not account for how the party controls the rest of the society or how it extracts rents from it. Nor does it account for hierarchical relationship within the party itself, or the quality of leaders it puts forward. Still, it elucidates the basic mechanism of how antagonistic factions can coordinate on rotating power, avoiding leadership stagnation. Other important elements can be added to these basic dynamics.

6 The Modernization Debate and the Future of Authoritarianism

While most of the literature on political economics of non-democracy that we surveyed in this paper deals with specific mechanisms of authoritarian power-sharing and control, the main empirical question about democracy is general: What is the relationship between democracy and development? There is no doubt that there is a high correlation between democracy and prosperity in historical data; save for some small exceptions, every rich country today is democratic. Yet is there a causal relationship between income and democratization?
Democracy and prosperity? As the methods of empirical inference improve, old arguments become obsolete and are replaced by new arguments, which rely on modern techniques.\footnote{\cite{Acemoglu2005} makes a strong argument for the use of proper instruments in establishing causal relationship between political institutions and economic development. \cite{Durlauf2020} surveys empirical literature on relationship between institutions, including political ones such as democracy, and growth, with particular emphasis on measuring institutions and hypothesizing the growth mechanism to estimate.}

In its classic form, the hypothesis that the economic development brings in democratization was formulated by Lipset (1959) who confirmed it by establishing the link between the level of per capita income and democracy in a global cross-section of nations. The 21st century discussion of the modernization theory started with Przeworski et al. (2000) affirming that democratization might be due to reasons unrelated to their level of economic development. Now, if economic prosperity precludes a country from slipping back to dictatorship, then the cross-country panel will show up a correlation between GDP per capita and democracy, even if the modernization theory was not true. In Przeworski et al. (2000), a country was classified authoritarian, as opposed to democratic, if one of the following conditions fails: (1) the chief executive is elected; (2) the legislature is elected; (3) there is more than one political party; and (4) an incumbent regime has lost power at least once. Using the Polity IV data, Epstein et al. (2006) categorized regimes as autocracies (Polity IV value 10 to 0), partial democracies (+1 to +7), or democracies (+8 to +10). Then, they established that higher GDP per capita significantly increased the likelihood of democratic regimes, both by enhancing the consolidation of existing democracies, thus confirming the findings of Przeworski et al. (2000) and by promoting transitions from authoritarian to democratic systems, which contrasts with it.

Murtin and Wacziarg (2014) use historical time series of income, education and democracy levels from 1870 to 2000 to show that primary schooling, and to a weaker extent per capita income levels, are strong determinants of the quality of political institutions. Seim and Parente (2013) argue theoretically, and calibrate using the Britain’s example, that elites democratize society only after the economy has accumulated enough wealth, and that the democratization date depends importantly on the history of rulers and distribution of land. Using data that covers two centuries, 1800-2000, Boix (2011) finds a positive and significant effect of income on the likelihood of democratic transitions and democratic consolidations. The results hold controlling for country and time fixed effects, and instrumenting for income. At the same time, the study confirms Acemoglu et al. (2008, 2009) findings that income has
no significant impact on democracy in the post-World War II period. Treisman (2020b) finds a strong and consistent relationship between higher income and both democratization and democratic survival in the medium term (10–20 years), but not necessarily in shorter time windows. The paper offers a conditional modernization theory, which can account for such lags: the effect of development on democracy is triggered by disruptive events such as economic crises, military defeats, or—most generally—leader change.

In addition to papers establishing, with the increasing degree of care about potential endogeneity in the recent years, that democracy follows prosperity, there are papers in which the relationship goes in the other direction. Using a panel of 100 countries covering 1960-1990, Barro (1996) found that controlling for the rule of law, the presence of free markets, small government consumption, high human capital, and the initial level of real GDP, the effect of democracy on growth is weakly negative. In Tavares and Wacziarg (2001), democracy fosters growth by improving the accumulation of human capital and, less robustly, by lowering income inequality. On the other hand, democracy hinders growth by reducing the rate of physical capital accumulation. Papaioannou and Van Zanden (2015) present evidence that dictator’s long years in office reduce economic growth, increase inflation, and harm the quality of institutions. The negative effect is particularly strong in young states and in Africa and the Near East.

In the context of transition from socialism to capitalism that started in late 1980s, some researchers expected democracy to have a negative effect on growth as the need to coordinate with many factions impeded reforms. However, using the data on 25 years of transition, Becker and Olofsgård (2018) found no discernible difference between democratic and non-democratic countries.

Drawing on an extensive global dataset, with some time series going back to the early 19th century, Knutsen (2015) reports robust evidence that democracy increases not only technology-induced growth but also net economic growth rates. Notably, the results hold when accounting for the endogeneity of democracy, country-fixed effects, and sample-selection bias. Using a sample of 23,000 initial public offerings around the world, Duong et al. (2021) find a positive relation between democracy and shareholder protection proxied by the difference between the initial price and the eventual market price of firm’s shares. Using data for political regimes, income and human capital for a sample of 141 countries over 1500–2000, Madsen, Raschky and Skali (2015) find democracy (instrumented by linguistic
distance-weighted foreign democracy) to be a significant determinant of income and growth; human capital, among other key variables, is controlled for.

*de Kadt and Wittels (2019)* apply the synthetic control method to a sample of 28 Sub-Saharan African states, 19 of which remained autocratic throughout in 1975–2008 and nine of which experienced democratic reforms in the 1980s and 1990s. The country-specific effects were sufficiently heterogeneous to give fodder for both sides of the discussion. In four cases, democratization had a negligible or negative influence on economic performance. In Zambia, the most extreme country of this type, annual GDP per capita was reduced by an average of 19 percent between 1992 and 2008. In five countries, there was a significant positive effect. Mali, the most extreme country in this category, had the greatest return from democratization, equivalent to a 42 percent average annual increase in per capita output. In general, democratic reforms were most likely to increase economic output in countries where economic liberalization was already underway.

*Besley and Kudamatsu (2009)* provide evidence that non-democracies exhibit more volatile growth than democracies.\(^\text{14}\) (See Figure 6 that replicates the figure from Besley and Kudamatsu, 2009, using a longer time period.) Using a dataset recording all buildings exceeding 150 meters, globally, Gjerløw and Knutsen (2019) find that autocracies systematically build more new skyscrapers than democracies. That is, the autocrats’ projects are more excessive and wasteful. Using regression and matching methods on data for a broad cross-section of countries, Justesen (2012) show that democracy on average increases access to treatment of HIV/AIDS. At the same time, Mulligan, Gil and Sala-i Martin (2004), using data on 142 countries over the years 1960–1990, concluded that democracy has not affected policies of redistribution or enhanced efficiency.

There is some evidence that autocracies discourage innovation along a variety of dimensions. Acemoglu and Robinson (2005) argued that the concerns about losing control slowed down the embrace industrialization and railroads by the absolutist Russian and Hapsburg Empires in the early 19\(^\text{th}\) century. Focusing on the experience of the Russian transition, Lamberova and Sonin (2018) show that an autocrat prefers to appoint an incompetent crony who would discourage market competition, fearing that a rules-based regime would result in new businesses supporting regime opponents. The hypothesis that non-democracies stifle

\(^{14}\text{Luo and Przeworski (2019) attribute autocratic “growth miracles” to the fact that to grow very fast, a country needs to have a low starting point, and poor countries are typically autocratic.}\)
innovation gets a new relevance now that the Chinese economy, after forty years of rapid
growth under an authoritarian regime, becomes close to the technological frontier. For
decades, the Chinese growth exploited the advantage of technological backwardness and
access to the enormous reserves of cheap labor. By the third decade of the 21\textsuperscript{th} century,
both resources have been depleted. Could China count on productivity-led growth without
dramatic expansion of democratic institutions?

An additional twist to the “China’s technological frontier without a democracy” challenge
comes from the fact that the recent political developments are the opposite of democrati-
ization. In late 20\textsuperscript{th} century, the decentralization drive of the federal center provided local
governments and business with pro-market incentives (Montinola, Qian and Weingast, 1995;
Roland, 2000). Recently, the Chinese government abolished term limits for the country’s
leader; the country seems to be shifting from a regime structured around an institutional-
ized ruling party to more personalistic, archaic regime. If there is an accumulated wisdom
in the literature on nondemocratic politics, it is that the emerging regime is inconsistent
with fast growth and development. Academically, what could be more exciting than having
a natural experiment in a country that comprises a quarter of the world population and
produces a fifth of the world GDP? In real life, we hope, the comparative lessons of the past
will be taken into account.

7 Conclusion

In this paper, we surveyed recent literature on political economics of non-democracies, fo-
cusing on information control mechanisms such as censorship and propaganda, authoritar-
ian elections, and other forms of informational control, and non-informational mechanisms
such as repressions, purges, and strategic disenfranchisement. All these mechanisms allow dictators to choose policies that would not be supported by citizens in open and competitive elections, be it a personal enrichment or an ideological quest. The collateral damage of mechanisms of authoritarian control is the erosion of incentives to innovate and grow, which, together with unpopular policies of dictators, result in protests, coups, and revolutions. The resulting dynamics are more stable at the initial stages of ascendant dictatorship, and more volatile and unpredictable than those of mature democracies when the dictators’ tenure in power spans decades. Recent advances in theoretical modeling and the technique of econometric inference have helped to clarify, refine, and understand many phenomena in nondemocratic politics. Yet there is still a lot to learn here.
References


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


