



A theory of autocratic transition by government leaders: prerequisites to self-enforcing democracy

Thomas Apolte¹

Received: 13 April 2021 / Accepted: 20 June 2022 / Published online: 14 July 2022
© The Author(s) 2022

Abstract

This paper aims at contributing to a better understanding of the conditions of self-enforcing democracy by analyzing the recent wave of autocratic transitions. Based on a game-theoretic framework, we work out the conditions under which governments may induce the diverse public authorities to coordinate on extra-constitutional activities, eventually transforming the politico-institutional setting into one of autocratic rule. We find three empirically testable characteristics that promote this coordination process, namely: populism and public support, corruption, and a lack in the separation of powers. By contrast, low degrees of corruption and strongly separated powers can be viewed as prerequisites to self-enforcing democracy.

Keywords Self-enforcing democracy · Political regimes · Autocratic transition

JEL Classification D02 · D72 · D74 · P48

1 Introduction

Over more than two decades following the collapse of the Soviet Union and the East-European communist world, a wave of democratization has shaped a remarkably optimistic view toward the future of democratic rule around the world. Indeed, almost 50 newly-established democracies have so far retained democratic rule.¹ The cases of failed transitions have long been outshined by the success stories, particularly those in Central and Eastern Europe. This notwithstanding, a non-negligible share of newly established democracies have not been successful. Some did not even survive their first years of existence. Uzbekistan and Kazakhstan went more or less

¹ Appendix A gives an overview of those countries listed in the database of the Polity Project that had democratized following 1989 and has still been democratic by 2018.

✉ Thomas Apolte
Apolte@wwu.de

¹ Center for Interdisciplinary Economics, University of Münster, Scharnhorststrasse 100, 48151 Münster, Germany

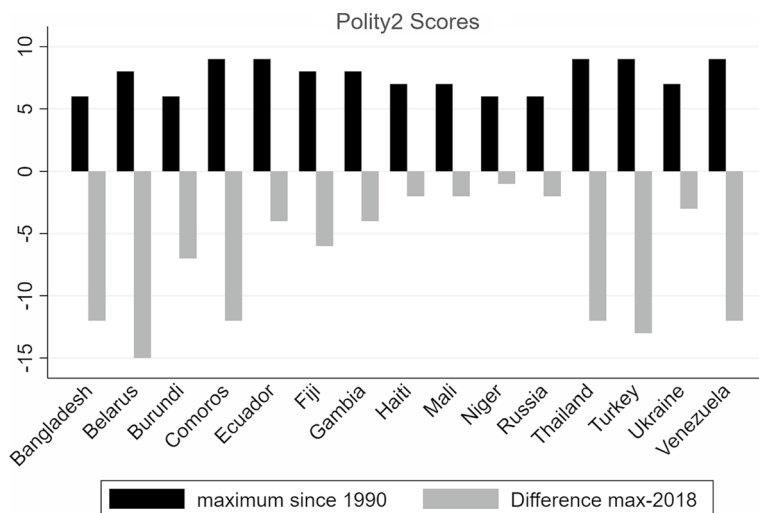


Fig. 1 Autocratic transition. Source: See Appendix A

straight into plain dictatorship following the Soviet collapse, and Belarus followed shortly after. Three Central Asian former Soviet republics, the countries of Caucasia as well as the Ukraine remained partly democratic at best. And on a worldwide basis, a number of countries that initially seemed to evolve into successful young democracies eventually backslid into autocratic structures. Perhaps the most prominent examples are Russia, Venezuela, Thailand and Turkey. Recently, even Hungary and Poland became candidates.² By now, these latter countries symbolize the end of the post-1989 democratization wave.

The Polity Project database labels a country as “democratic” if it scores 6 or higher in the polity2 data series which ranges from -10 to $+10$. Figure 1 gives an overview of countries that had reached at least a score of 6 in polity2 in some year following 1989 but later on fell below 6 and thus lost their status as “democratic”. The black bars give the highest score between 1990 and 2017, while the grey bars give the loss in polity2 scores between their respective maximum and the level in 2018. The largest loss was experienced by Belarus that reached a maximum of 8 in 1994 and ended at -7 . Thailand, Turkey, and Venezuela all scored 9 in the early 1990s and ended at -3 or -4 (Turkey) in 2018. All countries in Fig. 1, except Comoros, Burundi and Mali had reached their respective maximum in the decade between 1990 and 2000, and most of them in the early 1990s.

A high rate of failure of democratic constitutions is by no means only a recent phenomenon. To the contrary, the history of modern-age democracy is full of failing endeavors to establish democratic constitutions. Of ten newly established

² The latter is not yet visible in the Polity Project database, but in the database of Freedom House, Poland has been downgraded from 1 to 1.5 and Hungary even to 2.5 on a scale between 1 to 7, where 1 is full-fledged democracy. See details at: www.freedomhouse.org/report/freedom-world (December 16, 2020).

democracies following World War I,³ six failed until 1938, not to speak of what happened next.

Perhaps more surprising is the recent development of seemingly established, though still young, democracies in Europe, like those in Hungary and Poland. Their governments became notorious for openly challenging widely held convictions about indispensable institutional traits of democracy; remarkably, these governments have enjoyed extensive public support while doing so. This applies in an even more pronounced way to the government leaders of Turkey and Russia, who have already turned their countries into plain autocracies. These leaders started by attacking standard matters of course in established democracies, such as the freedom of the press and an independent judiciary. At the same time, they denounced such indispensable components of democracy as instruments secretly established by some internal or external conspirators. Further, they claim that these instruments are a threat to national values as well as to vital national interests and sovereignty. This, then, is the basis for how these governments justify restricting activities by national and international NGOs as well as by insubordinate media representatives. While some democracies have been able to sustain attacks like these, some others have already fallen victim to them. Russia and Turkey are clear-cut cases of the fallen democracies while the fate of the Hungarian and Polish democracies is still open to date.

A distinctive feature of these as well as of most other post-1989 autocratic transitions is that they emanated from deliberate activity of elected governments rather than from “classical” roots, such as military coups, foreign interventions or domestic revolts.

These observations represent the point of departure for this paper. It derives empirically testable conditions under which democracies sustain or fall victim to autocratic transition performed by elected governments. In doing so, the paper aims at contributing to the heretofore relatively narrow literature on self-enforcing democracy and self-enforcing constitutions as it has been developed by authors like Przeworski, Weingast, Fearon and others. Based on a simple game-theoretic analysis, we carve out the conditions under which a democratic constitution can be expected to sustain autocratic attacks by sitting government leaders. We start with democracy as an initial institutional equilibrium in order to establish the conditions under which a government leader may be willing and able to trigger a critical mass of further government actors to coordinate on an autocratic equilibrium. With this approach, we hope to shed light on the question of under what conditions governments will adhere to the constitutional rules by virtue of their individual incentives and, by contrast, under what conditions they may embark on a process toward autocratic transition.

Our central hypothesis rests on a strategic interaction between a government leader on the one hand and a number of further key government officials on the other, like heads of executive offices, judges, chief commanders of the police or the military, and so forth. The public may intervene into this interaction by challenging a government leader who seeks autocratic transition by vote or by public protests.

³ Either countries that were newly established as democracies or countries that democratized after 1914; see details in Table 2 in Appendix A. See also Ferguson (2001) and Weingast (2005).

The latter requires (1) the public's capability to solve its collective-action problem and (2) a critical mass of the government officials to defect to the public. We will capture these contingencies by the assumption of some—possibly terminal—election to be called for during the transition process which might, with non-zero probability, both stop the transition process and end the unfaithful government leader's tenure.

The rest of the paper is organized as follows: We give a brief overview of the relevant literature in Sect. 2. We then develop our model in two steps. In order to first focus on the problem of coordination on extra-constitutional activity, we present a restricted version of the model in Sect. 3. We then fully unfold our model in Sect. 4 by embedding the government leader's credibility problem regarding the distribution of autocratic rents. In Sect. 5, we discuss our results and derive some normative as well as some empirical implications. We conclude in Sect. 6.

2 Self-enforcing democracy

A large body of literature evolved over the last two decades mainly on three questions germane to the topic of this article. The first concerns the evolution of democratic rule in the Western world and beyond. Some of the most inspiring, but also disputed, contributions have been published by Acemoglu, Robinson and co-authors.⁴ Though less ambitious in the construction of venturous hypotheses, but still rather encompassing is the widely acknowledged book by North et al. (2009). A second question raised in the literature goes back to Lipset's (1959) modernization hypothesis. It investigates the direction of causality between high income levels on the one hand and democratic institutions on the other.⁵

In this paper, autocratic transition is pursued by the government leader. Another important channel of autocratic transition is through coups or, more generally, through disloyalty of the security elite around a government. This has been analyzed in important contributions by Bueno de Mesquita et al. (2006, chap. 7) as well as in the selectorate theory by de Mesquita et al. (2005); Bruce and Smith (2009). Svolik (2012), Boix and Svolik (2013) and Casper and Tyson (2014) analyze decision making and collective-action problems within the security elite and are, thus, close to our approach.

Still, our topic is not coups but autocratic transition pursued by the government leader. Hence, our model is situated in somewhat more sparse literature on a third question: How stable will democratic rule turn out to be under alternative institutional settings? Central to that literature is whether democratic constitutional rules are, or are not, self enforcing. The concept of self-enforcing democracy was introduced by Przeworski in a book on democratization in Eastern Europe and Latin America (see Przeworski 1991). According to this initial concept, a democratic

⁴ See Acemoglu and Robinson (2000), Acemoglu and Robinson (2001), Acemoglu et al. (2001) and Acemoglu and Robinson (2006). For a critical view, see Apolte (2012).

⁵ See, inter alia, Acemoglu (2006), Acemoglu and Robinson (2012), Acemoglu et al. (2019), Gründler and Krieger (2016), Gundlach and Paldam (2008) and Gundlach and Paldam (2009).

constitution is self-enforcing if an incumbent expects to be better off when stepping down following a lost election rather than sticking with his position by force and, hence, breaching the rules of the constitution. If the incumbent has a sufficiently good chance of being reelected in the due course of a subsequent election and/or will be entitled to a sufficiently generous pension, then this condition is likely to be satisfied. If not, the constitution may lose its self-enforcing character.

Weingast (1997, 2005) models self-enforcing democracy as a game, whereby the players of the game are a group of citizens that face a government leader who may transgress against the citizens' constitutional rights. In a one-shot game, the citizens are trapped in a prisoners' dilemma situation that precludes coordinated action against the government leader. This is different in a sequence of repeated games, since a sequence allows for revolting behavior against the incumbent in equilibrium strategies (see also Mittal and Weingast 2011).

Some 15 years after having introduced his initial concept of self-enforcing democracy, Przeworski came up with an entirely different approach (see Przeworski 2005; Benhabib and Przeworski 2006). The point of departure was the observation of a strong correlation between per-capita income and democratic resilience. At the heart of this approach is the assumption that per-capita income is subject to decreasing marginal utility, while utility of living in a democracy is assumed to be independent of income levels. Above some income level, then, marginal utility of the poor from income redistribution will drop below the utility of living under democracy. As a result, the poor will shy away from a revolt against the rich since there is a risk that such a revolt might terminate democracy. The rich, in turn, have an incentive to provide redistribution of income in order to reduce the incentive of the poor to revolt.

This more recent approach has not remained undisputed. Traversa (2015) argues that the results cannot be generalized because the model rests on too narrow a specification of the utility function of the poor. He demonstrates that the central finding disappears altogether when the utility function is slightly different.

Fearon (2011) introduced a model in which the general public implicitly threatens that it will revolt if the government transgresses against the constitutional rules of a democracy. Different from Weingast's approach, he assumes that the general public's strategic interaction with the government is a coordination game rather than a conflict game like the prisoners' dilemma. As a consequence, all the public needs to coordinate on a revolt equilibrium is a distinctive signal in combination with some "warm-glow benefits" of participation. In Fearon's view, a sufficiently distinctive signal might be electoral fraud.

Like Fearon, Przeworski's (2005) more recent approach reduces the issue to a coordination problem. It is indeed akin to approaches that explain institutional or constitutional stability by assuming that the population will coordinate on mutually consistent patterns of behavior (Calvert 1995; Hardin 2006; Przeworski 2006). Note, however, that coordination on a set of mutually benefiting constitutional rules is different from coordination on a rebellion equilibrium. This applies at least if there is a positive expected value of individual participation costs even when such a rebellion is successful. In a long-established democracy, participation costs might be negligible. Electoral fraud is then very likely to

spark an instant wave of protests that sweeps away the fraudulent government (Hyde and Marinov 2014). However, in the case of a democracy on the edge of autocratic transition, things are likely to be different. Electoral fraud may be perceived as a signal to rebel by only a part of the population, while another segment of the population might even appreciate when the security forces violently suppress protests. In such an environment, participation in rebellious activity implies a considerable risk of being injured or arrested; this turns a successful rebellion into a public good. Public resistance to an autocratic transition can then better be described as a prisoners' dilemma rather than as a pure coordination problem.

Weingast (1997) acknowledges the public-goods problem and solves it by referring to repeated games. However, almost any equilibrium can be generated in repeated games, and this is why we follow a different approach still. We model autocratic transition versus democratic stability within the structure of a game, the players of which are the government leader on the one hand and a number of key government officials on the other. Autocratic transition evolves within a certain time period. During this period, a government leader always faces the risk of being removed from office by intervention of the general public, either via vote or via mass protests. Note, however, that his efforts toward autocratic transition may indeed enjoy considerable support by at least part of the public.

We depart from most of the literature around autocratic decision making and self-enforcing democracy in the following way. Different from Acemoglu and Robinson (2006, chap. 7), we do not model the strategic interaction between a political elite on the one hand and the citizens on the other, although we consider the contingency of a public rebellion. However, we focus on the micro-structure within the political and administrative elite of a country. In this regard, our approach technically resembles the coup models in Svoboda (2012) and Boix and Svoboda (2013). Still, we do not analyze coups but autocratic transition pursued by the government. Hence, our model is most closely related to those of self-enforcing democracy, in particular to the seminal contribution by Przeworski (1991) as well as to the approach taken by Fearon (2011).

Once a democratic equilibrium is established, each member of the elite finds herself entrapped in a structure of mutually enforcing control mechanisms within the elite, which does not make it worthwhile for each of the individual elite members to depart from the established constitutional equilibrium strategy. As long as this equilibrium is unique, a government leader aiming at an autocratic transition cannot expect the other key government officials to follow him on a path of extra-constitutional action. There are, however, conditions under which such an equilibrium is not unique, which implies the risk of a switch away from an existing democratic equilibrium.

3 The transition game

Consider a country whose political system starts out as a democracy. All players are government actors. For simplicity, we assume that the government G as the top of the political hierarchy consists of only a single individual that is elected and might eventually be voted out of office. We refer to G as the *government leader*. Below the government leader, we have a number N of key government officials O_i with $i \in \{1, 2, \dots, N\}$. These individuals are appointed civil servants on lifetime positions that cannot be voted out of office. Upon having acted in an illegal way, however, a government official can still find himself dismissed. The number of all government actors taken together is $\bar{N} = N + 1$.

Among the government officials, we may consider the leading representatives of the different branches of government as well as chiefs of the police, the military, or some secret service. While G is the formally inaugurated head of the country's government, it is the players O_i that effectively run the country; each of these players decides to do so strictly on the basis of the existing legal system and, within that restriction, on the basis of the government leader's orders. However, each of these players may also decide to transgress against some of these rules. Moreover, each of them can, at least to a certain extent, refuse to follow G , again either within or beyond the limits of the constitutional rules.

We assume the government leader to signal his intentions by either abiding by the constitutional rules or by violating them. Upon having observed extra-constitutional conduct by the government leader, a share $h \in [0, 1]$ of the government officials effectively follows G in violating the constitution, while a share $1 - h$ does not. A necessary condition for a full autocratic transition is that $h \geq h^c \in (0, 1)$. Note that h^c is equivalent to the critical mass in the multi-equilibrium setting of the "Granovetter type" (Granovetter 1978; Marwell and Oliver 1993), in which the expectation of a certain minimum share of a population to participate motivates individual participation, thus turning the expectation into a self-fulfilling prophecy.

At the end of a completed autocratic transition process, the government leader G will no longer be challenged in his position by democratic means, though he will at least potentially be challenged by the government officials, specifically the military or the police, for example. However, we assume the transition process to take some time until its full completion. During this time of autocratic transition, (parts of) the public may still intervene, either by voting or by mass protest. With a non-zero probability, such interventions force all government actors back into a democratic equilibrium.

If a government official follows along with the extra-constitutional activities of a government leader, he may gain career benefits, but he may also suffer significant costs, such as losing his job or facing additional formal penalties. Furthermore, in this situation government officials face two major sources of uncertainty. First, to determine what decisions their colleagues are making and effectively executing, they must rely on sufficiently informative signals; second, they face a non-zero probability that the government leader will be deposed during the

autocratic transition period via public intervention, which may backfire on their decision to follow the government leader. Hence, a government leader who signals his intentions by purposefully violating the democratic rules of the game creates a delicate environment full of ambiguities for the government officials.⁶

To be precise, holding an office in a democratic constitutional environment provides a payoff of $U_G = U_i = 1$ to both the government leader and each of the government officials. The payoff can be broadly understood as an indicator of individual income or wealth. But it can also be understood as the level of influence one has on certain political programs or the like. On top of that, autocracy generates an additional pie R of payoffs for all government officials taken together, which we shall refer to as the autocratic rent. In the case where government actors share the autocratic rents, each of them would receive a share $\frac{R}{N}$ and hence an autocratic payoff of $U_{G,i} = s := \frac{R}{N} + 1$. However, we also allow for an unequal distribution of the autocratic rents. Such an unequal distribution of the autocratic rent may, for example, end up in autocratic payoffs $U_i = 1 \forall i$ and $U_G = R + 1$.

Each player may either respect the (initially) democratic constitution, in which case his choice is referred to as a_i or a_G , respectively (“abide”); or he might start mixing extra-constitutional measures into his activities, in which case his choice is referred to as b_i or b_G , respectively (“breach”). Players O_i can observe G 's choice prior to their own choice, but they cannot mutually observe their respective choices.

We assume the government officials to have less than perfect control over their respective subordinates. In particular, the execution of their decisions may effectively be undermined from within the bureaucracy. This assumption reflects two crucial aspects of reality: First, there is usually some managerial discretion left to subordinates in bureaucracies. Second, bureaucrats that are in a leading position but still subordinate to the respective government official may be reluctant to execute extra-constitutional decisions, and for two reasons: it violates social norms and it might bear unforeseeable and hence hardly calculable personal consequences. In brief, orders to conduct extra-constitutional activities cause potential loyalty conflicts which at least some of the bureaucrats might solve by effectively undermining the respective order's intention.

Borrowing from Bénabou and Tirole (2011), we may indeed conceive of these leading bureaucrats as the *alter ego* of the respective government official himself. Given, then, that the strictly rational ego finds transgressing against the constitution to be the optimal choice, the rule-following *alter ego* might simply shy away from actually taking the implied illegal action.

In order to capture these aspects, we introduce a ratio of tautness in the number of government officials that not only intentionally but effectively follow the government leader. To that end, we let \tilde{h} be the share of government officials that choose b_i in principle, and h the share of government officials that *effectively* choose b_i , where effectively means that the respective government official's decision will indeed be executed and become effective. We then assume the tautness ratio to be determined by $h/\tilde{h} = \epsilon^{\tilde{h}N}$ with $\epsilon \in (0, 1)$, so that the tautness ratio drops in the number $\tilde{h}N$ of government officials that want to follow the

⁶ For an early but still instructive non-formal analysis, see Tullock (1987).

government leader in principle. In a structure of $N = 1$, ϵ is the probability that the only existing government official effectively follows the government leader, conditional on his principle willingness to do so. For the more relevant cases of several government officials, then, ϵ will be adjusted by $\tilde{h}N > 1$. We henceforth refer to ϵ as the *reliability indicator* of a government official's intention to follow the government leader.

The adjustment of ϵ by $\tilde{h}N$ accounts for the rising difficulties to coordinate the actions of individuals away from an established equilibrium when the number of individuals rises. We will refer to this as the size effect. As an illustration of the size effect, assume that it takes only a few individuals to effectively coordinate in order to reach a critical mass. Then coordination is much easier and less risky compared to an environment where a critical mass requires coordination of a large number of individuals. As a result, knowing that successful coordination is less likely when the number of individuals rises, even individuals that would almost certainly join in if coordination were guaranteed become increasingly reluctant when the likelihood of successful coordination drops due to a rising number of individual actions that need to be coordinated.

We interpret N , \tilde{h}^c , and ϵ as indicators of the effective degree of the separation of powers in the pre-transition period. The formal definition of the separation of powers typically refers to the three branches of government, the representatives of which face checks and balances that make it difficult for them to cartelize their respective share in power. We supplement this formal definition by our effective definition in the following way: The higher the number N of government officials over all three government branches and the higher the share \tilde{h}^c necessary for them to coordinate on extra-constitutional action, the more difficult it becomes for the government officials to undermine the constitutional rules. Hence, the higher that both N and \tilde{h}^c are, the higher that the effective degree in the separation of powers is for any given formal degree.

Having chosen a_G leaves the government leader G with an expected payoff of $\omega \in (0, 1)$, which is his reelection probability under democratic rule. By contrast, upon having chosen b_G , he will be impeached unless at least a share h^c of the government officials O_i effectively follows him in choosing b_i . Should that happen, however, the general public may still intervene, either by vote or by mass protest (Fearon 2011). Technically, we model this as a call for some referendum in which the public may either endorse the autocratic transition process or bring it to a halt and end the unfaithful government leader's tenure. The latter happens with probability $1 - \sigma$. Hence, the government leader's chance of being reelected in its broadest sense is σ if he transgresses against the constitution, but it is ω if he abides by its rules.

The level of σ clearly depends on the severeness of collective-action problems. We hence define a severeness indicator $\kappa \in [0, 1]$ of collective action, with $\kappa = 0$ implying no collective-action problems whatsoever and $\kappa = 1$ indicating prohibitively severe collective-action problems. However, σ is also very likely to be correlated with the government leader's pre-transition popularity, as indicated by the reelection probability ω . This popularity to survive autocratic transition requires the population to tolerate transgressions of the democratic rules, which implies

a low valuation of democracy. We hence define an autocracy-tolerance indicator $\alpha \in [0, 1]$ with $\alpha = 0$ indicating zero tolerance and $\alpha = 1$ full tolerance. On the basis of these definitions, we posit a function $\sigma(\alpha, \kappa, \omega)$ of the following form:

$$\sigma(\alpha, \kappa, \omega) = (\alpha\omega)^{1-\kappa}. \tag{1}$$

Independently of ω and α , prohibitively severe collective-action problems ($\kappa = 1$) translate into a probability $\sigma = 1$ of the government leader to survive the referendum, since the population has no chance to counter anti-democratic activity in any way. With no collective action problems $\kappa = 0$, pre-transition popularity directly translates into the probability σ to the extent α to which the population tolerates autocratic rule.

As O_i are not elected politicians but rather appointed civil servants, they will not necessarily be fired but rather tried for extra-constitutional activity and then dismissed with probability $1 - \pi$.

In the “reelection” case, the autocratic-transition process is completed and the constitution loses all of its hitherto existing binding character, if any. No government activity will henceforth be restricted by constitutional rules of the game. In the stand-alone transition game, we assume G to be capable of credibly committing to an announcement according to which he equally shares the autocratic rent with those hN government officials that had effectively participated in the autocratic transition. That would imply payoffs $U_{G,i} = s_h := \frac{R}{hN+1} + 1 \geq s \forall i$. In the full game of Sect. 4, we will endogenize the distribution of the autocratic rent R .

Finally, should more than $(1 - h^c)N$ government officials O_i fail to effectively follow G in acting extra-constitutionally, the autocratic-transition attempt fails, G will be impeached and left with payoff zero. Those government officials that effectively followed G will again be tried for extra-constitutional activity and fired with probability $1 - \pi$. By contrast, those government officials that had effectively refrained from extra-constitutional activity will stay in office with payoff $U_i = 1$.

The timing of the transition game is as follows:

1. Government leader G chooses among actions $\{a_G, b_G\}$. Should G choose a_G , he will be reelected with probability ω , the government officials remain in their respective position, and the game ends with payoffs $U_G = \omega$ and $U_i = 1 \forall i$.
2. Upon having observed choice b_G by G , players O_i choose among actions $\{a_i, b_i\}$.
3. If $h < h^c$, the autocratic-transition attempt fails. G will be deposed and the government officials O_i that chose b_i will lose their positions with probability $1 - \pi$. The game ends with payoffs $U_G = 0, U_i = \pi$, and $U_{j \neq i} = 1$.
4. If $h \geq h^c$, a referendum is called for. G will win with probability σ and lose with $1 - \sigma$.
5. If G loses the referendum, the game ends with G being fired for sure, all O_i that chose b_i will be fired with probability $1 - \pi$, and all $O_{j \neq i}$ that chose a_j will stay in office, implying payoffs $U_G = 0, U_i = \pi \forall i$ and $U_j = 1 \forall j \neq i$.
6. If G wins the referendum, G equally shares the autocratic rent R with all O_i that chose b_i , implying $U_G, U_i = s_h$. By contrast, those $O_{j \neq i}$ that effectively chose a_j will have $U_j = 1$.

This gives us the following expected payoff functions for G as well as for each O_i :

$$U_G = \begin{cases} \omega & \text{for } a_G; \\ 0 & \text{for } b_G \wedge h < h^c; \\ \sigma s_h & \text{for } b_G \wedge h \geq h^c; \end{cases} \tag{2}$$

and:

$$U_i = \begin{cases} 1 & \text{for } a_i; \\ \pi & \text{for } b_G, b_i \wedge h < h^c; \\ \sigma s_h + (1 - \sigma)\pi & \text{for } b_G, b_i \wedge h \geq h^c. \end{cases} \tag{3}$$

Considering $\sigma = (\alpha\omega)^{1-\kappa}$ from Eq. 1, this leads to:

Proposition 1 *There are either one or two Nash-equilibria in pure strategies ζ_i, ζ_G with strategy profile:*

$$\zeta_i, \zeta_G = \begin{cases} b_i, b_G & \text{if } s > s^c := \max\left(\frac{1-(1-\epsilon^{\tilde{h}^c N})\pi}{e^{\tilde{h}^c N}(\alpha\omega)^{1-\kappa}}, \frac{\omega^\kappa}{e^{\tilde{h}^c N} \alpha^{1-\kappa}}\right); \\ a_i, a_G. & \end{cases}$$

Proof see Appendix B □

We refer to s^c as the critical value of the autocratic payoff s and hence distinguish two cases $s \leq s^c$ and $s > s^c$.

Case $s \leq s^c$: Self-enforcing Democracy

There is only one Nash-equilibrium $\{a_G, a_i\}$ in this case. All government actors abstain from extra-constitutional activity and accept the constitutional rules of the game. We refer to this case as *self-enforcing democracy*. For any given autocratic payoff s , the case of self-enforcing democracy is the more likely the higher is its critical value s^c . From Proposition 1, we get the following partial derivatives of s^c :

$$s^{c'}(\pi, \alpha, \kappa, \epsilon) < 0; s^{c'}(\tilde{h}^c, N) > 0. \tag{4}$$

The partial derivatives determine the influence of key institutional parameters on the likelihood of a democratic system to be self-enforcing. We will return to the results further below.

Case $s > s^c$: Tenuous Democracy There are two Nash-equilibria $\{a_i, a_G; b_i, b_G\}$ in this case. There is hence a problem of equilibrium selection (see Harsanyi and Selten 1988). The government leader and the government officials might still coordinate on an equilibrium (a_i, a_G) , but they might as well coordinate on an equilibrium (b_G, b_i) . Whether they would want to do the former or the latter cannot be determined unambiguously.⁷ Since the government actors might coordinate on

⁷ In order to technically generate a unique equilibrium, one could embed our model into the structure of a global game (Carlsson and Van Damme 1993; van Damme 2002; Morris and Shin 2010). However, this would require some demanding further assumptions, and it would not yield additional insights. More

extra-constitutional action (b_i, b_G) and hence initiate a process of autocratic transition, we refer to a democracy under the conditions of case $s > s^c$ as a *tenuous democracy*. The degree to which a tenuous democracy is indeed vulnerable to attempts of autocratic transition depends on the same institutional characteristics that determine whether a democracy is self enforcing or not. This can be demonstrated as follows.

It is straightforward that if (b_i, b_G) is a Nash-equilibrium, then it is also the payoff-dominant Nash-equilibrium. By contrast, the Nash-equilibrium (a_i, a_G) may be risk dominant. Consider, for a moment, a two-player game. Then, define $E(b_i|\delta)(E(a_i|\delta))$ as the expected payoff for any O_i from playing $b_i(a_i)$, conditional on a probability δ of player $O_{i \neq i}$ to play b_j . Then we have $E(b_i|\delta) = \delta \left[e^{\tilde{h}^c N} (\alpha \omega)^{1-\kappa} s + (1 - e^{\tilde{h}^c N}) \pi \right] + (1 - \delta) \pi$ and $E(a_i|\delta) = 1$. In this case, the critical probability δ_r that equalizes $E(b_i|\delta)$ and $E(a_i|\delta)$ and hence makes O_i indifferent between a_i and b_i is the risk factor of equilibrium (b_i, b_G) . Applied to our N -player setting and a critical mass $\tilde{h}^c N$ of players, the risk factor is $\delta_r^{\tilde{h}^c N}$. In equilibrium (b_i, b_G) , we hence find:

$$\delta_r = \left[\frac{1 - \pi}{e^{\tilde{h}^c N} ((\alpha \omega)^{1-\kappa} s - \pi)} \right]^{\frac{1}{\tilde{h}^c N}} \tag{5}$$

By Harsanyi and Selten’s (1988, 216) definition, a Nash-equilibrium (a_i, a_G) is risk dominant if $\delta_r^{\tilde{h}^c N} > 0.5$. Note, however, that (b_i, b_G) remains the payoff-dominant equilibrium independently of whether (a_i, a_G) is risk dominant, and that rational actors do not necessarily coordinate on the risk-dominant equilibrium. By contrast, we know from experiments and other empirical observations that real-world actors are the more likely to choose the risk-dominant strategy the higher is the risk factor $\delta_r^{\tilde{h}^c N}$.⁸ We are safe to assume, hence, that a tenuous democracy is the more vulnerable to attempts of autocratic transition the lower the risk factor. As a result, the partial derivatives of δ_r in Eq. 5 give us some institutional determinants of the degree of vulnerability:

$$\delta_r'(\pi, s, \alpha, \kappa, \omega, \epsilon) < 0; \delta_r'(\tilde{h}^c, N) > 0. \tag{6}$$

The implications of these partial derivatives are fully in line with those that determine the case of self-enforcing democracy, as given by Condition 4. For both cases, taken together, we hence find: A democracy is less likely to be self-enforcing and, if not, the government actors are more likely to coordinate on an extra-constitutional action, when:

Footnote 7 (Continued)

importantly, however, is that it would fade out the ambiguous structure of the tenuous democracy, which is an important implication of our model.

⁸ See, inter alia, Cooper et al. (1992), Straub (1995), Clark et al. (2001) and Février and Linnemer (2006, 170).

- there is a high probability π that a government official will remain in his position after having acted extra-constitutionally, even in the case of a failed transition attempt;
- there are high autocratic rents $s = \frac{R}{N} + 1$;
- there is a high level α of autocratic toleration and, hence, a low valuation of democratic rule by the population;
- there is a high level of severeness κ of collective-action problems in organizing protest;
- the government officials have few problems in enforcing intentions to break the constitutional rules, indicated by a high reliability indicator ϵ ;
- there is a low number N of government officials and/or a high critical share \tilde{h}^c of government officials necessary for autocratic transition.

The final two criteria are of particular importance since the probability of government officials to coordinate on extra-constitutional action quickly drops in $\tilde{h}^c N$ for any given ϵ , and it rises in ϵ for any given $\tilde{h}^c N$. Recall that N indicates the number of further government officials, while \tilde{h}^c gives the critical share of government officials necessary for autocratic transition. If \tilde{h}^c is low, then regardless of how many government officials there are, only a few of them—say a single army commander plus a single commander of the police—would really count. Additionally, a high reliability indicator ϵ implies that a government official has few problems in enforcing his intention to break the democratic rules within his bureaucracy.

These are the reasons why \tilde{h}^c , N , and ϵ , taken together, determine the effective degree of the separation of powers for any given formal degree in the pre-transition period. Democracies that exhibit little effective separation of powers are less likely to be self-enforcing, and, in tenuous democracies, little effective separation of powers makes coordination on the payoff-dominant equilibrium b_i, b_G and, hence, autocratic transition more likely.

An interesting further aspect is the ambiguous effect of the government leader's reelection probability ω . On the one hand, it raises the threshold value s^c via the right-hand side of the s^c -equation in Proposition 1. On the other hand, it lowers s^c via the left-hand side of that equation, which carries over to Eq. 6. The reason is as follows: On the one hand, a high pre-transition reelection probability may undermine a government leader's motivation to pursue autocratic transition since he has a good chance to be reelected by due process anyway. On the other hand, if a high pre-transition popularity of the government leader survives an attempt of autocratic transition, this makes it less likely for him to be deposed by popular protest or similar action. This makes autocratic transition more likely.

In the limiting case of $\kappa = 1$ and, hence, prohibitively severe collective-action problems, the reelection probability cancels out of the right-hand side in Proposition 1. As a result, the effect of the government leader's popularity becomes unambiguous, since a high popularity only has the effect of enhancing the government leader's chances of surviving the referendum.

Note, however, that our results are preliminary since we still operate within a restricted model setting. In this setting, the government leader is assumed to be

capable of committing to his promise of equally sharing the autocratic rent with the government officials. In the following section, we relax this assumption.

4 The post-transition game

In this section, we endogenize the potential struggle for autocratic rents among the government officials under the conditions of an abolished constitution. To that end, we relax the assumption that the government leader equally shares the autocratic rents with the government officials. In order to account for the underlying commitment problem, we nest a “post-transition game” into our game and expand it to our full game.

As demonstrated in the previous section, the government leader’s payoff will be s_h if he shares the autocratic-government rents with the government officials. But, if he fails to share and leaves the government officials with their initial payoff of $U_i = 1$, the government leader’s payoff will instead be $S := R + 1 > s_h$. Hence, an announcement that the government leader intends to share the autocratic rents with the government officials lacks credibility. Indeed, in the environment of a freshly abolished constitution, it is not immediately clear who would have the power to allocate government rents.

One might naturally think that this power would go to the government leader himself, since, after all, he is now the dictator. But the power of a dictator does not fall like manna from heaven. Rather, it rests with his capacity to play off the government officials against each other, particularly those who administer government decisions (de Mesquita et al. 2005; Svobik 2012; Tullock 1987). It is therefore crucial to add the post-transition game to the transition game, since rational actors will build expectations about their respective future positions in the post-transition struggle for autocratic rents when making their choices in the transition period.

In order to consider the post-transition interactions, we add the following steps to our game. Upon a successful transition, the government leader may decide to equally share the autocratic rents with the further government officials, thus raising the payoff of each of his followers along with his own payoff to a level of $s_h > 1$. Alternatively, he may as well renege on his promise, leaving all further government officials with their initial payoff $U_i = 1$ and raising his own payoff to $S = R + 1 > s_h$.

If the government leader reneges on his promise, each of those hN government officials that effectively chose b_i in the transition game have two further options, namely option c_i to continue supporting the government leader and option d_i to withdraw support. We define $\gamma \in [0, 1]$ as the share of hN government officials that effectively opt for d_i to withdraw support of G in the post-transition game. As in the transition game, we assume limited control of the government officials over their respective subordinates or, for that matter, over their *alter ego*. We define $\tilde{\gamma}$ as the share of government officials that intend to opt for d_i . We then define ρ as the probability of a government official to effectively withdraw support, conditional on the given intention to do so. Then, $\gamma = \rho^{\tilde{\gamma}hN} \tilde{\gamma}$ is the share of government officials that effectively rebel against the government leader, conditional on having effectively

chosen b_i in the transition game. We refer to ρ as the reliability indicator of the post-transition period.

Define now $g := \gamma h$ as the share of government officials that first effectively follow the government leader by choosing extra-constitutional action in the transition game *and* then effectively withdraw their support for the government leader in the post-transition game, if necessary. In the post-transition game, the government leader will be deposed and substituted by some individual from outside if he reneges on his sharing obligation *and* if γ reaches at least some critical value $\gamma^c \in (0, 1)$. In the latter case, his successor will be forced to share all rents R equally with the gN government officials.

Taken together, N , $\tilde{\gamma}^c$, and ρ represent the effective degree of the separation of powers in the post-transition period in much the same way as N , \tilde{h}^c , and ϵ represent the effective degree of the separation of powers in the pre-transition period.

By contrast, should the government officials fail to mobilize a share $\gamma \geq \gamma^c$ against the cheating government leader, then the leader will stay in office along with those government officials that continued to support him. The disobedient government officials, in turn, will be removed from office and replaced by newly appointed individuals from outside. The disobedient officials will be left with payoff zero, while G continues to claim all autocratic rents and hence a payoff $S = R + 1$, so that all remaining government officials, along with the newly appointed ones, will be left with payoff $U_i = 1$.

Complementing steps 1 to 5 of the transition game in Sect. 3, the time line of the post-transition game continues as follows:

6. Upon having won the referendum, G chooses among actions $\{c_G, d_G\}$; c_G (“comply”) implies meeting his promise to equally share the autocratic rent with the hN officials that followed him in the transition. By contrast, d_G (“defect”) implies seizing all autocratic rents R so as to reap a payoff $S = R + 1$ and leave the government officials with their initial payoff $U_i = 1$.
7. In the case of G choosing c_G , the game ends with payoffs $U_G = U_i = s_h$ and $U_j = 1$ with $j \neq i$ indicating those government officials O_j that had failed to follow G in the transition game.
8. Upon having observed G to choose d_G , by contrast, each government official chooses among actions c_i and d_i , where c_i implies continuing support of G while d_i implies draw back or withdraw support of G .
9. In the case of $\gamma \geq \gamma^c$, G will be deposed and substituted by an outsider G' . The game ends with payoff $U_G = 0$ as well as $U_{G'}, U_i = s_g := \frac{R}{h\gamma N + 1} + 1 \geq s$.
10. In the case of $\gamma < \gamma^c$, all rebelling government officials will be replaced by outsiders. The game ends with payoff $U_G = S, U_i = 0$ for all rebelling officials and $U_j = 1$ for all obedient government officials and all newly appointed government officials.

Considering the full game, players i and G decide over a strategy ζ_i, ζ_G consisting of a sequence $k \in \{1, 2\}$ of actions $q_{i,G}^k \in \{a_i, a_G, b_i, b_G, c_i, c_G, d_i, d_G\}$. For the government leader, the action profile is $q_G \in \{a_G, bc_G, bd_G\}$, while it

is $q_i \in \{a_i, bc_i, bd_i\}$ for each further government official O_i . The payoffs are as follows:

$$U_G = \begin{cases} \omega & \text{for } a_G; \\ 0 & \text{for } b_G \wedge h < h^c; \\ \sigma s_g & \text{for } bc_G \wedge h \geq h^c; \\ \sigma S & \text{for } bd_G \wedge h \geq h^c \wedge g < g_c; \\ 0 & \text{for } bd_G \wedge h \geq h^c \wedge g \geq g_c. \end{cases} \tag{7}$$

$$U_i = \begin{cases} 1 & \text{for } a_i; \\ \pi & \text{for } b_i \wedge h < h^c; \\ \sigma s_g + (1 - \sigma)\pi & \text{for } b_i \wedge h \geq h^c \wedge bc_G; \\ \sigma + (1 - \sigma)\pi & \text{for } bc_i \wedge h \geq h^c \wedge bd_G; \\ \sigma s_g + (1 - \sigma)\pi & \text{for } bd_i \wedge h \geq h^c \wedge bd_G \wedge \gamma \geq \gamma^c; \\ (1 - \sigma)\pi & \text{for } bd_i \wedge h \geq h^c \wedge bd_G \wedge \gamma < \gamma^c. \end{cases} \tag{8}$$

Considering $\sigma = (\alpha\omega)^{1-\kappa}$ from Eq. 1, this leads to:

Proposition 2 *There are either one or two Nash-equilibria in pure strategies ζ_i, ζ_G with strategy profile:*

$$\zeta_i, \zeta_G = \begin{cases} bd_i, bc_G & \text{if } s > s^c := \max\left(\frac{1-(1-e^{\tilde{h}^c N})\pi}{e^{\tilde{h}^c N} \rho^{\tilde{c} \tilde{h}^c N} (\alpha\omega)^{1-\kappa}}; \frac{\omega^\kappa}{e^{\tilde{h}^c N} \alpha^{1-\kappa}}\right); \\ a_i, a_G. \end{cases}$$

Proof see Appendix B □

We have again two cases $s \leq s^c$ and $s > s^c$.

Case $s \leq s^c$: Self-enforcing Democracy

There is again only one Nash-equilibrium $\{a_i, a_G\}$ in this case. All government actors abstain from extra-constitutional activity and accept the constitutional rules of the game. The institutional determinants of a realization of this case follow from the partial derivatives of s^c in Proposition 2:

$$s^{c'}(\pi, \alpha, \kappa, \rho, \epsilon) < 0; s^{c'}(\tilde{h}^c, \tilde{\gamma}^c, N) > 0. \tag{9}$$

Case $s > s^c$: Tenuous Democracy

There are again two Nash-equilibria in this case. Similar to the transition game, we find that if $(bd_i; bc_G)$ is a Nash-equilibrium, it is also the payoff-dominant equilibrium. To determine the risk factor, we define $E(bd_i | \delta^{\tilde{\gamma}^c \tilde{h}^c N})$ ($E(a_i | \delta^{\tilde{\gamma}^c \tilde{h}^c N})$) to be the expected payoff of government official O_i from choosing bd_i (a_i), conditional on the probability $\delta^{\tilde{\gamma}^c \tilde{h}^c N}$ of a critical mass $\tilde{\gamma}^c \tilde{h}^c N$ to choose b_i in the transition game and d_i the post-transition game. Then, the risk factor is the probability $\delta_r^{\tilde{\gamma}^c \tilde{h}^c N}$ for which $E(bd_i | \delta^{\tilde{\gamma}^c \tilde{h}^c N}) = E(a_i | \delta^{\tilde{\gamma}^c \tilde{h}^c N})$, which is given by:

$$\delta_r^{\tilde{\gamma}^c \tilde{h}^c N} \left[e^{\tilde{h}^c N} (\alpha \omega)^{1-\kappa} \rho^{\tilde{\gamma}^c h N} s + (1 - e^{\tilde{h}^c N}) \pi \right] + (1 - \delta_r^{\tilde{\gamma}^c \tilde{h}^c N}) \pi = 1. \tag{10}$$

Solving for δ_r , yields:

$$\delta_r = \left[\frac{1 - \pi}{e^{\tilde{h}^c N} ((\alpha \omega)^{1-\kappa} \rho s - \pi)} \right]^{\frac{1}{\tilde{\gamma}^c \tilde{h}^c N}} \tag{11}$$

Remember that the likelihood of the government actors to coordinate on an equilibrium $\{bd_i; bc_G\}$ drops with a rise of the risk factor $\delta_r^{\tilde{\gamma}^c \tilde{h}^c N}$. The partial derivatives of Eq. 11 hence determine the likelihood of autocratic transition. They are as follows:

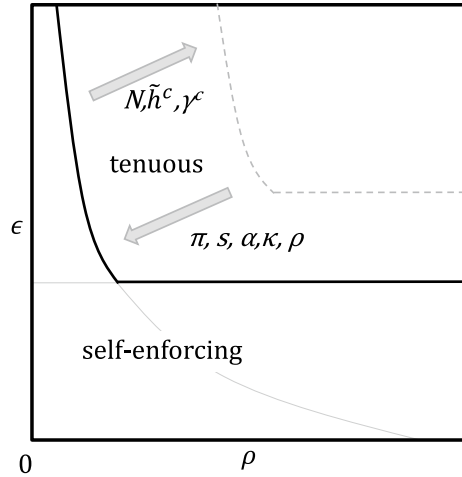
$$\delta_r'(\pi, s, \alpha, \kappa, \rho, \epsilon) < 0, \quad \delta_r'(\tilde{h}^c, \tilde{\gamma}^c, N) > 0. \tag{12}$$

The partial derivatives are again fully in line with those that determine the case $s \leq s^c$ of a self-enforcing democracy, as given by Condition 9. The main result of the transition game thus carries over to the full game. The differences in the result of the full game as compared to the transition game are determined by the following variables: first the critical share $\tilde{\gamma}^c$ of government officials that choose, if necessary, to rebel against the government leader in the post-transition period; and second the probability ρ of an individual government official to have a decision d_i in the post-transition period executed by their respective public authority (or by their *alter ego*, for that matter). These differences do not turn the results of the model upside down. But they are meaningful in another sense: The full game encompasses two significant collective-action problems that the actors face and that hence pinpoint the core difference between a democracy that is self-enforcing and one that is not. The difference arises because the government officials face problems in coordinating *twice* in the full model: first in the transition game where they need to coordinate on extra-constitutional activity, and second in the post-transition game where they need to potentially coordinate in order to credibly threaten to depose the newly established dictator for failing to share the autocratic rents.

What appears as a negligible modification of the conditions in the model may turn out to be decisive in reality. A potential autocratic government leader that signals his intention to switch to extra-constitutional activity must rely on his cooperation with the government officials. They, in turn, have a coordination problem, since failing to unite to follow the leader may have painful consequences for each of them. Additionally, however, they need to trust in the government leader’s promise to share the autocratic rents, although they do not have any natural reason to do so. As an alternative, they would have to trust in their own capability to keep the government leader in check, which requires them to solve a second and no less severe collective-action problem. Short of that, it is a better choice for each of the government officials to stick to constitution-abiding behavior.

In order to better grasp the intuition behind these results, we use the definition of the threshold value s^c of the government officials from Proposition 2:

Fig. 2 Stability regions



$$s^c := \max\left(\frac{1 - (1 - e^{\tilde{h}^c N})\pi}{e^{\tilde{h}^c N} \rho^{\tilde{\gamma}^c h N} (\alpha\omega)^{1-\kappa}}; \frac{\omega^\kappa}{e^{\tilde{h}^c N} \alpha^{1-\kappa}}\right). \tag{13}$$

The left-hand term in the brackets represents a necessary condition for the government officials to coordinate on extra-constitutional action. The right-hand term in the brackets gives a necessary condition for the government leader to switch to extra-constitutional action as a first-mover. Note that the latter requires the government leader to expect the government officials to follow suit, so that he must always see both conditions satisfied for even considering extra-constitutional action. The condition that is numerically higher is thus the binding one. Solving the left-hand term and the right-hand term of Condition 13 for ϵ yields:

$$\epsilon = \left[\frac{1 - \pi}{s^c (\alpha\omega)^{1-\kappa} \rho^{\tilde{\gamma}^c h N} - \pi} \right]^{\frac{1}{\tilde{h}^c N}}; \quad \epsilon = \left[\frac{\omega^\kappa}{s^c \alpha^{1-\kappa}} \right]^{\frac{1}{\tilde{h}^c N}}, \tag{14}$$

respectively. Each of the two equations represents combinations of the reliability indicators ρ and ϵ that form a threshold. If a given combination of ρ and ϵ is beyond both thresholds, then the constitution ceases to be self-enforcing.

The region to the north-east of the downward-sloping solid line in Fig. 2 represents combinations of ϵ and ρ that satisfy the left-hand side of Condition 14. The region above the horizontal solid line represents levels of ϵ that satisfy the right-hand side of Condition 14. If one of these conditions does not hold, then there is only one Nash-equilibrium (a_i, a_G) , and that applies to the region to the south-west of the bold and solid line. Hence, any point to the south-west of that line represents a self-enforcing democracy.

By contrast, any point that lies both to the north-east of the downward-sloping solid line *and* to the north of the horizontal solid line represents combinations of ϵ and ρ that satisfy both conditions in Eq. 13. Hence, we have a tenuous democracy in this region of the graph. Note that this does not mean that the government actors will in any case

coordinate on extra-constitutional action and thus launch a process of autocratic transition with certainty. What it does mean, however, is that they face two Nash-equilibria rather than one as soon as they cross the bold and solid line. Whenever that happens, each government actor needs to evaluate the likelihood of each equilibrium to be established prior to his individual choice between obeying or breaking the rules of the constitution. This kind of environment makes the constitution vulnerable to autocratic transition, since it might happen that a government leader attempts a first step into an autocratic-transition process, and it might then happen that the government officials coordinate on following the leader in his attempt. As a consequence, the bold and solid line in Fig. 2 can be viewed as some sort of a constitutional firewall. The lack of a second equilibrium, as it is illustrated in the lower left area, makes that firewall effective and democracy self-enforcing.

Some comparative statics can immediately be derived. According to Eq. 14, the bold and solid line in Fig. 2 is shifted outward by the following parameters: the number N of government officials; and the threshold shares \tilde{h}^c as well as $\tilde{\gamma}^c$ necessary for the government officials to coordinate in both the transition game and the post-transition game. By contrast, the following parameters shift the line inward: probability π of the government officials to remain in office in the case of a failed autocratic-transition attempt; the government officials' autocratic payoff s ; the level α of autocracy toleration by the population; and the severeness indicator κ of collective-action problems. Note that an inward shift of the bold and solid line in Fig. 2 shrinks the "self-enforcing" area and enhances the "tenuous" area.

The probability of autocratic transition to actually happen in the "tenuous" region is lowest at points close to the threshold line. It rises when we move from the threshold line toward the north-east of the diagram and hence when the reliability indicators ϵ and ρ rise. The latter indicates a rise in the capability of the individual government officials to have both of the following decisions executed by their respective subordinates—or accepted by their *alter ego*: a decision to break the constitutional rules in the transition period (ϵ); a decision to (potentially) rebel against the government leader in the post-transition game (ρ).

5 Discussion

For a democracy to be self enforcing or prone to autocratic transition, the following parameters are key according to our theory:

- Probability ω represents the government leader's pre-transition reelection probability.
- A high level of α represents a high autocracy toleration and, hence, a low valuation of democracy by the population.
- A high level of κ represents severe collective-action problems of the population in organizing protest activities following an autocratic-transition attempt.
- One of the most important determinants of the probability π is corruption. In an environment of the rule of law, government officials that had seriously transgressed against the constitutional rules will have to be sued and typically found personally unsuitable for their government position. Corruption is among

the most important diluting forces of this important principle and precisely so because corruption raises the probability that government officials will remain in office despite having transgressed against constitutional or legal rules.

- The number N , the threshold value \tilde{h}^c and the reliability indicator ϵ represent the effective degree of the separation of powers in the pre-transition period. The higher this degree, the more difficult it becomes for the government officials to coordinate with each other.
- The number N , the threshold value $\tilde{\gamma}^c$, and the reliability indicator ρ represent the effective degree of the separation of powers in the post-transition period. The higher the government officials expect this degree to be, the less likely they would expect to keep the future dictator in check, and the more reluctant they will be to coordinate on extra-constitutional action right away.

The pre-transition reelection probability ω has an ambiguous effect. On the one hand, a high probability of being reelected within the rules of democracy may undermine any motivation for autocratic transition right away. On the other hand, ω also represents popularity. If that popularity carries over to the transition period, it may help the government leader survive in office. Whether this plays any role depends on the question of whether the population is able to do something against autocratic transition. If κ is sufficiently high and, hence, collective-action problems are so severe that the population can do nothing, the government leader's popularity plays a role only as long as democratic rule still exists. In any case, a wannabe dictator is well advised to exacerbate potential collective-action problems in a sufficiently early period of autocratic transition, and that is precisely what most of them do.

The level α of autocracy toleration plays a similar role. If collective-action problems are not severe and autocratic toleration is low, then the public might be able to depose a government leader that attempts to turn his country into an autocracy. However, if the autocratic toleration is high and if the government leader is popular, then $\alpha\omega$ is high and the population abstains from ousting him, even if they face negligible collective-action problems.

Indeed, a high $\alpha\omega$ may indicate that it is not just the personal popularity of some individual politician that counts. What also counts is the policy such a politician pursues. Should the public plainly accept some strongman, possibly in times perceived as particularly threatening to the nation, then $\alpha\omega$ is high and so is popular support of a government leader that pursues autocratic transition. Modern populists typically may even gain popularity by referring to loftier goals which they—and only they—are not willing to sacrifice for the sake of some “fussy” statutory stipulations in the constitution, or they gain popularity by pursuing an allegedly desperate battle against external conspirators who exploit constitutional rules for the sake of hollowing out the cohesive bonds of the nation⁹. However, for such claims to resonate, specific institutional or political backgrounds are required.

Russia in the late 1990s may be an example. Vladimir Putin became President in 1999 following a rather chaotic period under Putin's predecessor. The privatization of former Soviet companies quickly spiraled out of control and led to an oligarchic structure that skyrocketed the wealth of a few but left the masses with virtually

⁹ For an extensive discussion, see Mounk (2018).

nothing. Corruption, crime, and Mafia structures flourished in parallel to rising poverty and unemployment. The face behind this undesirable development was that of then President Boris Yeltsin, whose character as *bon vivant*, and increasingly so as an alcoholic became symbols of both his personal incapability and the chaos of the country he led. And, becoming more obvious was the deeper problem behind the transition from a former superpower to a socially dissolving country shaken by corruption, financial distress, crime, economic stagnation and inequality: the surrendering of formerly honorable values to the primary antagonist of the now-perished glorious empire, namely to Western democracy.

Putin was the man to help right the ship, but he quickly made clear that his help would not be grounded in Western democratic rules. What is more, as the oligarchs, the corruption, the financial chaos and the lost empire came ostensibly out of Western democracy, a return to the *status quo ante* appeared almost natural to many. Indeed, Putin was able to curb the symptoms of many issues related to the allegedly decadent Western system: He stopped the disintegration of the Russian federation, he arrested prominent oligarchs, and he brought at least central parts of the economy back under his control. And he did all this with measures decidedly different from what Western consultants and representatives of liberal democracy repeatedly recommended. It was this policy that made him extremely popular with a large part, if not most, of the Russian population. This policy was hence an important precondition for Putin's path to autocratic transition, and populist rule was successful not least because of the particular unfortunate developments that Russia underwent during the 1990s under Boris Yeltsin.

That alone, however, was not sufficient for Russia; and it is generally not a sufficient condition for any country to embark on such a path. Rather, what is additionally needed is a high degree of mutual trust of the government elites regarding their ability to coordinate on a particular equilibrium; this factor relates to the effective degree of the separation of powers. Note that this is important for two reasons. First, it exacerbates the difficulty of collusion of the government officials against the constitution in the transition game; second, it aggravates the collective-action problem that the government officials face if the government leader fails to share the autocratic rents. Hence, while a populist policy might be capable of safeguarding public support on the path to autocracy, such a path might still remain impassable for the government officials if the degree of the separation of powers is too high.

Choosing to transgress or not to transgress against the constitution implies a far-reaching collective-action problem for the government officials. They must choose between two competing provisions that could protect their wealth: one provision stems from the existing set of constitutional rules, at least as long as these rules remain self-enforcing; the other stems from the presumed collective capability of the government officials to keep a future dictator in check by permanently threatening a coup in an environment lacking constitutional rules. If a sufficiently high share of the government officials expects that a sufficiently high share of their respective colleagues will not cooperate in both stages of the game, then individually keeping clear of the autocracy path is each government official's best choice. At the collective level of all government officials, then, this implies that the entire government sector will remain within the boundaries of the existing democratic constitution. By

contrast, should at least a share g^c of the government officials be expected to effectively cooperate on both stages, then it is individually rational for each individual government official to embark on the autocratic path. The constitution will fall victim to autocratic transition.

A truly fateful historical example of how a collective choice beyond the limits of self-enforcing democracy can go wrong is that of the highest-ranking officials of the German army (Wehrmacht) in the early 1930s. Feeling deeply degraded and undervalued under the conditions of the first German democracy following the disastrous World War I, the army officials trusted in their own apparent capability to deliberately exploit the political movement of a devilishly charismatic leader whom they did not like either. Although they disrespected Hitler because of his lack in military rank and nobility, they nevertheless pursued a hands-off approach regarding the Nazis' path to autocracy following Hitler's appointment as chancellor of the Weimar Republic by president Hindenburg on January 30, 1933. The army officials' aim was clear: Let Hitler do the dirty work of destroying the hated Weimar Constitution, eventually removing the Nazis in general, and Hitler in particular, from office, and then installing themselves to allegedly rescue the political culture of the German nation. In that way, they hoped for the resurrection of the pre-WW I German empire; or at least for some successor that was adequate according to their elitist views.

But the army officials dramatically underestimated Hitler's capability to undermine their capacity to coordinate on coup activities following the destruction of the Weimar democracy. Para-military organizations like "SS" and "SA" under the firm control of the Nazi party enhanced the complexity of the newly established system of security forces (which raised N); in combination with a set of additional measures, that split the security forces into numerous competing groups and subgroups (which raised $\tilde{\gamma}^c$). The result was a vastly reduced trust among a hitherto homogeneous elite of army commanders that were once bound by family ties and nobility membership.

After Hindenburg's death, Hitler urged the army commanders to transfer their vow of fidelity from the deceased president to Hitler himself (which, by activating the commanders' elitist *alter ego*, dramatically lowered ρ). The binding power of that vow grossly added to the commanders' difficulties in conspiring against Hitler within the complexity of competing security forces, eventually resulting in the failure of each further coup event all the way up to the end of WW II. It is safe to say that the army officials would have likely been more reluctant to exploit Hitler for the sake of destroying the Weimar Republic had they realized they would not collectively be able to keep the dictator in check.

The most important normative implication of our analysis is this: Constitutional design should pay particular attention to aspects like the formal and effective disentanglement of government authorities (separation of powers) as well as to federalism (Figueiredo Jr and Weingast 2005); competition among the different branches of government and among different federal levels; independent rules for appointing heads of the executive branch, the members of parliament and, most importantly, the judges of high-ranking courts. Failure to consider these factors may have been the most important driver of autocratic transition in the aftermath of the latest wave of democratization. As such, many of the newly established democratic constitutions

did not survive attacks by popular leaders like Lukashenko in Belarus, Nazarbayev in Kazakhstan, Karimov in Uzbekistan or, more recently, Putin in Russia.

By contrast, the constitution of the United States, with its deeply rooted separation of powers in various dimensions, has survived numerous attacks from leaders that were certainly not democrats but that did not even dare to openly confess their reservations against the pluralist democratic system; this even applies to very recent developments. In a similar fashion, Hungary and Poland may have a much better chance of surviving as democracies despite the strong attacks by their government leaders and by further decisive politicians because these countries are members of the EU, which enhances the complexity of their separation of powers. This is admittedly somewhat speculative, but our analysis would at least support this supposition.

As the most important empirical implication of our analysis, we have identified three main characteristics that either make a democratic constitution vulnerable to autocratic transition or transition-proof. These are the effective degree of the separation of powers, the level of corruption, and the popularity of transgressions against constitutional rules by (populist) government leaders. The one characteristic that can best be influenced in constitutional choice is the effective degree of the separation of powers. While this is indeed not a new aspect, it has possibly attracted less attention as a fundamental characteristic of sustainable democracies than it deserves. The principle *divide et impera* was meant to protect autocratic leaders against competitors that strive to attack the leader's regency. Its capacity to protect democracy against attacks on their constitutions, however, appears to rest on precisely the same mechanism.

A normative implication of our model is that constitutions matter. In this regard, it departs from Przeworski (2005, 2006). To him, a "constitution is neither sufficient nor necessary for democracy to survive" (Przeworski 2005, 267). In this view, a constitution is not necessary because actors would agree to the constitutional rules if they were an equilibrium anyway; and it is not sufficient because actors would break the rules if they did not describe an equilibrium. This implication follows from a specific modeling in which democracy is a unique equilibrium, conditional on the relevant parameters. Given these parameters, none of the relevant actors could be better off in any feasible alternative. Depending on the respective parameter setting, then, democracy is either a unique equilibrium or no equilibrium at all, implying in the latter case that democracy is not sustainable.

By contrast, there are potentially two equilibria in our model, of which democracy is but one, and not necessarily the one that is most preferred by the government officials. It is the task of the constitutional rules to structure mutually enforcing control mechanisms of all government actors that leave room for only one equilibrium on which all actors shall eventually coordinate—the democratic equilibrium (Calvert 1995; Hardin 2013). In that sense, then, a constitution can indeed be capable of making democracy "the only game in town" (Przeworski 2006, 324). Hence, our approach defines the conditions under which democracy is self-enforcing as it structures the constitutional rules in such a way as to allow for only one, namely the democratic equilibrium; by the same token, our approach defines the conditions under which democracy—if it exists—is not self-enforcing as it structures the constitutional rules in such a way as to allow for two equilibria on each of which the government actors might coordinate.

Note that a deep and diverse separation of powers is key among the rules that exclude the non-democratic equilibrium; as such, we may refer to our model as a post-Montesquieu approach. Since the depth and diversity of the separation of powers is obviously correlated with the level of per-capita income, our model provides an alternative explanation of Przeworski's empirical point of departure, namely the startling correlation between per-capita income and the probability of democracy to survive. However, while per-capita income is doubtlessly strongly correlated with the survival of democracy, it is still not causal in our model. It is rather correlated with those forces that are indeed causal for the survival of democracy.

6 Conclusions

We have developed a model of autocratic transition pursued by a sitting government leader. The aim was to identify the conditions under which autocratic transition is likely and, conversely, under which conditions democratic constitutions are effectively protected by mechanisms of self-enforcement. As a key element, each government official needs to expect coordination in two critical stages of the transition process in order to find their own participation worthwhile: Firstly, in an early period of transition, each government official must expect a sufficiently large share of the officials to participate in extra-constitutional activity. Second, the further government officials must, if necessary, expect a sufficient share of the government officials to participate in a coup against the newly established dictator.

Failure of this second condition to be satisfied leads to the non-credibility of the government leader's initial promise to share the autocratic rents with the government officials. Put in general terms, if the government officials fail to establish a credible and permanent collective coup threat, then they will not see themselves as protected against the exploitative power of the government leader in an autocracy.

As a result, we have identified three major testable empirical implications of our analysis. In particular, the likelihood of an autocratic transition is higher when (1) the effective degree of the separation of powers is low, (2) the degree of corruption is high, (3) collective action of the population against a transgressing government leader is difficult; and (3) the government leader that transgresses against constitutional rules for the sake of some allegedly higher goal—like protecting the nation against external or internal enemies or creating some glorious empire or the like—is highly popular.

As a normative implication, we find that constitutional design in young democracies should focus on measures that effectively divide the different powers both personally and institutionally, including separating the procedures for appointing government officials in different branches of the government and—if possible—on different federal levels.

Appendix A

See Tables 1 and 2.

Table 1 Stable newly-established democracies following 1989

Country	Polity2 max	Year max	Polity2 2018	Drop polity2
Albania	9	2005	9	0
Armenia	7	2018	7	0
Benin	7	2006	7	0
Bhutan	7	2018	7	0
Bulgaria	9	2001	9	0
Cape Verde	10	2001	10	0
Chile	10	2006	10	0
Croatia	9	2005	9	0
Czech Rep.	10	1993	9	1
Dominican Rep.	8	1996	7	1
Estonia	9	2000	9	0
Georgia	7	2005	7	0
East Germany	10	1990	10	0
Ghana	8	2004	8	0
Guatemala	8	1996	8	0
Guyana	7	2015	7	0
Hungary	10	1990	10	0
Indonesia	9	2014	9	0
Kenya	9	2013	9	0
South Korea	8	1998	8	0
Kyrgyzstan	8	2017	8	0
Latvia	8	1991	8	0
Lesotho	8	2002	8	0
Liberia	7	2018	7	0
Lithuania	10	1991	10	0
Macedonia, FYR	9	2002	9	0
Malaysia	7	2018	7	0
Mexico	8	2000	8	0
Moldova	9	2005	9	0
Mongolia	10	1996	10	0
Montenegro	9	2006	9	0
Myanmar	8	2018	8	0
Nepal	7	2018	7	0
Nigeria	7	2015	7	1
Pakistan	8	1990	7	1
Panama	9	1994	9	0
Paraguay	9	2013	9	0
Philippines	8	2002	8	0
Poland	10	2002	10	0
Romania	9	2004	9	0
Senegal	8	2000	7	1
Serbia	8	2006	8	0

Table 1 (continued)

Country	Polity2 max	Year max	Polity2 2018	Drop polity2
Sierra Leone	7	2007	7	0
Slovak Rep.	10	2006	10	0
Slovenia	10	1991	10	0
South Africa	9	1994	9	0
Taiwan	10	2004	10	0
Timor Leste	8	2014	8	0
Tunisia	7	2014	7	0

polity2 max: maximum polity2-value over period 1990-2018;

year max: first occurrence of maximum polity2-value, starting from 1990;

drop polity2: polity2 max minus polity2 2018.

Source: www.systemicpeace.org/inscrdata.html (December 16, 2020)

Table 2 Newly established democracies following World War I

Country	Polity2 max	Year max	Polity2 1938	Still democratic in 1938?
Austria	8	1920	0	No
Czechoslovakia	7	1918	7	Yes
Estonia	10	1919	-6	No
Finland	10	1919	4	No
Germany	6	1919	-9	No
Ireland	10	1927	8	Yes
Latvia	7	1920	-9	No
Netherlands	10	1917	10	Yes
Poland	8	1918	-6	No
Sweden	10	1917	10	Yes

polity2 max: maximum polity2-value from 1914 until 1938;

year max: first occurrence of maximum polity2-value, starting from 1914.

Source: www.systemicpeace.org/inscrdata.html (December 16, 2020)

Appendix B

Proof of Proposition 1 Player G will play b_G if and only if he expects $\tilde{h} \geq \tilde{h}^c$, and $\epsilon^{\tilde{h}^c N} \sigma_S > \omega$.¹⁰ Since players O_i can observe player G 's choice prior to their respective choice being due, their choice problem reduces to the question as to whether at least $h^c N$ of them effectively coordinate on b_i upon having observed b_G .

¹⁰ By convention, b_i, b_G is chosen iff $U(b) > U(a)$.

Now suppose that $\tilde{h} \geq \tilde{h}^c$. Upon having observed b_G , government official O_i expects $U_i = 1$ if he chooses a_i but $U_i = e^{\tilde{h}^c N} \sigma s_h + (1 - e^{\tilde{h}^c N}) \pi$ if he chooses b_i . He hence prefers $(b_i | b_G, \tilde{h} \geq \tilde{h}^c)$ over $(a_i | b_G, \tilde{h} \geq \tilde{h}^c)$ if and only if $e^{\tilde{h}^c N} \sigma s_h + (1 - e^{\tilde{h}^c N}) \pi > 1$, or if $s_h > \frac{1 - (1 - e^{\tilde{h}^c N}) \pi}{e^{\tilde{h}^c N} \sigma}$. At the same time, G has $U_G = e^{\tilde{h}^c N} \sigma s_h$ if he chooses b_G but $U_G = \omega$ if he chooses a_G . He hence prefers $(b_G | \tilde{h} \geq \tilde{h}^c)$ over $(a_G | \tilde{h} \geq \tilde{h}^c)$ if and only if $e^{\tilde{h}^c N} \sigma > \frac{\omega}{s_h}$. As a result, an action profile $\{b_i, b_G\}$ is a Nash-equilibrium if and only if $s_h > \max(\frac{1 - (1 - e^{\tilde{h}^c N}) \pi}{e^{\tilde{h}^c N} \sigma}, \frac{\omega}{s_h})$. Note that $\tilde{h} = 1$ in a Nash-equilibrium, which implies $s = s_h$. Considering $\sigma = (\alpha\omega)^{1-\kappa}$ from Eq. 1, this leads to $s > s^c := \max(\frac{1 - (1 - e^{\tilde{h}^c N}) \pi}{e^{\tilde{h}^c N} (\alpha\omega)^{1-\kappa}}, \frac{\omega^\kappa}{e^{\tilde{h}^c N} \alpha^{1-\kappa}})$.

Alternatively, suppose that $\tilde{h} < \tilde{h}^c$. Upon having observed b_G , government official O_i has $U_i = 1$ if he chooses a_i but $U_i = \pi$ if he chooses b_i . He thus always prefers $(a_i | b_G, \tilde{h} < \tilde{h}^c)$ over $(b_i | b_G, \tilde{h} < \tilde{h}^c)$. The government leader G , in turn, has $U_G = 0$ if he chooses b_G but $U_G = \omega$ if he chooses a_G . He thus always prefers $(a_G | \tilde{h} < \tilde{h}^c)$ over $(b_G | \tilde{h} < \tilde{h}^c)$. As a result, an action profile $\{a_i, a_G\}$ is always a Nash-equilibrium \square

Proof of Proposition 2 Suppose G to choose a sequence bc_G , and a share $\tilde{g} \geq \tilde{g}^c$. A player O_i will then have $U_i(a_i | \tilde{g} \geq \tilde{g}^c; bc_G) = 1$ as well as $U_i(bc_i | \tilde{g} \geq \tilde{g}^c; bc_G) = e^{\tilde{h}^c N} \rho^{\tilde{y}^c h N} \sigma s_g + (1 - e^{\tilde{h}^c N}) \pi$ and $U_i(bd_i | \tilde{g} \geq \tilde{g}^c; bc_G) = e^{\tilde{h}^c N} \rho^{\tilde{y}^c h N} \sigma s_g + (1 - e^{\tilde{h}^c N}) \pi$. O_i will then prefer both $(bd_i | \tilde{g} \geq \tilde{g}^c; bc_G)$ and $(bc_i | \tilde{g} \geq \tilde{g}^c; bc_G)$ over $(a_i | \tilde{g} \geq \tilde{g}^c; bc_G)$ if and only if $e^{\tilde{h}^c N} \rho^{\tilde{y}^c h N} \sigma s_g + (1 - e^{\tilde{h}^c N}) \pi > 1$ or if $s_g > \frac{1 - (1 - e^{\tilde{h}^c N}) \pi}{e^{\tilde{h}^c N} \rho^{\tilde{y}^c h N} \sigma}$. G , in turn, will have $U_G(bc_G | \tilde{g} \geq \tilde{g}^c) = e^{\tilde{h}^c N} \sigma s_g$ as well as $U_G(bd_G | \tilde{g} \geq \tilde{g}^c) = 0$ and $U_G(a_G | \tilde{g} \geq \tilde{g}^c) = \omega$. Player G will hence never choose $(bd_G | \tilde{g} \geq \tilde{g}^c)$, and he will prefer $(bc_G | \tilde{g} \geq \tilde{g}^c)$ over $(a_G | \tilde{g} \geq \tilde{g}^c)$ if and only if $s_g > \frac{\omega}{e^{\tilde{h}^c N} \sigma}$.

On the other hand, G has $U_G(bd_G | \tilde{h} \geq \tilde{h}^c; \tilde{y} < \tilde{y}^c) = e^{\tilde{h}^c N} \sigma S$, but he will only have $U_G(bc_G | \tilde{h} \geq \tilde{h}^c; \tilde{y} < \tilde{y}^c) = e^{\tilde{h}^c N} \sigma s_g < e^{\tilde{h}^c N} \sigma S$. His announcement to play a sequence bc_G is hence only credible in the case that $g \geq g^c$. As a result, an action profile $\{bd_i, bc_G\}$ is a subgame-perfect Nash-equilibrium if and only if $s_g > s^c := \max(\frac{1 - (1 - e^{\tilde{h}^c N}) \pi}{e^{\tilde{h}^c N} \rho^{\tilde{y}^c h N} (1 - \alpha) \omega}, \frac{1}{e^{\tilde{h}^c N} (1 - \alpha)})$. Note that $\tilde{g} = 1$ in the subgame-perfect Nash-equilibrium and hence $s = s_g$. Considering $\sigma = (\alpha\omega)^{1-\kappa}$ from Eq. 1, the condition for a subgame-perfect Nash-equilibrium can also be written as $s > s^c := \max(\frac{1 - (1 - e^{\tilde{h}^c N}) \pi}{e^{\tilde{h}^c N} \rho^{\tilde{y}^c h N} (\alpha\omega)^{1-\kappa}}, \frac{\omega^\kappa}{e^{\tilde{h}^c N} \alpha^{1-\kappa}})$.

Alternatively, suppose that $h < h^c$. Then each O_i has $U_i(bd_i | \tilde{h} < \tilde{h}^c) = U_i(bc_i | \tilde{h} < \tilde{h}^c) = \pi$ and $U_i(a_i | \tilde{h} < \tilde{h}^c) = 1$, so that each O_i prefers $(a_i | \tilde{h} < \tilde{h}^c)$ over any alternative. G , in turn, has $U_G(bd_G | \tilde{h} < \tilde{h}^c; bd_G) = U_G(bc_G | \tilde{h} < \tilde{h}^c) = 0$ and $U_G(a_G) = \omega$ and will thus prefer $(a_G | \tilde{h} < \tilde{h}^c)$ over any alternative. As a result, an action profile $\{a_i, a_G\} \forall i$ is always a Nash-equilibrium in pure strategies \square

Acknowledgements I am indebted to Celeste Brenneka, Lena Gerling, Mario Gilli, Helena Helfer and Kim Kellermann for many valuable comments and suggestions.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acemoglu D (2006) A simple model of inefficient institutions. *Scand J Econ* 108(4):515–546
- Acemoglu D, Johnson S, Robinson JA (2001) The colonial origins of comparative development: an empirical investigation. *Am Econ Rev* 91(5):1369–1401
- Acemoglu D, Naidu S, Restrepo P, Robinson JA (2019) Democracy does cause growth. *J Polit Econ* 127(1):47–100
- Acemoglu D, Robinson JA (2000) Why did the west extend the franchise? Democracy, inequality, and growth in historical perspective. *Q J Econ* 115(4):1167–1199
- Acemoglu D, Robinson JA (2001) A theory of political transitions. *Am Econ Rev* 91(4):938–963
- Acemoglu D, Robinson JA (2006) Economic origins of dictatorship and democracy. Cambridge University Press, Cambridge
- Acemoglu D, Robinson JA (2012) Why nations fail: the origins of power, prosperity and poverty. Profile Books, London
- Apolte T (2012) Why is there no revolution in North Korea? *Public Choice* 150(3–4):561–578
- Bénabou R, Tirole J (2011) Identity, morals, and taboos: beliefs as assets. *Q J Econ* 126(2):805–855
- Benhabib J, Przeworski A (2006) The political economy of redistribution under democracy. *Econ Theor* 29(2):271–290
- Boix C, Svolik MW (2013) The foundation of limited authoritarian government: institutions, commitment, and power-sharing in dictatorships. *J Polit* 75(2):300–316
- Bruce BM, Smith A (2009) Political survival and endogenous institutional change. *Comp Pol Stud* 42(2):167–197
- Calvert R (1995) The rational choice theory of social institutions: cooperation, coordination, and communication. In: Banks JS, Hanushek EA (eds) *Modern political economy*. Cambridge University Press, Cambridge, pp 216–268
- Carlsson H, Van Damme E (1993) Global games and equilibrium selection. *Econometrica* 5:989–1018
- Casper BA, Tyson SA (2014) Popular protest and elite coordination in a coup d'état. *J Polit* 76(02):548–564
- Clark K, Kay S, Sefton M (2001) When are Nash equilibria self-enforcing? An experimental analysis. *Internat J Game Theory* 29(4):495–515
- Cooper R, DeJong DV, Forsythe R, Ross TW (1992) Selection criteria in coordination games: some experimental results. *Q J Econ* 107(2):739–771
- de Mesquita B, Bruce AS, Siverson RM, Morrow JD (2005) *The logic of political survival*. The MIT Press, Cambridge
- Fearon JD (2011) Self-enforcing democracy. *Q J Econ* 126(4):1661–1708
- Ferguson N (2001) *The cash nexus: money and power in the modern world, 1700–2000*. Basic Books, New York
- Février P, Linnemer L (2006) Equilibrium selection: payoff or risk dominance?. The case of the Weakest Link. *J Econ Behav Organ* 60(2):164–181
- Figueiredo Jr RJP, Weingast BR (2005) Self-enforcing federalism. *J Law Econ Organ* 21(1):103–135
- Granovetter M (1978) Threshold models of collective behavior. *Am J Sociol* 1420–1443
- Gründler K, Krieger T (2016) Democracy and growth: evidence from a machine learning indicator. *Eur J Polit Econ* 45:85–107
- Gundlach E, Paldam M (2008) Income and democracy: a comment on Acemoglu, Johnson, Robinsn, and Yared 2008. Technical report. Economics working paper
- Gundlach E, Paldam M (2009) A farewell to critical junctures: sorting out long-run causality of income and democracy. *Eur J Polit Econ* 25(3):340–354

- Hardin R (2006) Constitutionalism. In: Weingast BR, Wittman DA (eds) *The Oxford handbook of political economy*, pp 289–311
- Hardin R (2013) Why a constitution. In: Galligan DJ, Versteeg M (eds) *The social and political foundations of constitutions*. Cambridge University Press, New York, pp 51–72
- Harsanyi JC, Selten R (1988) *A general theory of equilibrium selection in games*. The MIT Press, Cambridge
- Hyde SD, Marinov N (2014) Information and self-enforcing democracy: the role of international election observation. *Int Organ* 68(2):329–359
- Lipset SM (1959) Some social requisites of democracy: economic development and political legitimacy. *Am Polit Sci Rev* 53(1):69–105
- Marwell G, Oliver P (1993) *The critical mass in collective action*. Cambridge University Press, Cambridge
- Mittal S, Weingast BR (2011) Self-enforcing constitutions: with an application to democratic stability in America's first century. *J Law Econ Organ* 29(2):278–302
- Morris S, Shin HS (2010) Global games: theory and application. In: Dewatripont M, Peter LH, Turnovsky S (eds) *Advances in economics and econometrics*. Cambridge University Press, Cambridge, pp 57–114
- Mouk Y (2018) *The people versus democracy: why our freedom is in danger and how to save it*. Harvard University Press, Cambridge
- North DC, Wallis JJ, Weingast BR (2009) *Violence and social orders: a conceptual framework for interpreting recorded human history*. Cambridge University Press, Cambridge
- Przeworski A (1991) *Democracy and the market: political and economic reforms in Eastern Europe and Latin America*. Cambridge University Press, Cambridge
- Przeworski A (2005) Democracy as an equilibrium. *Public Choice* 123(3):253–273
- Przeworski A (2006) Self-enforcing democracy. In: Weingast BR, Wittman DA (eds) *Oxford handbook of political economy*. Oxford University Press, New York, pp 312–328
- Straub PG (1995) Risk dominance and coordination failures in static games. *Q Rev Econ Finance* 35(4):339–363
- Svolik MW (2012) *The politics of authoritarian rule*. Cambridge University Press, Cambridge
- Traversa F (2015) Income and the stability of democracy: pushing beyond the borders of logic to explain a strong correlation? *Const Polit Econ* 26(2):121–136
- Tullock G (1987) *Autocracy*. Kluwer, Dordrecht
- van Damme E (2002) Strategic equilibrium. In: Aumann R, Hart S (eds) *Handbook of game theory*, vol 3. North-Holland, Amsterdam, pp 1521–1596
- Weingast BR (1997) The political foundations of democracy and the rule of the law. *Am Polit Sci Rev* 91(2):245–263
- Weingast BR (2005) The constitutional dilemma of economic liberty. *J Econ Perspect* 19(3):89–108

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.