Shaping Institutions*

William Fuchs† Satoshi Fukuda‡

December 19, 2023

Abstract

We propose a simple model of the evolution of institutions, where leaders' actions have a persistent effect by shaping the norms of the institutions they lead. This can lead to different long-run behaviors even for institutions with the same formal rules. The early history of leaders plays a crucial role in determining which outcome prevails. Every period, a leader decides to respect or abuse their position. Respect strengthens the norms; abuse weakens them. Leaders' type and current norms determine the benefit/cost of abusing the position. Norms also determine the replacement probability of leaders. We elucidate democratic backsliding and corporate-board capturing.

JEL Codes: D02; D7; C73; G3; M14
Keywords: Institutions; Norms; Leadership; Long-Run Dynamics; Democratic Backsliding; Corporate Board Capturing

*We would like to thank Aydogan Alti, Simon Board, Alessandro Bonatti, Matteo Escudé, Scott Gehlbach, Ole Jann, Niccolò Lomys, Ester Manna, Suraj Prasad, Andrea Prat, and Thomas Wiseman for their insightful comments and discussions. We would also like to thank seminar participants at various conferences and universities. We thank Francesco Bilotta, Kevin Mei, and Mahyar Sefidgaran for their excellent research assistantship. Fuchs gratefully acknowledges the support from ERC Grant 681575, Grant PGC2018-096159-B-I00 financed by MCIN/AEI/10.13039/501100011033, and Comunidad de Madrid (Spain), Grant EPUC3M11 (V PRICIT) and Grant H2019/HUM-5891.

†McCombs School of Business, UT Austin, Universidad Carlos III Madrid, CEPR, and FTG.
‡Department of Decision Sciences and IGIER, Bocconi University.

1 Introduction

The legal framework detailing the duties and responsibilities of institutional leaders—for example, presidents and CEOs—is naturally incomplete. Thus, political and corporate norms play an important complementary role in shaping the behavior of leaders. For example, as Renan (2018) puts it: "The nature of the presidency in American constitutional governance cannot be understood without reference to norms.... Presidential power is both augmented and constrained by these unwritten rules of legitimate or respectworthy behavior."

In turn, norms are malleable, and an important component in the evolution of norms is the behavior of those leading the institutions. Thus, the actions of past leaders have long-lasting effects on the actions of future leaders. As President George Washington wisely observed, "There is scarcely any part of my conduct which may not hereafter be drawn into precedent" (Greenstein, 2009). As a result, two economies or organizations with the same formal institutions can have very different long-term outcomes because of the examples set by their early leadership. Also, importing elements from the formal legal framework of a successful nation is likely to fail, if the local norms and customs are not properly accounted for.

We present a parsimonious model to capture these ideas. In our model, the legal framework determines the initial institutional strength. Thereafter, it evolves endogenously as a function of the leader's actions. A leader can either abuse or respect the institution. Abuse weakens the norms. For example, what used to cause a scandal can become "normal" behavior.[1] Conversely, norms are also strengthened after they have been respected.

In turn, the institutional strength influences the behavior of the leader in two dimensions. Firstly, the weaker the institution, the larger the payoff the leader can reap from abusing it. Secondly, abusing power can affect the possibility of staying in office in two ways. First, misbehavior can be scandalous and increase the likelihood of being replaced. Second, in the opposite direction, abuse can allow for more political patronage, capturing the corporate board, or election meddling favoring the incumbent. Which of these effects dominates can itself be a function of the norms.

The leader's behavior also depends on her type. The leader's type determines the relative flow benefit of being in office under both actions. In our model, higher types are either more moral or less skilled at cheating and thus reap less benefit from abusing their position.

We parametrize how responsive institutional strength is to prior behavior. This allows us

[1]President Trump was severely criticized for not revealing his tax records, but it is likely that similar actions by future candidates will not be similarly frowned upon since there is now a precedent. It is important to note as well that, before President Ford shared his tax records in the seventies, there did not use to be a norm of doing this, so we can easily return to that benchmark.

to establish our main result: endogenous norm response is crucial for obtaining two possible long-term absorbing states for a given initial condition.

For a given set of leader types, when norms do not evolve, there are three possibilities: (i) weak norms which always promote abuse; (ii) strong norms which always promote respect; or (iii) intermediate norms which elicit a different behavior depending on the leader's type. Instead, when norms evolve endogenously, a fourth possibility arises. While a sequence of moral leaders can steer the economy into a steady state of strong norms and respect, a sequence of unethical leaders can steer the economy into a steady state of weak norms and abuse. This case is essential for understanding why several countries, such as Argentina, even though they modeled their constitutions after the United States, seem to be in a very different steady state (see, for instance, Alston and Gallo, 2010). Of course, there can be many factors explaining such long-term differences, but the US might also have been somewhat lucky with the leaders it has had. Remarking on President Trump's damage to American democracy, Kamarck (2021) points out "Fortunately, we haven't had many of those in our 200-plus years of history." In the corporate setting, the importance of early leadership in shaping organizational culture has been pointed out by Schein (1983).

This endogenous evolution of norms also helps rationalize the concerns about the long-term effects of President Trump's disregard for several institutional traditions. This widely-held sentiment was captured by Foran (2016): "Growing tolerance for conflicts of interest in government, limitations on media access and accountability, and harsh treatment of minority groups can accumulate.... Each norm that falls is one fewer safeguard against executive over-reach than we had before. Even if we never become an authoritarian state, our governance will suffer as a result. For now, we should recognize the precedents that are already being set and try to prevent them from becoming the new normal."

The long-lasting effect of President Trump's actions can be observed in the fact that a large fraction of the Republican electorate continued to believe the election lies years after the event. A poll requested by Newsweek on October 30th 2022 found that "40 percent of Americans believe that the 2020 presidential election ... was rigged or stolen."[2]

Related to the point above, it is important to note that the transition from a democracy to an autocracy can take place gradually. This gradual process is known as democratic backsliding. Cháves in Venezuela, Erdogan in Turkey, Orbán in Hungary, and Putin in Russia are recent prominent examples. The slow erosion of different democratic safeguards from freedom of the press, independence of the courts, corruption and abuse of state resources are clearly observed in all of these countries (see Figure 5). By interpreting norms broadly

[2]https://www.newsweek.com/40-americans-think-2020-election-stolen-days-before-midterms-1756218 (Date of Access: January 9th, 2023).

as including respect for these institutions, our model captures this important phenomena.

Several empirical papers provide additional support to these concerns by pointing out the importance of path dependence in shaping institutions.[3] La Porta et al. (1999) demonstrate the role of exogenous political historical factors in explaining government performance. Acemoglu et al. (2008) argue that cross-sectional relationship between democracy and income today is the result of societies embarking on divergent development paths at certain historical critical junctures (e.g., the founding of a nation in the context of our paper). Papers such as Acemoglu et al. (2001), Glaeser et al. (2004), and Acemoglu and Robinson (2008) demonstrate persistence of institutional outcomes. Syverson (2004) and Hsieh and Klenow (2009) report persistent performance differences among seemingly similar enterprises. Our paper suggests that, when evaluating the quality of governance in political or corporate settings, it is important to condition on the history of past leaders. The corporate setting is a more promising venue for empirical studies since the panel nature of the data and the fact that CEOs change firms (albeit not exogenously) might allow for controlling for leader type with a CEO fixed effect.

Consistently with this evidence, our paper shows (i) how countries or corporations with similar formal institutions may end up diverging due to the early leaders' behavior; and (ii) that there is a level of norm below (above) which the norm level persistently goes down (up). In this sense, our paper provides a dynamic micro-foundation for multiple steady states.

An important element determining the long-run properties of the model is the distribution of leader types. In the corporate context, Bertrand and Schoar (2003) document the importance of "types" by showing that individual managers' characteristics affect corporate behavior and performance. On a related point, Graham et al. (2020) show that the sudden death of an entrenched CEO vis-a-vis a non-entrenched one leads to a 3% higher abnormal return. Even with very strong norms, a very extreme type might still prefer to abuse. As Klein (2016) observed "The normal constraints, meanwhile, are failing this year. Trump does not have enough shame to check himself.... He doesn't care if he's condemned, or called a bigot, or shown to be a liar." Looking forward, Pfiffner (2021) points out: "The broader impact of President Trump's behavior will depend crucially on the character of future presidents."

The distribution of possible leader types can itself be responsive to the state of the system. We can extend our model to accommodate this possibility. For example, when there is a very poor institutional environment and abuse is more profitable, types that seek office with this purpose are more likely to enter. In addition, once a system becomes very corrupt it might become very hard to rise to power within a party/organization if one is not willing to engage in (or tolerate) corruption. As a result, the distribution of types can shift down

[3]For a broad overview, see, for instance, North (1990), Pierson (2000), and Acemoglu et al. (2021).

making it less likely to select a sufficiently honest leader who would not abuse their position.

Another possibility we can capture with an endogenous distribution of leader types is what might arise after a long period of abuse when the current leader becomes so powerful (weak norms) that it is very unlikely to be replaced. In the political context, there could be two types replacing the current leader, another despot (perhaps the son/daughter of the current leader, as it has occurred in North Korea, or some rival) or we could have the emergence of patriotic heroes, types that are willing to risk their life to help restore institutions.[4] Since institutions are very weak at this juncture, if a regime falls and there is no clear benevolent patriot to take over, it is very tempting to abuse the position and drift into an authoritarian regime once again. This might help explain the poor record of the Arab Spring or why a regime change from the outside is prone to failure.

In the corporate context, the set of types that are considered for the CEO position can also respond to a previous scandal. For instance, after more than 150 years of German leadership, Siemens hired its first non-German CEO following a bribery scandal. Also, the more entrenched a CEO is, the more likely he/she will have a strong influence on the choice of its successor.

While, for analytical convenience, our base model does not consider term limits, it is interesting to study them from a policy perspective. While the main features of our long-run dynamics are robust to the introduction of term limits, we point out two opposing effects of term limits on the behaviors of a leader. Term limits create a natural end-of-term effect which increases the incentive for the leader to behave myopically. In many circumstances this implies that the leader will be more likely to abuse at the end of the term. This might suggest that increasing the term limit would be a good thing from an institutional design perspective. However, we show that although this force is indeed there, there is a countervailing consideration that calls for shorter term limits. Institutional erosion might not be profitable in the short run and a leader might only want to engage in it if it has enough time to reap its benefits. Thus, shortening the term limit can be a disincentive for abuse. The less the leader can erode the institutions, the longer the optimal term limits.

In our base model, there is a persistent effect of formal rules on the evolution of institutional norms. While in our base model the formal rules are fixed throughout, in an extension we allow for these to change over time. They could change in two ways. First, when norms are not too weak, change can be reactionary. That is, after some norm gets abused, that action is codified in the written rule, thus making it stronger. A good example of this is term limits in the United States. There used to be no formal term limits, as Alexander

[4]This is similar in spirit to the "prominent agents in the model by Acemoglu and Jackson (2015) that can restore cooperative behavior.

Hamilton had even written in Federalist No. 69: "That magistrate is to be elected for four years; and is to be re-eligible as often as the people of the United States." Despite this, after George Washington and Thomas Jefferson served for just two terms, this became effectively the norm. As a result of Franklin D. Roosevelt serving beyond two terms the Twenty-Second Amendment was introduced to limit the presidency to two terms. Conversely, the written rules can also be made weaker. Nicaraguan strong man Daniel Ortega managed to remove the term limits in Nicaragua in 2014 so that he could serve for another term (and is still in power). As illustrated by the examples above, relative to our benchmark model, when changing the formal rules is possible, the economy is more likely to get absorbed into a degenerate steady state.

The rest of the paper is structured as follows. The rest of this section discusses the related literature. Section 2 lays out the model. Section 3 provides our main analysis. Specifically, Section 3.1 characterizes a leader's decision, and Section 3.2 studies norm dynamics. Section 4 discusses democratic backsliding and term limits. Section 5 studies extensions of the model. Section 6 provides concluding remarks. Proofs are relegated to Appendix A.

Related Literature

The legal and political-science literature has long studied the roles of informal rules and norms that political leaders face, as early as Bryce (1888 [1995]). More recently, in the legal scholarship literature, Renan (2018) and Ahmed (2022), for instance, study how such "presidential norms" or "constitutional norms" augment and constrain presidential powers. In the political science literature, Azari and Smith (2012) and Levitsky and Ziblatt (2018), for instance, study the roles of informal rules and norms on democratization and autocratization. Levitsky and Way (2015), Huq and Ginsburg (2018), and Diamond (2021) point out that democratic backsliding in the world has been caused not by coups but by elected governments, suggesting the importance of constitutional norms. As discussed, Cháves in Venezuela, Erdogan in Turkey, Orbán in Hungary, and Putin in Russia are recent prominent examples.

In the literature on democratic consolidation, a process through which democracies consolidate lowering the risk of reverting to authoritarianism (e.g., O'Donnell and Schmitter, 1986; Linz, 1990), O'Donnell (1996) argues the importance of informal rules, and Linz (1978) studies how political leaders' behavior can either reinforce or diminish democracy. Almond (1956) argues the role of "political culture" on the functioning of government (see also Almond and Verba, 1963; Diamond, 1999). Our paper contributes to these strands of literature by providing a micro-founded process in which institutional norms are gradually eroded or re-

inforced. This allows us to obtain endogenous and possibly differing long-run configurations of institutional norms.

There is now an emerging literature on democratic backsliding in economics and political science.[5] For instance, Helmke et al. (2022) study two parties that try to gradually tilt the electoral rules (e.g., gerrymandering). Grillo and Prato (2023) show in their static model that democratic backsliding can occur even if most voters and most politicians intrinsically value democracy, when minorities of both groups are willing to accept violations of democratic norms and politicians value popular support.

Howell et al. (2023) present a dynamic model in which, an executive, while subject to judicial review, slowly but surely undermines checks and balances. In the long run, there is a persistent accumulation of authority and a degradation of judicial checks. While our paper is related to theirs in that the behavior of an officeholder today affects that of a successor tomorrow, their paper does not address path dependence nor it allows for different long-run outcomes. In Luo and Przeworski (2023), an incumbent can choose whether to take a step to increase the probability of remaining in office. They show that backsliding occurs when either voters find the incumbent appealing even if the incumbent erodes the democracy or the incumbent is so unappealing that she can remain in office only by backsliding. Unlike norms in our model, the actions of past leaders have no effect on new leaders' actions and probabilities of remaining in office. To the best of our knowledge, our paper provides the first formal model that elucidates the role of the evolution of norms on democratic backsliding. We also make a novel contribution by characterizing long-run dynamics and path dependence in democratic backsliding.

Political scientists and economists have analyzed determinants of corruption. In the theoretical literature, Andvig and Moene (1990) present a static model of corruption with multiple equilibria which tries to explain why the same socio-economic structure can give rise to different levels of corruption. Shleifer and Vishny (1993) study a static model in which economic and political competition can reduce the level of corruption. In the empirical literature, Tanzi (1998) points out the role of the example provided by the political leadership. Our paper formalizes the role that the current political leadership plays on the behavior of the future leaders and thus a rationale for the persistence of corruption. Also, as pointed out by Paldman (2002), countries with similar backgrounds can drift into very different corruption regimes (e.g., Argentina and Chile).

Although in our model norms of checks and balances are disembodied, one could interpret the abusive action of the leaders as placing "yes-men" in supervisory or control positions.

[5]See, for instance, Lust and Waldner (2015) and Grillo et al. (2023) for surveys on democratic backsliding in the political science literature.

In the political setting, supporting election deniers to the attorney general position is such an example.[6] In the corporate setting, this would correspond to the CEO "capturing" the board. Discussing the Volkswagen emissions scandal of 2015, Alexander Juschus, director at IVOX, the German proxy adviser, remarked "The scandal clearly also has to do with structural issues at VW.... There have been warnings about VW's corporate governance for years, but they didn't take it to heart and now you see the result."[7] There is a large strand of literature in accounting and corporate finance discussing this as a concern and its potentially negative effects.[8]

There is also a broad literature on leadership. See, for instance, Gibbons and Henderson (2013), Bloom et al. (2014), and Bandiera et al. (2020) for the context of organizational economics and Jones and Olken (2005) and Myerson (2011) political economy. The role of leadership in our model is quite different from the one in the leadership literature, as our focus is on norms through which the behaviors of the current leader have a permanent effect on those of the future leaders. In this way we also distinguish ourselves from the broad literature that discusses leadership and culture, which is mostly focused on the contemporaneous influence of the CEOs on other employees of the organization (e.g., Ashforth and Anand, 2003; Biggerstaff et al., 2015; Guiso et al., 2015). Our paper also suggests the importance of conditioning on the history of past leaders in evaluating the quality of governance.

There are strands of literature that view norms as equilibrium objects. Here, focusing on institutional norms, Invernizzi and Ting (Forthcoming) consider a model in the spirit of Dixit et al. (2000). They view norms as the expected play in the efficient subgame perfect equilibrium of a policy game. Although they can study the interaction of formal or informal rules, they do not really have a focus on the norm dynamics. Unlike our model, there are no changing leader types and thus on path norms do not weaken or strengthen. Hence, they cannot address democratic backsliding nor the importance of path dependence for different steady states.

Finally, we model the evolution of norms as a form of capital. In this sense, we are close to the literature that studies the role of "social capital" on the functioning of governments (e.g., Putnam, 1993; Guiso et al., 2016). Persson and Tabellini (2009) theoretically and empirically study the notion that they formulate as "democratic capital," measured by a nation's historical experience with democracy and the incidence of democracy in its neighborhood.

[6] 47% of Republican candidates in the midterms for House, Senate, governor, secretary of state and attorney general did not accept the legitimacy of the 2020 election (https://fivethirtyeight.com/videos/the-number-of-election-denying-republicans-defined-the-2022-midterms/?cid=rrfeaturedvideo, Date of Access: January 9th, 2023).

[7] https://www.ft.com/content/e816cf86-6815-11e5-a57f-21b88f7d973f, Date of Access: January 14th, 2023.

[8] For instance, Laux (2008) studies the effects of board independence on CEO's rent-seeking.

They demonstrate that democratic capital reduces the exit rates from democracy and raise the exit rates from autocracy. Besley and Persson (2019) model democratic values, defined as the proportion of citizens who may fight for democracy against autocracy. In this respect, higher democratic values play a similar (complementary) role to having stronger norms. Yet, there are a few important differences with respect to our work. Chiefly, we highlight the importance of leader types in determining future outcomes. In particular, our main point is that leaders can have a permanent effect. In contrast, in their model the long-run outcomes are not history dependent. Their model does not lend itself to analyze (nor do they discuss) democratic backsliding. Moreover, our focus on the role of leaders allows us to study "organizational capital" in a broader context that includes both polities and corporations.[9] Turning to the corporate context, Dessein and Prat (2022) study a model of "organizational capital," an intangible asset that has to be maintained by a leader. Organizational capital affects firm performance. The leader faces whether to increase organizational capital or boost short-term profit. They characterize a steady state distribution of organizational capital in which otherwise similar firms may have persistent performance differences. Although similar long-run dynamics can arise in our model, mechanisms are very different. In particular, in our setting, a more patient leader may have less incentives to improve norms. In contrast, in their model, a more patient leader has stronger incentives to invest in organizational capital.

2 Model

Every period $t \in \{1, 2, \dots\}$, the incumbent leader must decide on one of two actions $a_t \in \{0, 1\}$. The action $a_t = 1$ represents the leader abusing her position or cheating. In contrast, $a_t = 0$ represents the leader abiding by or respecting the rules. The leader's time-t payoff from taking either of these actions is determined by two elements: (i) the type, h, representing the leader's level of honesty (e.g., the leader's sense of fiduciary duty) or ability to cheat; and (ii) the norm level N_t determining the institutional strength. We assume

$$u(a_t, N_t, h) := b - a_t(N_t + h).^{10}$$

[9] Besley and Persson (Forthcoming) study a different model in which organizational culture is formulated as the distribution of "types" in an organization which affect project choices.

[10] In the main application, we implicitly assume that violating a norm is bad for society. One could easily modify the model to a situation in which respecting a norm does not necessarily bring a positive value: for example, whether the supreme court adheres to precedents or not and whether a president issues an executive order or not.

The first term $b \geq 0$ is the benefit of being in power (e.g., non-pecuniary benefits from holding office and pecuniary benefits such as wages and office perks).[11] Thus, if the leader respects the rules (i.e., $a_t = 0$), then the payoff is b. If the leader abuses her position (i.e., $a_t = 1$), then the payoff is $b - (N_t + h)$. The second term measures the incremental benefit/cost of abusing power. Note that this payoff is decreasing in the strength of norm N_t and in the honesty parameter h.[12] A natural extension of the model would be to consider the case in which both actions and norms are a vector (see Section 5.3). For ease of exposition, we constrain our analysis to the scalar case.

Importantly, when deciding which action to take, the leader also takes into consideration how her actions affect her probability of remaining in power. We denote the replacement probability at time t by the function $\lambda(a_t, N_t)$. We assume: (i) $0 \leq \lambda(a_t, N_t) \leq 1$; (ii) $\lambda_0(N_t) := \lambda(0, N_t)$ is non-increasing and continuous; and (iii) $\lambda_1(N_t) := \lambda(1, N_t)$ is non-decreasing and continuous. The first assumption is needed since λ is a probability. Assumptions (ii) and (iii) imply that $\lambda_1(N_t) - \lambda_0(N_t)$ is non-decreasing in N_t. This is meant to capture the idea that the higher the norms, the more likely it is that abusing power will lead to losing the position. As we do not impose any assumption on the relative magnitudes of λ_1 and λ_0, our setup allows for the existence of some norm level $\tilde{N}$ such that for $N_t < \tilde{N}$ abusing power enhances the probability of remaining in office $\lambda_1(N_t) - \lambda_0(N_t) < 0$ while for $N_t > \tilde{N}$ abusing power lowers the probability of remaining in office $\lambda_1(N_t) - \lambda_0(N_t) > 0$. The interpretation of this is that, in the political context, when the norms are low, abusing power allows the politician to engage in activities that might help her get re-elected such as: patronage, bread and circuses, or directly meddling with the elections, while facing little risk of a scandal. See Section 5.1 for discussions on how λ could capture political competition, independence of media or judicial independence.[13] The model also allows us to contemplate the extreme possibility that when norms are sufficiently weakened a leader can guarantee remaining in power by abusing her position, i.e., $\lambda_1(N_t) = 0$.

The strength of the norms at time t, N_t, is a function of some formal set of rules, $\overline{N}$, and the history of actions by past leaders. Formally, we assume norms evolve according to:

$$N_{t+1} = (1 - \delta)N_t + \delta\overline{N} + (1 - 2a_t)\gamma, \tag{1}$$

with initial condition $N_1 = \overline{N}$. If the leader respects the rules, $a_t = 0$, then the strength of

[11] Our main results extend to the case in which the benefit from being in power depends on the norm level N_t, e.g., bN_t. See Section 5.5.

[12] Our main results extend straightforwardly to the case in which the payoff from abusing is separable and decreasing in N_t and h.

[13] In the context of a private organization, we can think of this as depicting the possibility that the CEO can "capture" the board.

the norms increases by $\gamma \geq 0$. If, in contrast, the leader abuses her position, $a_t = 1$, then the strength of the norms decreases by γ. Thus, γ measures the short-run sensitivity of norms to behavior. The parameter $\delta \in (0, 1]$ is akin to a rate of depreciation in capital accumulation models. Lower δ implies a longer lasting impact of current actions on future norms. The first two terms, $(1-\delta)N_t + \delta\overline{N}$, have the effect of mean reversion to $\overline{N}$. This highlights the sense in which the formal written rules have a more persistent role. In fact, with the absence of the effect of the leaders' actions on norms (i.e., the third term), the norm level converges to $\overline{N}$ in the long run. Despite this force, however, we will demonstrate that the norm level may be absorbed into different levels when norms are affected by the leaders' actions.

Remark 1. Assume $\delta < 1$.

1. $N_t \in \left(\overline{N} - \frac{\gamma}{\delta}, \overline{N} + \frac{\gamma}{\delta}\right)$ for all $t \in \mathbb{N}$.
2. If $a_t = 0$, then $N_{t+1} > N_t$.
3. If $a_t = 1$, then $N_{t+1} < N_t$.

Remark 1 implies that $N_{t+1} < N_t$ if and only if $a_t = 1$.[14] We denote by $N_L := \overline{N} - \frac{\gamma}{\delta}$ and $N_H := \overline{N} + \frac{\gamma}{\delta}$.

For a leader of type h, if we denote the leader's strategy at time t by $a = (a_t, a_{t+1}, \dots)$, then the discounted value from following strategy a given norm level N_t is:

$$V(N_t, h \mid a) := \sum_{s=t}^{\infty} \beta^{s-t} \Pi_s u(a_s, N_s, h),$$

where $\beta \in (0, 1)$ is the leader's discount rate and Π_s denotes the probability that the leader is still in power in a given future period s. It can be defined recursively as:

$$\Pi_s := \begin{cases} 1 & \text{if } s = t \\ (1 - \lambda(a_{s-1}, N_{s-1}))\Pi_{s-1} & \text{if } s > t \end{cases}.$$

Lastly, if the leader gets replaced at the end of time t, then a new leader is drawn from a distribution with full support $H_t = \left[\underline{h}_t, \overline{h}_t\right]$. Although the evolution of H_t plays no role in determining the optimal actions of the leader at time t, it can have implications for the long-run properties of the institution. We will first consider the case $H_t = H$ for all t and postpone to Section 4.4 the case of an endogenous evolution of H_t.

[14] Generally, the main results of the paper continue to hold for the dynamics of norms specified differently from Equation (1) as long as $(N_t)_{t\in\mathbb{N}}$ satisfies the last two properties of Remark 1 (the first property, which states that N_t is bounded, would be needed to obtain the characterization of the norm dynamics in the same way as Theorem 2). Also, the results go through with any other initial value N_1 than $N_1 = \overline{N}$.

3 Main Analysis

We divide our main analysis into two subsections. Section 3.1 studies the optimal sequence of actions for a given leader facing a given level of honesty. In principle the leader could choose an arbitrary sequence of actions but, importantly, we are able to show that it is optimal for the leader not to switch from one action to another. This allows us to derive an explicit closed-form characterization for the cutoff type for a given norm level N which we denote $\tilde{h}(N)$ and also the leader's value function. With that important property established, Section 3.2 then studies the dynamics of the norms.

3.1 Characterization of a Leader's Decision

Consider a leader with honesty $h \in H$ when the norm level is $N \in (N_L, N_H)$. The leader's problem can be stated recursively by the following Bellman equation:

$$V(N,h) = \max_{a \in \{0,1\}} b - a(h+N) + \beta\left(1 - \lambda(a,N)\right) V(N',h)$$
$$\text{subject to } N' = (1-\delta)N + \delta\overline{N} + (1-2a)\gamma.$$

Note that the value function V that satisfies the above Bellman equation exists uniquely. To see this, the right-hand side of the Bellman equation is well-defined, as the maximum is taken over the binary values. Then, existence and uniqueness follow from the fact that the operation that defines the right-hand side of the Bellman equation is a contraction mapping, as the usual Blackwell conditions are satisfied.

To characterize the leaders' optimal actions, consider the effects of choosing abuse versus choosing respect. Firstly, the flow payoff changes. If the leader abuses at time t, then she gets an extra $-(N_t + h)$ flow payoff at t. Secondly, there are two additional effects on the continuation payoffs. First, by abusing, the probability of being in power in the next period changes from $1 - \lambda_0(N_t)$ to $1 - \lambda_1(N_t)$. Second, conditional on remaining in power, the continuation value changes from $V((1-\delta)N_t + \delta\overline{N} + \gamma, h)$ to $V((1-\delta)N_t + \delta\overline{N} - \gamma, h)$. Given these various effects, it is hard to solve for the optimal policy directly.

Instead, we rely on the property that, for any N_t, $N_{t+1} > N_t$ if and only if $a_t = 0$ (Remark 1) to make progress. This property implies that if there exists a non-increasing threshold function $\tilde{h}$ such that the policy function a^* satisfies the following: the leader of type h takes $a^*(N,h) = 1$ if $h < \tilde{h}(N)$ and $a^*(N,h) = 0$ if $h > \tilde{h}(N)$, then the optimal action sequence is constant. This follows because if it is optimal for the leader to abuse today, i.e., $h < \tilde{h}(N_t)$, then $N_{t+1} < N_t$ and thus $h < \tilde{h}(N_t) \leq \tilde{h}(N_{t+1})$. Hence, it is optimal for the leader to abuse tomorrow as well.

Next, we guess and verify that the threshold function $\tilde{h}$ is non-increasing. Given the conjecture, if $h < \tilde{h}(N)$, then the leader abuses the position forever, as $h < \tilde{h}(N) \leq \tilde{h}(N_t^1)$, where N_t^1 denotes the decreasing path of norms when $a = (1, 1, \dots)$ with $N_1^1 = N$. Thus,

$$V(N, h \mid (1, 1, \dots)) = \sum_{t=1}^{\infty} \beta^{t-1} \left(\prod_{s=1}^{t-1} (1 - \lambda_1(N_s^1)) \right) \left(b - \left(N_t^1 + h \right) \right).$$

On the other hand, if $h > \tilde{h}(N)$, then the leader respects forever, as $h > \tilde{h}(N) \geq \tilde{h}(N_t^0)$, where N_t^0 denotes the increasing path of norms when $a = (0, 0, \dots)$ with $N_1^0 = N$. Thus,

$$V(N, h \mid (0, 0, \dots)) = \sum_{t=1}^{\infty} \beta^{t-1} \left(\prod_{s=1}^{t-1} (1 - \lambda_0(N_s^0)) \right) b.$$

Then, the threshold function can be simply computed by solving for

$$V(N, \tilde{h}(N) \mid (0, 0, \dots)) = V(N, \tilde{h}(N) \mid (1, 1, \dots)). \tag{2}$$

This is because, if the leader's type is $\tilde{h}(N)$ when the norm level is N, then she is indifferent between abusing forever and respecting forever. On the one hand, since the replacement probability λ_0 is non-decreasing in N and the flow payoff is constant when $a = (0, 0, \dots)$, the left-hand side of Expression (2) is non-decreasing in N and does not depend on h. On the other hand, since the replacement probability λ_1 is non-increasing in N and the flow payoff are non-increasing in N and h when $a = (1, 1, \dots)$, the right-hand side of Expression (2) is non-increasing in N and h. Therefore, it must be the case that $\tilde{h}$ is non-increasing in N.

This allows us to obtain closed-form solutions for the value function in both cases and verify that indeed the implied optimal policy threshold function is non-increasing in N as conjectured. Formally we have:

Theorem 1. *The leader's optimal action is constant over time. For any given $N \in (N_L, N_H)$, there exists $\tilde{h}(N) \in \mathbb{R}$ such that if $h < \tilde{h}(N)$ the leader abuses her position and if $h > \tilde{h}(N)$ the leader respects the rules. The threshold $\tilde{h}(N)$ is non-increasing in N and is given by:*

$$\begin{aligned} \tilde{h}(N) = & \left(1 - \frac{\sum_{t=1}^{\infty} \beta^{t-1} \left(\prod_{s=1}^{t-1} (1 - \lambda_0(N_s^0)) \right)}{\sum_{t=1}^{\infty} \beta^{t-1} \left(\prod_{s=1}^{t-1} (1 - \lambda_1(N_s^1)) \right)} \right) b - N_L \\ & - \frac{\sum_{t=1}^{\infty} (\beta(1-\delta))^{t-1} \left(\prod_{s=1}^{t-1} (1 - \lambda_1(N_s^1)) \right)}{\sum_{t=1}^{\infty} \beta^{t-1} \left(\prod_{s=1}^{t-1} (1 - \lambda_1(N_s^1)) \right)} (N - N_L), \end{aligned} \tag{3}$$

and the value function satisfies:

$$V(N,h) = \begin{cases} \sum_{t=1}^{\infty} \beta^{t-1} \left(\prod_{s=1}^{t-1} (1 - \lambda_0(N_s^0)) \right) b & \text{if } h \geq \tilde{h}(N) \\ \sum_{t=1}^{\infty} \beta^{t-1} \left(\prod_{s=1}^{t-1} (1 - \lambda_1(N_s^1)) \right) (b - (N_t^1 + h)) & \text{if } h \leq \tilde{h}(N) \end{cases},$$

where N_t^0 denotes the increasing path of norms when $a = (0, 0, \dots)$, $N_{t+1}^0 = (1-\delta)N_t^0 + \delta\overline{N} + \gamma$, with $N_1^0 = N$; and N_t^1 the decreasing path of norms when $a = (1, 1, \dots)$, $N_{t+1}^1 = (1-\delta)N_t^1 + \delta\overline{N} - \gamma$, with $N_1^1 = N$.

If the replacement probability λ is constant, then the threshold function $\tilde{h}$ reduces to a simple linear equation.

When there are term limits, we can characterize the optimal action sequence by backward induction. As we will discuss in Section 4.2, it may no longer be the case that the optimal action sequence is constant. This is because the problem is not stationary. In particular, the leader is less deterred from abusing in the last period since there is no impact of her action on the replacement probability.

3.2 Dynamics of Norms

We now study the dynamics of norms. We assume that the set from which leader types are drawn, H, is a compact interval $H = [\underline{h}, \overline{h}]$ and that, when a leader is replaced, the next leader's type is drawn (independently of histories) from a distribution F_H with full support H. At time $t = 1$, the norm level starts with $N_1 = \overline{N}$. A leader with type h_1 is chosen according to the distribution F_H. Then, the leader makes her decision $a_1 = a^*(\overline{N}, h_1)$, which leads to the norm level $N_2 = \overline{N} + (1 - 2a_1)\gamma$ at the beginning of the next period. In period $t \geq 2$, with probability $1 - \lambda(a_{t-1}, N_{t-1})$, the incumbent stays in power: $h_t = h_{t-1}$. Otherwise, with probability $\lambda(a_{t-1}, N_{t-1})$, a new leader with type $h_t \in H$ is chosen. In either case, the leader at time t takes $a^*(N_t, h_t)$, which determines the level of norm N_{t+1} at the beginning of the next period and so on. For ease of presentation, this subsection assumes $\lambda(\cdot, \cdot) \in (0, 1)$.[15] Below we characterize the long-run properties of norms.

To highlight the importance of the endogenous norms, we first discuss the case in which norms are constant, i.e., $N_t = \overline{N}$. This is the case when $(\delta, \gamma) = (1, 0)$. In this case, there are three possibilities: (i) if $\tilde{h}(\overline{N}) > \overline{h}$ then all types want to abuse power and that is the only outcome observed; (ii) if $\tilde{h}(\overline{N}) < \underline{h}$ then no type abuses power and rules are always respected; and (iii) $\overline{h} > \tilde{h}(\overline{N}) > \underline{h}$ then there is a subset of types that would abuse power and a subset that wouldn't. As a result, we will observe transitions from abuse to respect

[15] Theorem 2 can be easily modified when we allow for $\lambda(\cdot, \cdot) \in \{0, 1\}$.

and vice versa as the type of a leader changes.

Importantly, note that with constant norms it is not possible for two countries to have very different long-run outcomes if they start with the same initial condition. In contrast, when norms evolve endogenously this arises as a possibility. To see this, consider a situation as in (iii) above with $\overline{h} > \tilde{h}(\overline{N}) > \underline{h}$. Now suppose in one country the initial sequence of elected leaders has $h_1 > \tilde{h}(\overline{N})$ and thus no abuse takes place. This implies that the norm gets stronger and as a result, the cutoff type decreases $\tilde{h}(\overline{N}) > \tilde{h}(N_1) > \tilde{h}(N_2) \ldots$. If the string of good leaders is sustained sufficiently long, then we might reach a point in which $\tilde{h}(N_t) < \underline{h}$ and, at this point, the norms are so strong that even if the worst possible leader is elected she will still respect the rules. As a result, norms will just keep getting stronger and $N_t \to N_H$ and the rules will always be respected from then on. Yet, for the same initial condition, the opposite might also be possible. A draw of bad leaders early on, who choose to abuse the norms, can lead the norms to weaken to a point at which $\tilde{h}(N_t) > \overline{h}$. If that happens, from that point on, not even the best possible leader would respect the rules. Thus, rules are never again respected and norms just keep on drifting down: $N_t \to N_L$. Thus, we can have two very different absorbing steady states.

Our main result formalizes this discussion. For ease of notation, denote by $\tilde{h}(N_L) := \lim_{N \downarrow N_L} \tilde{h}(N)$ and $\tilde{h}(N_H) := \lim_{N \uparrow N_H} \tilde{h}(N)$.

Theorem 2. *The long-run norm dynamics are characterized by the following four cases.*

1. *If (i) $\underline{h} < \tilde{h}(N_H)$ and (ii) $\overline{h} < \tilde{h}(N_L)$, then $N_t \downarrow N_L$ almost surely along any path.*
2. *If (i) $\underline{h} > \tilde{h}(N_H)$ and (ii) $\overline{h} > \tilde{h}(N_L)$, then $N_t \uparrow N_H$ almost surely along any path.*
3. *If (i) $\underline{h} < \tilde{h}(N_H)$ and (ii) $\overline{h} > \tilde{h}(N_L)$, then there exists a full-support limit distribution on $N_\infty \in (N_L, N_H)$.*
4. *If (i) $\underline{h} > \tilde{h}(N_H)$ and (ii) $\overline{h} < \tilde{h}(N_L)$, then almost surely along any path, either $N_t \downarrow N_L$ or $N_t \uparrow N_H$. There exists a limit distribution on $N_\infty \in \{N_L, N_H\}$.*

In Case 1, depicted in the top left panel of Figure 1, almost surely along any path, the level of norm converges to the lowest level. Put differently, the leaders' actions satisfy $a_t = 1$ for all but finitely many times. For this to be the case we need two conditions to hold: (i) for any norm level, there are some types who want to abuse the position; and (ii) once the norm level is sufficiently low (i.e., below N_* in the figure), even the most honest type wants to abuse. The first condition implies there is a path that takes us to lowest norm with positive probability and the second that once that point is reached it is absorbing.

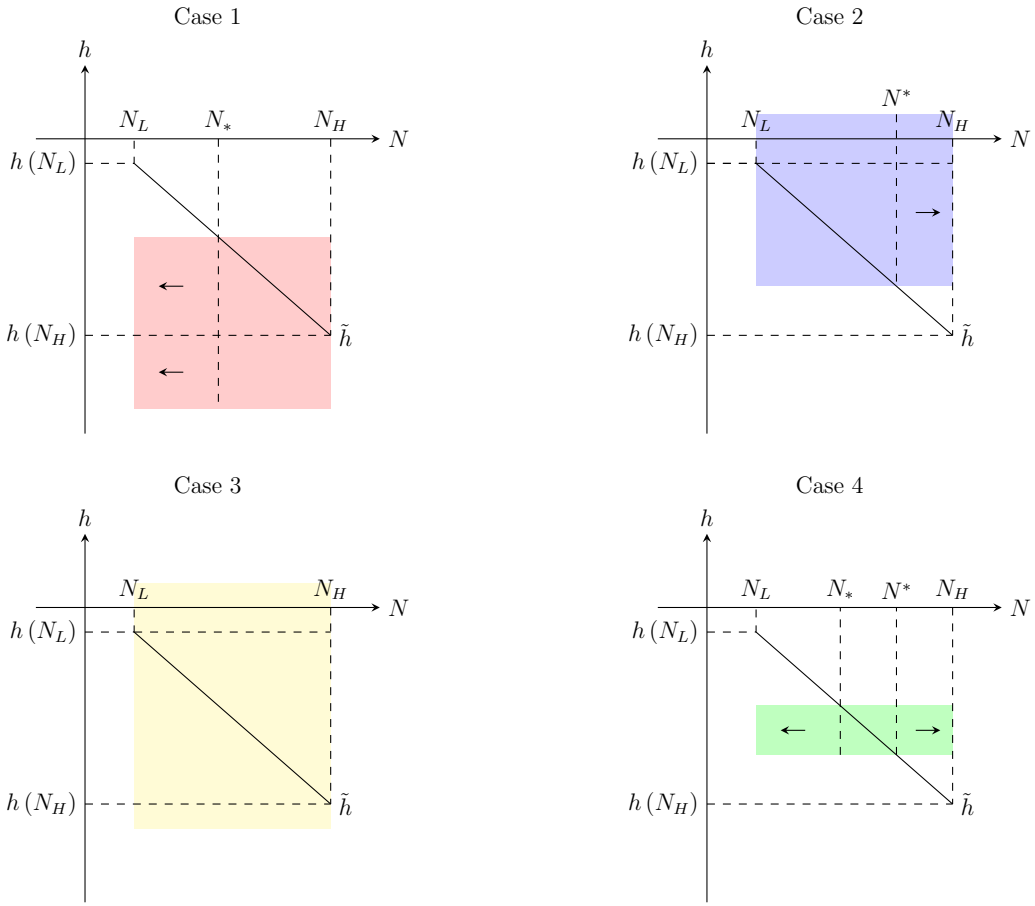

Figure 1: Theorem 2. The top left panel depicts Case 1; the top right panel Case 2; the bottom left panel Case 3; and the bottom right panel Case 4.

In contrast, in Case 2, depicted in the top right panel of Figure 1, almost surely along any path, the level of norm converges to the highest level. The leaders' actions satisfy $a_t = 0$ for all but finitely many times. The conditions for this are the exact opposite to Case 1: in words, (i) there must always be a type willing to respect the rules; and (ii) once the norms are sufficiently strong (i.e., above N^* in the figure), no type wants to abuse.

For a non-degenerate limiting distribution to exist, Case 3, depicted in the bottom left panel of Figure 1, it must be the case that (i) even when norms are strong there are always types willing to abuse; and that (ii) even when norms are weak there are always some types willing to respect the rules. In this case, the leaders take $a_t = 0$ and $a_t = 1$ infinitely often. In their study of democratic reversals, Kapstein and Converse (2008) document that, among those democracies that were reversed, some such as Pakistan, Peru, and Thailand, experienced second and even third democratization episodes.

Case 4, depicted in the bottom right panel of Figure 1, is perhaps the most interesting and highlights the importance of how leaders can shape the institutions. For this case to

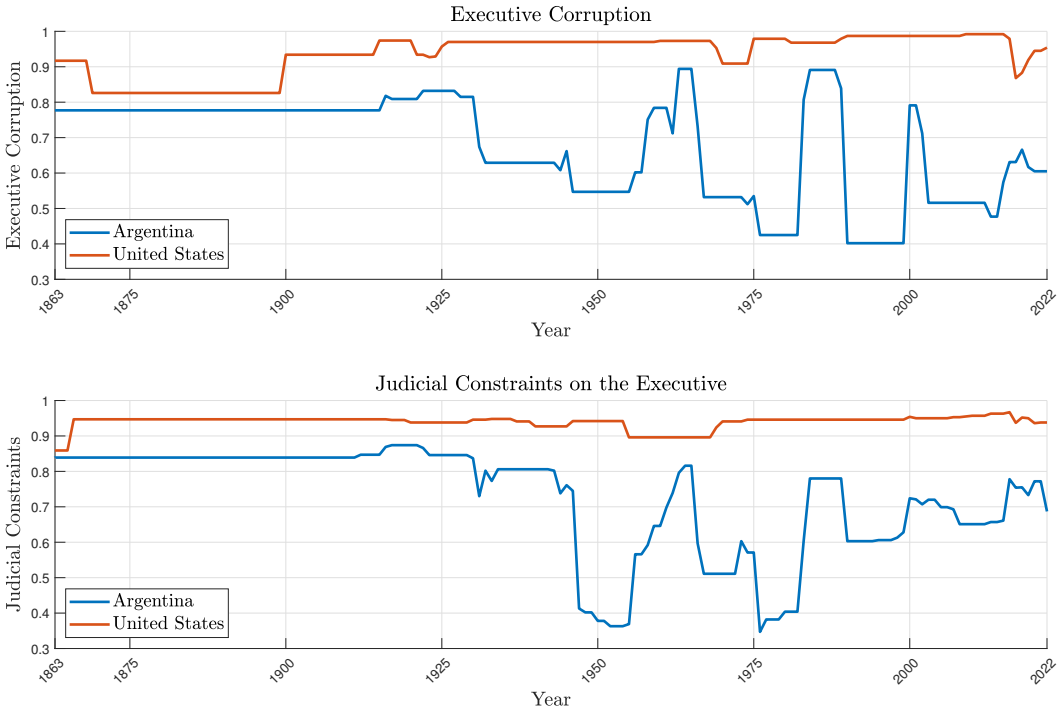

Figure 2: Path Dependence. The two panels depict Executive Corruption Index and Judicial Constraints on the Executive Index from V-Dem (Coppedge et al., 2023), respectively.

arise, (i) once the norms are sufficiently strong (i.e., above N^* in the figure), no type wants to abuse, and (ii) once the norm level is sufficiently low (i.e., below N_* in the figure), even the most honest type wants to abuse. In this case, almost surely along any path, the norm level either converges to N_L or converges to N_H. Put differently, almost surely along any path, either $a_t = 1$ for all but finitely many times or $a_t = 0$ for all but finitely many times. This happens because a string of very honest leaders, who respect the rules, can raise through their actions the norm level to a sufficiently high point such that when this level is reached, a leader, however low her type is, never abuses the norms. Conversely, if a sequence of bad leaders abuse the norms, then norms might become so weak that even if a better leader is elected she will still be tempted to abuse the norms.

Figure 2 illustrates how, while both started with similar scores on both measures and similar institutions, Argentina and the United States have followed very different paths in terms of executive corruption and judicial constraints on the executive.[16] The US has slightly improved over time. Instead Argentina has followed a more volatile path and seems to have converged to a lower steady state. The two measures naturally move together since

[16]The starting point in the figure is 1863, ten years after the ratification of the Constitution of Argentina, which was modeled after the United States Constitution.

the executive has an incentive to undermine the courts as a way to be able to get away with their illegal actions. They also have incentives to remain in power to make use of the immunity it grants. For example, the current Vice-President and former President Cristina Fernández de Kirchner has been found guilty of corruption and sentenced to 6 years in prison. Her political immunity has so far kept her out from jail. Her government has tried very hard to manipulate the courts to be able to steer her legal proceedings.

Theorem 2 speaks to the persistent effect that early leaders can have on institutions or the culture of organizations. Thus, young organizations must devote extra care in the selection of their leaders. In the political context, Keefer (2007) and Kapstein and Converse (2008) document that young democracies are especially at risk of reversal, and suggest that the absence of checks and balances such as political competition is among the most powerful predictors of democratic failure. In the context of our model, a young democracy whose norm level is between N_* and N^* is indeed at risk of reversal and the democratic failure may occur as the level of norms is deteriorated. The absence of checks and balances corresponds to a lower λ_1, under which leaders tend to abuse more and norms are more likely to deteriorate.

Evidence consistent with the particularly important role that early leadership can have in shaping institutions is presented by Simons (1994) in the corporate context. His longitudinal study suggests that new CEOs use their first 18 months to define and measure critical performance variables and to overcome organizational inertia. It is also consistent with the fact that young or new institutions and organizations make important investments to try to select the right leadership. Indeed, our model suggests that, to the extent that institutions can spend resources to screen for leader types, it is particularly crucial that they do so in their infancy.

To highlight path dependence further, it is important to reiterate that Case 4 only obtains when institutions, i.e., norms, endogenously respond to the leaders' actions. If we let $(\delta, \gamma) = (1, 0)$, then we get $N_t = \overline{N}$ for all t. In this case, Cases 1-3 are still possible but not Case 4. In fact, if we were in Case 4 before and we made $(\delta, \gamma) = (1, 0)$ then we would find ourselves in Case 3.

Corollary 1. *If $(\delta, \gamma) = (1, 0)$, then an optimal action sequence cannot converge to two constant actions, i.e., Case 4 is not possible.*

Thus, endogenous norm formation provides us with a way to conceive how two institutions with seemingly equal formal rules can converge to two very different steady states.

Figure 3 depicts the dynamics of norms.[17] The left panel depicts Case 1: along each path, eventually the norm level converges to the lowest level. The central panel depicts Case 3, in

[17]For our numerical simulations, we discretize $H = \{h^1, \ldots, h^n\}$ with $h^1 = \underline{h}$ and $h^n = \overline{h}$ and assume that F_H is a uniform distribution.

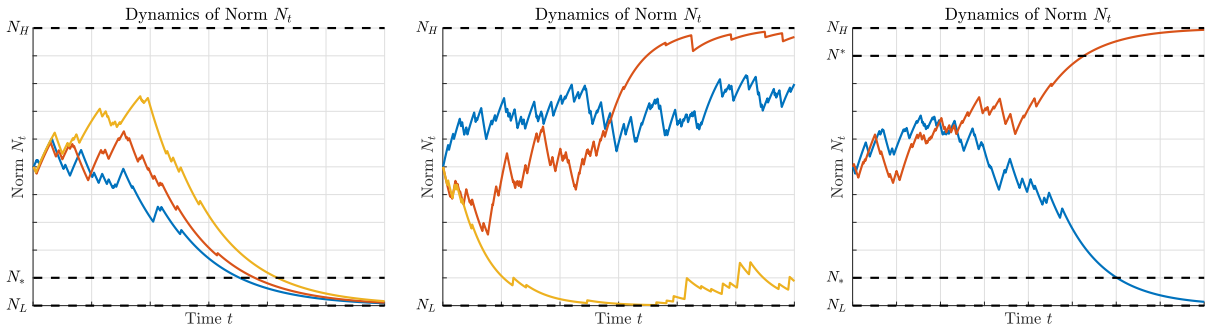

Figure 3: Dynamics of Norm. The left panel depicts Case 1: along any path, eventually the norm level converges to the lowest level. The central panel depicts Case 3: there exists a stationary distribution on N_∞. The right panel depicts Case 4: with the same initial norm level, the norm level converges to either the lowest level or the highest level.

which the norm level remains ergodic. Note, however, that for a significant amount of time, the norm level is close to one of the extremes. This is because once the norm reaches such a level, the mass of types that switch the action relative to their predecessor, even though it is positive, is relatively small. In addition, $\lambda_1(N_L)$, the replacement probability after abuse at N_L, is small.[18]

The right panel depicts Case 4: starting from the initial formal rules, an institution can converge to two very different steady states. Distinguishing Case 3 from Case 4 might not be empirically very easy since, in practice, the type space could have long but thin tails. Yet it is important that in both cases the key is that we will have path dependence and persistence close to the extremes. Figure 4 depicts the simulated long-run distribution of norms for Case 3. Beyond the fact that most of the density is in the extremes, it is worth noting that this distribution is likely to be skewed towards the left. This is the case because when norms are very low regimes can last longer. Instead, when institutions are strong, turnover is higher.

Mainwaring and Bizzarro (2019) document 91 transitions (in 79 countries) to democracy from 1974 to 2012. Consistently with our theory, their result suggests extreme long-run outcomes of either full breakdown or restoration. Namely, out of these 91 cases of "third-wave" democratization, 62 experienced either breakdowns or stagnation at a low level; 27 either achieved major democratic advances or attained high levels of democracy from their first year of democracy to 2017; and in 2 (Ecuador and Poland), levels of democracy eroded substantially while the regime remained a democracy per their classification.

[18] Our discussion in Section 4.1.1 is also closely related to this point.

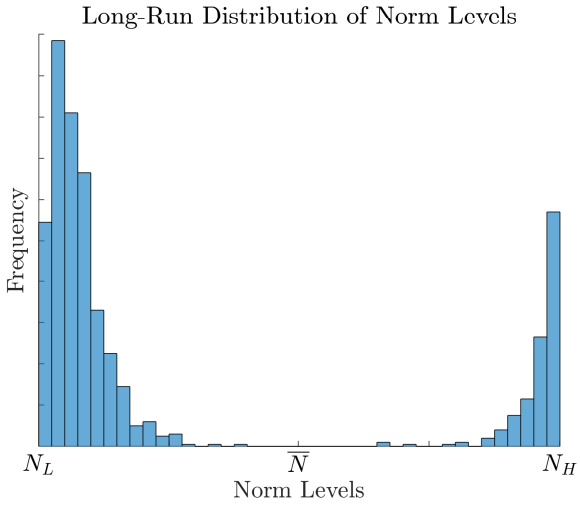

Figure 4: Simulated Long-Run Distribution of Norm Levels for Case 3.

4 Discussions

This section discusses democratic backsliding as an important application of our model. It also discusses term limits. Finally, this section provides comparative statics results and discusses endogenous leader types.

4.1 Democratic Backsliding

As discussed in the law and political science literature, many autocracies are the result of a slow erosion of institutions rather than a rapid wholesale shift (e.g., Huq and Ginsburg, 2018). As illustrated in Figure 5, recent examples include Cháves in Venezuela, Erdogan in Turkey, Orbán in Hungary, and Putin in Russia. High-income countries and older democracies may also experience a slow erosion of institutions, even if they do not transition into an autocracy.

For this, it is useful to consider the "abuse" action as including ones such as replacing key figures that might play an important role in limiting the leader's power. In the political arena, three relevant examples are: (i) changing the composition of the courts, for example by expanding the supreme court; (ii) changing the people in charge of running/supervising elections from honest brokers to puppets; and (iii) manipulating public debate through media. Former President Trump's attempts to overturn the 2020 election were to a large extent derailed by Department of Justice leaders that were unwilling to do his bidding. Indeed, Trump has strongly endorsed many candidates in the 2022 election largely on the basis that they denied the outcome of the 2020 election. Under Prime Minister Orbán in Hungary, "elections rules have been modified 20 times, paralysing opposition parties; and

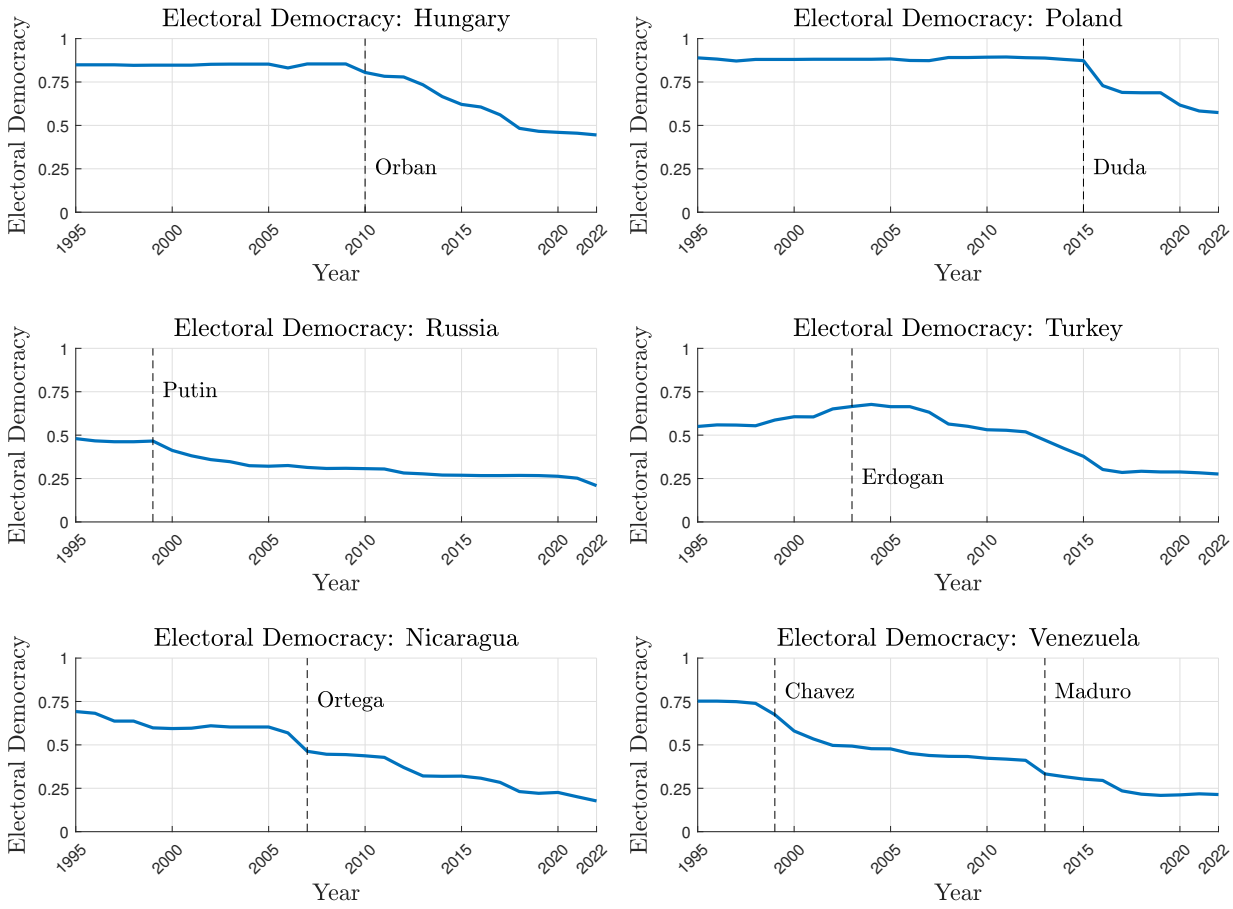

Figure 5: Democratic Backsliding. Each panel depicts Electoral Democracy Index from V-Dem (Coppedge et al., 2023).

Fidesz has heaped pressure on the independent judiciary" (Szelényi, 2022).

A common justification for the consolidation of power is the existence of internal or external threats. As Napoleon Bonaparte reflected during his exile,

> "When I acquired the supreme direction of affairs, it was wished that I might become a Washington.... Had I been in America, I would willingly have been a Washington, and I should have had little merit in so being. But had Washington been in France, exposed to discord within, and invasion from without, I would have defied him to have been what he was in America; at least, he would have been a fool to attempt it, and would only have prolonged the existence of evil. For my own part, I could only have been a Crowned Washington. It was only in a congress of kings, in the midst of kings, yielding or subdued, that I could become so. Then and there alone, I could successfully display Washington's moderation, disinterestedness, and wisdom. I could not reasonably attain to this but by means of the universal Dictatorship. To this I aspired; can that be thought a crime?"

(Las Cases, 2017 [1855])

In the corporate world, the analog to democratic backsliding is known as board capturing: the CEO being able to influence the composition of the board to people she can control. This board will fail to provide proper oversight and control over the CEO's actions.

We can capture these processes in our model if the replacement probability $\lambda_1(N)$ decreases as the leader undermines the institutional safeguards piece by piece. Such piecemeal subversion of norms is less visible and attracts less resistance than a wholesale shift such as a coup. As in Venezuela, Turkey, Hungary, and Russia, many leaders who have subverted democratic norms have retained the support of a majority or a ruling coalition through several election cycles. In the limit, we could reach the situation in which the leader consolidates herself as an autocrat and is never replaced, i.e., $\lambda_1(N) = 0$, once the norms are sufficiently eroded. In this sense, while Hungary and Poland have suffered a significant erosion of their democratic institutions their leaders are not quite as cemented in power as those of Russia, Nicaragua, and Venezuela. Finally, it is worth noting that even if the economy has experienced a long history of high norm levels, such a process of democratic backsliding can occur in our model. This speaks to the current concerns of what a return of Trump to power can entail for American institutions. As Kagan (2023) puts it, "There is a clear path to dictatorship in the United States, and it is getting shorter every day. In 13 weeks, Donald Trump will have locked up the Republican nomination."

4.1.1 Restoration of Democratic Practices

Once an economy has fallen into a despotic regime rather than $\lambda_1(N) = 0$, we might think that there is still a very small probability of replacing the current leader. In this case, it is natural to think that the set of possible replacement types would also differ. For ease of presentation, consider $H_t = \left\{h_{t-1}, h^{\mathrm{h}}\right\}$, where h_{t-1} can represent the despot replaced by a family member (as in North Korea) or a political rival that would continue with the current practices (as when one war lord deposes another). Instead, h^{h} represents a hero type that is willing to potentially risk her life to depose the current leader. There are several historic figures that we might associate with such a type. Furthermore, assume h^{h} is such that this type would not cheat once in power. This would give institutions a chance to recover and reestablish the necessary checks and balances for a proper functioning of democracy.

This, of course, is not easy and can help explain the difficulty in restoring democratic practices in former autocratic regimes. Since institutions are very weak, the temptation for a new leader to abuse is very high. This is particularly hard when such a heroic figure is absent or replaced too soon, such as evidenced recently in Egypt, Libya, and Yemen in the political

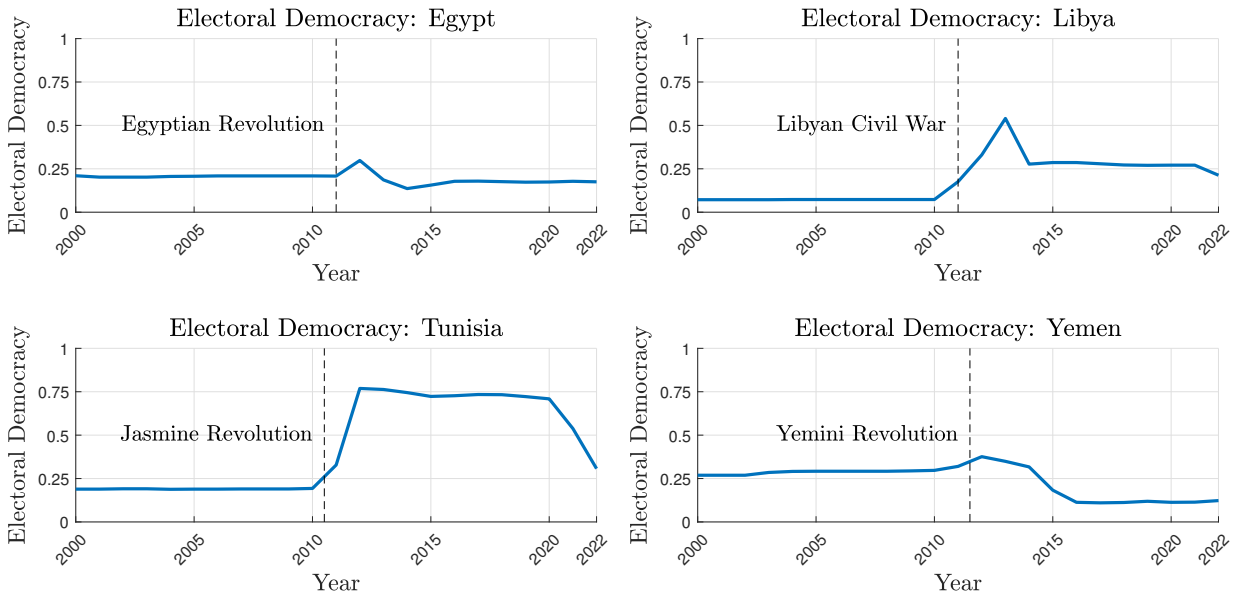

Figure 6: Restoration of Democratic Practices. Each panel depicts Electoral Democracy Index from V-Dem (Coppedge et al., 2023).

context. This might also help explain why regime changes from the outside tend to fail. Diamond (2021) presents a list of 20 countries where mass public protests or an unexpected defeat of an authoritarian incumbent might have resulted in a transition to democracy for the period of 2009-2020. At the time the paper was written, only 2 out of 20 countries in the list (one of which is Tunisia to be discussed below) had so far resulted in democratic transitions.[19]

Figure 6 illustrates several examples of these failed attempts at restoring democratic practices. A particularly interesting case to highlight is Tunisia which until recently it seemed to have managed to succeed. Former Prime Minister and President Essebsi played an important role on that initial success. Unfortunately, upon his death on 2019, he was succeeded by Kais Saied who in 2021, dismissed the parliament and carried out a self-coup. He has ruled by decree and passed a new constitution since then.

4.2 Term Limits

For analytical convenience, we focused on a stationary model with no explicit term limits. We consider here the role of term limits: there exists a time T such that a leader will be

[19]When the norm level is low, the replacement probability λ_0 after respect may be high. In the corporate context, Michael Woodford was ousted within two weeks of being promoted to CEO of Olympus. This happened when he raised questions about a series of irregular acquisition payouts, which turned out to be known as the Olympus scandal.

replaced for sure if she had served for T periods.[20] First, constant actions may no longer be optimal. In particular, a term limit may encourage a leader to switch her action from respect to abuse toward the end of the term.[21] This is likely to arise when the benefit b is sufficiently high, the leader's type is low, and the replacement probabilities satisfy $\lambda_1(N) - \lambda_0(N) > 0$. Consider the two-period model. In this case, abusing in the first period is costly because of the loss of b in the second period. In the second period, the effect of λ is irrelevant, and thus the leader would take a myopically best action.

Second, the following remark states that extending the term may have opposing effects.

Remark 2. Consider an extension from two to three periods. On the one hand, not to lose the benefit b of being in office, a leader may go from $(a_1, a_2) = (0, 1)$ to $(a_1, a_2, a_3) = (0, 0, 1)$. On the other hand, the leader may have an incentive to undermine the norms earlier since now she can reap the benefits from abusing longer. Thus, the leader may go from $(a_1, a_2) = (0, 1)$ to $(a_1, a_2, a_3) = (1, 1, 1)$.

Furthermore, the effect that a leader may abuse her position at the end of her term may affect the dynamics characterization. For Case 2, while the norm level still converges towards the upper bound, for some parameters the leaders would abuse at the last period of their term, introducing a momentary reduction in norms. For Case 4, term limits generate a regression-towards-the-mean effect in terms of types and norms, and the speed of convergence may become slower. Furthermore, it increases the likelihood of the convergence to the steady state of abuse. While these effects are interesting, they do not qualitatively change the conclusion of Theorem 2, and highlight the advantage of using the stationary model for our main analysis.

4.3 Comparative Statics

Our closed-form characterization of the threshold function $\tilde{h}$ (Equation (3)) allows us to perform comparative statics.[22]

A stronger set of formal rules (i.e., a higher $\overline{N}$) implies (i) that leaders would be more likely to respect the rules; and (ii) a higher probability of getting absorbed into the steady state of respect.[23] Although this clearly suggests we would want to start with a strong

[20] The origin of term limits dates back to Aristotle: "no office should ever be held twice by the same person" (Barker, 1946, p. 258).

[21] Ferraz and Finan (2011) empirically study the effect of a term limit on corruption using a dataset from an anti-corruption audit program by the Brazilian government.

[22] Sharper results can be obtained when the replacement probability λ_0 after respect does not depend on the norm level.

[23] In our model, there is no countervailing force favoring a weaker set of formal rules. One could imagine

set of formal rules, this is not easy. Determining what a good set of formal rules is from observed outcomes is hard since it requires to condition on the sequence of leader types. This is particularly the case with governments. For example, while the US constitutional framework is usually regarded as being strong, there have been many examples of countries adopting very similar frameworks yet experiencing very different outcomes. In the corporate setting, the existence of a large number of firms simultaneously coexisting with different governance provisions and switches of CEOs across firms allow for some insights into what constitutes good corporate governance.[24]

In the political setting, an increase in δ, which governs the reversion to $\overline{N}$, may be construed as conferring less flexibility to the interpretation of the constitution and thus allowing less room for the role of informal rules. When the leaders abuse the position, the norm level decreases more slowly from the initial level. Thus, leaders are less able to influence their future replacement probability and their flow payoff, while the replacement probability λ_0 after respect stays the same. Hence, they are less likely to abuse the position. Thus, conferring less flexibility to the interpretation of the constitution may deter democratic backsliding.

When norms are more malleable, which corresponds to a higher γ, leaders are able to decrease the replacement probability and increase the flow payoff in the future, while the replacement probability λ_0 after respect stays the same. Thus, they have more incentives to abuse their position.

In the political setting, one can interpret λ_1 as the scrutiny of media, political competition, or the independence of the supreme court. In the corporate setting, one can interpret λ_1 as the independence of the corporate board or the strength of the minority shareholder rights. As oversight increases, the likelihood of abuse decreases. See also our discussion in Section 5.1.

Next, we consider the effect of the benefit b of being in office. It can be seen from Equation (3) that the effect of b on the leader's behavior depends on the replacement probability λ. Thus, for simplicity, let us focus on the case in which each $\lambda(a_t, \cdot)$ is constant and $\lambda_1 > \lambda_0$: the abuse action is more likely to lead to losing the position. Then, an increase in b leads to a decrease in the threshold $\tilde{h}(N)$, meaning that the leader is more likely to respect the institution. In other words, the leader is more likely to respect the institution for the "re-election" motives. In the political context, this comparative statics result is consistent with

that, when flexibility is valuable, a stronger set of rules may hinder it and backfire. Thus, an inter optimum level of formal rules might arise. See, for example, Gratton and Lee (Forthcoming) and Invernizzi and Ting (Forthcoming).

[24]See, for instance, Shleifer and Vishny (1997), Gompers et al. (2003), Bebchuk et al. (2009), and Adams et al. (2010).

the empirical findings of Ferraz and Finan (2010) and Gagliarducci and Nannicini (2013) that a salary increase for politicians improves political performance, although they also point out another effect that a salary increase leads to a better selection of politicians.[25]

Next, we consider the effect of the discount factor β.[26] Let us suppose first that the leader's action has no effect on the replacement probability, i.e., $\lambda_1 = \lambda_0$. It is important to note that even in such a case, the leader's problem is not static because the future payoff from abusing is affected by its actions today. In particular, suppose that the leader is currently indifferent between abusing forever or always respecting. If we increase β, that would increase the benefit of abusing because the benefits of abusing are increasing over time due to the weakening of the norms while the payoff from respect is constant over time. If, in addition, $\lambda_1 \neq \lambda_0$, then there is a further consideration arising from the change in the replacement probability. If abusing lowers the replacement probability, $\lambda_1(N) < \lambda_0$, then this effect reinforces the leaders' incentive to abuse as we increase β. Instead, if respect lowers the replacement probability, $\lambda_1(N) > \lambda_0$, then there is a countervailing force. This effect can dominate when b is sufficiently large.

Thus, when the replacement probability is constant over time, an increase in β leads to a higher threshold. However, in general, the sign of the comparative statics with respect to β depends on a particular functional form of λ.

Finally, we could also consider the effect of distributions on H (for simplicity, assume $H = [\underline{h}, \overline{h}]$). When a distribution first-order stochastically dominates another, under the former distribution, the leaders are likely to be of a higher type and thus the long-run outcome is more likely to be an absorption into the high steady state. When it comes to second-order stochastic dominance, however, the effect of variance is not straightforward anymore, as for a weak institution, higher variance might give a higher chance for the reversal, i.e., the absorption into the high steady state.

4.4 Endogenous Leader Types

In the main analysis, we have assumed that the type distribution on H is constant over time and, in particular, independent of the history and the current norm level. It is natural to think that this might not be the case. For example, when the norm level is high, the internal process of emerging as a leader inside of a political party would favor higher types. In the opposite direction, when norms deteriorate significantly, those types more willing to cheat

[25] In our model, if H_t is endogenized and if a higher b leads to a "higher" H_t (in a set-theoretical sense or putting a larger mass on higher $h \in H_t$), then the leader is more likely to respect the institution as well.

[26] While we can interpret β literally as the discount factor, β may also be affected by the prevalence of political assassinations.

or use patronage to buy support are more likely to enter the political process or succeed at early stages and thus be more relevant, moving the distribution of types down. An example of this could be seen in the Republican party primaries for the 2022 mid-term elections. Many of the primary winners continued to question the 2020 presidential election. Thus, the lower the norm, the lower the probability that a potential new leader is of a higher type.

We denote the support of the distribution at time t by H_t. It is important to note that the endogeneity of H_t will not change the optimal response of the current leader. This implies that the only effects will be observed on the long-run properties of the institution. The endogeneity of H_t will give more "inertia" to the system: if the norm deteriorated from time t to $t+1$, then with an endogenous H_t, it would be more likely to continue deteriorating (and vice-versa for an improvement).[27]

If the change does not affect the support of the distributions, then Theorem 2 will continue to hold as stated. The only difference is that convergence will be faster for the cases with absorbing regions and for Case 3 with a long-run stationary distribution, we will observe more mass on the extremes of the long-run distribution. If the support moves, then, in addition, previous parametrizations that lead to having a long-run distribution (Case 3) will instead now fall into Case 4 in which the economy gets absorbed into either the high norm steady state or the low norm steady state. Thus, for a given legal framework, the early realization of its leader's types will have more important long-term consequences. Historians debate to what extent individuals play an outsize role in shaping outcomes relative to broad forces. In our model, both play a role. Yet, the possibility of Case 4 suggests that the relative importance is time dependence, where individuals play particularly important roles early on.

5 Extensions

Our parsimonious model allows for various extensions. First, we discuss the possibility to micro-found the replacement probability λ. We then discuss other possible extensions of our main model.

[27] When λ is micro-founded as the replacement probability at an election (see Section 5.1 for such discussions), this may justify the assumption that, as the norm level decreases, the replacement probability λ_1 decreases, as the candidates are more concentrated at lower honesty types. On a related point, as one mechanism through which democratic backsliding occurs, one of the main insights of Luo and Przeworski (2023) is that backsliding occurs when voters knowingly consent to democratic erosion when they find incumbent appealing. This would correspond to the dynamics in which the incumbent would keep being reelected and keep taking the abusive action.

5.1 Endogenizing Accountability and Leader Replacement

In our main analysis, the simple assumptions on the replacement probability λ allowed us to sharply characterize the decision of a leader and the norm dynamics. While we have not provided a particular micro-founded mechanism through which λ is derived, the simplicity and generality of λ in fact makes it possible to capture different environments through which the leader's replacement probability could be micro-founded. Specifically, we will briefly discuss how one could interpret the role of media, political patronage, or political competition through the lens of our model. The proper study of each of such mechanism itself would call for a separate paper. The aim of this subsection is to argue that such micro-foundations would be possible and an interesting avenue for future research.

5.1.1 Media

The role of media on the quality of democratic governance has been well-recognized. In a letter to James Currie, Thomas Jefferson eloquently wrote "Our liberty depends on the freedom of the press, and that cannot be limited without being lost." The First Amendment to the United States Constitution protects the freedom of press. The press has been recognized as the fourth estate or the fourth power. A strong and independent media is important to hold leaders accountable for their actions.

In our model, the stronger the media, the higher the likelihood that the leader would be criticized and lose the elections when taking an abusive action. Of course, leaders understand this, and they can take actions to limit the power of media. The independence of media can be eroded by granting access and other favors (particularly if they form part of larger business conglomerates) or by more direct threats and limitations to what they are allowed to publish. In the process of democratic backsliding, an autocratic leader would typically undermine the independence of media. In such contexts, it is not uncommon to see newspapers and other outlets shut down, journalists imprisoned, and forced take overs of media organizations. For example, after the 2016 failed coup attempt, Turkey has closed around 150 media organizations, including major newspapers such as Zaman, and has jailed around 160 journalists.[28] In another example, Nicaraguan journalist Carlos F. Chamorro in his 2023 Reuters Memorial Lecture reports: "This double-sided criminalisation of both freedom of the press and freedom of expression with the purpose of silencing journalists, news sources, and freedom of opinion, represents the latest stage in a long process of demolishing the rule

[28] https://www.bbc.com/news/world-europe-36910556 and https://www.reuters.com/article/us-turkey-security-newspaper/turkish-court-orders-release-of-journalists-during-their-trial-idUSKCN1GL2OR (Date of Access: June 12th, 2023).

of law in Nicaragua."[29]

As the media acts as a check on leaders' actions, we can interpret λ as capturing the independence of media. There are two ways in which media plays a role in the model.

In a static sense, suppose that leaders' actions do not affect the independence of media. Thus, λ can be interpreted as the degree of scrutiny of media. In this case, there are two effects of the independence of media on λ. First, as the degree of scrutiny increases, for any action and norm level (a_t, N_t), the replacement probability $\lambda(a_t, N_t)$ would be higher. Second, as the degree of scrutiny increases, for any norm level N_t, the replacement probability $\lambda_1(N_t)$ increases more than $\lambda_0(N_t)$, so that the difference $\lambda_1(N_t) - \lambda_0(N_t)$ is higher.

In a dynamic sense, suppose that a leader's abuse action is an action that undermines the independence and scrutiny of media. For instance, the abuse action may correspond to giving privileged access to media that supports the leader. In December 2015, the Poland's PiS party enacted a media law that required all broadcasters to have a board controlled by the government, while sidelining the constitutional body that is charged with guarding media independence (Fomina and Kucharczyk, 2016). In this second specification, as the norm level decreases, the media is less independent. Thus, as N_t decreases, $\lambda_1(N_t) - \lambda_0(N_t)$ decreases. Gratton and Lee (Forthcoming) consider a model in which there are two underlying states, where in one the population prefers an illiberal government and in the other they don't. The trade-off is that once an illiberal government is elected, they can manipulate the signal about the state thus increasing the chance they get reelected.

Our assumptions on λ are consistent with the properties of λ described under either specification. Note that we could consider the independence of judiciary similarly to that of media.

In the literature on the role of media on political accountability broadly construed, Besley and Prat (2006), Fearon (2011), and Guriev and Treisman (2020) indeed provide a model in which the more scrutiny disciplines politicians more.[30] When the independence of media is violated, the incumbent politician is more likely to be reelected. Our assumptions on λ_1 captures this effect.

[29]https://reutersinstitute.politics.ox.ac.uk/news/2023-reuters-memorial-lecture-how-report-under-dictatorship-lessons-nicaragua-and-beyond (Date of Access: June 12th, 2023).

[30]For surveys on theoretical and empirical work on media capture, see Prat and Strömberg (2013) and Strömberg (2015). The model of Besley and Prat (2006) can also be applied to a setting in which the agent who engages in media capture is not a government, such as a corporation. In the empirical literature, Dyck and Zingales (2004) and Dyck et al. (2008) study the role of the media in corporate governance.

5.1.2 Political Patronage

Politicians can distort state resources to reward wealthy voters and interest groups for their electoral support. When the incumbent leader is less constrained to distort state resources to get political support, the incumbent would be more likely to be reelected (see, for instance, Wantchekon, 2013).

Within the context of our model, this can rationalize why, when the norms N_t are weak, the probability of replacement when the politician engages in excessive patronage ($a_t = 1$) could be smaller than when it doesn't: $\lambda_1(N_t) < \lambda_0(N_t)$. Acemoglu et al. (2004), for instance, study a model explaining the survival of a kleptocracy under a weak-institutionalized state such as the Democratic Republic of the Congo (Zaire) and the Dominican Republic. Also, most regimes that drift towards autocracy are characterized by the existence of an important elite or oligarchy that benefits from the regime in exchange for their help in sustaining it. This symbiotic relationship is an extreme form of patronage.[31]

5.1.3 Political Competition

Competition among political parties is vital to democracy, as it creates a system of checks and balances. The institutional arrangements that protect and respect political opponents, the fairness of elections, and the freedom of associations are important elements of a well-functioning democracy. Like the role of media, when there are strong alternative candidates, if the incumbent abuses her position, the opposition would have a higher chance to remove her from power. In the context of our model, this would imply that for high N_t, $\lambda_1(N_t)$ is large. Thus, in equilibrium, the incumbent would be more inclined to respect the rules.

In cementing their regimes, autocratic leaders work hard to handicap or eliminate the potential competition. Russia presents a number of examples in this respect. Mikhail Borisovich Khodorkovsky was a wealthy and successful oligarch until he dared start Open Russia, an organization seeking to "build and strengthen civil society." He was swiftly charged, stripped of most of his assets and sent to prison for ten years. More recently Alexei Navalny, an anti-corruption and emerging politician, was first poisoned and then imprisoned after a trial that was described as a sham by Amnesty International. In our model, this is captured by the property that when N_t is low, the leader can abuse the rule of law to cement her grip on power. Also, understanding this, the leader has an incentive to progressively undermine the judiciary. When the political competition is suppressed by the incumbent, the incumbent's reelection probability would be higher. Our assumption on λ_1 captures this effect.[32]

[31] See Brancati (2014) for a survey on patronage distribution/clientelism in autocratic regimes.

[32] These actions also would affect the set H which, as we discussed in Section 4.4, can also evolve endogenously.

In his model of self-enforcing democracy, Fearon (2011) demonstrates that an unfair "playing field" for party competition can undermine the coordinated threat by the citizens and thus poses a greater challenge to democracy.

A more direct way to undermine political competition is by outright electoral fraud or by creating an unfair playing field. Again, the weaker the institutions, the higher the benefits and lower the costs of engaging in such activities which imply the same parametric assumptions.

It is worth noting that the type of the incumbent or competitor need not be observable. Thus, it is not obvious that competition would naturally select higher types. Also, as discussed above, those interested in receiving patronage might prefer to support a less scrupulous type.

5.2 Evolving Formal Rules

In our base model, formal rules have two roles. First, formal rules $\overline{N}$ determine the initial norm level N_1 of an institution. Second, formal rules possess the reversion-to-the-mean effect on the evolution of institutional norms through Expression (1). For the first role, our main results would not change when the initial norm level of the institution is given arbitrarily by $N_1 \in (\overline{N} - \frac{\gamma}{\delta}, \overline{N} + \frac{\gamma}{\delta})$.

For the second role, we have shown that even if the formal rules have a persistent effect on norms, institutions may exhibit path dependence. Here, we consider the case in which formal rules may change over time: $(\overline{N}_t)_{t\in\mathbb{N}}$. While it is an interesting research avenue to study the evolution of formal rules themselves and their effects on the long-run dynamics of norms, here we provide a simple and general extension under which our main results hold.[33] To that end, we first assume that $(\overline{N}_t)_{t\in\mathbb{N}}$ is bounded. When the variation is small enough, our main results would straightforwardly hold. More generally when such an assumption is not made, instead of providing a specific form for the evolution of formal rules, we make two additional assumptions on them: (i) there exists $\overline{N}^*$ such that if $\overline{N}_t \geq \overline{N}^*$ for some $t \in \mathbb{N}$ then $\overline{N}_{t+1} \geq \overline{N}^*$; and (ii) there exists $\overline{N}_*$ such that if $\overline{N}_t \leq \overline{N}_*$ for some $t \in \mathbb{N}$ then $\overline{N}_{t+1} \leq \overline{N}_*$. The first assumption states that there exists a level of formal rules such that if the level of formal rules is at least as high as the threshold level, then the level of formal rules is never below the threshold in the future. That is, we assume that there exists a level of formal rules which can sustain its level. The second assumption is the opposite. When there exists a level of formal rules such that if the level of formal rules is at least as low as the threshold level, then the level of formal rules is never above the threshold in the future.

[33] For the studies of the evolution of legal rules, see, for instance, Gennaioli and Shleifer (2007) and Niblett et al. (2010).

That is, we assume that there exists a level of formal rules which can never improve upon its level. Under these three assumptions on $(\overline{N}_t)_{t\in\mathbb{N}}$, the long-run evolution of norms (Theorem 2) would still hold.

Note that the assumptions allow for an intermediate region in which formal rules can react to abuses. For example, after the norm on term limits was broken by Presient Roosevelt in 1940, the Twenty-Second Amendment to the United States Constitution was introduced in 1951 to limit a President to two terms.

Conversely, autocrats can change the formal rules to allow for indefinite tenure. For example, in Venezuela, Chávez managed to abolish term limits in 2009. It is worth noting that Chávez illegally used the resources of the State to accomplish his goals. According to opposition figure Leopoldo López, "we aren't competing against a political party, we're competing against an entire state and all of the power it can wield" (Walter and Cancel, 2009).

5.3 Richer Action Sets

While we focused on the simplest case in which a leader faces binary actions, we can consider extensions in which the effect of a leader's actions on the norm is more nuanced. We consider two ways in which one can extend our baseline model.

First, there is a degree to which a leader's action affects the norm. Thus, the leaders' available actions are in $[0, 1]$, where a_t denotes the degree to which the leader abuses the position. The norm dynamics would be given by the same equation (i.e., Equation (1)). The characterization of the leader's decision then would depend on the assumptions on the replacement probability $\lambda(a_t, N_t)$. There are cases in which the decision of the leader is still characterized by Theorem 1. For instance, the degree of abuse would not affect the replacement probability for any $a_t > 0$. Formally, (i) $\lambda(a_t, N_t) = \lambda_1(N_t)$ for all $a_t \in (0, 1]$ and it is non-decreasing in N_t; and (ii) $\lambda_0(N_t)$ is non-increasing in N_t as in our baseline model. Otherwise, there would also be cases in which the leader gradually increases the degree of abuse (the size of a_t) as the norm level decreases, the mechanism of which is somewhat akin to the adjustment-cost models of investment.

Second, one can capture different kinds of norms, for instance, the strengthening or undermining the independence of media and that of the supreme court. Thus, a leader's action is a vector $a_t \in \{0, 1\}^n$, and the norm level also takes an n-dimensional vector. The replacement probability depends on the vector of actions and norms. This could allow the possibility for the leader to abuse certain norms first, for example eroding the freedom of the press or augmenting patronage and taking more extreme actions such as intervening the

courts once the overall level of scrutiny is low.

5.4 Asymmetry between Norm-Destruction and Norm-Building

We could easily extend the model to capture the possibility that it is easier to destroy norms than to build them up. This could be done by allowing for $\gamma_A > \gamma > \gamma_R$, where A and R denote abuse and respect, respectively. Since the decision to cheat only depends on γ_A, we have two observations. First, the higher γ_R (resp. γ_A) is, the faster the absorption into the highest (resp. lowest) norm level is. Second, we can analyze the effect of γ_A on the leaders' decisions just by the comparative statics of γ on $\tilde{h}$. A higher γ_A leads to a higher threshold $\tilde{h}$.

5.5 Endogenous Payoff from Being in Power

While we have assumed that the benefit b of being power is constant over time, it may be natural to consider the case in which the benefit depends on the strength of norm N_t: the higher the norm level N_t is, the higher the benefit from being in power. This consideration has two opposing effects on the continuation play after abusing. First, abusing today makes abusing tomorrow less attractive as it decreases the payoff from abusing tomorrow.[34] Second, abusing today makes abusing tomorrow more attractive as it decreases the payoff from law-abiding tomorrow. Which effect dominates depends on the specification of the payoff of being in power, the replacement probability, and the current norm level. As long as the second effect dominates (for instance, when the benefit of being in power would be given by bN_t), our theorems still hold. Thus, our main results are robust to such consideration.

6 Conclusion

This paper provides a parsimonious model of the evolution of institutional norms and the behavior of a leader that they induce. The leader's action has a persistent effect on the behaviors of the future leaders. This can lead to different long-run behaviors even for institutions with the same level of formal rules. The evolution of norms plays a crucial role in path dependence. Especially, the early history of leaders may play a crucial role in determining which outcome prevails. Thus, the paper suggests the importance of conditioning on the history of past leaders in evaluating the quality of governance. This may explain why a regime change from the outside tends to fail. Our model can capture democratic backsliding and corporate board capturing, whereby institutional norms are gradually eroded.

[34] A related idea is explored in the reputation literature. See, for instance, Phelan (2006).

We believe that our simple model admits many other interesting extensions for future research. In the previous section, we have sketched some of them. For others, for instance, one may consider multiple organizations in which the action of a leader in one organization may affect the incentives of the leaders of the other organizations. In the political context, it corresponds to cross-diffusion of anti-democracies: Rydgren (2005) studies the emergence of the party family of extreme right-wing populist parties in Western Europe, beginning with the electoral breakthrough in 1984 of the French *Front National* led by Jean-Marie Le Pen.

A Proofs

Proof of Theorem 2. 1. In each period, with positive probability, h falls into $\left(\underline{h}, \tilde{h}\left(N_H\right)\right)$ and the norm level decreases. Also, there exists a threshold norm level N_* below which N_t deterministically converges to N_L. Thus, almost surely along any path, $N_t \to N_L$.

2. In each period, with positive probability, h falls into $\left(\tilde{h}\left(N_L\right), \overline{h}\right)$ and the norm level increases. Also, there exists a threshold norm level N^* above which N_t deterministically converges to N_H. Hence, almost surely along any path, $N_t \to N_H$.

3. For each t and for any $N_t \in (N_L, N_H)$, we have $N_{t+1} = (1-\delta)N_t + \delta\overline{N} + \gamma$ with strictly positive probability and $N_{t+1} = (1-\delta)N_t + \delta\overline{N} - \gamma$ with strictly positive probability. This shows that a limit distribution, which exists, has full support.

4. There is N_* such that if $N_t \leq N_*$ for some t then N_t deterministically converges to N_L. Likewise, there is N^* such that if $N_t \geq N^*$ for some t then N_t deterministically converges to N_H. In each period t, if $N_t \in (N_*, N^*)$, then with positive probability, either N_t decreases over time and is below N_* in some finite time or N_t increases over time and is above N^* in some finite time. Thus, the measure of paths $(N_t)_t$ such that $N_t \in (N_*, N^*)$ for infinitely many t is zero. This establishes the statement.

□

Proof of Corollary 1. Assume $(\delta, \gamma) = (1, 0)$. First, if $\tilde{h}(\overline{N}) \geq \overline{h}$, then, almost surely along any path, the optimal action sequence is always to abuse, i.e., Case 1 obtains. Note that if $\tilde{h}(\overline{N}) > \overline{h}$, then the optimal action sequence is deterministically always to abuse. Second, if $\tilde{h}(\overline{N}) \leq \underline{h}$, then, almost surely along any path, the optimal action sequence is always to abide by the rules, i.e., Case 2 obtains. Note that if $\tilde{h}(\overline{N}) < \underline{h}$, then the optimal action sequence is deterministically always to abide by the rules. Third, if $\tilde{h}(\overline{N}) \in (\underline{h}, \overline{h})$, then there exists a limit distribution on the set of action sequences, i.e., Case 3 obtains. Since these cases are exhaustive, the proof is complete. □

References

Acemoglu, D., G. Egorov, and K. Sonin (2021): "Institutional Change and Institutional Persistence," in *The Handbook of Historical Economics*, Elsevier, 365–389.

Acemoglu, D. and M. O. Jackson (2015): "History, Expectations, and Leadership in the Evolution of Social Norms," *Review of Economic Studies*, 82, 423–456.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001): "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, 1369–1401.

Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared (2008): "Income and Democracy," *American Economic Review*, 98, 808–842.

Acemoglu, D. and J. A. Robinson (2008): "Persistence of Power, Elites, and Institutions," *American Economic Review*, 98, 267–293.

Acemoglu, D., T. Verdier, and J. A. Robinson (2004): "Kleptocracy and Divide-and-Rule: A Model of Personal Rule," *Journal of the European Economic Association*, 2, 162–192.

Adams, R. B., B. E. Hermalin, and M. S. Weisbach (2010): "The Role of Boards of Directors in Corporate Governance: A Conceptual Framework and Survey," *Journal of Economic Literature*, 48, 58–107.

Ahmed, A. (2022): "A Theory of Constitutional Norms," *Michigan Law Review*, 120, 1361–1418.

Almond, G. and S. Verba (1963): *The Civic Culture: Political Attitudes and Democracy in Five Nations*, Princeton University Press.

Almond, G. A. (1956): "Comparative Political Systems," *Journal of Politics*, 18, 391–409.

Alston, L. J. and A. A. Gallo (2010): "Electoral Fraud, the Rise of Peron and Demise of Checks and Balances in Argentina," *Explorations in Economic History*, 47, 179–197.

Andvig, J. C. and K. O. Moene (1990): "How Corruption May Corrupt," *Journal of Economic Behavior and Organization*, 13, 63–76.

Ashforth, B. E. and V. Anand (2003): "The Normalization of Corruption in Organizations," *Research in Organizational Behavior*, 25, 1–52.

Azari, J. R. and J. K. Smith (2012): "Unwritten Rules: Informal Institutions in Established Democracies," *Perspectives on Politics*, 10, 37–55.

Bandiera, O., A. Prat, S. Hansen, and R. Sadun (2020): "CEO Behavior and Firm Performance," *Journal of Political Economy*, 128, 1325–1349.

Barker, E., ed. (1946): *The Politics of Aristotle*, Oxford University Press.

Bebchuk, L., A. Cohen, and A. Ferrell (2009): "What Matters in Corporate Governance?" *Review of Financial Studies*, 22, 783–827.

Bertrand, M. and A. Schoar (2003): "Managing with Style: The Effect of Managers on Firm Policies," *Quarterly Journal of Economics*, 118, 1169–1208.

Besley, T. and T. Persson (2019): "Democratic Values and Institutions," *American Economic Review: Insights*, 1, 59–76.

——— (Forthcoming): "Organizational Dynamics: Culture, Design,and Performance," *Journal of Law, Economics, and Organizations*.

Besley, T. and A. Prat (2006): "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability," *American Economic Review*, 96, 720–736.

Biggerstaff, L., D. C. Cicero, and A. Puckett (2015): "Suspect CEOs, Unethical Culture, and Corporate Misbehavior," *Journal of Financial Economics*, 117, 98–121.

Bloom, N., R. Lemos, R. Sadun, D. Scur, and J. Van Reenen (2014): "JEEA-FBBVA Lecture: The New Empirical Economics of Management," *Journal of the European Economic Association*, 12, 835–876.

Brancati, D. (2014): "Democratic Authoritarianism: Origins and Effects," *Annual Review of Political Science*, 17, 313–326.

Bryce, J. (1888 [1995]): *The American Commonwealth*, vol. I, Liberty Fund.

Coppedge, M., J. Gerring, C. H. Knutsen, S. I. Lindberg, J. Teorell, D. Altman, M. Bernhard, A. Cornell, M. S. Fish, L. Gastaldi, H. Gjerløw, A. Glynn, S. Grahn, A. Hicken, K. Kinzelbach, K. L. Marquardt, K. McMann, V. Mechkova, A. Neundorf, P. Paxton, D. Pemstein, O. Rydén, J. von Römer, B. Seim, R. Sigman, S.-E. Skaaning, J. Staton, A. Sundström, E. Tzelgov, L. Uberti, Y. Wang, T. Wig, and D. Ziblatt (2023): "V-Dem Codebook v13," Varieties of Democracy (V-Dem) Project.

Dessein, W. and A. Prat (2022): "Organizational Capital, Corporate Leadership, and Firm Dynamics," *Journal of Political Economy*, 130, 1477–1536.

Diamond, L. (2021): "Democratic Regression in Comparative Perspective: Scope, Methods, and Causes," *Democratization*, 28, 22–42.

Diamond, L. J. (1999): *Developing Democracy: Towards Consolidation*, Johns Hopkins University Press.

Dixit, A., G. M. Grossman, and F. Gul (2000): "The Dynamics of Political Compromise," *Journal of Political Economy*, 108, 531–568.

Dyck, A., N. Volchkova, and L. Zingales (2008): "The Corporate Governance Role of the Media: Evidence from Russia," *Journal of Finance*, 63, 1093–1135.

Dyck, A. and L. Zingales (2004): "Private Benefits of Control: An International Comparison," *Journal of Finance*, 59, 537–600.

Fearon, J. D. (2011): "Self-Enforcing Democracy," *Quarterly Journal of Economics*, 126, 1661–1708.

Ferraz, C. and F. Finan (2010): "Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance," Working paper.

——— (2011): "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments," *American Economic Review*, 101, 1274–1311.

Fomina, J. and J. Kucharczyk (2016): "Populism and Protest in Poland," *Journal of Democracy*, 27, 58–68.

Foran, C. (2016): "An Erosion of Democratic Norms in America," *Atlantic*, november 22.

Gagliarducci, S. and T. Nannicini (2013): "Do Better Paid Politicians Perform Better? Disentangling Incentives From Selection," *Journal of the European Economic Association*, 11, 369–398.

Gennaioli, N. and A. Shleifer (2007): "The Evolution of Common Law," *Journal of Political Economy*, 115, 43–68.

Gibbons, R. and R. Henderson (2013): "What Do Managers Do?" in *The Handbook of Organizational Economics*, Princeton University Press, chap. 17, 680–731.

Glaeser, E. L., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2004): "Do Institutions Cause Growth?" *Journal of Economic Growth*, 9, 271–303.

Gompers, P., J. Ishii, and A. Metrick (2003): "Corporate Governance and Equity Prices," *Quarterly Journal of Economics*, 118, 107–156.

Graham, J. R., H. Kim, and M. Leary (2020): "CEO-board Dynamics," *Journal of Financial Economics*, 137, 612–636.

Gratton, G. and B. E. Lee (Forthcoming): "Liberty, Security, and Accountability: The Rise and Fall of Illiberal Democracies," *Review of Economic Studies*.

Greenstein, F. I. (2009): *Inventing the Job of President: Leadership Style from George Washington to Andrew Jackson*, Princeton University Press.

Grillo, E., Z. Luo, M. Nalepa, and C. Prato (2023): "Theories of Democratic Backsliding," Working paper.

Grillo, E. and C. Prato (2023): "Reference Points and Democratic Backsliding," *American Journal of Political Science*, 67, 71–88.

Guiso, L., P. Sapienza, and L. Zingales (2015): "The Value of Corporate Culture," *Journal of Financial Economics*, 117, 60–76.

Guiso, L., L. Zingales, and P. Sapienza (2016): "Long-term Persistence," *Journal of the European Economic Association*, 14, 1401–1436.

Guriev, S. and D. Treisman (2020): "A Theory of Informational Autocracy," *Journal of Public Economics*, 186, 104158.

Helmke, G., M. Krorger, and J. Paine (2022): "Democracy by Deterrence: Norms, Constitutions, and Electoral Tilting," *American Journal of Political Science*, 66, 434–450.

Howell, W. G., K. A. Shepsle, and S. Wolton (2023): "Executive Absolutism: The Dynamics of Authority Acquisition in a System of Separated Powers," *Qarterly Journal of Political Science*, 18, 243–275.

Hsieh, C.-T. and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *Quartrely Journal of Economics*, 124, 1403–1448.

Huq, A. and T. Ginsburg (2018): "How to Lose a Constitutional Democracy," *U.C.L.A. Law Journal*, 65, 78–169.

Invernizzi, G. M. and M. M. Ting (Forthcoming): "Institutions and Political Restraint," *American Journal of Political Science*.

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

Kagan, R. (2023): "A Trump Dictatorship is Increasingly Inevitable. We should Stop Pretending," *Washington Post*, november 30.

Kamarck, E. (2021): "Did Trump Damage American Democracy?" *Brookings Institution*, july 9.

Kapstein, E. B. and N. Converse (2008): *The Fate of Young Democracies*, Cambridge University Press.

Keefer, P. (2007): "Clientelism, Credibility, and the Policy Choices of Young Democracies," *American Journal of Political Science*, 51, 804–821.

Klein, E. (2016): "Donald Trump and the Destruction of America's Political Norms," *vox.com*, june 7.

La Porta, R., F. Lopez-de Silanes, A. Shleifer, and R. Vishny (1999): "The Quality of Government," *Journal of Law, Economics, and Organization*, 15, 222–279.

Las Cases, E.-A.-D. (2017 [1855]): *Memoirs of the Life, Exile, and Conversations of the Emperor Napoleon*, vol. I.

Laux, V. (2008): "Board Independence and CEO Turnover," *Journal of Accounting Research*, 46, 137–171.

Levitsky, S. and L. Way (2015): "The Myth of Democratic Recession," *Journal of Democracy*, 26, 45–58.

Levitsky, S. and D. Ziblatt (2018): *How Democracies Die*, Crown Publishing.

Linz, J. J. (1978): *The Breakdown of Democratic Regimes: Crisis, Breakdown and Reequilibration*, Johns Hopkins University Press.

——— (1990): "Transitions to Democracy," *Washington Quarterly*, 13, 143–164.

Luo, Z. and A. Przeworski (2023): "Democracy and its Vulnerabilities: Dynamics of Democratic Backsliding," *Quarterly Journal of Political Science*, 18, 105–130.

Lust, E. and D. Waldner (2015): "Unwelcome Change: Understanding, Evaluating, and Extending Theories of Democratic Backsliding," U.S. Agency for International Development.

Mainwaring, S. and F. Bizzarro (2019): "The Fates of Third-Wave Democracies," *Journal of Democracy*, 30, 99–113.

Myerson, R. B. (2011): "Toward a Theory of Leadership and State Building," *Proceedings of the National Academy of Sciences*, 108, 21297–21301.

Niblett, A., R. A. Posner, and A. Shleifer (2010): "The Evolution of a Legal Rule," *Journal of Legal Studies*, 39, 325–358.

North, D. C. (1990): *Institutions, Institutional Change, and Economic Performance*, Cambridge University Press.

O'Donnell, G. and P. C. Schmitter (1986): *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies*, Johns Hopkins University Press.

O'Donnell, G. A. (1996): "Illusions about Consolidation," *Journal of Democracy*, 7, 34–51.

Paldman, M. (2002): "The Cross-Country Pattern of Corruption: Economics, Culture and the Seesaw Dynamics," *European Journal of Political Economy*, 18, 215–240.

Persson, T. and G. Tabellini (2009): "Democratic Capital: The Nexus of Political and Economic Change," *American Economic Journal: Macroeconomics*, 1, 88–126.

Pfiffner, J. P. (2021): "Donald Trump and the Norms of the Presidency," *Presidential Studies Quarterly*, 51, 96–124.

Phelan, C. (2006): "Public Trust and Government Betrayal," *Journal of Economic Theory*, 130, 27–43.

Pierson, P. (2000): "Increasing Returns, Path Dependence, and the Study of Politics," *American Political Science Review*, 94, 251–267.

Prat, A. and D. Strömberg (2013): "The Political Economy of Mass Media," in *Advances in Economics and Econometrics: Tenth World Congress*, ed. by D. Acemoglu, M. Arellano, and E. Dekel, Cambridge University Press, 135–187.

Putnam, R. (1993): *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton University Press.

Renan, D. (2018): "Presidential Norms and Article II," *Harvard Law Review*, 131, 2187–2282.

Rydgren, J. (2005): "Is Extreme Right-Wing Populism Contagious? Explaining the Emergence of a New Party Family," *European Journal of Political Research*, 44, 413–437.

Schein, E. H. (1983): "The Role of the Founder in Creating Organizational Culture," *Organizational Dynamics*, 12, 13–28.

Shleifer, A. and R. W. Vishny (1993): "Corruption," *Quarterly Journal of Economics*, 108, 599–617.

——— (1997): "A Survey of Corporate Governance," *Journal of Finance*, 52, 737–787.

Simons, R. (1994): "How New Top Managers Use Control Systems as Levers of Strategic Renewal," *Strategic Management Journal*, 15, 169–189.

Strömberg, D. (2015): "Media and Politics," *Annual Review of Economics*, 7, 173–205.

Syverson, C. (2004): "Product Substitutability and Productivity Dispersion," *Review of Economics and Statistics*, 86, 534–550.

Szelényi, Z. (2022): "Viktor Orbán's Machiavellian Genius: Will His Illiberal Experiment Return to Haunt Him?" *UnHerd*.

Tanzi, V. (1998): "Corruption Around the World: Causes, Consequences, Scope, and Cures," *IMF Staff Papers*, 45, 559–594.

Walter, M. and D. Cancel (2009): "Chavez Wins Venezuela Term Limit Vote, Opens Campaign for 2012," *Bloomberg News*, february 16.

Wantchekon, L. (2013): "Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin," *World Politics*, 55, 399–422.