

Charismatic Leaders and Democratic Backsliding*

James R. Hollyer[†]

Marko Klašnja[‡]

Rocío Titiunik[§]

October 27, 2025

Abstract

An important role of political parties in democracies is to constrain the abuse of power, including by their members. Parties also prize winning elections, and nominating charismatic party leaders tends to be electorally beneficial. We show theoretically that when parties become reliant on a leader's charisma, they grow less able to sanction their behavior in office *ex post*. Parties recognize this, and may try to preempt it by screening out charismatic leadership candidates *ex ante*. However, they sometimes willingly sacrifice horizontal accountability in exchange for the electoral gains that charisma brings, foregoing their gate-keeping role. Consequently, charismatic leaders are given more room than less charismatic leaders to subvert democracy. This dynamic is more pronounced in polarized and less stable democracies, where charismatic leaders are more frequent. Associations between leaders' charisma and patterns of party personalization, party illiberalism, democratic backsliding, and democratic breakdown in cross-national data are consistent with these predictions.

*We thank Caitlin Andrews-Lee, Raúl Aldaz, Elena Barham, Alexandra Cirone, John Huber, John Marshall, Carlo Prato, Maria Popova, Stephanie Rickard, Arturas Rozenas, Julian Waller, audiences at seminars at American University, Georgetown, George Washington University, McGill, the Quality of Government Institute, Stockholm Institute of Transition Economics, and Toronto Metropolitan University, as well as participants in the Global Research in Political Economy seminar, the Virtual Formal Theory Workshop, the Political Economy Workshop at the Norwegian Business School, the Columbia Conference on Elites in Historical and Comparative Perspective, the European Political Science Association conference, the American Political Science Association conference, the Center for the Study of Democratic Politics Conference on Democratic Frontsliding at Princeton University, and the APSA Pre-Conference on the Political Economy of Europe for valuable comments and suggestions. Thanks also to Matt Eichinger, Ezgi Özçelik, Gus Beringer, and Matthew Swedin for excellent research assistance. Klašnja and Titiunik gratefully acknowledge the support of the National Science Foundation via research award [SES-2116224](#).

An online appendix with supplemental materials is available at [LINK](#).

Replication files are available in the JOP Dataverse (<https://dataverse.harvard.edu/dataverse/jop>). The empirical analysis has been successfully replicated by the JOP replication analyst

[†]University of Minnesota, jhollyer@umn.edu

[‡]Georgetown University, marko.klasnja@georgetown.edu

[§]Princeton University, titiunik@princeton.edu

The January 6th attack on the U.S. Capitol, following President Trump's 2020 electoral loss, presented a critical juncture for the Republican Party. Initially, figures like the former UN ambassador Nikki Haley condemned Trump's role in instigating the attack, proclaiming that Republicans "can't let that ever happen again" ([Alberta, February 2021](#)). Yet, Haley reversed her stance within months and, echoing many Republican elites, asserted: "I don't want us to go back to the days before Trump" ([McCormick, October 5, 2021](#)). This reversal is striking: a party would normally be expected to abandon a candidate who unsuccessfully tried to overturn a democratic election and publicly threatened his own Vice President and party members ([Zengerle and Cowan, June 17, 2022](#)). Instead, the GOP renominated Trump.

We observe similar dynamics elsewhere. For example, following Prime Minister Indira Gandhi's nearly two-year-long "Emergency" (1975-1977), which saw the suspension of civil liberties, widespread censorship of the media, and the imprisonment of many opposition politicians and activists ([Plys, 2020](#)), her Indian National Congress party lost power for the first time since India's independence. Several months later, the party elites moved to expel her; "The cancer is out and we are not carrying the burden of Mrs. Indira Gandhi any more," proclaimed one member of the party's executive committee ([The New York Times, January 4, 1978](#)). Yet, by the next election two years later, most of those elites flocked to Gandhi's newly formed faction called Congress (Indira), epitomizing the words of one of her loyalists: "She is the Congress" ([Kidwai, June 8, 2019](#)).

Why do political parties fail to screen out or sanction leaders who seek to erode democratic norms and institutions? We argue that parties are particularly prone to such inaction when their leaders are charismatic. The modern concept of charisma was introduced by Weber, who defined it as "a certain quality of individual personality by virtue of which he is considered extraordinary and treated as endowed with supernatural, superhuman, or at least specifically exceptional powers or qualities" ([Weber, 1922 \[1978\]](#), p. 241). Both Trump and Gandhi have commonly been portrayed, by followers and critics alike, as possessive of such powers, having been likened to Biblical figures (Trump) and Hindu Goddesses (Gandhi) ([Malhotra, 1991](#); [Worthen, 2025](#)). Such

charismatic appeal implies strong and direct connections with voters and party activists that translate into a personal electoral advantage that is *separate* from and *non-transferable* to the party's own appeal. We formalize a theory where this unmediated appeal can make the party unable to credibly commit to imposing ex post sanctions on a charismatic leader, foregoing horizontal accountability because the costs of a standoff with such a leader may outweigh the electoral dividends from their charisma. Anticipating such an inversion of power, a party might ex ante choose to screen out potentially recalcitrant leaders. However, we show that parties will sometimes nominate a charismatic leader knowing full well that they may not be able to control them once in office, even when they could control a leader lacking such appeal. This rational acquiescence represents an equilibrium personalization of the party—its willingness to prioritize the goals of its leader.

The key consequence of party reticence in the face of a leader's charisma is that democratic backsliding (and potentially democratic breakdown) is more likely under charismatic than under less charismatic leaders. This is not because charismatic leaders inherently possess a greater proclivity to authoritarianism,¹ but because their parties grant them more room to engage in anti-democratic actions. As a corollary, we demonstrate why “strong” leaders like Trump and Gandhi are often charismatic, and why parties with charismatic leaders frequently become little more than their personal vehicles. This reliance on personalistic ties may weaken the party (Hollyer, Klačnja and Titiunik, 2022), but we show in our framework that charisma affects the incentives and choices of both weak and strong parties.

We further use this theoretical framework to show that a party is less likely to control a charismatic leader as elite polarization increases. Hurting electoral chances by screening out or sanctioning a popular leader becomes increasingly unpalatable as competitors grow distant in policy and ideology. We also demonstrate that as an attempted *autogolpe* grows more likely to succeed—as in less stable democracies—the value of discipline to a party drops, increasing the

¹We are agnostic about any such association and none is assumed in our theory.

incentive to nominate charismatic politicians. Consequently, charismatic leaders tend to be more frequent in less consolidated democracies and rarer (though not absent) in more consolidated ones. While this implies that the risk of backsliding is lower in consolidated than in less consolidated democracies, we demonstrate that when backsliding does occur in stable democracies, it is more likely to be because parties voluntarily forsake control in pursuit of electoral gains from charisma.

We empirically illustrate our model's predictions using cross-national proxies for leader charisma, democratic breakdown, quality of democracy, party illiberalism and personalization, polarization, and democratic stability. The observed patterns generally align with our expectations: backsliding and democratic breakdown are more likely under more charismatic leaders, who are more frequent in less stable democracies. Consistent with our theoretical mechanism, incumbent parties with charismatic chief executives are indeed more illiberal and personalized, and these associations are more pronounced in more polarized societies.

While the recent wave of democratic backsliding motivates our work, we do not suggest that this trend should be attributed to an increase in the frequency or intensity of political charisma. Charismatic politicians are not a new phenomenon (as the example of Gandhi illustrates), and our model is not designed to explore the emergence or rise in the salience of charisma. Instead, we take charisma as given, focusing our contribution on the analysis of its consequences rather than its causes.

The rest of the manuscript is organized as follows. The next section discusses in detail the concept of charisma, and the connections of our work to prior literature. We then introduce our theoretical framework, followed by an illustration that assesses some of the empirical implications of our model. The last section concludes.

Charisma and Parties

Since Weber's (1922 [1978]) foundational treatise on charisma, the literature has debated whether it is an objective quality of a politician's personality or a subjective phenomenon emanating from followers' perceptions (e.g., Willner and Willner, 1965; Tucker, 1968). Social scientists have generally adopted the latter conception, as it better explains a leader's ability to elicit a strong following (Willner, 1985). While individual attributes like physical attractiveness or oratory skills may increase the likelihood of being perceived as charismatic, what ultimately matters for charismatic appeal is not objective skills but the masses' belief that the leader is extraordinary and uniquely suited to solve their problems. Thus, we conceptualize a leader as charismatic if they are able to elicit a devoted and direct following.

While our understanding of charisma is rooted in voter perceptions, our theoretical model operationalizes it as a politician-level parameter that generates an electoral boost, similarly to our approach in Hollyer, Klašnja and Titiunik (2022). This electoral advantage makes charisma a 'valence' parameter. The concept of valence, introduced by Stokes (1963), describes the non-policy characteristics universally coveted by voters such as "competence, trustworthiness, charisma, courage, imagination, and intelligence" (Dewan and Shepsle, 2011, p. 313). However, we argue that charisma distinguishes itself from other forms of valence in two key ways. First, while attributes like integrity or competence can confer an advantage, they do not inherently ensure popularity.² By contrast, since charisma implies a leader's ability to elicit a strong, unmediated bond with his followers, it translates directly into higher popularity. Therefore, while not all popular leaders are charismatic (and may be popular due to other valence characteristics), all charismatic leaders are popular by virtue of their devoted following. Put differently, not all valence is charisma, but all charisma is valence.

²For example, Jimmy Carter was widely perceived as honest and Great Britain's Gordon Brown as competent, yet both leaders were relatively unpopular for most of their tenure.

The second, and related, distinguishing quality of charisma is the non-fungibility of such popularity. A leader who adopts widely preferred policies or effectively employs clientelistic practices may achieve popularity and boost his and his party's electoral prospects regardless of whether he enjoys a special bond with voters. Such popularity is to a degree fungible: it can be transferred to other candidates in the party, by having them adopt the same popular policy positions or the same effective clientelistic practices. The support that stems from charisma, in contrast, lies in the direct connection between the leader and his followers, and is therefore not so easily transferable. If a charismatic leader is replaced by another politician, the electoral boost from their charisma is mostly lost.³

Our assertion is that these two properties of charisma—inherent popularity and non-fungibility—give charismatic candidates greater independence from their parties. In our framework, this independence can lead to greater leverage over the party, in that it can create incentives for the party to acquiesce to a charismatic leader's actions, including when they are antidemocratic, in order to continue enjoying such leader's non-fungible appeal. With such party acquiescence, charisma can lead to the personalization of the leader's party and, in turn, democratic erosion.

The idea that personalism threatens democracy is not new; scholars have long argued that a democratic leader's reliance on personalistic ties undermines pluralism and democratic values (for a review, see [Urbinati, 2019](#)). These accounts often suggest that a democracy's ability to withstand backsliding depends on parties restraining a leader's anti-democratic impulses ([Levitsky and Ziblatt, 2018](#); [Weyland, 2001](#)). [Frantz, Kendall-Taylor and Wright \(2024\)](#) attribute the recent wave of democratic erosion to political parties whose primary goal is to act as a vehicle for their leader's ascension to power rather than the pursuit of programmatic policies.

Our contribution is to explain the conditions under which parties become personalized, and the role of charisma in this process. Existing work typically assumes that a party that possesses

³Although successors may sometimes continue to successfully harness a charismatic leader's appeal ([Andrews-Lee, 2021](#)).

the necessary resources to constrain an authoritarian leader will do so. Others take party personalization as a given, and examine its consequences. By contrast, in the theory we propose, circumstances arise when parties *choose* to nominate a charismatic leader whose behavior they know they cannot control, even when they would be able to constrain a less charismatic politician. Therefore, we show that a party's ability does not always equate to its willingness to exercise horizontal accountability, and personalization may arise endogenously as a consequence of leader charisma.

That said, we emphasize that in our theoretical framework charisma and party personalization are distinct. In equilibrium, charismatic leaders do not arise only among parties that are unable or unwilling to constrain their leader; they also arise in parties that are able and willing to rein in leaders regardless of their charisma. This is because charismatic appeal is electorally beneficial; a party that exercises control over any type of leader stands to gain electorally from a politician's charismatic appeal, without risking the costs of democratic backsliding.

Another corollary of this logic is that charisma in our theory is distinct from party weakness. While defined in different ways in the empirical literature (see, e.g., [Mainwaring and Torcal, 2006](#)), if party strength is understood as the collective control over its members' actions, including the leader, then our model implies that reliance on charisma is not purely the provenance of weak parties.⁴ The existing literature highlights the importance of party discipline for accountability ([Schattschneider, 1942](#); [Cox and McCubbins, 2007](#)), including in restraining populist, anti-pluralistic leaders ([Levitsky and Ziblatt, 2018](#); [Weyland, 2001](#)). The implication is that illiberal leaders emerge primarily when parties are weak. Our model challenges this notion by showing that charisma influences the incentives and choices of both weak and strong parties.

More broadly, our study complements the existing literature on democratic consolidation, sur-

⁴That said, charisma may weaken parties by undermining the party's willingness to nominate candidates who promote programmatic goals, thus weakening incentives for programmatic discipline ([Hollyer, Klačnjak and Titiunik, 2022](#)).

vival, and backsliding (for recent reviews, see [Grillo et al., 2023](#); [Waldner and Lust, 2018](#)). The strategic interactions that sustain or undermine democratic rule are typically studied by emphasizing the role of the populace (vertical constraints) or institutional actors such as political parties (horizontal constraints). We primarily contribute to the latter. Though, in an extension, we find that – for some parameter values – the presence of vertical accountability may make parties more willing to jeopardize horizontal control and nominate charismatic leaders. The existing work on horizontal constraints primarily emphasizes inter-party dynamics—how opposition parties constrain anti-democratic executives through mutual deterrence ([Helmke, Kroeger and Paine, 2022](#)), opposing interests ([Howell, Shepsle and Wolton, 2023](#)), or electoral competition ([Luo and Przeworski, 2023](#); [Miller, 2021](#)). In contrast, we highlight the crucial role of intra-party interactions.

While our theoretical account elides the strategic considerations of the electorate, it shares similarities with [Svolik \(2020\)](#), who shows how polarized electorates may tolerate democratic subversion to obtain their desired policies (see also [Haggard and Kaufman, 2021](#)). However, our model demonstrates that backsliding can occur without polarization, and that a leader's charisma can amplify the link between ideological polarization and democratic backsliding.

Our contribution is also related to the literature on the rise of populism and its threat to liberal democracy ([Berman and Snegovaya, 2019](#); [Grzymala-Busse, 2019](#)). Populism has been characterized as a type of political rule where the leader establishes uninstitutionalized personalistic connections with large masses of followers, often adopting anti-elite stances and claiming to be the sole legitimate representative of the people (e.g. [Müller, 2016](#); [Weyland, 2001](#)). Charisma facilitates intense connections with voters, which is why populist leaders are often charismatic; however, charisma is not considered a defining feature of populism ([Weyland, 2001](#), p. 13). By inducing the personalization of parties, our theory suggests how highly charismatic candidates may enable and sustain populism.

Theoretical Model

We present a model wherein a party both selects a leader and decides whether to sanction him for the actions the leader takes once in office. Potential leaders may have more or less of an electoral following independent of the party, a feature which we attribute to the leader's charisma.⁵ Thus, the party has an electoral incentive to select a charismatic leader, but this very incentive makes the leader less subject to party control. The party may remove a recalcitrant leader from office, but suffers an electoral cost from doing so. Alternatively, the party may opt to forgo charismatic leadership in favor of selecting a less charismatic leader who is easier to control. We thus examine the trade-offs between ex ante screening and ex post control, incorporating concepts of gate-keeping (screening) and horizontal accountability (sanctioning). The origins of charisma, and the fluctuations in the returns to a charismatic style of politics, are exogenous to the model.

Primitives

Consider an interaction between an incumbent party I and its leader L . There is also an opposition party O , which is non-strategic. We index actors with $j \in \{L, I, O\}$. Because our interest is in the ability and willingness of the incumbent party to discipline its incumbent leader, we focus on the strategic considerations involving only I and L .

L is characterized by two binary type parameters. First, L may belong to one of two parties, $\theta \in \{\theta_I, \theta_O\}$. Second, L may be either non-charismatic or charismatic, $\nu \in \{\underline{\nu}, \bar{\nu}\}$, where $\bar{\nu} > \underline{\nu} > 0$. ν represents the appeal of a leader to voters independent of the appeal of the party.

The interaction takes place over two periods $t \in \{1, 2\}$. At time $t = 1$, the incumbent party is in power. At the start of this stage, the incumbent party makes a choice between a charismatic and non-charismatic leader, both of type $\theta_t = \theta_I$. The potential leaders' values of ν are common

⁵Critical for our argument is that this source of electoral support is fixed and not subject to strategic choices by the leader, at least in the short term.

knowledge, as charisma is by its very nature observable. With some abuse of notation, we equate this action to choosing the leader's type ν , $c \in \{\underline{\nu}, \bar{\nu}\}$.

Following his selection, the sitting L makes a binary policy choice $a \in \{0, 1\}$, which captures the degree of 'authoritarianism.' In line with our motivating examples, choosing $a = 1$ reflects policies such as the concentration of power or the undermining of institutions and civil liberties in ways that hinder L 's opponents. An authoritarian turn in policy may lead to a complete breakdown of democracy, but need not do so. Specifically, if $a = 1$ and L is retained in power, democracy is overthrown with probability $1 - \sigma$, where $\sigma \in (0, 1)$ denotes the degree of democratic consolidation. We intend σ to represent political institutions, elite actors, or structural features of the polity that inhibit an *autogolpe*. For instance, σ may be rising in the duration of democratic rule. We alternatively refer to σ either as democratic stability or consolidation.

After L 's choice in the first period, I decides whether to retain or remove him, $r \in \{0, 1\}$, where $r = 1$ means retaining; $r = 0$ may represent the party deciding not to renominate the leader in subsequent electoral contests, expel the leader, or stage a vote of no-confidence or support such a move by the opposition. The assumption that I makes a binary retention decision is an analytical simplification. A variety of sanctions are likely available to the party other than removal, such as the withdrawal of campaign funds, ousting the leader's allies from party positions, etc. Our insights extend to such other forms of sanctioning, as long as they impose a cost on the incumbent party and party members.

I thus has two tools through which it can influence the behavior of L : ex ante screening (the choice of c), and ex post sanctioning (the choice of r). Ex ante screening is captured by an equilibrium choice of a non-charismatic leader, $c = \underline{\nu}$. As further discussed below, this takes place only if the party anticipates that a charismatic leader will prove too difficult to control. We view this choice as capturing, for instance, the Democratic Party's decision to eschew George Wallace's nomination for president in 1964 and 1968 in favor of Lyndon Johnson and Hubert Humphrey, respectively (for a discussion of this case, see [Levitsky and Ziblatt, 2018](#)).

Ex post sanctioning is captured by the choice to remove the leader. Instances of parties punishing their own leaders are rare—as one would expect in equilibrium—but can be found in a variety of contexts. For example, Jacob Zuma chose to resign from the South African presidency following a string of corruption scandals when it became clear that his party, the African National Congress, would not support him in an upcoming no-confidence vote (Campbell, February 15, 2018).⁶

If I removes the leader following $a = 1$, the threat to democracy is removed. That is, we assume that the current leader's attempted backsliding does not directly open the door to similar maneuvers by his successor. In this way, the incumbent party can act as a guardrail of democracy. Following $r = 0$, the incumbent party chooses a new leader with type $\theta_t = \theta_I$ to lead them into the next election. Contrastingly, if $r = 1$ following a backsliding attempt, democracy survives with probability $\sigma \in (0, 1)$.

To capture histories in which democracy collapses, denote the state variable $B \in \{0, 1\}$ (for breakdown), with $B = 1$ denoting democratic collapse. If $B = 0$, an election follows, pitting I against the opposition O . I wins with probability $\rho(r \times \nu) \in [0, 1]$. Critically, $\rho(r \times \nu)$ is a function both of I 's leader retention decision and the leader's charismatic appeal. $\rho(\cdot)$ is a monotonically increasing function, where $\rho(0) > 0$ and $\lim_{x \rightarrow \infty} \rho(x) \leq 1$. This functional form encapsulates three assumptions:

- (a) The probability that the incumbent party wins is rising in the leader's charisma, so long as that leader is retained. If the leader is removed, the probability the incumbent party wins is invariant in the (ex-)leader's charisma, $\rho(r \times \nu) = \rho(0) \forall \nu$.
- (b) Whenever the leader is ousted, the probability the incumbent party holds on to power ($\rho(0)$)

⁶Other examples include Richard Nixon's resignation, brought on by indications from Senate co-partisans that they would no longer back his case; and the recent impeachment of Yoon Suk Yeol by a coalition including members of his own party.

is lower than if the leader was retained ($\rho(\nu)$, $\nu \geq 0$). That is, removing a sitting leader is costly for the incumbent party. There is no guarantee the leader will depart quietly; some may seek revenge by withholding their endorsement of the successor, launching whisper campaigns to get donors to deny funding, or even splitting off to form new parties. The example Jacob Zuma is again instructive (as is the earlier example of Indira Gandhi). After resigning, Zuma backed the formation of a breakaway party (the Spear of the Nation, or MK) that campaigned against the ANC in the 2024 general elections.

- (c) The electoral cost of removing the leader, given by $\rho(\nu) - \rho(0)$, is falling as the charismatic appeal of the leader declines.

If I wins the election when $B = 0$, its current leader with type θ_I remains in power. By contrast, if O wins, it chooses a new leader from among its ranks, $\theta_2 = \theta_O$. Define an indicator $E \in \{0, 1\}$ which captures the outcome of the election. $E = 1$ corresponds to a victory for the incumbent party I , which takes place with probability $\rho(\cdot)$; $E = 0$ otherwise. If $B = 1$, democracy has collapsed and so the incumbent leader remains in power. Therefore, $\theta_2 = \theta_I$ in time $t = 2$.

We assume that all actors prefer to have a leader of their own type in power. To capture this, we define the indicator function $g_t(\theta_t) = \mathbb{1}(\theta_t = \theta_I)$. For I and L , the expression $Dg_t(\theta_t)$, will appear as a component of their utility.⁷ Each player gets a utility ‘kick’ of magnitude $D > 0$ from having a leader selected from among their own ranks, as opposed the ranks of their opponents. We interpret D primarily as partisan polarization. For instance, parties are better placed to advance their policy objectives when a co-partisan is the chief executive.⁸ But, D may also encompass affective polarization—an attachment to in-group members and aversion to members of other groups, whether defined by partisan or some other identity, like ethnicity, religion, or race.

⁷For O , it would be $D[1 - g_t(\theta_t)]$.

⁸An alternative version of the model where leaders set policy on a univariate policy dimension consistent with this interpretation is available from the authors on request.

In addition to receiving the utility from a leader's identity, I suffers a cost from backsliding, $\kappa > 0$.⁹ We assume that κ is incurred only if (1) L attempts to backslide in the first period; and (2) I fails to check this abuse of power and instead retains L . This parametrization of κ most directly reflects the cost members of the ruling party suffer to their reputation, in the eyes of history or perhaps in their own moral accounting (insofar as they are committed to democracy), for complicity in democratic backsliding. In an extension discussed below and in the Supplementary Appendix, we explore a variety of other costs that the members and leader of I may suffer as a result of backsliding.

The value of κ is likely to be heterogeneous across members of I . Of relevance for us is whether a sufficient number of party members suffer high enough costs from democratic subversion that they are induced to sanction the leader. As a shorthand, we represent this cost as a single parameter, which may be thought of as the cost to the pivotal member of I . To capture the idea that such costs may vary, we assume that κ is drawn from density $F_\kappa(\cdot)$ with support on \mathbb{R}_+ , and with an associated pdf $f_\kappa(\cdot)$. The realization of κ is observed by all actors at the start of the interaction.

I thus cares about (1) the return from having a leader of its preferred type $\theta_t = \theta_I$ in power (given by $Dg_t(\theta_t)$), and (2) the costs from backsliding (given by $a r \kappa$). We assume that L shares I 's concern about the partisan identity of the leadership. But, L further derives an ego rent from holding power himself, $\phi > 0$.¹⁰

⁹Members of O may also suffer a cost from backsliding, but as O is non-strategic, such a cost does not affect the dynamics of horizontal accountability we examine.

¹⁰We assume this rent is the same in a democracy and autocracy. However, our results would be unchanged if L derived greater rents from office as a dictator.

I 's utility at the conclusion of both periods of play is thus given by:¹¹

$$u_I = D[1 + E + B(1 - E)] - ar\kappa.$$

L 's utility is given by:

$$u_L = \phi[1 + rE + B(1 - rE)] + D[1 + E + B(1 - E)].$$

The term $\phi[1 + rE + B(1 - rE)]$ captures the leader's ego rents: he gains ϕ with certainty in time $t = 1$, and in time $t = 2$ enjoys them either following a democratic breakdown or following retention and reelection. If he is either removed or defeated in an election (with probability $1 - \rho(\nu)$), L gains an ego rent of zero.

The order of play is as follows:

- 1 κ is drawn and made common knowledge to all players. I makes a choice of L 's type $c \in \{\underline{\nu}, \bar{\nu}\}$.
- 2 L determines $a \in \{0, 1\}$.
- 3 I determines whether or not to retain the leader $r \in \{0, 1\}$. If $r = 0$, L is replaced by a new leader such that $\theta_1 = \theta_I$.
- 3a If $a = 1$ and $r = 1$, backsliding occurs. L successfully overthrows democracy with probability $1 - \sigma$.
- 3b If either $a = 0$ or $r = 0$, or if L 's backsliding bid fails (with probability σ), an election takes place. I wins with probability $\rho(r \times \nu)$. If O wins, a new leader is selected and $\theta_2 = \theta_O$. If I wins $\theta_2 = \theta_I$.

¹¹We assume that players do not discount the future. Results would be qualitatively unchanged if a discount factor was included.

4 All payoffs are realized and the game ends.

Equilibrium

We consider subgame perfect equilibria. A strategy for L consists of a mapping from his realization of ν into $a : \mathbb{R}_+ \rightarrow \{0, 1\}$. A strategy for I is (1) a mapping from the charisma typespace (the values of $\underline{\nu}$ and $\bar{\nu}$) and κ into the party's choice of candidate $c : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \{\underline{\nu}, \bar{\nu}\}$; and (2) a choice of $r \in \{0, 1\}$, which is a mapping from the first period levels of authoritarianism and the realization of κ , $r : \{0, 1\} \times \mathbb{R}_+ \rightarrow \{0, 1\}$.

We characterize the unique equilibrium in this game. We start by defining two inequalities in κ that dictate I 's strategy. If the cost the party suffers from backsliding is sufficiently high, it will act to forestall any attempt at such behavior, regardless of leader's charisma. Contrastingly, if κ is sufficiently low, the party will never play this role as it is not committed to democracy, and so horizontal accountability will fail to hold even against a non-charismatic leader. We term these two thresholds $\tilde{\kappa}$ and $\underline{\kappa}$, respectively.

Lemma 1. *Let $\tilde{\kappa} = D[1 - \sigma + \sigma\rho(\bar{\nu}) - \rho(0)]$ and $\underline{\kappa} = D[1 - \sigma + \sigma\rho(\underline{\nu}) - \rho(0)]$, where $\underline{\kappa} < \tilde{\kappa}$. If $\kappa > \tilde{\kappa}$ or $\kappa < \underline{\kappa}$, I always selects the charismatic leader, $\nu = \bar{\nu}$.*

Putting aside any effect of charisma on party control, I 's utility is strictly rising in the charisma of its leader. Charismatic leaders increase the likelihood of the party's reelection in any subgame where democracy survives (which occurs with positive probability at all points in the game). If $\kappa > \tilde{\kappa}$, the party maintains control over their leader even when he is charismatic. Since the opportunity cost of replacing a charismatic leader ($\rho(\bar{\nu}) - \rho(0)$) is always greater than that of replacing a non-charismatic one ($\rho(\underline{\nu}) - \rho(0)$), when the party chooses to sanction a charismatic type it will also choose to sanction a non-charismatic type. Hence for $\kappa > \tilde{\kappa}$, the party can rein in a leader regardless of their charisma. Given this, the party has nothing to lose and stands to receive an electoral boost from selecting a charismatic leader. Therefore, in equilibrium, $c = \bar{\nu}$.

Contrastingly, if $\kappa < \tilde{\kappa}$, horizontal accountability fails regardless of leader charisma. In that case, a backsliding attempt is inevitable. However, this attempt will not always succeed in subverting future elections and so the party has nothing to lose, and may stand to gain, from selecting a leader who provides the greatest electoral boost—a charismatic one. As before, in equilibrium, $c = \bar{\nu}$.

For realizations of $\kappa \in [\tilde{\kappa}, \tilde{\kappa}]$, I is able to control an uncharismatic leader but not a charismatic one. It therefore faces a trade-off. If it nominates a charismatic leader it will face backsliding, but if democracy endures this threat, I will be more electorally competitive. On the other hand, if I selects a non-charismatic leader, he will never attempt to subvert democratic rule, but will provide a lower return in an ensuing election. The magnitude of this trade-off will be dictated by κ . We define a value of κ , denoted $\hat{\kappa}$, such that when $\kappa > \hat{\kappa}$, the incumbent party opts for the uncharismatic leader; otherwise, it selects a charismatic one.

Lemma 2. *Define $\hat{\kappa} \equiv D[1 - \sigma + \sigma\rho(\bar{\nu}) - \rho(\underline{\nu})]$. If $\hat{\kappa} > \tilde{\kappa}$, then for $\kappa \in [\hat{\kappa}, \tilde{\kappa}]$, I nominates a non-charismatic leader ($\nu = \underline{\nu}$) in equilibrium, the leader never engages in backsliding, and democracy is maintained. For $\kappa \in [\tilde{\kappa}, \hat{\kappa})$, I nominates a charismatic leader ($\nu = \bar{\nu}$) who engages in backsliding with certainty. Democracy survives with probability σ . If $\hat{\kappa} < \tilde{\kappa}$, then I nominates a non-charismatic type in the region $[\tilde{\kappa}, \tilde{\kappa}]$ who never engages in backsliding and democracy is maintained.*

It is always the case that $\hat{\kappa} < \tilde{\kappa}$. Hence, there is always a portion of the parameter space wherein I opts to preserve its control—and democracy—by deemphasizing charisma. In stable democracies, $\hat{\kappa}$ is likely to be greater than $\tilde{\kappa}$, and so interior to $(\tilde{\kappa}, \tilde{\kappa})$. (Proposition 2 formalizes the relationship between the willing abandonment of party control and democratic consolidation.) In that case, I chooses to forgo the possibility of control and selects a charismatic leader for $\kappa \in [\tilde{\kappa}, \hat{\kappa})$, despite the fact that it would be able to avoid democratic backsliding by nominating a non-charismatic leader. Thus the model depicts a situation in which parties willingly sacrifice democratic stability for the electoral returns that charismatic leaders offer.

Comparative Statics

We turn our attention to the empirical implications of the model. Before turning to comparative statics, we examine empirically useful features of the equilibrium.

Remark 1. *Backsliding takes place in equilibrium ($a = 1$) if and only if $\kappa < \max\{\tilde{\kappa}, \hat{\kappa}\}$. In such settings, I will not sanction a charismatic leader, regardless of their actions in office. For such realizations of κ , I always nominates a charismatic leader ($\nu = \bar{\nu}$).*

Remark 1 characterizes an equilibrium relationship between leader charisma, horizontal accountability, and backsliding. In the model, horizontal accountability fails and democracy backslides (and possibly breaks down) only under a charismatic leader. This relationship is in part driven by selection, since the decision to elevate a charismatic (rather than a non-charismatic) politician to a leadership position lies with the party—charisma is thus (in part) endogenous.

Moreover, having a charismatic leader correlates positively with the failure of horizontal accountability—party personalization—but this correlation is not one to one. As defined in Lemma 1, whenever $\kappa > \tilde{\kappa}$, the party selects a charismatic leader but it does not become personalized since it maintains control over the leader and democracy survives. However, when $\kappa < \max\{\tilde{\kappa}, \hat{\kappa}\}$, horizontal accountability breaks down in equilibrium and the party becomes personalized, choosing a charismatic leader whom it subsequently does not sanction for anti-democratic actions. Therefore, the presence of a charismatic leader does not imply a breakdown of horizontal accountability and a threat to democracy; however, horizontal accountability only breaks down in the presence of a charismatic leader.

When $\kappa \in [\tilde{\kappa}, \hat{\kappa}]$, charisma leads directly to the party's loss of control (and thus its personalization). The party would be able to control an uncharismatic leader, yet it chooses to forgo this control in pursuit of the electoral advantage conferred by charisma. Thus, in equilibrium, the party abandons its gate-keeping role. When $\kappa < \tilde{\kappa}$, however, it is the inability to control what leads to selection of a charismatic candidate: the party cannot exercise control regardless of

leader charisma, and it chooses a charismatic leader simply to boost its electoral chances should the backsliding attempt fail to bring about a complete democratic collapse. Thus, charismatic leaders arise in equilibrium under strong control ($\kappa > \tilde{\kappa}$), lack of control ($\kappa < \tilde{\kappa}$), and in the intermediate range ($\kappa \in [\tilde{\kappa}, \hat{\kappa}]$). However, charismatic leaders only give rise to personalization and backsliding in equilibrium in the latter two cases.

Either because charisma leads to loss of control or vice-versa, party personalization gives rise to the possibility of democratic backsliding and (with probability $1 - \sigma$) the wholesale collapse of the democratic regime. Hence, the rise of charismatic leaders is associated with democratic backsliding.

We now turn our attention to comparative static predictions from the model, beginning with a discussion of partisan polarization (D). This parameter governs, in part, the cost to the party of ousting L . If it removes the leader, I will find itself disadvantaged in the subsequent election ($\rho(0) < \rho(\nu)$), increasing the risk that the opposition comes to power. The greater the polarization between I and O , the more reluctant the party will be to take any steps that increase the likelihood of this outcome. This means that I 's ability to credibly sanction falls. Both $\tilde{\kappa}$ and $\hat{\kappa}$ rise. Also, the increased fear of electoral failure increases I 's willingness to risk democratic backsliding for electoral gain. That is, I 's willingness to engage in *ex ante* screening falls. This is captured as a rise in the value of $\hat{\kappa}$. Therefore, all three thresholds $\{\tilde{\kappa}, \hat{\kappa}, \tilde{\kappa}\}$ rise in D . Recall that the incumbent party becomes personalized (and backsliding takes place) whenever $\kappa < \max\{\tilde{\kappa}, \hat{\kappa}\}$. The probabilities of incumbent party personalization, democratic backsliding, and autocratic reversion are thus all rising in polarization. To confer these intuitions, we present graphs of the parameter space for different values of polarization in Figure 1(a).

Interestingly, the model also gives rise to an interactive relationship between polarization and charisma. This effect, however, is associational, since charisma is endogenous. We can characterize the *ex ante* probability of a backsliding episode given that a charismatic leader is chosen:

$$Pr(a = 1 | \nu = \bar{\nu}) = \frac{F_{\kappa}(\max\{\tilde{\kappa}, \hat{\kappa}\})}{1 - F_{\kappa}(\tilde{\kappa}) + F_{\kappa}(\max\{\tilde{\kappa}, \hat{\kappa}\})}. \text{ (Notice that } Pr(a = 1 | \nu = \underline{\nu}) = 0, \text{ from Remark 1.) As}$$

noted above, higher levels of polarization raise all three thresholds in κ , $\{\tilde{\kappa}, \tilde{\kappa}, \hat{\kappa}\}$. Because the stakes of winning are higher for the party, it is less able to control any leader, and the probability that it chooses a charismatic candidate because it knows it will remain in control ($1 - F_{\kappa}(\tilde{\kappa})$) declines. But, because the stakes are high, it is more willing to nominate a charismatic candidate it knows it cannot control— $F_{\kappa}(\max\{\tilde{\kappa}, \hat{\kappa}\})$ rises. Thus, $Pr(a = 1|\nu = \bar{\nu})$ rises in polarization, as does the marginal association between charisma and backsliding (and democratic collapse), given by $Pr(a = 1|\nu = \bar{\nu}) - Pr(a = 1|\nu = \underline{\nu})$.

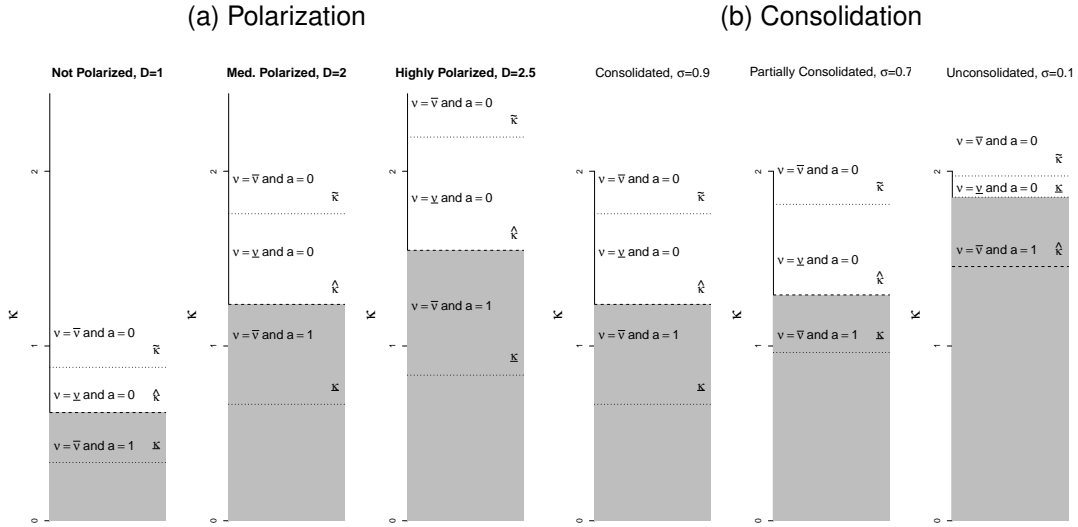
Proposition 1. (a) *The thresholds $\{\tilde{\kappa}, \hat{\kappa}, \tilde{\kappa}\}$ are all rising in polarization, as is the probability of democratic backsliding and democratic breakdown.*

(b) *The conditional probability of backsliding given the selection of a charismatic leader $Pr(a = 1|\nu = \bar{\nu})$ is rising in polarization. The probability of backsliding given the selection of a non-charismatic leader is fixed and equal to zero for all values of polarization. Therefore, the relationship between leader charisma and backsliding, given by $Pr(a = 1|\nu = \bar{\nu}) - Pr(a = 1|\nu = \underline{\nu})$, is rising in D .*

We turn next to the level of democratic consolidation (σ). By assumption, democratic consolidation plays a direct role in the probability of democratic breakdown; we assume that backsliding leads to democratic breakdown with probability $1 - \sigma$. But this parameter plays a more subtle role in determining the extent of horizontal accountability. Specifically, the incumbent party and its leader have a shared interest insofar as each wishes any attempt at backsliding to be a success. Thus, I 's distaste for backsliding is lower when the leader's anti-democratic actions are more likely to be successful. This is reflected in the thresholds $\{\tilde{\kappa}, \hat{\kappa}, \tilde{\kappa}\}$, all of which are falling in consolidation.

Democratic consolidation also has a more nuanced effect through its influence on the party's gate-keeping role. Specifically, as democracy becomes fragile ($\sigma \rightarrow 0$) the range of κ for which the party chooses to nominate an uncharismatic politician collapses to the empty set. As democ-

Figure 1: Comparative Statics



Plots of the parameter space where $\bar{\nu} = 2$, $\underline{\nu} = 0.3$, and $\rho(\cdot)$ is a gamma density with shape and scale parameters 1. Plots to the left depict three different values of polarization. Plots on the right depict three different levels of consolidation. $\sigma = 0.9$ for the polarization plots, and $D = 2$ for the consolidation plots. Shading depicts regions in κ -space where backsliding ($a = 1$) takes place in equilibrium.

racy grows more unstable, democratic collapse grows more likely following a backsliding attempt. If democracy collapses, free and fair elections cease—the incumbent party is entrenched in power, regardless of their leader’s charisma. So, as stability declines, the opportunity cost to the party of removing a charismatic leader converges to the cost of removing an uncharismatic leader—both go to zero. The party is thus equally (un)willing to control either type, and so more inclined to set $c = \bar{\nu}$. Figure 1(b) plots the intuition in the relevant κ -space.

An expansion of the range of κ for which an uncharismatic type is selected does not directly imply that the probability of selecting an uncharismatic leader rises in democratic consolidation.¹² Notice that $\tilde{\kappa}$ is also falling in σ , and the party always selects (and maintains control over) a charismatic leader when $\kappa > \tilde{\kappa}$. The overall effect on the probability of selecting a charismatic versus a non-charismatic leader depends on which effect dominates. That said, we can unam-

¹²Though, the fact that as $\sigma \rightarrow 0$ this space collapses does imply that this probability must be rising for some realizations of σ .

biguously say that the probability that the incumbent party selects an uncharismatic leader is higher in relatively consolidated democracies than in the most unstable ones.

Proposition 2. (a) *All three thresholds $\{\kappa, \hat{\kappa}, \tilde{\kappa}\}$ are falling in democratic consolidation. Hence, the probability of party personalization, democratic backsliding, and autocratic reversion are falling in σ .*

(b) *The range of values of κ for which I chooses a non-charismatic leader is expanding in democratic consolidation.*

(c) *$\hat{\kappa} - \tilde{\kappa}$ is increasing in consolidation. In other words, if a backsliding episode takes place in a consolidated democracy, it is more likely to result from the party's deliberate abandonment of its gate-keeping role than if a backsliding episode takes place in an unconsolidated democracy.*

(d) *Conditional on the party being able to control an uncharismatic but not a charismatic leader, the party 'gives up' and nonetheless selects a charismatic leader for a wider range of values of κ as consolidation increases. $\frac{\tilde{\kappa} - \max\{\hat{\kappa}, \tilde{\kappa}\}}{\tilde{\kappa} - \kappa}$ is falling in σ .*

Notice further from Figure 1(b) that another region expands as σ rises. Of the three thresholds in κ , it is $\tilde{\kappa}$ which is most sensitive to σ . For values of κ below this threshold, the party cannot credibly control either charismatic or non-charismatic leaders. Given this, it may as well opt to select a charismatic type, who at least conveys an advantage at the polls.

In highly unconsolidated regimes we have $\tilde{\kappa} > \hat{\kappa}$, meaning that a party that selects a charismatic leader and experiences backsliding would have experienced similar outcomes even if the leader had been non-charismatic. In that case, the association between charismatic leadership and democratic backsliding is driven entirely by selection.

As consolidation rises, however, $\hat{\kappa}$ begins to exceed $\tilde{\kappa}$, opening up a region $[\tilde{\kappa}, \hat{\kappa}]$ of voluntary surrender of party control. As Proposition 2 states, the region $\hat{\kappa} - \tilde{\kappa}$ is expanding as consolidation

rises. Thus, if one sees a backsliding episode in a consolidated democracy, it is more likely to be the result of a party's voluntary abandonment of its gate-keeping role rather than that party's inability to rein in its leadership.

In other words, in consolidated democracies, parties that could control an uncharismatic leader are willing to gamble on nominating a charismatic leader they cannot control for a wider range of values of κ . One can interpret this as parties behaving with the understanding that they have a safety net—greater democratic consolidation. Because such parties anticipate that elections will likely continue regardless of the leader's anti-democratic behavior, they are willing to bank on the electoral benefits of a leader's charisma in the future. This echoes a finding in a model extension discussed below (and presented in the Supplementary Appendix), where for some parameter values parties may completely offload the responsibility for restraining leaders to the electorate.

In Section A.4 of the Supplementary Appendix, we also examine the effects of increasing the electoral returns to charisma, $\rho(\bar{v})$. Such returns may increase, for instance, through new technologies like social media which may facilitate direct connections with followers. Higher returns to charisma expand the range for which the party voluntarily surrenders control. We also investigate the effect of changes in $\rho(0)$, which may be interpreted as a baseline level of popularity for the party, even after a leadership fight breaks out. As this popularity rises, both $\tilde{\kappa}$ and $\tilde{\kappa}$ shift down. The net effect on the probability of nominating a charismatic leader is ambiguous but, conditional on selecting such a leader, a party that is popular in its own right is more likely to keep him in control and less likely to experience backsliding. However, $\hat{\kappa}$ is invariant to $\rho(0)$. This implies that for some parties, an increase in popularity may enable them to rein in a recalcitrant leader who is uncharismatic, but in equilibrium they willingly forgo this possibility and opt to nominate a charismatic leader who threatens democracy instead. This portion of the parameter space expands with party popularity.

Model Extensions

We have expressed the comparative statics in terms of κ —a cost incurred by members of the incumbent party from not checking their leader’s authoritarian actions, regardless of whether those actions result in an *autogolpe* or not. This interpretation, however, may not capture the full extent of costs that a party may suffer as a result of democratic backsliding. In the Supplementary Appendix, we explore model extensions with additional cost parameterizations. Perhaps most significantly, these extensions allow us to examine a potential trade off between vertical and horizontal accountability.

As a reduced form means to capture vertical accountability, we consider an electoral penalty associated with participating in backsliding. This penalty arises whenever a backsliding episode takes place ($a = r = 1$) but democracy survives ($B = 0$). We assume that, in the ensuing election, I ’s probability of winning (with the backsliding leader L at the helm) is given by $[1 - ar\lambda]^\rho(\nu)$, where $\lambda \in (0, 1)$ represents the extent of vertical accountability—the propensity of the electorate to hold incumbents accountable for violations of democratic norms. This cost is incurred by both I and L .

The results with this cost structure are directly analogous to those in our baseline model. For an electoral penalty above some threshold in λ , backsliding never takes place and charismatic leaders are always selected, regardless of κ . This happens because L prefers not to violate the democratic equilibrium, with or without horizontal accountability. That is, vertical accountability (λ) is sufficiently strong that the role of the party as a check on the leader is superfluous. Therefore, I can always choose a charismatic leader with no repercussions. For lower levels of λ , the model results remain the same as in our baseline model, with the additional insight that threshold values of κ fall as λ increases—horizontal and vertical accountability act as substitutes.

Empirically, these insights suggest that our comparative static predictions should hold in all settings where vertical accountability is insufficiently strong. Existing results on the fragility of

electoral checks on backsliding, even when voters hold pro-democracy beliefs (e.g. [Svolik, 2020](#)), suggest the possibility that this will be true in a broad range of cases.¹³

We further consider a possibility wherein party members also suffer a cost $\chi > 0$ from living under autocratic rule. That is, party elites gain from having their preferred type of leader in power, but lose from having to cope with strongman rule. This cost may, for example, stem from the risk of being politically persecuted without due process. χ is only incurred in the event of democratic breakdown ($B = 1$).

The introduction of this cost may alter some of our predictions. Specifically, while the results with respect to polarization remain unchanged, those with respect to democratic consolidation cease to hold for a sufficiently high value of χ . Namely, in the baseline model, a decrease in democratic consolidation (σ) increases the likelihood that I enjoys the policy benefits of a dictatorship under their chosen leader following a backsliding episode. Thus, parties in weak democracies are less likely to forestall such an episode. However, if a party suffers some cost that outweighs these policy gains under autocracy, the reverse phenomenon holds. Knowing that backsliding is likely to lead to strongman rule, parties in less consolidated democracies will be particularly motivated to avoid this cost both through horizontal accountability and by selecting uncharismatic leaders.

Whether partisan elites are more likely to value the policy gains from authoritarian rule or fear its potential costs (beyond the reputational and moral costs captured in the parameter κ) is ultimately an empirical question. Our empirical results, presented below and in the Supplementary Appendix, are consistent with our baseline model, suggesting that partisan elites are motivated more by reputational, moral, and electoral concerns about backsliding than the fear of life under strongman rule.

¹³Charisma may also weaken vertical accountability directly, since (some) voters may be more forgiving of charismatic than uncharismatic leaders ([Merolla and Zechmeister, 2011](#)).

Empirical Illustration

We now illustrate how our theoretical predictions can inform the empirical study of democratic backsliding. To capture democratic breakdown (taking place with probability $(1-\sigma)F_{\kappa}(\max\{\tilde{\kappa}, \hat{\kappa}\})$), we use the binary variable *Democracy* from Boix, Miller and Rosato (2013, updated through 2020), indicating a minimal level of suffrage and free and fair elections. An exit in a given country-year from *Democracy* = 1 represents democratic collapse. Given our theory, we focus on the hazard that a democracy spell ends through *autogolpe*.¹⁴

Regime transitions are extreme, and rare, events.¹⁵ To capture the less drastic antidemocratic actions (with probability $F_{\kappa}(\max\{\tilde{\kappa}, \hat{\kappa}\})$ in the model), we use the *Liberal Democracy Index* from the Varieties of Democracy (V-Dem) project (Coppedge et al., 2022). The index measures constraints on the executive—protection of civil liberties, the rule of law, and checks and balances.

The precondition for breakdown or backsliding in our theory is either the strategic acquiescence or abdication by the ruling party to antidemocratic actions of their chief executive (again, with probability $F_{\kappa}(\max\{\tilde{\kappa}, \hat{\kappa}\})$). We proxy for this party-level backsliding with the *Anti-Pluralism Index* from the Varieties of Party Identity and Organization (V-Party) dataset (Lindberg et al., 2022), indicating the degree of parties' "lacking commitment to democratic norms prior to elections." Since party-level backsliding entails accommodating the leader's will, we use the *Personalization of Party* variable from V-Party as another plausible measure mapping onto the same theoretical quantity. This variable indicates the degree to which parties are "a vehicle for the personal will and priorities of one individual leader."

Guided by Remark 1 and Propositions 1 and 2, we examine the relationship between these variables and proxies for leader charisma (ν), ideological polarization (D), and democratic stabil-

¹⁴For breakdown to be considered an autogolpe, we require that the leader remain the same before and after the collapse, as coded by Goemans, Gleditsch and Chiozza (2009) and other sources.

¹⁵Only 1.64 percent of country-year observations in our data involve a democratic breakdown, and 0.45% involve an autogolpe.

ity (σ). To proxy for leader charisma, we use the *Person of the Leader* variable from V-Dem, indicating the degree to which the chief executive is “portrayed as being endowed with extraordinary personal characteristics and/or leadership skills.” This measure echoes Weber’s foundational definition of charismatic authority discussed above. We show in the Supplementary Appendix Section B.3 that this proxy has face validity, indicating higher appeal for leaders we mentioned in the paper and the appendix as having high charisma (Figure B3). It also strongly correlates with parties’ vote shares, capturing the popularity that charismatic leaders tend to enjoy (and is thus unlikely to only capture institutional features that make regimes personality-centric). That said, we also show similar results with two alternative proxies: (a) a leader’s outsider status, reflecting that charisma often leads to a meteoric rise in politics (Section B.4), and (b) the (inverse of the) degree to which party linkages with voters focus on programmatic policies as opposed to more personalized appeals (Section B.5).

We further use *Political Polarization*, also from V-Dem, indicating the degree to which “supporters of opposing political camps are reluctant to engage in friendly interactions,” to capture the disutility to one partisan group from the possibility of victory by another political camp. In Supplementary Appendix Section B.6, we show similar results with an alternative measure based on the spatial distances on parties’ policy positions. Finally, to capture democratic stability, we use the length of the ongoing spell of democracy based on data from Boix, Miller and Rosato (2013).

When examining democratic breakdown and backsliding, our sample consists of all democratic-spell years between 1950 and 2020.¹⁶ For breakdown, the unit of observation is country-year; for backsliding, it is the leader spell. When examining party anti-pluralism and personalization, our sample consists of democratic-spell-election years between 1970 and 2020 (limited to governing parties).

¹⁶Since multiple leaders may serve in the same calendar year, we keep only the longest-serving leader in that year. The exceptions are leaders of caretaker governments, who are eliminated even if they served longer than the permanent outgoing or incoming leader in that year.

Hypotheses

Our theoretical predictions can be translated into the following three empirical implications:

Hypothesis 1 (From Remark 1). *More charismatic leaders (higher values of ν) are associated with (a) a higher probability of autogolpe; (b) lower quality of democracy; (c) an increase in the governing parties' anti-pluralism; and (d) greater governing party personalization.*

Hypothesis 2 (From Proposition 1(b)). *The associations in hypothesis 1 are more pronounced when partisan polarization is high (higher values of D).*

Hypothesis 3 (From Proposition 2(b)). *More stable democracies (higher values of σ) have fewer charismatic leaders.¹⁷*

As outlined in our model, the links between leader charisma and our outcomes may arise through two channels. The first is direct: it is because of a leader's charisma that an incumbent party strategically acquiesces to the leader's anti-democratic actions and becomes personalized. The second occurs via selection: backsliding and personalization are both a byproduct of a party that is unable to control any type of leader in the first place, and chooses a charismatic one to at least boost its electoral prospects. While it is difficult to empirically disentangle the two channels, they both imply the presence of clear and testable associations, which we evaluate next. Naturally, even if we find empirical associations consistent with our theoretical predictions, it is possible that such associations arise because of confounders rather than our hypothesized mechanisms. Although we do attempt to account for potentially confounding factors in parametric specifications in the Supplementary Appendix, the inability to firmly distinguish pre-treatment from post-treatment variables and the possibility of unobserved confounders poses limitations to this approach. Our results are thus meant to be illustrative rather than conclusive.

¹⁷These are not the only implications arising from our model, but they are novel and refer to leader charisma. Our empirical illustrations, however, will accommodate the other predictions from Propositions 1 and 2; we briefly discuss them in the Supplementary Appendix.

Results

Figure 2 shows descriptively the associations predicted by hypothesis 1. Panel (a) plots the probability of a democracy surviving past time t (indicated on the horizontal axis), separately for leaders with below-average charisma per our proxy (dashed line) and above-average charisma (solid line). While autocratic reversions are rare (the probability of survival is mostly above 50 percent), democratic survival is visibly lower under more charismatic leaders.

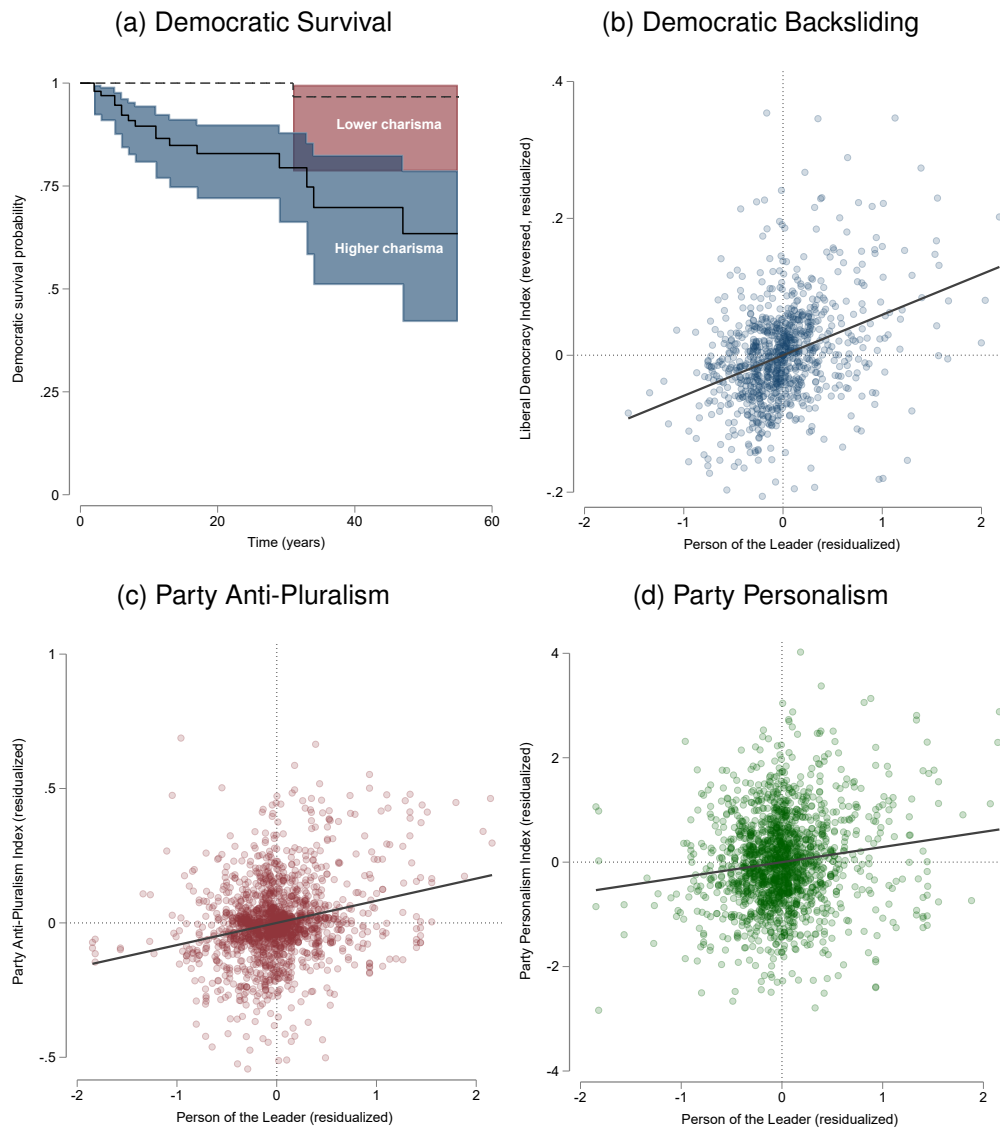
Panels (b)-(d) plot the association between leader charisma (x -axis) and, respectively, democratic backsliding (the reversed Liberal Democracy Index, so that higher values indicate poorer-quality democracies), party anti-pluralism, and party personalism. All three panels plot residuals after netting out country and year fixed effects; linear best-fit lines are also shown. While there is a fair amount of variability in each panel, the raw data indicate plausibly positive associations, as predicted.

In Table B1 in the Supplementary Appendix Section B, we present parametric analyses that complement the descriptive evidence presented in the plots, where we include several other potentially important correlates identified by the prior literature. The regression results conform with the descriptive patterns in Figure 2: higher leader charisma is statistically significantly associated with all four outcomes—higher hazard rate of an autogolpe, lower quality of democracy, greater incumbent party illiberalism, and higher party personalism.

Figure 3 assesses descriptively our expectation in hypothesis 2, by plotting the association between democratic backsliding and leader's charisma (both net of country and year fixed effects) separately in polities with below-average polarization (panel a) and above-average polarization (panel b).¹⁸ The patterns in the raw data appear consistent with our expectation: there is a

¹⁸Similar figures for party anti-pluralism and personalization are shown in Supplementary Appendix Section B.1. We cannot evaluate hypothesis 2 with respect to democratic survival given the rarity of autogolpe in the data.

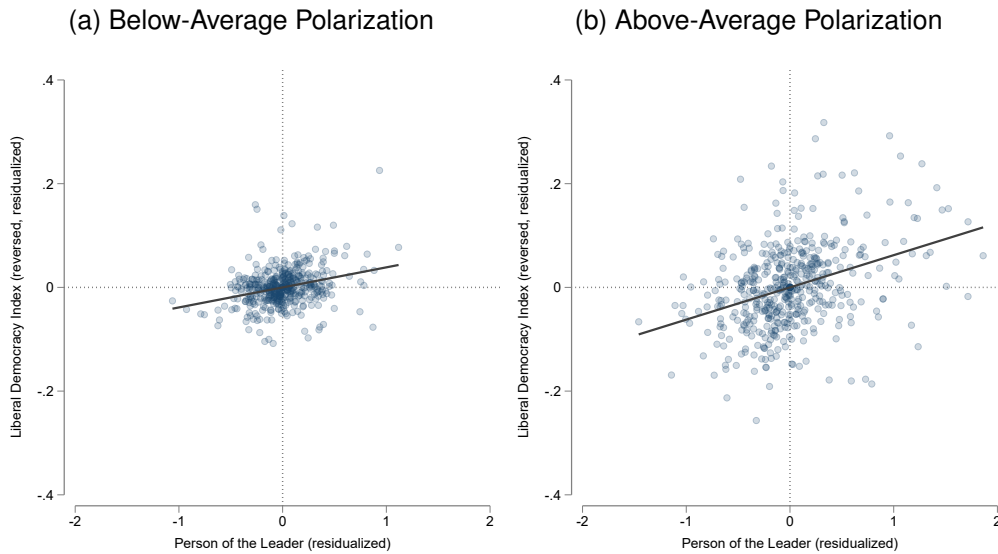
Figure 2: Leader Charisma and Democratic Survival, Backsliding, Anti-Pluralism, and Personalism



stronger positive association between leader's charisma and backsliding when partisan polarization is higher.

Table B2 in Section B of the Supplementary Appendix shows the results of parametric analyses where the models are augmented by an interaction between leader charisma and polarization. The results are in line with hypothesis 2 in that higher polarization statistically significantly magni-

Figure 3: Leader Charisma and Backsliding, by Political Polarization

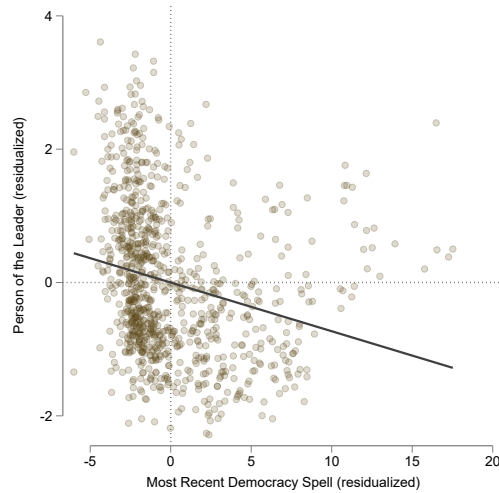


fies the positive association between the leader charisma for both the polity-level and party-level backsliding (the results are less clear for party personalism, as the interaction term estimate is positive but not statistically significant).

Finally, Figure 4 evaluates descriptively the raw patterns with respect to hypothesis 3, plotting the *Person of the Leader* variable (*y*-axis) against the length of the most recent spell of democracy as a proxy for stability (in years, divided by 10). The figure plots the residuals after accounting for year fixed effects. The raw patterns are again consistent with our expectation: per our charisma proxy, more durable democracies have on average less charismatic leaders. Table B3 in the Supplemental Appendix presents the accompanying parametric results, which show a statistically significant negative association between charisma and democratic stability.

In sum, while only illustrative, the patterns in the data are consistent with the predictions laid out in hypotheses 1-3. Both democratic breakdown and democratic backsliding are more likely under more charismatic chief executives, and as predicted, these associations are accompanied by greater illiberalization and personalization of incumbent parties. These patterns are more pronounced in more polarized societies, and charismatic leaders are more likely in less stable

Figure 4: Democratic Stability and Leader Charisma



democracies.

Conclusion

When political parties advance charismatic candidates, they may increase their short-term electoral success. However, such short-term gains may come at a (potentially longer-term) cost. Leader charisma can lower the parties' willingness to constrain their behavior, weakening an important institutional check on the potential abuse of power. In this way, even if charismatic leaders are no more authoritarian than their less charismatic counterparts, they will be given more leeway to engage in democratic subversion should they be so inclined.

Parties anticipate this ex post challenge to horizontal accountability, and may be willing to screen out charismatic politicians from leadership positions ex ante. However, such gate-keeping behavior does not always take place in our theoretical model. Parties that are committed to continued democratic rule have no need to rely on such screening, and can be sure that the shadow of party discipline is sufficient to keep even charismatic demagogues in line. For parties that are unable to keep even the uncharismatic politicians in check, screening out charismatic

leaders provides no benefit.

However, sometimes parties capable of exerting control over their leaders voluntarily forgo their gate-keeping role knowing full well that this poses a risk to the democratic order. In equilibrium, such parties gamble with the fate of democracy in order to secure the electoral gains that the charismatic appeal of their leader provides. In our model, such behavior is most common in consolidated democracies where elites are highly polarized. (Contrastingly, in unconsolidated democracies, few parties are committed enough to exercise control with or without ex-ante screening.) This finding emphasizes a novel threat to polities often considered relatively insulated from regime instability. We believe that these risks warrant further research on the causes and consequences of charismatic politicians in democratic politics.

References

- Alberta, Tim. February 2021. “Nikki Haley’s Time for Choosing.” *The Politico Magazine* . <https://www.politico.com/interactives/2021/magazine-nikki-haleys-choice/> (accessed February 26, 2025).
- Andrews-Lee, Caitlin. 2021. *The emergence and revival of charismatic movements: Argentine Peronism and Venezuelan Chavismo*. Cambridge University Press.
- Berman, Sheri and Maria Snegovaya. 2019. “Populism and the decline of social democracy.” *Journal of Democracy* 30(3):5–19.
- Boix, Carles, Michael Miller and Sebastian Rosato. 2013. “A Complete Data Set of Political Regimes, 1800–2007.” *Comparative Political Studies* 46(12):1523–1554.
- Campbell, John. February 15, 2018. “Finally, Jacob Zuma Resigns as President of South Africa.” *Foreign Affairs* . <https://www.cfr.org/blog/finally-jacob-zuma-resigns-president-south-africa> (accessed 02/12/2025).

- Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerløw, Adam Glynn, Sandra Grahn, Allen Hicken, Katrin Kinzelbach, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Pamela Paxton, Daniel Pemstein, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Eitan Tzelgov, Luca Uberti, Yi-ting Wang, Tore Wig and Daniel Ziblatt. 2022. "VDEM [Country–Year/Country–Date] Dataset v12." Varieties of Democracy (V-Dem) Project. <https://doi.org/10.23696/vdemds22>.
- Cox, Gary W. and Matthew D. McCubbins. 2007. *Legislative Leviathan: Party Government in the House*. Cambridge University Press.
- Dewan, Torun and Kenneth A Shepsle. 2011. "Political economy models of elections." *Annual Review of Political Science* 14(1):311–330.
- Frantz, Erica, Andrea Kendall-Taylor and Joseph Wright. 2024. *The Origins of Elected Strongmen: How Personalist Parties Destroy Democracy from Within*. Oxford University Press.
- Goemans, Hein E., Kristian Skrede Gleditsch and Giacomo Chiozza. 2009. "Introducing Archigos: A Data Set of Political Leaders." *Journal of Peace Research* 46(2):269–283.
- Grillo, Edoardo, Zhaotian Luo, Monika Nalepa and Carlo Prato. 2023. "Theories of Democratic Backsliding." *Annual Review of Political Science* 27.
- Grzymala-Busse, Anna. 2019. "How populists rule: The consequences for democratic governance." *Polity* 51(4):707–717.
- Haggard, Stephan and Robert Kaufman. 2021. *Backsliding: Democratic Regress in the Contemporary World*. Cambridge, UK: Cambridge University Press.
- Helmke, Gretchen, Mary Kroeger and Jack Paine. 2022. "Democracy by deterrence: Norms, constitutions, and electoral tilting." *American Journal of Political Science* 66(2):434–450.

- Hollyer, James R., Marko Klačnja and Rocío Titiunik. 2022. "Parties as Disciplinarians: Charisma and Commitment Problems in Programmatic Campaigning." *American Journal of Political Science* 66(1):75–92.
- Howell, William G., Kenneth A. Shepsle and Stephane Wolton. 2023. "Executive absolutism: the dynamics of authority acquisition in a system of separated powers." *Quarterly Journal of Political Science* 18(2):243–275.
- Kidwai, Rasheed. June 8, 2019. "How Indira Faced Crises in Congress." *The Tribune*. <https://www.tribuneindia.com/news/archive/comment/how-indira-faced-crises-in-congress-784804/> (accessed June 23, 2025).
- Levitsky, Steven and Daniel Ziblatt. 2018. *How Democracies Die*. Crown.
- Lindberg, Staffan I., Nils Düpont, Masaaki Higashijima, Yaman Berker Kavasoglu, Kyle L. Marquardt, Michael Bernhard, Holger Döring, Allen Hicken, Melis Laebens, Juraj Medzihorsky, Anja Neundorf, Ora John Reuter, Saskia Ruth-Lovell, Keith R. Weghorst, Nina Wiesehomeier, Joseph Wright, Nazifa Alizada, Paul Bederke, Lisa Gastaldi, Sandra Grahn, Garry Hindle, Nina Ilchenko, Johannes von Römer, Steven Wilson, Daniel Pemstein and Brigitte Seim. 2022. "Varieties of Party Identity and Organization (V-Party) Dataset V2." Varieties of Democracy (V-Dem) Project. <https://doi.org/10.23696/vpartydsv2>.
- Luo, Zhaotian and Adam Przeworski. 2023. "Democracy and its Vulnerabilities: Dynamics of Democratic Backsliding." *Quarterly Journal of Political Science* 18(1):105–130.
- Mainwaring, Scott and Mariano Torcal. 2006. "Party system institutionalization and party system theory after the third wave of democratization." *Handbook of party politics* 11(6):204–227.
- Malhotra, Inder. 1991. *Indira Gandhi: A Personal and Political Biography*. Hay House.

- McCormick, John. October 5, 2021. "Nikki Haley Embraces Trump in Her Version of GOP Future." *The Wall Street Journal* . <https://www.wsj.com/articles/nikki-haley-embraces-trump-in-her-vision-of-gop-future-11633424400?mod=e2twp> (accessed February 26, 2025).
- Merolla, Jennifer L. and Elizabeth J. Zechmeister. 2011. "The Nature, Determinants, and Consequences of Chávez's Charisma: Evidence from a Study of Venezuelan Public Opinion." *Comparative Political Studies* 44(1):28–54.
- Miller, Michael K. 2021. "A republic, if you can keep it: Breakdown and erosion in modern democracies." *The Journal of Politics* 83(1):198–213.
- Müller, Jan-Werner. 2016. *What is populism?* University of Pennsylvania Press.
- Plys, Kristin Victoria Magistrelli. 2020. *Brewing Resistance: Indian Coffee House and the Emergency in Postcolonial India*. Cambridge University Press.
- Schattschneider, E. E. 1942. *Party Government*. Holt, Rinehart and Winston.
- Stokes, Donald E. 1963. "Spatial models of party competition." *American political science review* 57(2):368–377.
- Svolik, Milan W. 2020. "When Polarization Trumps Civic Virtue: Partisan Conflict and the Subversion of Democracy by Incumbents." *Quarterly Journal of Political Science* 15(1):3–31.
- The New York Times. January 4, 1978. "Congress Party Ousts Mrs. Gandhi." *The New York Times* . <https://www.nytimes.com/1978/01/04/archives/congress-party-ousts-mrs-gandhi.html> (accessed June 23, 2025).
- Tucker, Robert C. 1968. "The theory of charismatic leadership." *Daedalus* 97(1):731–756.

- Urbinati, Nadia. 2019. "Political theory of populism." *Annual review of political science* 22(1):111–127.
- Waldner, David and Ellen Lust. 2018. "Unwelcome change: Coming to terms with democratic backsliding." *Annual Review of Political Science* 21(1):93–113.
- Weber, Max. 1922 [1978]. *Economy and Society*. In *Vols. 1 and 2*, ed. G. Roth and C. Wittich. University of California press.
- Weyland, Kurt. 2001. "Clarifying a Contested Concept: Populism in the Study of Latin American Politics." *Comparative Politics* 34(1):1–22.
- Willner, Ann Ruth. 1985. *The spellbinders: Charismatic political leadership*. Yale University Press.
- Willner, Ann Ruth and Dorothy Willner. 1965. "The Rise and Role of Charismatic Leaders." *The Annals of the American Academy of Political and Social Science* 358(1):77–88.
- Worthen, Molly. 2025. *Spellbound: How Charisma Shaped American History from the Puritans to Donald Trump*. Forum Books.
- Zengerle, Patricia and Richard Cowan. June 17, 2022. "Trump Pressed, Threatened Pence to Overturn Election, Panel Hears." *Reuters* . <https://www.reuters.com/world/us/us-capitol-jan-6-panel-turns-attention-pence-thursdays-hearing-2022-06-16/> (accessed June 23, 2025).