Demagogues and the Fragility of Democracy*

Dan Bernhardt† Stefan Krasa‡ Mehdi Shadmehr§

July 27, 2021

Abstract

We investigate the susceptibility of Democracy to demagogues, studying tensions between representatives who guard voters' long-run interests and demagogues who cater to voters' short-run desires. Parties propose consumption and investment. Voters base choices on current-period consumption and valence shocks. Younger/poorer economies and economically-disadvantaged voters are attracted to the demagogue's dis-investment policies, forcing far-sighted representatives to mimic them. This electoral competition can destroy democracy: if capital falls below a critical level, a death spiral ensues with capital stocks falling thereafter. We identify when economic development mitigates this risk and characterize how the death-spiral risk declines as capital grows large.

*We are grateful for comments of Odilon Camara, Wioletta Dziuda, and the Editor, Roland Benabou, and seminar participants at Northwestern Kellogg, University of Vienna, University of Bonn, and 2019 APSA.

†University of Illinois and University of Warwick, danber@illinois.edu

‡University of Illinois, skrasa@illinois.edu

§University of North Carolina at Chapel Hill, mshadmeh@gmail.com

“The republican principle,” wrote Hamilton in Federalist No. 71, “does not require an unqualified complaisance to every sudden breeze of passion, or to every transient impulse which the people may receive from the arts of men, who flatter their prejudices to betray their interests.” To the contrary, Hamilton argued, when “the interests of the people are at variance with their inclinations, it is the duty of the persons whom they have appointed to be the guardians of those interests, to withstand the temporary delusion.... conduct of this kind has saved the people from very fatal consequences of their own mistakes, and procured...their gratitude to the men who had courage and magnanimity enough to serve them at the peril of their displeasure.” Still, if such magnanimous representatives cause too much displeasure, they would lose election to those who will implement those “temporary delusions”, paying “obsequious court to the people; commencing demagogues, and ending tyrants.”

Our paper studies the tension highlighted by Hamilton between far-sighted, magnanimous representatives who guard the long-run interests of voters and office-seeking demagogues who cater to voters’ short-run desires.[1] We characterize the long-run outcomes of democracy in a country populated by a short-sighted majority. Demagogues and short-sighted voters have long been considered inter-related vices of republican governments. For example, “Madison’s [belief] about democracy was based on [one] about human beings: man, by nature, preferred to follow his passion rather than his reason; he invariably chose short-term over long-term interests” (Middlekauff (2007), p. 678). Indeed, researchers define demagogues by this characteristic. According to Guiso et al. (2018), parties led by demagogues “champion short-term policies while hiding their long-term costs.”[2] What is not well-understood is how demagogues distort the behavior of far-sighted parties, and how a democratic country emerges from the long-run confrontation between selfish demagogues and socially benevolent, but pragmatic parties.

We analyze the dynamic political competition between a far-sighted, benevolent party that seeks to maximize voter welfare, and an office-motivated demagogue who only cares about winning. The two parties and a representative short-sighted median voter interact over time in an infinite horizon setting. To capture tensions between short- and long-term considerations, we model the political decision process as an investment problem in which parties propose how to allocate existing resources between current consumption and investment in capital that facilitates future consumption, where dis-investment of up to some fraction of the current capital stock is feasible. Voters care about a party’s investment policy and its valence, which captures

[1] As Burman et al. (2010) observe, “The basic problem is that policymakers want to make people happy, which means more spending and lower taxes...Political leaders perceive that their reelection depends on short-term results, even if the short-term expedients may be disastrous over the long term.”

[2] Historically, populists were referred to as demagogues but now these terms are often used interchangeably.

the party's non-policy attributes. The benevolent party's net valence is either high or low—even if the benevolent party is far more likely to have a higher valence, there is still a chance that the demagogue has the higher valence and wins election. The myopic median voter assesses parties based solely on current period utility. Given the proposed investments and realized valences, the voter picks the winner who implements its investment policy.

Absent a demagogue, the benevolent party acts like a social planner, internalizing voters' utility from future consumption, and capital stocks grow without bound over time. Demagogues, even those likely to have low valences, change this. Demagogues design their policies to appeal maximally to short-sighted voters. The benevolent party's dilemma is that if it ignores the demagogue in its policy choices, it imperils its electoral support, while "trying to beat a...populist insurgency by becoming one...turns out to be a fool's errand...[as it] has a huge...economic cost."[3] The policy choices of Huey Long and President Roosevelt illustrate this dilemma. In the midst of the Great Depression, Long proposed a high progressive tax, and distributing the revenue to every American family, 5,000 dollars each—supposedly enough for a home, a car, and a radio—plus shorter working hours, pensions and many other benefits. In response, FDR proposed a Second New Deal that included a Wealth Tax Act designed "to save our system" from the "crackpot ideas" of demagogues. FDR knew that his proposal was bad for the economy, but "I am fighting Communism, Huey Longism, ...," he said, indicating that the consequences of losing were worse (Kennedy (1999), p. 241-7).

We focus on CRRA utility with coefficients of relative risk aversion that exceed one. These preferences have the feature that when capital falls, the utility differences in implied consumptions associated with fixed investment rates rise. In the Markov perfect equilibrium, the demagogue proposes to prop up short-term consumption by proposing to disinvest as much as possible, while the benevolent party chooses an investment level that ensures that the median voter supports it when its valence is high.

As a benchmark, we analyze a scenario in which there is exogenous stochastic selection of the winner, so that policy choices do not affect who wins. We contrast this setting with that where the winner is endogenously determined. We show that the benevolent party proposes to invest less than it would in the benchmark if and only if the capital stock is below a threshold. When capital is above that threshold, the benevolent party proposes to invest even more than it would in the benchmark, converging to the benchmark investment level from above as capital stock grows larger. That is, the benevolent party's proposed investment, as a share of capital, is

[3] NY times Op Ed, https://www.nytimes.com/2018/07/16/opinion/to-defeat-far-right-nationalists-dont-try-to-imitate-them.html

non-monotone in the level of capital. In particular, when the capital stock is below a threshold, the benevolent party's proposed investment, as a share of capital, is increasing in capital stock.[4] In contrast, when capital stocks are sufficiently high, the benevolent party's optimal investment, as a share of capital, trends downward with capital.

The benevolent party understands that it needs to win in order to invest, and that worse outcomes would obtain were the demagogue to win. When capital, and hence proposed consumption, is low, the salience of differences in proposed consumptions is high relative to that in candidate valences—whenever voters have relative risk aversions above one. In essence, when voters are hungry, valence becomes more of a secondary concern, raising the relative attraction of the demagogue's dis-investment policy, making poorer economies more prone to the consequences of demagoguery. When capital is low, the benevolent party proposes investment rates just low enough to avoid losing when it has high valence. In contrast, once capital is sufficiently high, the benevolent party can propose the benchmark investment without losing when it has high valence. Despite this, it proposes even higher investments to insure against the risk that the demagogue may win and reduce capital stocks toward levels where its electoral competition severely constrains investment. Only when capital levels are so high that such electoral threats could only arise in the distant future do its proposed investments converge to the benchmark.

Our second set of results identify the possibility of death spirals for democracy: We identify a critical cutoff on capital (the point of no return) below which capital stocks shrink at accelerating rates toward zero. If capital ever falls below this level, then the benevolent party, itself, must propose to dis-invest to preserve a chance of winning. As a result, capital stocks shrink, forcing the benevolent party to increasingly mimic the demagogue. As capital stocks spiral downward, the benevolent party's disinvestment converges toward the demagogue's, albeit never reaching it. This death spiral is self enforcing, leading to zero capital and consumption in the long run.

This downward cycle is not a poverty trap in which people are too poor to invest, thereby perpetuating economic misery. Rather, regardless of the economy's intrinsic productivity, the death spiral is driven by the heightened electoral pressure from demagogues when capital is low. Moreover, a demagogue maximizes his chances of winning an election by maximizing spending, reducing future capital. But, reducing future capital then amplifies the utility difference between the demagogue's dis-investment policy and any fixed investment policy. Thus, by damaging capital stocks and destroying social capital associated with institutions and property rights, a demagogue raises the future relative attraction of its policies to voters.

[4] Thus, when capital stock is low, e.g., during a severe depression, the benevolent party's optimal investment implies a form of consumption smoothing, or counter-cyclical government "spending".

We then characterize the risk that democracy enters a death spiral. Even when capital stocks exceed the critical death spiral threshold, the demagogue always has a chance of winning, and hence capital stocks can always fall. Thus, an economy is always just some bad draws away from having capital drop below the death spiral threshold and hence from an inevitable meltdown, underscoring the shocks to the real economy generated by election outcomes that take the form of a victory by a demagogue that reduces capital.[5]

This death spiral risk can be described as a gambler's ruin problem. In this problem, (i) one gambler starts with finite wealth and the other gambler is infinitely wealthy, (ii) the sizes of the steps up and down need not be equal, and (iii) ruin (zero wealth) corresponds to hitting a capital threshold. We provide conditions under which democracy is not doomed, i.e., there is a chance of never entering a death spiral. We derive upper and lower bounds on the probability of a death spiral as a function of the current capital stock. We show that the rate at which the probability of death spiral vanishes as capital grows very large is slower than exponential; in fact, the tail follows a power law. This heavy tail property means that the presence of a demagogue introduces significant disaster risk to the economy. We show that death spirals tend to be less likely when capital stocks are higher at the outset, the benevolent party selects high valence candidates with higher likelihood, economic productivity is higher, and the institutional constraints on a demagogue's dis-saving are tighter.

Our results highlight close connections between democracy and development. Democracies in developing economies with less capital are more susceptible to meltdowns, because fewer shocks are needed to drop capital below the point where meltdowns become inevitable. Similarly, young democracies tend to have less social capital in the form of trust in institutions, making them more susceptible to negative shocks in the form of demagogues. Lower productivity, less institutional constraints on demagogic policies, and ineffective parties that fail to frequently select high valence candidates all exacerbate the problem. Our results also point to the value of good leaders in young democracies who can build enough of a cushion of capital stock that a country can withstand the negative shock of a perhaps rare, but inevitable future demagogue.

Finally, our analysis points to the value of investment in (physical and social) stock in good times when capital stock is already high. It is exactly at such times that the country can build a cushion of capital to prepare for the inevitable rise of future demagogues. Hard times, in contrast, call for compromise and some mimicry of demagogic policies as the lesser evil. Thus, for example, it is exactly during economic boom or periods of cultural tolerance when demagogues

[5] Adding macroeconomic shocks such as real business cycle shocks or wars can also help drive capital below the death spiral threshold.

are least appealing and hence may seem least dangerous or relevant that democracy builds its resilience against systemic demagogic risk.

Our base analysis focuses on a representative "median" voter, because this voter's preferences suffice to determine the strategic policy choices of the parties. That said, we show how our model can account for cross-sectional patterns in voting, and how these patterns vary with economic development. To do this, we enrich our base model so that (1) some voters receive more of the economic pie than others, and (2) the fraction of voters who view the benevolent party to be the high valence candidate is stochastic, either high or low. The logic of our base model extends. More specifically, when capital stocks are low, its investment policy draws the support of just enough voters to win when a majority view it to be high valence. Moreover, its policy is designed to appeal to the economic elites, while more economically-disadvantaged voters support the demagogue regardless of valence. We then show how economic development mitigates a demagogue's appeal. Specifically, once the benevolent party's policy choices cease to be constrained by electoral competition, higher capital levels reduce (1) the demagogue's vote share; and (2) the ability of the demagogue to sway economically-disadvantaged voters with its short-sighted policies shrinks, with valence determining the choices of more voters. We also touch on how other real world considerations affect outcomes, allowing for business cycle fluctuations in the form of shocks to the productivity of capital, as well as the possibility that a demagogue who wins can alter institutions to improve its future electoral chances.

To place our work, it is useful to describe the technical challenges and how we address them. We characterize Markov perfect equilibria that satisfy a "no use of dominated actions in a subgame" refinement. We show that the benevolent party's equilibrium strategy solves a well-defined constrained optimization problem (Problem P). However, challenge arises because the "beat-the-demagogue" constraints in Problem P is non-linear, indeed, non-convex. We proceed in steps. We first consider the hypothetical benchmark problem (Problem BP) where we drop the constraint imposed by electoral competition, so that rotation of office-holders is exogenous, as in Aguiar and Amador (2011). We show that Problem BP is scalable in capital, permitting explicit solution. In this hypothetical scenario, death spirals never arise—underscoring the importance of explicitly modeling electoral competition. We then consider a modified problem (Problem MP) in which the "beat-the-demagogue" constraints are replaced by linear constraints. With these linear constraints, Problem MP is also scalable in capital, although not time invariant. We then relate the solutions of Problems MP and BP. Finally, we choose the linear constraint parameters of Problem MP to map it to a version of Problem P with more relaxed constraints. This mapping lets us extend our characterization of Problem MP's solution to Problem P.

Related Literature. An earlier draft (Bernhardt et al., 2019a) analyzed a setting in which parties could not dis-invest, but capital depreciated, so that capital would fall absent investment. That framework also featured a continuous valence shock on a support that contains zero, so that each candidate had a chance of having the higher valence. The benevolent party's proposed capital choice traded off on the margin between the gains from the greater investment when it won and the cost of the increased probability that the demagogue wins. Qualitatively similar results obtain, but focusing on a binary valence shock permits more detailed and explicit characterizations. That draft also considered log preferences over consumption, which permit explicit solution even when voters were less myopic, and it characterized non-Markov equilibria.

Bisin et al. (2015) build a three-date model with voters who use hyperbolic discounting. Voters understand their self-control issues, and can use illiquid assets to prevent overspending at date two. They show how two office-motivated candidates can undo this commitment device via excessive debt accumulation, hurting voters. In contrast, voters in our model are unaware of how investment affects future consumption. If we only had office-motivated candidates as in their model, no investments would be made. Our paper focuses on the extent to which Hamilton's notion of a good political leader can be effective in the presence of a demagogue and a short-sighted majority, how this effectiveness hinges on the economy's state of development, and the role of good leadership at the outset of a developing economy for its long-term prospects.

Guiso et al. (2018) define a party as populist if it champions short-term policies while hiding their long-term costs, and show empirically, that hard economic times lead to increased support for populists and populist policies, and that establishment parties' policies grow more populist in nature. Our theoretical model assumes that demagogues can hide the long-term consequences of economic policies from voters, as in their paper. Consistent with their findings, we show that following hard times, the need to mimic populist policies to appeal to voters rises, causing established parties to become more populist. Baccini and Sattler (2021) provide additional evidence, using a difference-in-difference analysis to show that "austerity increases support for populist parties in economically vulnerable regions, but austerity has little effect on voting in economically less vulnerable regions." With heterogeneous voters, we show that this result obtains in our model—the demagogue receives differential support from economically-disadvantaged voters especially when economic conditions are bad, while the benevolent party appeals to the elites, and the demagogue's policies have little impact on voter choices in advanced economies.

Levy et al. (2021) develop a model in which policies based on misspecified models of voters are implemented periodically in perpetuity. The successes of better, sophisticated policies cause voters with misspecified models of the world to infer that their preferred policies would perform even better, increasing their turnout. Acemoglu et al. (2013) build a model in which a lobby can try to bribe politicians to select policies that favor the wealthy. In their model, populists are not susceptible to bribes and signal this by choosing extreme left-wing policies. In contrast, consistent with the findings in Guiso et al. (2018), demagogues in our model maximize their electoral chances by championing short-term policies that appeal to short-sighted voters. We then characterize the long-term consequence of the electoral competition between such demagogues and strategic, far-sighted benevolent parties that aim to maximize voter welfare.

Our notion of a "death spiral" differs from intertemporal dissaving caused by a low return on assets relative to a consumer's discount factor. For example, consider an infinitely-lived consumer with CRRA utility and a linear savings technology. Then consumption is always a fixed fraction of current assets, and if the discount factor times the gross rate of return is less than one, the consumer dis-saves at a constant rate so assets decline over time, converging to zero. This decline would be accelerated if the consumer is present biased. Aguiar and Amador (2011) show how exogenous probabilistic replacement of an office-holder who cares more about citizen payoffs when in office than when out can be formulated as a self-control problem of a single present-biased agent. Halac and Yared (2014) build on this present-biased setting in a model where the office-holder is privately informed about the value of current consumption, which evolves according to a binary Markov process. They characterize the optimal mechanism design under different levels of commitment. Without ex-ante commitment, the stochastic nature gives rise to a precautionary savings motive, which somewhat diminishes the incentive to run savings down. However, if the present bias is sufficiently strong, capital and savings again go to zero. The dynamics are analogous to our benchmark setting without electoral competition, where there are stochastic fluctuations, but strategies and long-run outcomes do not hinge on the capital state of the economy. In contrast, death spirals in our model are driven by intensifying competition from the demagogue as capital stocks go low, and they can occur regardless of the discount factor and the productivity of the economy. While there is always a risk of a death spiral, they are not inevitable, and the probability of a future death spiral depends on the current level of capital.

A small macroeconomics literature embeds probabilistic voting à la Persson and Tabellini (2000) in a dynamic model (e.g., Persson and Svensson (1989), Cukierman and Meltzer (1989), Alesina and Tabellini (1990), Song et al. (2012), Battaglini (2014)). These models feature policy convergence: the political process generates ex-ante distortions when parties choose platforms, but with convergence it does not matter who wins. Battaglini and Coate (2008) analyze the effect of legislative bargaining on government debt and public good provision when debt constrains public good investment and pork provision, and each district is represented by a legislator. A first mover advantage in bargaining means that the proposer's identity determines which district receives the most pork, but public good investment and debt levels are unaffected. While these formulations of political competition facilitate many insights, who wins the actual election is irrelevant. In contrast, in our model the election itself generates uncertainty and dynamic economic distortions, capturing the observation that election results *do* matter (e.g., Kelly et al. (2016) analyze the impact of electoral outcomes on forward-looking financial markets).

Demagogues are politicians who appeal to the people solely to win power for themselves. The term populist is often used interchangeably, but contains aspects that we do not model. For example, according to Müller (2017), populists claim to represent the true people against an elite who controls the levers of government at the expense of the true people. As a result, populists believe it is legitimate to move away from pluralistic democracy, because they, and only they, are the legitimate representatives of the people. In contrast, in our model, majority rule is always preserved. Were a demagogue, instead, able to change the rules underlying fair political competition, making it harder for a benevolent party to regain power, it would strengthen our results.[6] In particular, because its costs of losing would rise, the benevolent party would have even stronger incentives to mimic populist policies.

In our model, increasing current consumption comes at the expense of reducing future consumption. One may argue that citizens should be able to come to understand this link. In reality, this link is less clear because governments can borrow for long periods of time without discernible impacts on consumption. We model this by assuming that citizens are short-sighted. Other papers that model voters in this way include Baron and Diermeier (2001) and Dal Bo et al. (2017). Our historical companion paper (Bernhardt et al., 2019b) documents the extensive concerns of founding fathers of American Democracy about precisely this short-sightedness.

1 Model

The model extends over infinitely many time periods, $t = 0, 1, 2, \ldots$ There is a consumption good and a capital good. If the period t capital stock is k_t, then output is ϕk_t, i.e., the rate of return on capital is $\phi > 0$. Output and capital together can be turned into consumption c_t and investment i_t. That is, i_t can be negative, but we limit this dis-investment to no more than a fraction $\delta > 0$ of the current capital stock. Thus, capital evolves according to $k_{t+1} = k_t + i_t$, where the budget constraint is $c_t + i_t \leq \phi k_t$ and $i_t \geq -\delta k_t$.

[6] One can also interpret a demagogue's efforts to weaken democratic institutions and norms in order to achieve short-term objectives as a form of reduced investment in social capital, with adverse future consequences.

There are two parties, a benevolent party and a demagogue, labeled b and d, respectively. A party's policy at date t is a proposed feasible investment i_t. The median voter's date t utility is given by

$$u(c_t) + v_{P,t},$$

where $u'(c_t) > 0$ and $u''(c_t) < 0$, and $v_{P,t}$ is a valence shock that measures the utility the voter derives if party $P = b, d$ is in power. We interpret $v_{P,t}$ as measuring the non-economic policy aspects associated with candidate P that enter a voter's utility that are out of a party's control. Without loss of generality, we set $v_{d,t} = 0$ and write v_t instead of $v_{b,t}$. We assume that there are two possible valence realizations, v_H and v_L, with $v_H > 0 > v_L$ that occur with probabilities p_H and p_L, respectively. Thus, either candidate can have the net higher valence. We allow for general voter preferences over consumption in our existence and general characterization in Section 2.1. Further characterizations focus on constant relative risk aversion preferences.

We consider myopic voters who base electoral decisions solely on period utility. This captures the idea that voters are unsophisticated and do not understand the long-term impacts of economic policy (Guiso et al., 2018). Our working paper shows how qualitative results extend if voters just underweight future payoffs (Bernhardt et al., 2019a).

In contrast to voters, parties are sophisticated and forward looking. Parties discount future payoffs by β, where $\beta \in (0, 1)$ is assumed to satisfy $\beta(1 + \phi) > 1$, so that absent the demagogue, some investment would always be optimal. We will also impose conditions that ensure expected lifetime payoffs are finite and hence well defined. The demagogue, d, only cares about winning; it receives a period payoff of 1 if it wins, and 0, otherwise. The benevolent party, b, is policy motivated, receiving the same period utility as the median voter. This framework nests a setting in which multiple benevolent parties compete: Each would offer the same economic policy, and when party d loses, the benevolent party with the highest valence would be elected.

A party's policy choice can be described by the proposed capital level for the next period, because it determines the levels of investment and consumption. That is, if k is the current capital stock and k'_j is the capital stock proposed by party $j = d, b$, then investment would be $i_j = k'_j - k$, and consumption would be $c_j = \phi k - i_j = (1 + \phi)k - k'_j$. Thus, in each period the game evolves as follows:

1. Both parties propose capital stocks for the next period.

2. The valence shock is realized.

3. The median voter selects the winning party, which implements its proposed policy.

2 Equilibrium

2.1 Basic Properties of Equilibria

We focus on Markov perfect equilibria of the game. In particular, this means that a party's strategy only depends on the current level of capital, k. A party's strategy determines next period's capital level k' as a function of k. The capital choice must be feasible, and capital disinvestment cannot exceed fraction δ. Thus, a Markov strategy in our game is defined as follows.

Definition 1 *Party j's Markov strategy is given by $s_j \colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $(1-\delta)k \le s_j(k) \le (1+\phi)k$, for all $k \ge 0$.*

Our first result, Proposition 1, details properties of all equilibria that satisfy a simple refinement criterion. The refinement, which we now detail, has a motivation similar to that for weak dominance. Consider Markov strategies s_j, $j \in \{b, d\}$, and any subgame that starts with some capital level k at time t. These strategies induce a one-shot game in period t. In this game, parties simultaneously propose capital levels for period $t+1$, and follow strategies s_j, $j \in \{b, d\}$, in all periods after $t+1$. Our equilibrium refinement requires that the proposed capital choices not be weakly dominated in any of these induced games. More formally, we have:

Definition 2 *A Markov perfect equilibrium s_j, $j \in \{b, d\}$, of the dynamic game uses weakly dominated actions if and only if the following holds:*

There exists a subgame starting at some period t, a capital level k, and a capital choice $k'_j \neq s_j(k)$ for a party $j \in \{b, d\}$ such that k'_j gives j a utility that is at least as high as $s_j(k)$ against any capital choice of its rival, and a strictly higher utility for at least one capital choice of its rival, assuming that both parties continue with strategies s_j, $j \in \{b, d\}$, in all future periods.

To see how the refinement works, consider an infinitely-repeated game with a stage game of:

		Player 2		
		L	M	R
	T	1,1	1,2	-1,-1
Player 1	M	4,2	1,0	-1,-1
	B	-1,-1	-1,-1	0,0

Playing (T, M) is an equilibrium of the stage game, and playing (T, M) in all periods is a Markov perfect equilibrium. Suppose player 1 deviates to M in the current period. This does not affect future payoffs, because both players are assumed to continue with their equilibrium strategies.

However, in the current period player 1's payoff remains the same if player 2 chooses M or R but it is strictly higher if player 2 chooses L. Thus, T is a weakly dominated action for player 1.[7]

We next provide a general characterization of equilibria of our electoral game.

Proposition 1 *Suppose there exists a pure strategy Markov perfect equilibrium that does not use weakly dominated actions. Then in period t,*

1. *The benevolent party wins election if and only if the median voter's preference shock is v_H.*
2. *The demagogue's strategy is $s_d(k) = (1 - \delta)k$.*
3. *The benevolent party's strategy $s_b(k)$ satisfies $u((1 + \phi)k - s_b(k)) + v_H \geq u((\phi + \delta)k)$. If this condition holds as an equality, then the median voter elects the benevolent party.*

To understand the intuition for these results, note that a Markov strategy in a subgame can only depend on the current capital stock. Thus, the benevolent party can always ensure that it wins when its valence is high by proposing the same investment level as the demagogue. The median voter prefers party b due to its high valence, and party b is also better off because it has the same period utility as the median voter. However, this deviation does not affect future payoffs, because next-period's capital would be exactly the same as when the demagogue wins. Therefore, future payoffs are unaffected. The strict increase in current payoff makes this deviation profitable for the benevolent party. Thus, party b must win when its valence is v_H. An analogous argument yields that the demagogue must win when the benevolent party's valence is low.

Next observe that it is a weakly dominant action for the demagogue to maximize the current consumption offered to voters, which it does by choosing the lowest possible capital level in the next period, i.e., $s_d(k) = (1 - \delta)k$. By doing this, the demagogue ensures victory for the largest possible set of proposed investment levels by his rival. Further, changing the action in the current period does not affect the demagogue's future payoffs—as established above, the demagogue wins in future periods if and only if the valence is v_L and is therefore independent of capital. Therefore, any other future capital choice by the demagogue is a weakly dominated action.

Having established that (i) the demagogue always proposes to disinvest maximally, and (ii) the benevolent party must win when its valence is v_H, the third statement of the Proposition

[7] Note that a strategy of playing T in all periods independent of histories is not a weakly dominated strategy. To show this, it suffices to find a (possibly non-Markov) strategy for player 2 such that any strategy other than always playing T makes player 1 strictly worse off. For example, player 2 could use a trigger strategy, choosing action M as long as player 1 chooses T, but selecting R if player 1 chooses M or B in any previous period. Then player 1's ex-ante utility strictly decreases when switching to any other strategy.

follows directly. In particular, the total output available in period t is ϕk. If k' is the capital level chosen by the winner, then investment is $k' - k$, so consumption is $\phi k - (k' - k) = (1 + \phi)k - k'$. The voter's utility from party b when it has high valence is therefore $u((1+\phi)k - s_b(k)) + v_H$. The voter's utility from the demagogue is $u((1+\phi)k - s_d(k)) = u((1+\phi)k-(1-\delta)k) = u((\phi+\delta)k)$. Thus, $u((1+\phi)k - s_b(k)) + v_H \geq u((\phi+\delta)k)$ must hold in order for party b to win when its valence is high.

2.2 The Benevolent Party's Strategy and Equilibrium Existence

Proposition 1 shows that, in equilibrium, the demagogue proposes to dis-invest by δk_t in any period t. Our analysis proceeds by determining the benevolent party's best response to this strategy. We show that this best response is Markovian, i.e., it depends only on the current capital stock k_t. In turn, this means that along an equilibrium path, the extant capital level only depends on the history of valence realizations. It follows that the history of valence realizations is a sufficient statistic for the equilibrium path of capital.

Let h_t denote the history of valence realizations up to but not including time period t. Thus, $h_0 = \emptyset$. Let $\mathcal{H}_t$ be the set of all such histories. Then $h_t = (h_{t-1}, v_{t-1}) \in \mathcal{H}_t$, when the period $t-1$ valence realization is v_{t-1}. Let $P(h_t)$ be the probability of history h_t. This probability can be defined inductively by $P(h_{t-1}, v) = p_v P(h_{t-1})$ for $v \in \{v_H, v_L\}$, where $P(h_0) = 1$.

We denote the capital level at time t by k_{h_t} to capture its dependence on the history of valence realizations. This history fully determines electoral outcomes and policy choices along the equilibrium path in all periods prior to t. The benevolent party proposes a capital level k' for the next period $t+1$. Proposition 1 yields that this policy will be implemented in period t if and only if the valence realization $v_t = v_H$. Thus, $k_{h_t, v_H} = k'$. If the valence realization is, instead, $v_t = v_L$, then Proposition 1 reveals that the demagogue wins and $k_{h_t, v_L} = (1-\delta)k_{h_t}$. The capital choices determine the current consumption level. We use this notation and drop the time index on the valence shocks where the context is clear.

In the analysis that follows, we assume the following sufficient conditions.

Assumption 1

1. $\sum_{t=0}^{\infty} \beta^t |u((1-\delta)^t)| < \infty$.

2. $\lim_{t\to\infty} \beta^t u((1+\phi)^t) = 0$.

Utility from consumption must be well defined when disinvestment is maximal, i.e., $\sum_{t=0}^{\infty} \beta^t u((\phi+\delta)k(1-\delta)^t)$ must converge. This condition is satisfied if and only if condition 1 of Assumption 1 holds. With CRRA preferences, $u(c) = \frac{c^{1-s}}{1-s}$, it is equivalent to requiring $\beta(1-\delta)^{1-s} < 1$. In addition, delaying consumption to infinity must result in zero utility, i.e., $\lim_{t\to\infty} \beta^t u((1+\phi)^t k) < \infty$. This holds if and only if condition 2 in Assumption 1 holds. For CRRA utility with $s > 1$ this is automatically satisfied, because utility is bounded from above by zero.

Lemma 1 *Party b's equilibrium strategy solves the following optimization problem:*

Problem P

$$\max_{\{k_{h_t}:\, h_t\in\mathcal{H}_t\}} \sum_{t=0}^{\infty} \sum_{h_t\in\mathcal{H}_t} \beta^t P(h_t) \left(p_H u\left((1+\phi)k_{h_t} - k_{h_t,v_H}\right) + p_L u\left((\phi+\delta)k_{h_t}\right)\right) \tag{1}$$

$$s.t. \quad u\left((1+\phi)k_{h_t} - k_{h_t,v_H}\right) + v_H \geq u\left((\phi+\delta)k_{h_t}\right), \textit{for all } h_t \in \mathcal{H}_t,\ t \geq 0 \tag{2}$$

$$k_{h_t,v_L} = k_{h_t}(1-\delta), \textit{for all } h_t \in \mathcal{H}_t,\ t \geq 0 \tag{3}$$

$$k_{h_t,v_H} \geq k_{h_t}(1-\delta), \textit{for all } h_t \in \mathcal{H}_t,\ t \geq 0 \tag{4}$$

$$k_{h_t,v_H} \leq (1+\phi)k_{h_t}, \textit{for all } h_t \in \mathcal{H}_t,\ t \geq 0. \tag{5}$$

This optimization problem is well-defined since the infinite sum of utilities in the objective converge by Assumption 1. Constraint (2) is the condition in statement 3 of Proposition 1 that party b wins if and only if its valence is high. In particular, the voter must weakly prefer the benevolent party given its proposed future capital stock when its valence is high to a demagogue that proposes to increase current consumption to the maximum extent possible. Recall that the benevolent party receives the same period utility as the consumer, which includes the valence of the winning candidate. Because the benevolent party wins if and only if its valence is high, this constraint means that we can drop these valence terms from its optimization problem.

Constraint (3) reflects statement 2 of Proposition 1, i.e., that the demagogue maximizes consumption in the current period, and hence determines next period's capital stock when the valence is v_L. Constraint 4 reflects the condition that capital cannot be reduced at a rate exceeding $(1-\delta)$, and (5) is a feasibility constraint ensuring that consumption is non-negative. We will prove that the last two constraints are slack.

The benevolent party's objective function is concave. However, the constraint set in Problem P is not necessarily convex because of constraint (2). Thus, the solution to this optimization problem does not need to be unique. Note, however, that party b' optimal investment strategy at t only depends on the current capital level k_{h_t}. Thus, the following result obtains.

Proposition 2 *There exists a Markov perfect equilibrium that does not use weakly dominated actions. In this equilibrium, the demagogue maximally disinvests in every period and wins if and only if the valence is v_L. The benevolent party's strategy solves optimization problem P.*

In the remaining analysis, we assume that reducing consumption by a fixed percentage has a higher utility impact on consumers when consumption is lower, i.e., $u(c) - u(\gamma c)$ is decreasing in c for $0 < \gamma < 1$. This assumption says that it is more difficult to convince short-sighted voters to save in bad times; this makes it harder to sell austerity in bad times. With constant relative risk aversion voter preferences, $u(c) = \frac{c^{1-s}}{1-s}$, this condition implies that $s > 1$, consistent with most macroeconomic calibrations.[8] We now focus on this case.

2.3 Benchmark: Exogenous Winning Probabilities

We begin with a benchmark problem in which there is exogenous stochastic rotation of office holders as in Aguiar and Amador (2011). That is, there is no electoral competition. Applied to our model this means dropping constraint (2) — the benevolent party and demagogue win with exogenous probabilities p_H and p_L, respectively, regardless of the median voter's preferences. We label this benchmark problem as Problem BP. Let $V_{BP}(k)$ denote the discounted payoff that the benevolent party expects in this benchmark problem when the current capital stock is k. We start by characterizing solutions of this benchmark problem.

Proposition 3 *In the benchmark problem, Problem BP,*

1. *Party b's value function takes the form $V_{BP}(k) = a_b k^{1-s}/(1-s)$, where $a_b > 0$ is a constant.*
2. *The benevolent party's capital choice for the next period is given by $k' = (1+\lambda^*)k$, where k is the current capital, and λ^* solves*

$$(1 - \beta p_L(1-\delta)^{1-s})(1+\lambda^*)^s = \beta p_H(1+\phi) + \beta p_L(\phi+\delta)^{1-s}(\phi-\lambda^*)^s. \tag{6}$$

3. *Party b always proposes to increase the capital stock: $\lambda^* \in (0, \phi)$.*
4. *λ^* decreases in the probability p_H that party b wins, but increases in the extent δ to which a demagogue can dis-invest, the discount factor β, and the productivity ϕ of the economy.*

Setting $p_H = 1$ yields the benevolent party's optimal investment policy when it does not face the threat of a demagogue. With no political competition from a demagogue, we have

$$\lambda^* = (\beta(1+\phi))^{\frac{1}{s}} - 1 > 0.$$

[8] $u(c) - u(\gamma c)$ decreases in c if $u'(c) - \gamma u'(\gamma c) < 0$. This condition holds if $\gamma u'(\gamma c)$ is strictly decreasing for $\gamma \in (0,1)$. Differentiating $\gamma u'(\gamma c)$ with respect to γ shows that this holds if and only if relative risk aversion strictly exceeds 1.

In problem BP, even though the likelihood that the benevolent party wins is exogenous, the presence of the demagogue still affects the party b's choice of λ^*. In fact, the more likely that the demagogue is to win, the more the benevolent party will save in order to insure against the possibility that the demagogue wins and spends down capital. Without the electoral constraint (2), the benevolent party is free to propose greater savings in order to provide greater insurance.

2.4 Characterization of the Benevolent Party's Equilibrium Strategy

We now analyze solutions of the benevolent party's optimization problem (problem P), in which constraint (2) captures electoral competition from the demagogue. However, this constraint is non-linear and non-concave, resulting in a non-convex constraint set, which significantly complicates the analysis.

Similarly to the benchmark problem, we now describe the investment rate by parameters $\lambda(k)$, where $k' = (1 + \lambda(k))k$. We saw that in the benchmark problem the optimal $\lambda(k)$ is constant. In contrast, the presence of constraint (2) will cause the optimal investment rate to depend on the capital stock. For CRRA utility, we rewrite constraint (2) as

$$(\phi - \lambda(k))^{1-s}k^{1-s} + (1 - s)v_H \leq (\phi + \delta)^{1-s}k^{1-s}.$$

Given the current capital stock k, let $\lambda_c(k)$ be the proposed investment rate at which the constraint binds. Then

$$\lambda_c(k) = \phi - \left((\phi + \delta)^{1-s} + v_H(s - 1)k^{s-1}\right)^{\frac{1}{1-s}}. \tag{7}$$

Note that, when $s > 1$, $\lambda_c(k)$ is strictly increasing in k. That is, when capital stocks are higher, the presence of the demagogue is less binding on the benevolent party's current period investment proposals. We now provide a general characterization of party b's equilibrium strategy.

Proposition 4 *Let k be the current capital stock, and let k^* solve $\lambda_c(k^*) = \lambda^*$, where $\lambda_c(\cdot)$ is defined in* (7)*.*

1. *If $p_H < 1$, then there exists $\hat{k} > k^*$ such that the benevolent party proposes future capital stock $k' = (1 + \lambda_c(k))k$ if $k \leq \hat{k}$, and it proposes $k' > (1 + \lambda^*)k$ if $k > k^*$.*

2. *If $p_H = 1$, then the benevolent party proposes future capital stock $k' = (1+\min\{\lambda^*, \lambda_c(k)\})k$.*

3. *The investment share $\lambda(k)$ converges to λ^* as $k \to \infty$.*

Proposition 4 conveys the essence of how the demagogue's presence affects the benevolent party's equilibrium policy choices. When capital levels are low, voters care primarily about policy vis-à-vis valence, so the demagogue's willingness to appeal to a short-sighted voter in its policy binds on the benevolent party. This constrains the benevolent party, causing it to mimic the demagogue's policy, but by the minimum amount needed to ensure that it wins. As capital rises toward k^*, voters care relatively less about policy, allowing the benevolent party to reduce the extent by which it mimics the demagogue, increasing investment.

Once capital levels exceed k^*, the benevolent party could offer the unconstrained investment λ^* associated with the benchmark problem in which there is no electoral competition. However, party b chooses to invest even more. In fact, it turns out the electoral competition constraint (2) still binds for some capital levels above k^*. This over-investment reflects a precautionary savings motive—party b imposes austerity in order to ensure that the electoral competition constraint would be somewhat less binding if the demagogue wins in the future and drives down capital. This precautionary savings motive against the possibility of future electoral constraints is over and above that in the benchmark problem, where λ^* already reflects insurance against the exogenous probability that the demagogue will win and dis-invest, driving down capital stocks.

As capital grows arbitrarily high, capital levels at which the competition constraint binds can only be reached in the distant future. Discounting ensures that the value of precautionary savings goes to zero, and hence that investment converges to λ^*. It follows that the benevolent party's investment policy, i.e., $\lambda(k)$, evolves non-monotonically: If $k < \hat{k}$ then $\lambda(k)$ is strictly increasing (and $\lambda(\hat{k}) > \lambda^*$), while $\lambda(k)$ approaches λ^* from above as $k \to \infty$.

In the rest of this section we explain the construction of the proof and its intuition. The main idea can be gleaned most easily when $p_H = 1$, so that party b always wins. Ignoring the linear constraints, which do not alter the intuition, party b's optimization problem takes the form

$$\max_{k_t} \sum_{t=0}^{\infty} \beta^k u((1+\phi)k_t - k_{t+1}) \text{ s.t. } k_{t+1} \leq h_t(k_t),\ t \in \mathbb{N}, \tag{8}$$

where $h_t(k_t)$ is strictly increasing in k_t.

Let $\{k_t^*\}_{t\in\mathbb{N}}$ be a solution to this optimization problem, and suppose that some of the constraints bind. Let $W_1(k_1)$ be the continuation value in period 1. Next define the investment shares $\bar{\lambda}_t = k_{t+1}^*/k_t^* - 1$ for each period. Then if investment shares are fixed at $\bar{\lambda}_t$ for all t and we start with capital level k_1 at $t = 1$, the continuation value in period 1 is given by

$$V_1(k_1) = \left(\sum_{t=1}^{\infty} \beta^{t-1} (\phi - \bar{\lambda}_t)^{1-s} \prod_{i=1}^{t-1} (1 + \bar{\lambda}_i)^{1-s} \right) u(k_1). \tag{9}$$

By definition, $V_1(k_1^*) = W_1(k_1^*)$. Now consider the experiment of increasing capital from its initial level, k_1^*. The constraints in optimization problem (8) are still satisfied, because each h_t is increasing in current capital. Thus, $V_1(k) \leq W_1(k)$ for $k \geq k_1^*$. Combining $V_1(k_1^*) = W_1(k_1^*)$ and $V_1(k) \leq W_1(k)$ for $k \geq k_1^*$ yields that if W is differentiable then $V_1'(k_1^*) \leq W_1'(k_1^*)$.

The next key step is to show that the marginal utility $V_1'(k^*)$ is larger than the marginal utility $V_{BP}'(k^*)$ from the unconstrained benchmark problem. Because we have dropped all linear constraints the benchmark problem is (8) without the constraints $k_{t+1} \leq h_t(k_t)$. As Proposition 3 shows, such unconstrained problems are scalable in capital, which implies that the value function takes the form $V_{BP}(k) = a_{BP}k^{1-s}$, where a_{BP} is a constant. With linear constraints, it can similarly be shown that $V_1(k) = a_M k^{1-s}$, for some constant a_M. If the constraints bind, then $V_1(k_1^*) < V_{BP}(k_1^*)$, i.e., a higher utility can be achieved without the constraints. Because $s > 1$, this implies $a_M < a_{BP} < 0$, which, in turn, implies

$$V_1'(k_1^*) = (1-s)a_M(k_1^*)^{-s} > (1-s)a_{BP}(k_1^*)^{-s} = V_{BP}'(k_1^*). \tag{10}$$

Above we have established that $W_1'(k_1^*) \geq V_1'(k_1^*)$. This, and (10) yield $W_1'(k_1^*) > V_{BP}'(k_1^*)$, i.e., the marginal product of capital in period 1 is higher in problem P than in problem BP. However, this means that if the constraint in the initial period $t = 0$ is slack, then the optimal investment is larger than in problem BP: $\lambda(k_0) > \lambda^*$. This observation has two key implications. First, if the constraint is slack then $\lambda(k_0) > \lambda^*$, i.e., there is over-investment relative to the benchmark model without electoral competition. Second, if $\lambda_c(k_0) < \lambda^*$ then the constraint must bind. Thus, investment for $k < k^*$, where k^* is defined by $\lambda_c(k^*) = \lambda^*$ is determined by the constraint. By continuity, when $p_H < 1$ the constraint continues to bind up to some level $\hat{k} > k^*$.

This intuition extends to any time period $t > 0$, and it does not rely on the assumption that $p_H = 1$. To deal with the more general case, we turn (9) into a separate optimization problem. In addition, we circumvent assuming that W is differentiable by proceeding by way of contradiction. This is necessary because standard results that establish the equivalence of the recursive approach fail because utility is not bounded from below and the constraint set is non-convex. In fact, the value function is not differentiable at the point where the constraint ceases to bind.

2.5 Death Spirals

We say that capital stocks exhibit a death spiral if there exists a capital level below which capital stocks decline thereafter converging monotonically to zero at an accelerating rate. From (7), $\lambda_c(k)$ is increasing in k with $\lambda_c(0) = -\delta < 0$, and hence there is a capital stock threshold $\bar{k}$ solving $\lambda_c(\bar{k}) = 0$. This means that once capital falls below $\bar{k}$, then even the benevolent party will propose dis-investment: $\lambda_c(k) < 0$. Thus, regardless of who wins the election, capital will continue to fall even further. Moreover, because $\lambda_c(k)$ decreases in capital, the fall accelerates as capital stocks shrink further, resulting in a death spiral.

Proposition 5 *Let $\bar{k}$ satisfy $\lambda_c(\bar{k}) = 0$, i.e.,*

$$\bar{k} = \left(\frac{\phi^{1-s} - (\phi + \delta)^{1-s}}{v_H(s-1)} \right)^{\frac{1}{s-1}}. \tag{11}$$

Then a death spiral occurs with probability 1 if and only if either $k < \bar{k}$ or $k = \bar{k}$ and $p_H < 1$.

The proposition highlights that once capital falls below $\bar{k}$, the benevolent party must propose to dis-save in order to have a chance of beating the demagogue. As a result, capital stocks fall further. This decline forces the benevolent party to increasingly mimic the demagogue's dis-investment policy even when p_L is small, i.e., even when the demagogue is typically unpopular with little chance of winning. As capital stocks spiral downward further toward zero, the benevolent party's investment policy converges toward the demagogue's, albeit never reaching it, as having a higher valence still provides the benevolent party some advantage.

Corollary 1 *The death spiral cutoff $\bar{k}$ declines with productivity ϕ, but rises with the demagogue's dis-saving capacity δ. Moreover, the marginal effect of higher productivity in reducing $\bar{k}$ is higher when the demagogue has a higher dis-saving capacity: $\frac{\partial^2 \bar{k}}{\partial \delta \partial \phi} < 0$.*

In more productive economies, capital stocks have to fall further before a death spiral ensues. However, when there are weaker institutional constraints on the destruction of capital stock, so that δ is higher, a death spiral occurs at a higher capital stock. Corollary 1 also shows that the marginal effect of higher productivity in reducing the death spiral critical threshold is higher when institutional constraints on the destruction of capital stocks are weaker, so that δ is higher. An implication is that the effect of productivity shocks (e.g., in the form of new technologies) in preventing death spirals is more pronounced in countries with weaker institutional constraints on policies (e.g., in newer democracies with weaker judiciaries). Inspection of $\lambda_c(k)$ indicates that it always exceeds $-\delta$ and that it has the same comparative statics as those for $\bar{k}$, implying, for example, that death spirals proceed more slowly in more productive economies.

2.6 Economic Development and the Risk of Death Spirals

We showed that once capital stock falls below a threshold, the presence of a demagogue leads to a death spiral. We now investigate the risk of entering a death spiral when the capital stock is higher. We first identify a lower bound on the probability of entering a death spiral given any current capital stock k. We then identify conditions under which, even when the benevolent party is able to grow the capital stock arbitrarily high, democracy is still eventually doomed to enter a death spiral with probability one. One might think that because the demagogue always has a chance of winning and driving capital stocks down below the critical level that a death spiral may be inevitable in the long run. We show that this is not so. We identify sufficient conditions for economic development to mitigate the threat of a demagogue and the risk of a future implosion of democracy. That said, we show that some risk to democracy always remains. Specifically, we establish that as capital grows large, the probability of a death spiral only declines according to a power law.

Proposition 6 *1. If $k > \bar{k}$, then a death spiral occurs with a probability of at least $p_L^{1+\alpha}$, where $\alpha = \log(\bar{k}/k)/\log(1-\delta)$.*

2. *If $(1-\delta)^{p_L}(1+\lambda^*)^{p_H} < 1$, then a death spiral occurs with probability 1 regardless of the current capital level k.*

3. *Suppose $(1-\delta)^{p_L}(1+\lambda^*)^{p_H} > 1$ and let $Q(k)$ be the probability of entering a death spiral starting from $k > k^* > \bar{k}$. Fix an arbitrary $\epsilon > 0$. There is a constant $C \in (0,1)$ such that*

$$C\left(\frac{k}{k^*}\right)^{-\frac{\log(y)}{\log(1-\delta)}-\varepsilon} \leq Q(k) \leq \left(\frac{k}{k^*}\right)^{-\frac{\log(y)}{\log(1-\delta)}}, \tag{12}$$

where y is the unique solution in the interval $(0,1)$ to

$$p_H y^{1-\frac{\log(1+\lambda^*)}{\log(1-\delta)}} - y + p_L = 0. \tag{13}$$

The first result reflects that even when the current capital stock k exceeds $k^* > \bar{k}$, the demagogue may win any given election when $p_L > 0$. If the demagogue wins enough times, then the demagogue's dis-saving will drive capital stocks below $\bar{k}$, at which point a death spiral ensues. This result reveals that poor democracies are more vulnerable, as when k is lower, a demagogue needs to win fewer times before a death spiral ensues. Interpreting k as including well-defined property rights and the social capital associated with the institutional norms of democracy, this result further suggests that younger democracies are more vulnerable.

The second and third results identify conditions under which a death spiral is avoidable and characterize the rate at which the probability of death spiral vanishes as capital grows very large. Two conditions determine whether a death spiral is avoidable: $k > \bar{k}$ and $(1-\delta)^{p_L}(1+\lambda^*)^{p_H} > 1$. Recalling how the environment influences λ^* from Proposition 3, this result established that *if* (i) a democracy is lucky enough at the outset to draw good leaders who grow capital to some level $k > k^*$, (ii) demagogues are sufficiently unlikely to win ($p_L = 1 - p_H$ is small), (iii) productivity (ϕ) is sufficiently high, and (iv) the demagogue's ability to dis-save is sufficiently limited (δ is small), then the probability of a death spiral is bounded away from one. In such circumstances, when k is large, electing a demagogue causes damage, but the democracy is likely resilient enough to recover. That said, the disaster risk of a death spiral only falls slowly as capital grows large, declining according to a power law. That is, the presence of a demagogue introduces significant tail risk to the economy.

Our proof casts the problem as a "gambler's ruin problem." The classic formulation, which dates back to a letter from Blaise Pascal to Pierre Fermat in 1656, considers two players who begin with fixed stakes and a possibly biased random walk that determines the direction of a unit transfer from one player to the other that continues until one of the players is "ruined" by reaching zero. Feller (1968) analyzes more general random walks and allows for one player to be infinitely rich. We characterize the generalized random walk that moves up k_H steps with probability p_H and moves down k_L steps with probability p_L. We then characterize the probability that the capital stock k ever falls to capital stock $\bar{k}$ ("ruin"). The power law tail arise because of the random proportional growth nature of the evolution of capital. In particular, $k_{t+1} = (1 + x)k_t$, where x is a random variable that takes the value of $-\delta$ with probability p_L (when the demagogue wins) and the value of $\lambda(k_t)$ with probability p_H (when the benevolent party wins), where $\lambda(k_t) \approx \lambda^*$ for large capital. As Gabaix (2016) highlights, proportional growth models often underlie the emergence of power laws, and consistent empirical evidence has been found in many settings.

To illustrate this result, suppose that there exists an integer j, such that $1-\log(1+\lambda^*)/\log(1-\delta) = 1 + 1/j$. For example, if $(1 + \lambda^*)(1 - \delta) = 1$ then (13) reduces to the quadratic equation $p_H y^2 - y + p_L = 0$, with solution $y = p_L/p_H$. Then when $p_L < p_H$, (12) implies

$$C\left(\frac{k}{k^*}\right)^{-\frac{\log(p_L)-\log(p_H)}{\log(1-\delta)}-\varepsilon} \leq Q(k) \leq \left(\frac{k}{k^*}\right)^{-\frac{\log(p_L)-\log(p_H)}{\log(1-\delta)}},$$

where ϵ can be made arbitrarily small. For example, if $p_H = 0.9$ and $\delta = 0.3$ then the exponent is approximately -6.16. Thus, a 50% increase in capital lowers the bound on the probability of falling below k^* by 92%, substantially reducing the chance of entering a death spiral.

Now suppose that productivity ϕ is lowered to the point where $(1 + \lambda^*)^2(1 - \delta) = 1$. Then equation (13) becomes $p_H y^{3/2} - y + p_L = 0$, with solution

$$y = \frac{p_L\left(1 + p_H + \sqrt{p_L + 3p_H p_L}\right)}{2p_H^2}.$$

Now, when $p_H = 0.9$ and $\delta = 0.3$, the exponent becomes -5.23. Thus, a 50% increase in capital reduces the bound on the probability of falling below k^* by 88%.

Proposition 3 provides comparative statics for λ^*, for example, that λ^* is increasing in the productivity, ϕ. The discussion above shows that such a productivity reduction may initially only have a small impact on the probability of entering a death spiral, provided that the established democracy has reached a high level of capital. However, if an economy suffers a sufficient permanent decrease of productivity, then the second statement of Proposition 6 shows that a death spiral becomes inevitable.

For example, suppose $p_H = 0.9$, $\delta = 0.3$ (as above), $s = 1.2$, and $\beta = 0.9$. If the productivity ϕ drops to 15% or lower then the second statement of Proposition 6 applies, and death spirals occur with probability 1, but for $\phi = 0.16$, statement 3 applies. In this latter case, $y = 0.983$ and a 50% increase in capital reduces the probability of dropping below k^* by only 2%, but death spirals are not inevitable. However, further increases in the return, even small ones, reduce the probability of death spirals substantially: if $\phi = 17\%$ then a 50% increase in capital reduces the probability by almost 30%, and at $\phi = 18\%$ that reduction is 45%. Thus, small changes in productivity can sometimes have very large impacts on the long-term stability of democracies.

Summarizing the content of these findings, if at the outset, a country has the good fortune of electing benevolent leaders, those leaders may grow the capital by so much that there is a good chance of forestalling a retreat below the current high level. Concretely, enlightened leadership by Washington, Adams, Madison, ... may be able to build enough institutional capital to forestall the adverse effects of later occasionally drawing a demagogue. If, instead, a country has the misfortune at the outset of drawing a few demagogues, then it may be doomed thereafter.

2.7 Heterogeneous Voters

To this point our analysis has focused on a single decisive voter because this voter's behavior fully pins down the strategic behavior of the two parties. In this section, we introduce voter heterogeneity, so that some voters receive more of the economic pie than others. This lets us draw insights into which voter types support the demagogue, and which ones support the benevolent party, and how the degree of development in the economy affects the support of each candidate.

We consider a continuum of measure one of voters. Voter j obtains a period payoff of $u(w_j(\phi k - i)) = \frac{(w_j c)^{1-s}}{1-s}$ when the current capital stock is k and investment is i, where $s > 1$. Here, $w_j \in [a, b]$, with $a > 0$, is a measure of voter j's claim to the economic pie—a voter with a higher w_j receives more. We assume that w_j is distributed in the population according to a strictly increasing cdf G. We also allow voters to disagree on the net valence attached to the benevolent party. Voters receive idiosyncratic, conditionally-independent, voter-specific shocks, where with probability p_H a measure $\alpha > 1/2$ of voters assigns (net) valence $v_H > 0$ to the benevolent party, and with probability $p_L = 1 - p_H$, a measure $\alpha > 1/2$ of voters assigns it a valence $v_L < 0$.

The multiplicative structure of period payoffs from consumption means that the benevolent party has the same optimal policy choice for each voter. This means that its objective is unchanged from our base case setting. Of note, this multiplicative structure preserves the scalability of Problem MP, allowing us to use our existing analysis to characterize individual voting choices. The key to characterization then reduces to identifying the decisive voter that the benevolent party must target in order to win when a majority of voters find it to be high valence.

From Proposition 2, the benevolent party wins if and only if a majority of voters finds it to be high valence. The relative preferences of different w_i voters over party b and demagogue are ordered in the same way regardless of the capital level—higher w_i voters with the same valence shock have a relatively higher preference for party b. The following result is immediate.

Lemma 2 *Voter w_m is decisive, where w_m solves $\alpha(1 - G(w_m)) = 1/2$. Party b wins when the capital stock is k_{h_t} if and only if a majority of voters find its valence to be high and its proposed future capital stock k_{h_t,v_H} satisfies*

$$u\left(w_m((1+\phi)k_{h_t} - k_{h_t,v_H})\right) + v_H \geq u\left(w_m(\phi+\delta)k_{h_t}\right). \tag{14}$$

If the decisive voter w_m weakly prefers party b, then so do all voters with $w_i > w_m$ who attach valence v_H to party b. If the decisive voter is indifferent between the candidates, then all voters with $w_i < w_m$ prefer the demagogue.

One can partition voters who find party b to have valence v_H into those with a large enough share w_i to support it, and those whose share w_i is too small. Thus, the model predicts that party b draws its support from the "economic elites," while relatively economically-disadvantaged voters support the demagogue. It follows that party b's proposed capital choices solve Problem P, with constraint (2) replaced by constraint (14). Thus, there exists a $\hat{k}$ such that for $k \leq \hat{k}$, party b's capital choice solves (14) at equality, mimicking the demagogue by the minimum extent needed to deliver victory, while for $k > \hat{k}$, constraint 14 is slack, and its vote share is higher.

Proposition 7 *Among voters who find the benevolent party to be high valence,*

- *Those voters with $w_i \geq w_m$ always select the benevolent party.*
- *For $k \leq \hat{k}$, only those voters support the benevolent party.*
- *For $k > \hat{k}$, voters with $w_i < w_m$ but sufficiently close to w_m select the benevolent party.*
- *As $k \to \infty$, all voters who attach a high valence to the benevolent party select it.*

These results indicate that the benevolent party draws its electoral support from "elites" who obtain sufficiently large shares of the economic pie and also attach a valence v_H to party b. In contrast, the demagogue wins all voters who find party b to be low valence plus those who attach a high valence to party b but receive sufficiently small shares of the pie. These findings do not reflect that economically-disadvantaged voters are more short-sighted than the elites or that they do not mind the fact that a demagogue has a low valence. Rather, they reflect that disadvantaged, low w_i, voters care more about the policies proposed, and relatively less about valence.

When capital stocks are small enough, party b is constrained by the need to design its policy to appeal to voter w_m who finds that it has a high valence. As a result, the demagogue wins support of all voters with $w_i < w_m$. However, as capital stocks grow large, the demagogue's policy appeals fade in importance relative to the considerations that voters place on valence. As a result, among voters who find party b to be high valence, the demagogue only wins support drawn from those at the bottom percentiles of the economic strata. Further, w_i is bounded away from zero, so that once development levels grow high enough, valence considerations rather than policy determine voting choices of all citizens, including the poor, limiting a demagogue's appeal.

The final statement of the Proposition indicates that the vote share of the demagogue shrinks when capital becomes large. In particular, for sufficiently high capital levels all voters who attach a lower valence to the demagogue vote for the benevolent party:

Corollary 2 *If capital levels are low enough that constraint (14) binds, then demagogue receives vote share* $1/2$ *with probability* p_H *and share* $\alpha + (1-\alpha)G(w_m)$ *with probability* p_L*. If capital levels are high enough that constraint (14) does not bind, then the demagogue's vote share is strictly lower, and for sufficiently high capital levels, the demagogue receives vote share* $1-\alpha$ *with probability* p_H *and vote share* α *with probability* p_L*.*[9]

2.8 The Role of Relative Risk Aversion $s > 1$

A central assumption of our model is that $u(c) - u(\gamma c)$ decreases in c for $0 < \gamma < 1$. With CRRA preferences, this implies that the coefficient of relative risk aversion satisfies $s > 1$. By raising the salience of investment policy differences when capital is low, $s > 1$ makes it harder to sell austerity in bad times. To understand how radically outcomes are affected if we change this assumption, suppose that $s < 1$, and write the electoral competition constraint on party b as:

$$\frac{(1-s)v_H}{k^{1-s}} \geq (\phi + \delta)^{1-s} - (\phi - \lambda)^{1-s}. \tag{15}$$

[9] When constraint (14) does not bind, a sufficient, but not necessary, condition for the demagogue's vote share to be decreasing in capital is for $\lambda(k)$ to be decreasing in k, which we find numerically.

Now in hard times, where capital stocks go very low, the left-hand side goes to infinity—policy choices by party b cease to affect the choices of voters, leaving party b completely unconstrained in its choice of $\lambda(k)$. In particular, if the economy starts at a small level of capital k, then it takes many periods until the constraint starts to bind. As a consequence, $\lambda(k)$ is close to λ^* for small k. This, in turn, means that as long as ϕ is sufficiently high that $(1+\lambda^*)^{p_H}(1-\delta)^{p_L} > 1$ then $k = 0$ is never an absorbing state, in sharp contrast to when $s > 1$. Further, (15) implies

$$k \leq \tilde{k} = \left(\frac{(1-s)v_H}{(\phi+\delta)^{1-s} - \phi^{1-s}} \right)^{\frac{1}{1-s}}. \tag{16}$$

For $k > \tilde{k}$, the competition constraint implies that party b must dis-invest to win. Thus, $\tilde{k}$ becomes a reflecting boundary when $p_H < 1$, as long as capital starts below $\tilde{k}$. It follows that whenever $(1-\delta)^{p_L}(1+\lambda^*)^{p_H} > 1$, capital levels move stochastically in an intermediate range, never becoming very small or very large.

In sum, when $s < 1$, death spirals never arise, but conversely the potential for growing the economy very large is destroyed by the electoral competition from the demagogue. In particular, $s < 1$ is at odds with real world observation, as it implies that the attraction of populists is smallest when capital is low, and highest when capital is high.

2.9 Changing the Rules

Our analysis presumes that a demagogue who wins office cannot alter the legal rules and institutions to increase its probability p_L of winning. In practice, the demagogue may be able to gerrymander electoral districts to favor those who support the demagogue, or to disenfranchise or raise the costs of voting to individuals who might be pre-disposed to voting against the demagogue. Such changes would increase the demagogue's future chances of winning.

One can incorporate such a possibility by modifying our model to introduce a probability $\zeta > 0$ that, if elected, the demagogue can increase its probability of 'drawing a high valence' from p_L to $\bar{p}_L$. After such a rule change, Lemma 6 and Proposition 4 yield that party b proposes increased investment when unconstrained, and it raises the probability of entering a death spiral. Further, prior to such a rule change, similar results follows directly from the fact that the possibility of such a future rule change reduces the value function of party b, *raising* the marginal value of increasing future capital stocks. In particular, such changes increase the analogue to λ^*, inducing party b to save more. Intuitively, the benevolent party internalizes that if the demagogue does change the laws down the road, the greater capital mitigates the consequences of the increased probability that the demagogue will win and drive capital down.

2.10 Productivity Fluctuations and Economic Downturns

Economic stagnation in our model corresponds to a low value of ϕ. The comparative statics of our model yield that both $\lambda_c(k)$ and λ^* increase in ϕ. These results imply that economic stagnation is associated with reduced investment in future capital stocks, and hence a greater probability of a democratic decline in the form of a death spiral.

At the expense of significantly more notation, one can introduce productivity shocks to the economy, and show that the qualitative results extend. Specifically, one can allow for i.i.d. productivity shocks $\phi \in \{\phi_L, \phi_H\}$, with $0 < \phi_L < \phi_H$ where ϕ_j occurs with probability q_j, $j = L, H$. The two parties see the current productivity shock before proposing policy. The same techniques still apply, permitting a similar characterization. With i.i.d. shocks, the future value of a given level of capital to party b does not depend on the current productivity shock, and hence neither does its marginal value. It follows that (i) party b's unconstrained optimal investment share is lower after ϕ_L than ϕ_H, and (ii) electoral competition from the demagogue places a lower bound on party b's proposed investment share after ϕ_L than ϕ_H. Thus, the benevolent party always proposes reduced investment shares in economic downturns.

3 Conclusion

Our paper investigates the long-run susceptibility of Democracy to demagogues. We analyze the dynamic political competition between a far-sighted, benevolent party that seeks to maximize voter welfare, and an office-motivated demagogue who only cares about winning. Parties propose how to allocate existing resources between current consumption and investment. Myopic voters base electoral choices on the current utilities derived from policies and a valence shock, and the winning candidate's policy is implemented.

Demagogues design their policies to appeal maximally to short-sighted voters, proposing to increase consumption by sacrificing some capital. We show that when voters have CRRA preferences with relative risk aversions above one, this electoral competition constrains a benevolent party's choices whenever capital levels are sufficiently small. This reflects that the salience of differences in proposed consumption levels rises relative to differences in valences when capital, and hence consumption, is lower. Relatedly, we establish that the benevolent party draws its electoral support from "elites" who obtain sufficiently large shares of the economic pie, while the demagogue draws support from the economically disadvantaged who care more about the policies proposed, and relatively less about valence. We show that as development levels increase, valence considerations rather than policy determine the voting choices of more citizens, including the poor, limiting a demagogue's appeal.

Demagogues have a chance of winning so there is always a risk that capital is reduced low enough that a benevolent party must also propose to dis-invest to win. When this happens, the economy enters a death spiral, with capital declining to zero. More optimistically, we identify sufficient conditions for development to mitigate the long-term threat of a demagogue. If (i) a democracy draws good leaders at the outset who grow capital sufficiently, (ii) demagogues are sufficiently unlikely to win and limited in their abilities to dis-invest, and (iii) productivity is high enough then death spirals become extremely unlikely. In such circumstances, electing a demagogue causes damage, but the democracy is typically resilient enough to recover.

4 Appendix

Proof of Proposition 1. *Statement 1.* Suppose that party d wins when the preference shock is υ_H. Then b can win by making the same offer. The future capital stock and hence the actions in the subgame is unaffected. However, party b wins, which makes b better off, because he receives the additional utility of υ_H. This contradicts the premise of optimization by b. Now suppose that party b wins when the preference shock is υ_L. Then using an analogous argument to that for party b, the demagogue can match the offer and win in the current period, without affecting future payoffs, a contradiction of optimization by d.
Statement 2. Suppose party d proposes $i_d(k) > -\delta k$. Note that deviations do not affect d's future payoffs, because d only cares about winning, which only depends on valence shocks (from Statement 1). If party d lowers its investment, then d still wins when the preference shock is υ_L. If the shock is υ_H then $i_d(k) = -\delta k$ weakly dominates all other actions, because d wins if party b chooses an investment $i_b(k)$ with $u(\phi k - i_b(k)) + \upsilon_H < u((\phi + \delta)k)$.
Statement 3. Because party b wins when the shock is υ_H, we have $u(\phi k - i_b(k)) + \upsilon_H \geq u((\phi + \delta)k)$. Suppose that the condition holds at equality. If the median voter does not choose party b, then the fact that $i_b(k) > -\delta k$ implies that b can choose a marginally smaller investment, which would make the median voter strictly better off. The future actions of party d are unaffected (from Statement 1). Thus, this deviation would make b strictly better off, a contradiction. Therefore, party b must be elected if the constraint holds with equality. ■

Proof of Proposition 3. *Statement 1.* We first prove that the value function is scalable and then derive its functional form. Given initial capital $k_{h_0} = \bar{k}$, party b's optimization problem is:

Problem BP

$$V_{BP}(\bar{k}) = \max_{\{k_{h_t}: h_t \in \mathcal{H}_t\}} \sum_{t=0}^{\infty} \sum_{h_t \in \mathcal{H}_t} \beta^t P(h_t) \left(p_H u\left((1+\phi)k_{h_t} - k_{h_t,\upsilon_H}\right) + p_L u\left((\phi+\delta)k_{h_t}\right)\right)$$

$$\text{s.t.} \qquad k_{h_t,\upsilon_L} = k_{h_t}(1-\delta), \text{ for all } h_t \in \mathcal{H}_t, \ t \geq 0$$

$$-\delta k_{h_t} \leq k_{h_t,\upsilon_H} - k_{h_t} \leq \phi k_{h_t}, \text{ for all } h_t \in \mathcal{H}_t, \ t \geq 0.$$

Let $\{\bar{k}_{h_t}\}_{t=0}^{\infty}$ be an optimal sequence for this problem. Now, multiply the initial capital by $\alpha > 0$, so that the initial capital stock is $\hat{k} = \alpha \bar{k}$, and consider the sequence $\{\hat{k}_{h_t}\}_{t=0}^{\infty}$, where $\hat{k}_{h_t} = \alpha \bar{k}_{h_t}$. This new sequence, $\{\hat{k}_{h_t}\}_{t=0}^{\infty}$, satisfies all the constraints because $-\delta \bar{k}_{h_t} \leq \bar{k}_{h_t,\upsilon_H} - \bar{k}_{h_t} \leq \phi \bar{k}_{h_t}$ if and only if $-\delta\hat{k}_{h_t} \leq \hat{k}_{h_t,v_H} - \hat{k}_{h_t} \leq \phi\hat{k}_{h_t}$. Because $V_{BP}(\hat{k})$ reflects optimization given $\hat{k}$ rather than $\bar{k}$,

$$\begin{aligned} V_{BP}(\hat{k}) = V_{BP}(\alpha\bar{k}) \quad &\geq \quad \alpha^{1-s}\sum_{t=0}^{\infty}\sum_{h_t\in\mathcal{H}_t}\beta^t P(h_t)\Big(p_H u\Big((1+\phi)\bar{k}_{h_t} - \bar{k}_{h_t,v_H}\Big) + p_L u\Big((\phi+\delta)\bar{k}_{h_t}\Big)\Big) \\ &= \quad \alpha^{1-s}V_{BP}(\bar{k}). \end{aligned} \tag{17}$$

Because α and $\bar{k}$ are both arbitrary, we can use $\frac{1}{\alpha}$ instead of α, and $\alpha\bar{k}$ instead of $\bar{k}$ to get

$$V_{BP}(\bar{k}) \geq \frac{1}{\alpha^{1-s}}V_{BP}(\alpha\bar{k}). \tag{18}$$

Suppose the inequality in (17) were strict. Then, substituting the right-hand side of (17) into (18) yields

$$V_{BP}(\alpha\bar{k}) > \alpha^{1-s}V_{BP}(\bar{k}) \geq \alpha^{1-s}\frac{1}{\alpha^{1-s}}V_{BP}(\alpha\bar{k}) = V_{BP}(\alpha\bar{k}),$$

a contradiction. Thus, we must have

$$V_{BP}(\alpha\bar{k}) = \alpha^{1-s}V_{BP}(\bar{k}).$$

Substituting $\bar{k} = 1$ and $\alpha = k$ yields

$$V_{BP}(k) = k^{1-s}V_{BP}(1),$$

where $V_{BP}(1)$ depends on s. It remains to show that $V_{BP}(1)$ is finite. First, consider the lower bound obtained if party b mimics party d. Then the consumption payoff becomes

$$\sum_{t=0}^{\infty}\beta^t\frac{((\phi+\delta)(1-\delta)^t k)^{1-s}}{1-s} = \frac{((\phi+\delta)k)^{1-s}}{1-s}\frac{1}{1-\beta(1-\delta)^{1-s}} > -\infty$$

as $\beta(1-\delta)^{1-s} < 1$, by Assumption 1. Further, a strict upper bound on $V_{BP}(1)$ is zero. Thus, $V_{BP}(1)$ and the value functions are well defined.

Statements 2 and 3. The benevolent party solves

$$\max_{k'} p_H\Big(u((1+\phi)k-k') + \beta V_{BP}(k')\Big) + p_L\Big(u((\phi+\delta)k) + \beta V_{BP}(k(1-\delta))\Big) \text{ s.t. } k' \geq (1-\delta)k. \tag{19}$$

Assuming that the constraint is slack, the first-order condition is

$$-u'((1+\phi)k-k') + \beta V'_{BP}(k') = 0. \tag{20}$$

Writing $k' = (1+\lambda^*)k$ and noting that $u'((1+\phi)k-k') = (\phi-\lambda^*)^{-s}u'(k)$, equation (20) implies

$$V'_{BP}(k') = \frac{(\phi-\lambda^*)^{-s}k^{-s}}{\beta} = \frac{1}{\beta}\left(\frac{1+\lambda^*}{\phi-\lambda^*}\right)^s u'(k'). \tag{21}$$

We can apply the envelope theorem to (19) to get

$$V'_{BP}(k) = p_H(1+\phi)u'((1+\phi)k-k') + p_L\Big((\phi+\delta)u'((\phi+\delta)k) + \beta(1-\delta)V'_{BP}(k(1-\delta))\Big). \quad (22)$$

Re-arranging and using Statement 1 yields

$$\left(1-p_L\beta(1-\delta)^{1-s}\right)V'_{BP}(k) = p_H(1+\phi)u'((1+\phi)k-k') + p_L(\phi+\delta)u'((\phi+\delta)k). \quad (23)$$

Substituting (21) into (23) yields

$$\left(1-\beta p_L(1-\delta)^{1-s}\right)\frac{1}{\beta}\left(\frac{1+\lambda^*}{\phi-\lambda^*}\right)^s u'(k) = \frac{p_H(1+\phi)}{(\phi-\lambda^*)^s}u'(k) + p_L(\phi+\delta)^{1-s}u'(k). \quad (24)$$

Dividing by $u'(k)$ and multiplying by β and $(\phi-\lambda^*)^s$ yields the first order condition.

To establish that $\lambda^* > 0$, we show that the derivative at $\lambda^* = 0$ is strictly positive. It suffices to show that the left-hand side of (24) is strictly less than its right-hand side at $\lambda^* = 0$. At $\lambda^* = 0$, multiplying both sides of (24) by $\beta\phi^s/u'(k)$ and re-arranging yields

$$1-\beta p_L(1-\delta)^{1-s} < \beta(p_H(1+\phi) + p_L(\phi+\delta)^{1-s}\phi^s),$$

which is equivalent to

$$1 < \beta p_H(1+\phi) + \beta p_L\left((\phi+\delta)^{1-s}\phi^s + (1-\delta)^{1-s}\right). \quad (25)$$

To show that (25) holds, it is sufficient to prove that

$$(\phi+\delta)^{1-s}\phi^s + (1-\delta)^{1-s} \geq \phi. \quad (26)$$

Then the right-hand side of (25) would be greater or equal to $\beta p_H(1+\phi)+\beta p_L\phi > \beta\phi$, and by assumption $\beta\phi \geq 1$. To establish this, note that the left-hand side of (26) is strictly increasing in δ when $s > 1$. In particular, the derivative of the left-hand side of (26) with respect to δ is given by

$$(s-1)\left(\left(\frac{1}{1-\delta}\right)^s - \left(\frac{\phi}{\phi+\delta}\right)^s\right) > 0.$$

Thus, it is sufficient to show that (26) holds for $\delta = 0$, which is immediate. That $\lambda^* > 0$ also implies that the constraint $k' \geq (1-\delta)k$ is slack. That $\lambda^* < \phi$ follows from $u(0) = -\infty$.

Statement 4. The value function is increasing in p_H, because party b can choose to imitate the demagogue. From Statement 1 and (19), party b's optimization problem is strictly concave with unique solution $k' = (1+\lambda^*)k$. From Statement 2, $\lambda^* > 0$. Thus, a higher δ lowers the payoff when the demagogue wins. Thus, the value function is decreasing in δ. From Statement 1, $V_{BP}(k) = a_b k^{1-s}/(1-s)$. Thus, a_b is decreasing in p_H and increasing in δ. Party b solves

$$\max_{\lambda} \frac{(\phi-\lambda)^{1-s}}{1-s} + \beta a_b \frac{(1+\lambda)^{1-s}}{1-s}$$

with the first-order condition

$$-(\phi-\lambda)^{-s} + \beta a_b (1+\lambda)^{-s} = 0 \Leftrightarrow \beta a_b = \left(\frac{1+\lambda^*}{\phi-\lambda^*}\right)^s. \tag{27}$$

The right-hand side of (27) is increasing in λ^* and hence λ^* is increasing in a_b. The comparative static properties for p_H and δ follow.

To obtain the comparative statics for β and ϕ, we use (6). Observe that the left-hand side of (6) is increasing in λ^*, while the right-hand side is decreasing. Next observe that, fixing λ^*, increases in β reduce the LHS, but increase RHS, implying that λ^* must rise with β.

Finally, ϕ does not appear on the LHS of (6), but, fixing λ^*, the RHS is increasing in ϕ since

$$(\phi+\delta)^{1-s}(\phi-\lambda^*)^s = (\phi+\delta)\left(\frac{\phi-\lambda^*}{\phi+\delta}\right)^s$$

increases in ϕ. Thus, λ^* increases in ϕ. ■

Proof of Proposition 4. We proceed in two steps. First, we introduce a modified problem with linear constraints, which we call problem MP. The linearity of constraints in the modified problem MP allows us to characterize its solution by comparing its value function with that of the benchmark problem BP. In the second step, we choose the (linear) constraint parameters of problem MP to map it to a version of problem P with more relaxed constraints. This mapping enables us to carry over our characterization of problem MP's solution to problem P.

Step 1. The modified problem MP corresponds exactly to problem P except that we modify constraints (2): for each $t \in \mathbb{N}$ we replace constraints (2) with constraints that are linear in k. We fix linear parameters $\{\bar{\lambda}(h_t) \colon h_t \in \mathcal{H}_t, t \in \mathbb{N}\}$, where $-\delta \leq \bar{\lambda}(h_t) < \phi$.

Problem MP

$$\max_{\{k_{h_t}\colon h_t \in \mathcal{H}_t\}} \sum_{t=0}^{\infty} \sum_{h_t \in \mathcal{H}_t} \beta^t P(h_t)\left(p_H u\left((1+\phi)k_{h_t} - k_{h_t,v_H}\right) + p_L u\left((\phi+\delta)k_{h_t}\right)\right) \tag{28}$$

$$\text{s.t. } k_{h_0,v_H} \leq k_{h_0}(1+\bar{\lambda}(h_0)); \tag{29}$$

$$k_{h_t,v_H} = k_{h_t}(1+\bar{\lambda}(h_t)), \text{ for all } h_t \in \mathcal{H}_t,\ t > 0; \tag{30}$$

$$k_{h_t,v_L} = k_{h_t}(1-\delta), \text{ for all } h_t \in \mathcal{H}_t,\ t \geq 0. \tag{31}$$

We next use the linearity of the constraints in Problem MP to prove that Problem MP is scalable in capital. Recall that the value function for problem BP is scalable, taking the form $V_{BP}(k) = a_b k^{1-s}/(1-s)$, where $a_b > 0$ is a constant, and the associated optimal investment strategy is given by $k_{h_{t,v_H}} = (1+\lambda^*)k_{h_t}$. One can similarly write Problem MP recursively, with associated value function $V_{MP}(h_t, k)$. The dependence on history h_t indicates that the optimization problem is not time invariant because parameters $\bar{\lambda}(\cdot)$ may depend on the history of past shocks.

Lemma 3 *Problem MP has a unique solution. Let $V_{MP}(h_t, k)$ be the continuation utility given history h_t and current capital level k, and let $k_{MP}(h_0, v_H)$ be the optimal capital choice given history (h_0, v_H). For all $k > 0$,*

1. *$V_{MP}(h_t, k) = a_{m,h_t}\, k^{1-s}/(1-s)$, where $a_{m,h_t} > 0$ is a constant.*
2. *$V'_{MP}(h_0, k) \geq V'_{BP}(k)$. Further, if $\bar{\lambda}(h_t) \neq \lambda^*$, for some $t > 0$, then $V'_{MP}(h_0, k) > V'_{BP}(k)$.*
3. *If $\bar{\lambda}(h_0) \leq \lambda^*$, then constraint (29) binds, i.e., $k_{MP}(h_0, v_H) = (1+\bar{\lambda}(h_0))k_{h_0}$.*
4. *If $\bar{\lambda}(h_t) \neq \lambda^*$, for some $t > 0$, and $\bar{\lambda}(h_0) > \lambda^*$, then $k_{MP}(h_0, v_H) > (1+\lambda^*)k_{h_0}$.*

Proof of Lemma 3. Problem MP has a unique solution, because the objective is strictly concave and the constraint set is convex. Because $\bar{\lambda}(\cdot) \geq -\delta$, the condition $\beta(1-\delta)^{1-s} < 1$, implied by Assumption 1, ensures that utility is finite.
Statement 1. Let $k_{h_0} = \bar{k}$ be the initial capital stock and let $\{\bar{k}_{h_t}\}_{t=0}^{\infty}$ be an optimal solution to problem MP. Now, multiply the initial capital by $\alpha > 0$, so that the initial capital stock is $\hat{k} = \alpha\bar{k}$, and consider the sequence $\{\hat{k}_{h_t}\}_{t=0}^{\infty}$, where $\hat{k}_{h_t} = \alpha\bar{k}_{h_t}$. This new sequence $\{\hat{k}_{h_t}\}_{t=0}^{\infty}$ satisfies all the constraints. Thus, the same proof for the scalability of the value function in problem BP applies.
Statement 2. Problem MP is more constrained than problem BP. Thus, $V_{MP}(h_0, k) \leq V_{BP}(k) < 0$ for all k. This, in turn, implies $a_b \leq a_{m,h_0}$, which implies $V'_{BP}(k) \leq V'_{MP}(h_0, k)$ for all k. If $\bar{\lambda}(h_t) \neq \lambda^*$ for any constraint that is reached with positive probability, then $V_{MP}(h_0, k) < V_{BP}(k) < 0$ for all k. Thus, $a_b < a_{m,h_0}$, which, in turn, implies that $V'_{BP}(k) < V'_{MP}(h_0, k)$ for all k.
Statement 3. From Part 1 of Lemma 3 and (19), the objective function in Problem BP is strictly concave. Thus, from Part 2 of Lemma 3 and (20),

$$-u'((1+\phi)k - k') + \beta V'_{BP}(k') > 0, \quad \text{for all } k' < (1+\lambda^*)k. \tag{32}$$

Moreover, from Part 2, $V'_{BP}(k) \leq V'_{MP}(h_0, k)$. Combining this with (32), we have

$$-u'((1+\phi)k - k') + \beta V'_{MP}(h_0, k') \geq -u'((1+\phi)k - k') + \beta V'_{BP}(k') > 0, \text{ for all } k' < (1+\lambda^*)k. \tag{33}$$

Thus, for Problem MP, we have $k_{MP}(h_0, v_H) = (1 + \bar{\lambda}(h_0))k_{h_0}$ for any $\bar{\lambda}(h_0) \leq \lambda^*$.
Statement 4. Because $\bar{\lambda}(h_t) \neq \lambda^*$, $t > 0$, Part 2 implies that $V'_{BP}(k) < V'_{MP}(h_0, k)$. Thus, we get a strict inequality in (33), which implies the result. ■

Step 2. We use Lemma 3 to prove the statements of the Proposition.
Statement 1. Let $\bar{k}$ be the initial capital level and let $k_P(h_t)$ be the optimal capital level given history h_t in Problem P. Suppose $\bar{k} \leq k^*$ and $k_P(h_0, v_H) < (1+\lambda_c(\bar{k}))\bar{k}$. Using this posited optimal solution, we now construct parameters for the constraints of Problem MP that we use to derive a contradiction. Consider Problem MP, starting at $t = 0$ with initial capital $\bar{k}$ and the constraints given by $\bar{\lambda}(h_0) = \lambda_c(\bar{k})$ and $\bar{\lambda}(h_t) = (k_P(h_t, v_H)/k_P(h_t)) - 1$, for $t > 0$. By Part 3 of Lemma 3, the optimal capital level at $t = 1$ is $k_{MP}(h_0, v_H) = (1+\lambda_c(\bar{k}))\bar{k}$, and hence $k_{MP}(h_0, v_H) > k_P(h_0, v_H)$. By construction, the optimal choices in Problem P, $\{k_P(h_t) \colon h_t \in \mathcal{H}_t\}$, satisfy the constraints of Problem MP. Thus, the capital levels $\{k_{MP}(h_t) \colon h_t \in \mathcal{H}_t\}$ generate a strictly higher value for the objective function of Problem MP than $\{k_P(h_t) \colon h_t \in \mathcal{H}_t\}$ that we posited solve Problem P. Because the objective functions of Problem MP and Problem P are identical, $\{k_{MP}(h_t) \colon h_t \in \mathcal{H}_t\}$ also generate a strictly higher value for the objective function of Problem P than $\{k_P(h_t) \colon h_t \in \mathcal{H}_t\}$. To obtain a contradiction, we show that $\{k_{MP}(h_t) \colon h_t \in \mathcal{H}_t\}$ is feasible in Problem P.

To see this, note that $k_{MP}(h_0, v_H) > k_P(h_0, v_H)$. This and constraint (30) imply that the capital level $k_{MP}(h_t)$ that solves Problem MP exceeds the $k_P(h_t)$ that solves Problem P. This implies

$$\frac{k_{MP}(h_t, v_H)}{k_{MP}(h_t)} = 1 + \bar{\lambda}(h_t) = \frac{k_P(h_t, v_H)}{k_P(h_t)} \leq 1 + \lambda_c(k_P(h_t)) \leq 1 + \lambda_c(k_{MP}(h_t)),$$

where the last inequality follows because $\lambda_c(k)$ is increasing in k.

The above argument uses statement 2 of Lemma 3, which establishes the strict inequality between the marginal products of capital for problems MP and BP, respectively. Thus, by continuity the argument immediately extends to capital levels k that are not too far above k^*.

The proof for $\bar{k} > \hat{k}$ is analogous, except that we use Part 4 of Lemma 3. In particular, because $\bar{k} > k^*$, we have $\bar{\lambda}(h_0) = \lambda_c(\bar{k}) > \lambda^*$. Because $p_H < 1$, the demagogue sometimes wins and hence the capital sometimes will fall. Thus, $\lambda_c(k_P(h_t)) < \lambda^*$ for some $t > 0$. Thus, by Part 4 of Lemma 3, $k_{MP}(h_0, v_H) > (1 + \lambda^*)\bar{k}$. The rest of the proof is identical to above.
Statement 2. If $k \leq k^*$, Part 1 implies $k' = (1 + \lambda_c(k))k$. Next, suppose $k > k^*$. Because $p_H = 1$, the demagogue never wins in equilibrium. Thus, the solution to Problem BP (i.e., $k' = (1+\lambda^*)k$) is always feasible in Problem P.
Statement 3. Let $k > k^*$. Define $n(k) = \min\{n \in \mathbb{N} | (1 - \delta)^n k > k^*\}$. If we start with capital k, then we can eliminate constraint (2) in the optimization problem for the first n periods.

From the second statement it follows that the result is immediate if $p_H = 1$; and from the first statement for $p_H < 1$, we have $\lambda(k) > \lambda^*$ for $k > k^*$. Suppose by way of contradiction that $\liminf_{k\to\infty} \lambda(k) > \lambda^*$. Consider Problem BP, which does not have constraint (2). Then the strict optimality of λ^* implies that party b's utility under λ^* exceeds that from investments $\lambda(k)$ by at least some amount $\varepsilon > 0$. Further, there exists a time period T such that, for any investment strategy of the infinite horizon model, the investment strategy restricted to a model with finite time horizon T results in a utility level that differs from that over the infinite horizon by at most $\varepsilon/2$. Thus, the utility from using λ^* for T periods if constraint (2) is slack for those periods exceeds that from using $\lambda(k)$ by at least $\varepsilon/2$. However, if k is large then constraint (2) is slack for T periods. This contradicts the posited optimality of $\lambda(k)$. ■

Proof of Proposition 6. *Statement 1.* From Proposition 5, death spiral occurs with probability 1 if capital drops below $\bar{k}$. Starting with a capital level, k, we reach $\bar{k}$ if we have α consecutive low valence realizations, where $(1-\delta)^\alpha k \le \bar{k}$, i.e., $\alpha = \log(\bar{k}/k)/\log(1-\delta)$. If α is not an integer, we need one additional low valence realization. Thus, the probability of reaching $\bar{k}$ is at least $p_L^{\alpha+1}$.
Statements 2 and 3. Suppose that when party b wins and current capital is k, it invests $\lambda^* k$. We will use this to find bounds on the probability of a death spiral. Consider a log-scale, so that with probability p_H, $\log(k') = \log(k) + \log(1 + \lambda^*)$; and with probability p_L, $\log(k') = \log(k) - |\log(1 - \delta)|$. Let $k_H = \log(1 + \lambda^*)$ and $k_L = |\log(1 - \delta)|$. Thus, we have a random walk with two potentially unequal steps: in each period, with probability p_H, the location moves up a step of size k_H; and with probability p_L, the location moves down a step of size k_L. We start from location $\log(k)$, with $k > k^*$, and we are interested in the probability that the location falls to $\log(k^*)$ (or below) at any future period. Changing the origin, this is equivalent to the probability that, starting from $\log(k/k^*)$, the location becomes non-positive (≤ 0) at any future period.

This process corresponds to a gambler's ruin problem in which one player is infinitely rich, analyzed, for example, in Chapter 14.8 of Feller (1968). First, suppose there exists a upper absorbing location $a > 0$, so that if the location z weakly exceeds a, then the process ends. Our analysis corresponds to the limit as $a \to \infty$. Let $Q(z)$ be the probability that the location becomes non-positive at any future period when the process starts from z at time 0. Then,

$$Q(z) = p_H Q(z + k_H) + p_L Q(z - k_L), \text{ for } 0 < z < a, \tag{34}$$

with boundary conditions

$$Q(x) = \begin{cases} 1 & \text{if } x \le 0; \\ 0 & \text{if } x \ge a. \end{cases} \tag{35}$$

The characteristic equation associated with equation (34) is

$$L(\sigma) \equiv p_H \sigma^{k_H} + p_L \sigma^{-k_L} = 1. \tag{36}$$

This equation always has a solution at $\sigma = 1$. Moreover, $L(\sigma)$ is strictly convex in $\sigma \in (0, \infty)$, with $\lim_{\sigma \to 0} L(\sigma) = \lim_{\sigma \to \infty} L(\sigma) = \infty$. Suppose the process does not have a zero mean, i.e., $p_L(-k_L) + p_H(k_H) \neq 0$. Then, $L'(1) \neq 0$ and equation (36) has exactly one other positive solution besides 1: mirroring Feller (1968)'s analysis on p. 366, for the purpose of finding bounds on the probability of ruin, we do not need to consider negative solutions to equation (36). Call this solution σ_1. If $L'(1) < 0$, then $\sigma_1 > 1$. If, instead, $L'(1) > 0$, then $\sigma_1 < 1$. Then, $Q(z) = A + B\sigma_1^z$, where A and B are constants, satisfies equation (34) for some A and B.

Next, observe that if we choose $A = \bar{A}$ and $B = \bar{B}$ such that $\overline{Q}(z = 0) \equiv Q(z = 0; A = \bar{A}, B = \bar{B}) = 1$ and $\overline{Q}(z = a + k_H) \equiv Q(z = a + k_H; A = \bar{A}, B = \bar{B}) = 0$, then

$$\overline{Q}(x) \geq \begin{cases} 1 & ; -k_L \leq x \leq 0 \\ 0 & ; a \leq x \leq a + k_H. \end{cases} \tag{37}$$

Thus, $\bar{A} + \bar{B}\sigma_1^z - Q(z)$ satisfies the difference equation (34) with non-negative boundary values (37). Thus, $\bar{A} + \bar{B}\sigma_1^z \geq Q(z)$. This, will be an upper bound on $Q(z)$. To find $\bar{A}$ and $\bar{B}$, observe that

$$\overline{Q}(0) = \bar{A} + \bar{B} = 1, \text{ and } \overline{Q}(a + k_H) = \bar{A} + \bar{B}\sigma_1^{a+k_H} = 0.$$

Thus,

$$\bar{A} = \frac{-\sigma_1^{a+k_H}}{1 - \sigma_1^{a+k_H}}, \text{ and } \bar{B} = \frac{1}{1 - \sigma_1^{a+k_H}},$$

which implies

$$Q(z) \leq \bar{A} + \bar{B}\sigma_1^z = \frac{\sigma_1^{a+k_H} - \sigma_1^z}{\sigma_1^{a+k_H} - 1}. \tag{38}$$

We can find a lower bound for $Q(z)$ in a similar manner. Choose $A = \underline{A}$ and $B = \underline{B}$, so that $\underline{Q}(z = -k_L) \equiv Q(z = -k_L; A = \underline{A}, B = \underline{B}) = 1$ and $\underline{Q}(z = a) \equiv Q(z = a; A = \underline{A}, B = \underline{B}) = 0$. Then,

$$\underline{Q}(-k_L) = \underline{A} + \underline{B}\sigma_1^{-k_L} = 1, \text{ and } \underline{Q}(a) = \underline{A} + \underline{B}\sigma_1^a = 0.$$

Thus,

$$\underline{A} = \frac{-\sigma_1^a}{\sigma_1^{-k_L} - \sigma_1^a}, \text{ and } \underline{B} = \frac{1}{\sigma_1^{-k_L} - \sigma_1^a},$$

which implies

$$\underline{A} + \underline{B}\sigma_1^z = \frac{\sigma_1^a - \sigma_1^z}{\sigma_1^a - \sigma_1^{-k_L}} \leq Q(z). \tag{39}$$

Combining (38) and (39) yields

$$\frac{\sigma_1^a - \sigma_1^z}{\sigma_1^a - \sigma_1^{-k_L}} \leq Q(z) \leq \frac{\sigma_1^{a+k_H} - \sigma_1^z}{\sigma_1^{a+k_H} - 1}. \tag{40}$$

If $L'(\sigma = 1) < 0$, so that $\sigma_1 > 1$, then, as $a \to \infty$, both the lower and the upper bounds in (40) converge to 1, and hence so does $Q(z)$.

If, instead, $L'(\sigma = 1) > 0$, so that $\sigma_1 < 1$, then, from (40), in the limit when $a \to \infty$, we have

$$\sigma_1^{k_L+z} \leq Q(z) \leq \sigma_1^z, \tag{41}$$

where we recall that $z = \log(k/k^*)$ and $\sigma_1 < 1$ is the unique positive solution, other than one, of the characteristic equation (36).

Thus, we have proven the following.

Result. *Starting from $k > k^*$ and assuming that party b always invests λ^* when it wins, the probability of falling to k^* or below in any future period, denoted by $P(k; k^*, \lambda^*)$, is such that:*

1. If $p_H \log(1 + \lambda^*) + p_L \log(1 - \delta) < 0$, then $P(k; k^*, \lambda^*) = 1$.

2. If $p_H \log(1 + \lambda^*) + p_L \log(1 - \delta) > 0$, then

$$\sigma_1^{\log(k/k^*)+|\log(1-\delta)|} \leq P(k; k^*, \lambda^*) \leq \sigma_1^{\log(k/k^*)},$$

where $\sigma_1 < 1$ is the unique positive solution to $p_H \sigma^{\log(1+\lambda^*)} + p_L \sigma^{\log(1-\delta)} = 1$. Using $y = \sigma^{-\log(1-\delta)}$, this equation is equivalent to $p_H y^{1-\frac{\log(1+\lambda^*)}{\log(1-\delta)}} - y + p_L = 0$.

We now apply these results to our setting, in which the benevolent party's investment decision depends on the capital stock k. Let $Q(k)$ be the probability that, starting from a capital level k, the economy enters a death spiral (i.e., the capital stock falls below $\bar{k}$) in some period.

From Proposition 4, when $k > k^*$, the equilibrium investment is $\lambda(k) \geq k^*$. This implies that the step up exceeds λ^*. Moreover, recall that $k^* > \bar{k}$. Thus, the upper bound on the probability of ruin in Statement 2 of the Result directly applies: If $p_H \log(1 + \lambda^*) + p_L \log(1 - \delta) > 0$, then $Q(k) \leq \sigma_1^{\log(k/k^*)}$, where $\sigma_1 \in (0, 1)$ solves $p_H \sigma^{\log(1+\lambda^*)} + p_L \sigma^{log(1-\delta)} = 1$. Substituting $y = \sigma^{-\log(1-\delta)}$ yields the the upper bound.

Next, observe that Statement 1 of the Result does not depend on k or k^*. However, it assumes that party b invests a fraction λ^* when it wins, implying a step up of size $\log(1 + \lambda^*)$. From Proposition 4, as k grows large, $\lambda(k)$ converges to λ^*. Thus, there exists a $\hat{k}$ such that for all $k > \hat{k}$, $\lambda(k)$ is close enough to λ^* that $p_H \log(1 + \lambda(k)) + p_L \log(1 - \delta) < 0$. Now, replace λ^* with $\sup_{k>\hat{k}} \lambda(k)$ and $\hat{k}$ instead of k^*. Then, starting from $k > \hat{k}$, the probability of falling below $\hat{k}$

is 1. From Statement 1 of this Proposition, starting from $k \leq \hat{k}$, there is a positive probability of going below $\bar{k}$ before going above $\hat{k}$. Let $\hat{p} = p_L^{1+\alpha}$, with $\alpha = \log(\bar{k}/\hat{k})/\log(1-\delta)$. Moreover, as we just showed, if we go above $\hat{k}$, then with probability 1, we go below $\hat{k}$ again. Thus, starting from $k > \hat{k}$, the probability that we never go below $\bar{k}$ does not exceed $\lim_{n\to\infty}(1-\hat{p})^n = 0$. Thus, if $p_H \log(1+\lambda^*) + p_L \log(1-\delta) < 0$, then k falls below $\bar{k}$ with probability 1.

We use a similar argument to find a lower bound for $Q(k)$ when k is sufficiently large. Let $\lambda_m = \sup_{k \geq \tilde{k}} \lambda(k)$. From Statement 2 of the Result, if we replace λ^* with λ_m, then the probability of falling from $k > \tilde{k}$ to $\tilde{k}$, $P(k; \tilde{k}, \lambda_m)$, is larger than $\tilde{\sigma}_1^{\log(k/\tilde{k})+|\log(1-\delta)|}$, where $\tilde{\sigma}_1 < 1$ is the unique positive solution to $p_H \sigma^{\log(1+\lambda_m)} + p_L \sigma^{-log(1-\delta)} = 1$. It is straightforward to show that $\sigma_1 > \tilde{\sigma}_1$. Thus,

$$P(k; \tilde{k}, \lambda_m) \geq \tilde{\sigma}_1^{\log(k/\tilde{k})+|\log(1-\delta)|} = \tilde{\sigma}_1^{|\log(1-\delta)|} \left(\frac{\tilde{\sigma}_1}{\sigma_1}\right)^{\log(k/\tilde{k})} \sigma_1^{\log(k/\tilde{k})} = \tilde{\sigma}_1^{|\log(1-\delta)|} \left(\frac{k}{\tilde{k}}\right)^{\log(\tilde{\sigma}_1/\sigma_1)} \left(\frac{k}{\tilde{k}}\right)^{\log(\sigma_1)}.$$

Because $\lim_{k\to\infty} \lambda(k) = \lambda^*$, it follows that for every $\varepsilon > 0$ there exists $\tilde{k}$ such that, for all $k > \tilde{k}$, $1 - \varepsilon < \tilde{\sigma}_1/\sigma_1 \leq 1$.

First, suppose $k \leq \tilde{k}$. The probability of dropping below $\bar{k}$ is bounded away from zero for $k \leq \tilde{k}$. Moreover, the terms on the left-hand side of the inequality in the Proposition is continuous on the compact set $[\bar{k}, \tilde{k}]$. Thus, there exists a $C > 0$ such that the lower inequality is satisfied.

Next, suppose $k > \tilde{k}$. Let $C_1(\tilde{k}) = \tilde{\sigma}_1^{|\log(1-\delta)|}$. Then

$$P(k; \tilde{k}, \lambda_m) \geq C_1(\tilde{k}) \left(\frac{k}{\tilde{k}}\right)^{\log(\sigma_1)-\varepsilon},$$

where $C_1(\tilde{k}) > 0$ is independent of k.

Now, starting from $\tilde{k}$, the probability of falling to $\bar{k}$ is a constant, which is strictly between 0 and 1. Call it $C_2(\tilde{k}) \in (0,1)$. Thus, for $k > \tilde{k} > \bar{k}$,

$$Q(k) \geq C(\tilde{k}) \left(\frac{k}{k^*}\right)^{\log(\sigma_1)-\varepsilon}, \tag{42}$$

and $C(\tilde{k}) = C_1(\tilde{k}) \cdot C_2(\tilde{k}) \cdot \left(\frac{k^*}{\tilde{k}}\right)^{\log(\sigma_1)-\varepsilon}$ is bounded away from 0 and 1. ■

References

Acemoglu, D., G. Egorov, and K. Sonin (2013). A political theory of populism. *The Quarterly Journal of Economics 128*(2), 771–805.

Aguiar, M. and M. Amador (2011). Growth in the shadow of expropriation. *Quarterly Journal of Economics 126*(2), 651–697.

Alesina, A. and G. Tabellini (1990). A positive theory of fiscal deficits and government debt. *The Review of Economic Studies 57*(3), 403–414.

Baccini, L. and T. Sattler (2021). Austerity, economic vulnerability, and populism. SSRN Working Paper 3766022.

Baron, D. P. and D. Diermeier (2001). Elections, governments, and parliaments in proportional representation systems. *The Quarterly Journal of Economics 116*(3), 933–967.

Battaglini, M. (2014). A dynamic theory of electoral competition. *Theoretical Economics 9*(2), 515–554.

Battaglini, M. and S. Coate (2008). A dynamic theory of public spending, taxation, and debt. *American Economic Review 98*(1), 201–36.

Bernhardt, D., S. Krasa, and M. Shadmehr (2019a). Demagogues and the fragility of democracy. working paper.

Bernhardt, D., S. Krasa, and M. Shadmehr (2019b). Demagogues in America: From the revolution to the Second World War. working paper.

Bisin, A., A. Lizzeri, and L. Yariv (2015). Government policy with time inconsistent voters. *American Economic Review 105*(6), 1711–37.

Burman, L. E., J. Rohaly, J. Rosenberg, and K. C. Lim (2010). Catastrophic budget failure. *National Tax Journal 63*, 561–584.

Cukierman, A. and A. H. Meltzer (1989). A political theory of government debt and deficits in a neo-ricardian framework. *The American Economic Review*, 713–732.

Dal Bo, E., P. Dal Bo, and E. Eyster (2017). The demand for bad policy when voters underappreciate equilibrium effects. *The Review of Economic Studies 85*(2), 964–998.

Feller, W. (1968). *An introduction to probability theory and its applications*. John Wiley & Sons, Inc.: New York, London, Sidney.

Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives 30*(1), 185–206.

Guiso, L., H. Herrera, M. Morelli, and T. Sonno (2018). Populism: Demand and supply. working paper.

Halac, M. and P. Yared (2014). Fiscal rules and discretion under persistent shocks. *Econometrica 62*(5), 1557–1614.

Kelly, B., L. Pástor, and P. Veronesi (2016). The price of political uncertainty: Theory and evidence from the option market. *The Journal of Finance 71*(5), 2417–2480.

Kennedy, D. M. (1999). *Freedom from Fear: The American People in Depression and War, 1929-1945*. New York, NY: Oxford University Press.

Levy, G., R. Razin, and A. Young (2021). Misspecified politics and the recurrence of populism. working paper.

Middlekauff, R. (2007). *The Glorious Cause: The American Revolution, 1763-1789*. New York, NY: Oxford University Press.

Müller, J.-W. (2017). *What is populism?* Penguin UK.

Persson, T. and L. E. Svensson (1989). Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences. *The Quarterly Journal of Economics 104*(2), 325–345.

Persson, T. and G. Tabellini (2000). *Political Economy: Explaining Economic Policy*. MIT Press.

Song, Z., K. Storesletten, and F. Zilibotti (2012). Rotten parents and disciplined children: A politico-economic theory of public expenditure and debt. *Econometrica 80*(6), 2785–2803.