Demagogues and the Instability of Democracy

Dan Bernhardt   Stefan Krasa   Mehdi Shadmehr*

December 31, 2018

Abstract

TBA

*University of Illinois, Department of Economics, Urbana, IL 61801 USA; skrasa@illinois.edu University of Chicago

1 Introduction

Our starting point are standard models of political competition. We first allow candidates to be both policy and office motivated, and the amount of office and policy motivation will differ between candidates. Second, we allow candidates to choose between policies that are beneficial in the long run vs policies that have a short-term gain, but are costly in the long run. We refer to the latter as an example of a populist policy. We assume that the electorate does not fully understand long-term costs and benefits. More formally, we assume that voters are myopic, while political parties discount the future at some rate $\beta > 0$.

The median voter derives utility from the policy chosen, as well from other unspecified characteristics of the candidates. This is a short cut for modeling other policy issues (e.g., cultural issues) that matter to voters. More formally, the median voter receives a utility ϵ from one of the candidates, where the realization of ϵ is not known at the time when candidates choose policies. Examples for forward looking vs instant gratification: Climate Change, Pensions, debt/deficit in lieu of higher taxes We can interpret party L as representing establishment parties that only care about policy and therefore converge to the same policy. This way, a result such that party L wins with almost 100% probability simply means that establishment parties always win.

2 Model

Two parties $P = L, R$ choose capital investment in each period, $t = 0, 1, 2, \ldots$. There is a consumption good and a capital good. If k_{t-1} is the amount of capital in period $t - 1$ then $f(k_{t-1}) = \phi k_{t-1}$ is the amount of the consumption good provided in period t, where $\phi > 0$. The consumption good can be invested and effectively turned into the capital good: the capital stock at time t is given by $k_t = \rho k_{t-1} + i_t$, where $i_t \geq 0$ is the investment, and $0 < 1 - \rho < 1$ is the depreciation rate of the capital stock. A party's policy in period t is a proposed investment level i_t.

The median voter's utility in period t is $u(c_t) + v_P$, where $c_t = f(k_{t-1}) - i_t$, and v_P is a valence shock that measures the utility the voter derives if party P is in power. We will focus on constant relative risk aversion utility preferences over consumption, i.e., $u(c) = c^{1-s}/(1 - s)$ for $s > 0$, $s \neq 1$ and $u(c) = \log(c)$ for relative risk aversion $s = 1$. We interpret $v_{P,t}$ as measuring other, non-economic policy aspects that are fixed. Without loss of generality we assume that $v_{L,t} = 0$ and write v_t instead of $v_{R,t}$. Let G_t be the cdf of the distribution of v_t, with density g_t. We require that $\lim_{x \to -\infty} xG(x) = 0$. [**explain**] The median voter is myopic, and bases electoral decision solely on period utility. This captures the idea that voters are not sophisticated and do not understand the long-term impacts of economic policy.[1]

In contrast, parties are sophisticated and forward looking. Parties discount future payoffs by a discount factor $0 < \beta < 1$. There are two parties, a populist party, R that only cares about winning. That is, the party receives a period payoff of 1 if it wins, and 0, otherwise. The other party, L, only cares about policy. The party's period utility from consumption is the same as that of the median voter, but the parties discounts future payoffs by β.

We make minimal assumptions about the distribution of valence. In particular, the expected net-valence does not need to be zero. We only require that $0 < G_t(0) < 1$, so that even if party L perfectly mimics party R's platform, then it has a strictly positive probability of losing, and that the density g_t is strictly positive on its support. To ease presentation, we typically assume that the valence distribution is i.i.d., and that the density g is strictly positive on $\mathbb{R}$.

Throughout the paper, we maintain the following assumption.

Assumption 1 $\beta(\rho + \phi)^{1-s} < 1 < \beta\phi$.

The first inequality ensures that L's discounted expected utility is finite. The second inequality means that the project's return exceeds the L's discount factor. If risk aversion $s \geq 1$ then $\beta\phi > 1$ implies that this second assumption is immediately satisfied.

The game proceeds as follows. At the beginning of each period, parties simultaneously propose policies $i_{P,t} \in [0, f(k_{t-1})]$, $P = L, R$. Then the valence shock is realized and the median voter selects his preferred party. The party that wins the election then implements its announced investment policy platform.

We consider subgame perfect equilibria of the game.

3 The Planner's Problem

We first determine the socially optimal level of investment, i.e., the investment that would maximize the median voter's discounted payoff. This outcome would arise absent competition from the populist party, because party L would then be unconstrained by re-election concerns.

Let $\bar{k} = k_{-1} > 0$ be the initial level of capital. Then the planner seeks to maximize

$$\max_{i_t} \sum_{t=0}^{\infty} \beta^t u(\phi k_{t-1} - i_t) \text{ s.t. } \quad k_t = \rho k_{t-1} + i_t,\ 0 \leq i_t \leq \phi k_{t-1}. \tag{1}$$

[1]This is a weak notion of bounded rationality. For example, voters do not understand the consequences of future debt or pension obligations on future consumption (explain more)—see Greece, or underinvestment in infrastructure.

A significant challenge with analyzing this dynamic optimization problem is that the flow payoff, $u(c)$, is unbounded. Thus, the usual results that ensure differentiability of the value function as well as the contraction mapping results that ensure uniqueness do not apply. Further, the problem does not map into the approaches used when one has constant returns to scale (Stokey and Lucas 1999, Ch. 4.3). One approach is to exploit the monotonicity of the Bellman operator, finding an upper bound for the value function, and showing that it converges to its fixed point (Stokey and Lucas 1999, Theorem 4.14). We take a simpler approach that exploits the scalability of the value function. This enables us to quickly arrive at the functional form of the value function. To show this, we cast the problem in the standard format with k_t as the state variable:

$$V(\bar{k}) = \max_{\{k_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(k_{t-1}, k_t) \text{ s.t. } \rho k_{t-1} \le k_t \le (\rho + \phi)k_{t-1}, \text{ with } k_{-1} = \bar{k}. \tag{2}$$

Let $\{\bar{k}_t\}_{t=0}^{\infty}$ be an optimal sequence. Now, suppose we multiply the initial capital by $\alpha > 0$, so that we begin from an initial capital stock of $\hat{k}_{-1} = \alpha\bar{k}$. Similarly, consider a sequence $\{\hat{k}_t\}_{t=0}^{\infty}$, $\hat{k}_t = \alpha\bar{k}_t$. Obviously, this new series, $\{\hat{k}_t\}_{t=0}^{\infty}$, satisfies all the constraints because $\rho\bar{k}_{t-1} \le \bar{k}_t \le (\rho + \phi)\bar{k}_{t-1}$ whenever $\rho\hat{k}_{t-1} \le \hat{k}_t \le (\rho + \phi)\hat{k}_{t-1}$. Moreover, because $V(\hat{k})$ is the optimal value of the series

$$V(\hat{k}) = V(\alpha\bar{k}) \ge \sum_{t=0}^{\infty} \beta^t u(\hat{k}_{t-1}, \hat{k}_t) = \begin{cases} \frac{\alpha^{1-s}}{1-s} V_s(\bar{k}) & ; s \ne 1 \\ \frac{log(\alpha)}{1-\beta} + V_s(\bar{k}) & ; s = 1 \end{cases} \tag{3}$$

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

$$V(\bar{k}) \ge \begin{cases} \frac{(1/\alpha)^{1-s}}{1-s} V_s(\alpha\bar{k}) & ; s \ne 1 \\ \frac{log(1/\alpha)}{1-\beta} + V_s(\alpha\bar{k}) & ; s = 1 \end{cases} \tag{4}$$

Now, suppose the inequality in (3) was strict. Then, combining (3) and (4), we would have $V(\alpha\bar{k}) > V(\alpha\bar{k})$, a contradiction. Thus, we must have

$$V(\alpha\bar{k}) = \begin{cases} \frac{\alpha^{1-s}}{1-s} V_s(\bar{k}) & ; s \ne 1 \\ \frac{log(\alpha)}{1-\beta} + V_s(\bar{k}) & ; s = 1 \end{cases}, \text{ that is, } V(k) = \begin{cases} \frac{k^{1-s}}{1-s} V_s(1) & ; s \ne 1 \\ \frac{log(k)}{1-\beta} + V_s(1) & ; s = 1 \end{cases},$$

where $V_s(1) \in \mathbb{R} \cup \{\pm\infty\}$. It remains to show that $V_s(1)$ is finite. First, consider a lower bound, which is obtained when the planner never invest, so that $k_{t-1} = \rho^t$, implying a consumption of $c_t = \phi k_{t-1} = \rho^t\phi$. When $s \ne 1$,

$$\sum_{t=0}^{\infty} \beta^t \frac{(\phi\rho^t)^{1-s}}{1-s} = \frac{1}{1-s} \sum_{t=0}^{\infty} (\beta(\phi\rho)^{1-s})^t = \frac{1}{1-s} \frac{1}{1-\beta(\phi\rho)^{1-s}} \in \mathbb{R},$$

because $\beta(\rho\phi)^{1-s} < \beta(\rho + \phi)^{1-s} < 1$ by Assumption 1 and the fact that $\rho < 1$. When $s = 1$,

$$\sum_{t=0}^{\infty} \beta^t log(\phi\rho^t) = \sum_{t=0}^{\infty} \beta^t (log(\phi) + tlog(\rho)) = \frac{1}{1-\beta} log(\phi) + \frac{\beta}{(1-\beta)^2} log(\rho) \in \mathbb{R}.$$

Next, observe that because $log(\rho) < 0$, for sufficiently large t, all the terms of the terms are negative, $\phi\rho^t < 0$. This implies that we have a bounded, decreasing sequence, and hence it converges. The same argument holds for the case of $s > 1$. For the case of $s < 1$, we use a similar argument, using the upper bound. Consider an upper bound, which is obtained from letting the planner invest all the output, so that $k_{t-1} = (\rho + \phi)^t$, but setting the consumption at the sum of principle capital and the output, so that $c_t = (\rho + \phi)k_{t-1} = (\rho + \phi)^{t+1}$. When $s \neq 1$,

$$\sum_{t=0}^{\infty} \beta^t \frac{((\phi + \rho)^{1+t})^{1-s}}{1-s} = \frac{(\rho + \phi)^{1-s}}{1-s} \sum_{t=0}^{\infty} (\beta(\phi + \rho)^{1-s})^t = \frac{(\rho + \phi)^{1-s}}{1-s} \frac{1}{1 - \beta(\phi + \rho)^{1-s}} \in \mathbb{R},$$

by Assumption 1.

We now show this result also implies that a constant fraction of the each period's capital is invested, and the remaining constant fraction is consumed. That is, the optimal investment in period t, is λk_t for some $\lambda > 0$, which depends on s. To see this, note that we showed that if $\bar{k}_t$ is the optimal sequence of capital stocks beginning from $\bar{k}_{-1}$, then $\hat{k}_t = \alpha\bar{k}_t$ is the optimal sequence of capital stocks beginning from $\alpha\bar{k}_{-1}$. Thus, if $\bar{i}_t = \bar{k}_t - \phi\bar{k}_{t-1}$ is the optimal investment sequence beginning from $\bar{k}_{-1}$, then $\hat{i}_t = \hat{k}_t - \phi\hat{k}_{t-1} = \alpha\bar{i}_t$ is the optimal investment sequence beginning from $\hat{k}_{-1} = \alpha\bar{k}_{-1}$.

In sum, the scalability of the initial stock of capital in our dynamic optimization problem generates a simple closed form value function, which is differentiable and concave. Thus, we can characterize the solution either by using the Euler equations, or by posing the problem in its recursive form, and then directly optimizing using the functional form of the value function. Proposition 1 characterizes the solution.

Proposition 1 *The planner's problem can be described recursively as*

$$V_P(k) = \max_{i \in [0, \phi k]} u(\phi k - i) + \beta V_P(\rho k + i) \tag{5}$$

and the value function takes the form $V_P(k) = k^{1-s} v_P/(1-s)$ *if* $s \neq 1$ *where* $v_P/(1-s) = V_P(1)$ *is the discounted lifetime payoff from starting with initial capital* $\bar{k} = 1$ *and* $V_P(k) = v_P + 1/(1-\beta)\log(k)$ *for* $s = 1$, *where* $v_P \in \mathbb{R}$. *If* k *is the capital at the beginning of a period, then the optimal investment is given by* $i(k) = \lambda k$, *where*

$$\lambda = (\beta(\rho + \phi))^{1/s} - \rho > 1 - \rho. \tag{6}$$

The planner always grows capital: $\lambda > 1 - \rho$. To see this, note that $\beta\phi > 1$ implies that $(\beta(\rho + \phi))^{1/s} > 1$. The optimal level of investment increases in ϕ and β. This simply reflects that the value of investment is higher if the future is discounted by less or if the project is more productive. The comparative statics with respect to ρ are more subtle. However, if $s \geq 1$ then increasing ρ, i.e., decreasing capital depreciation, reduces investment—the planner can achieve the same level of capital with reduced investment when capital depreciates more slowly.[2]

4 The Strategic Problem

4.1 Log Utility with Myopic Populist

Now suppose there are two parties, and that party R is the populist party who only care about winning the election. We first consider the case of a myopic populist party who only cares about winning the current period.

Electoral competition introduces uncertainty, which is captured by the sequence of wins and losses of party L. Let $\Omega = \{w, \ell\}^{\mathbb{N}}$ be the sequence representing these wins and losses, and let $W_t \colon \Omega \to \{0, 1\}$ be the random variable that is 1 if party L wins in period t, and is 0, otherwise. The history at time t is given by the sequence of wins and losses up to time $t - 1$. Let $\mathcal{F}_t$ be the filtration of Ω that corresponds to knowing that history. Then the capital stock and the investments $k_t, i_t \colon \Omega \to \mathbb{R}$ must be $\mathcal{F}_t$-measurable.

The probability distribution on Ω (i.e., the distribution over wins and losses) depends on k_t and i_t. To derive this distribution, let $\omega \in \Omega$. The median voter in state ω at time t is indifferent between the parties at a net-valence level ϵ_t at which $\log(\phi k_{t-1}(\omega) - i_t(\omega)) = \log(\phi k_{t-1}(\omega)) + \epsilon_t$. Let D_t be the utility difference from the two policy choices

$$D_t(\omega) = \log(\phi k_{t-1}(\omega) - i_t(\omega)) - \log(\phi k_{t-1}(\omega)) = log\left(\frac{\phi k_{t-1}(\omega) - i_t(\omega)}{\phi k_{t-1}(\omega)}\right). \tag{7}$$

Then the probability that L wins at time t in state ω is $G(D_t(\omega))$. Letting $E[\cdot; i, k]$ be the resulting expectation over Ω, party L's optimization problem is

$$\max_{i_t} E\left[\sum_{t=0}^{\infty} \beta^t \Big(W_t(\omega)\log(\phi k_{t-1}(\omega) - i_t(\omega)) + (1 - W_t(\omega))\log(\phi k_{t-1}(\omega))\Big); \{i_t(\omega), k_t(\omega)\}_{t\in\mathbb{N}}\right] \tag{8}$$
$$\text{s.t. } k_t(\omega) = \rho k_{t-1}(\omega) + i_t(\omega),\ 0 \leq i_t(\omega) \leq \phi k_{t-1}(\omega), \text{ where } k_t, i_t \text{ are } \mathcal{F}_t \text{ measurable.}$$

A key observation is that, with log utility, the winning probabilities remain unchanged after scaling: $D_t = log\left(\frac{\phi k_{t-1} - i_t}{\phi k_{t-1}}\right) = log\left(\frac{\phi\alpha k_{t-1} - \alpha i_t}{\phi\alpha k_{t-1}}\right)$. This implies that, as in the case of the social planner,

[2] Why the heck is this not true when $s < 1$ and β is small?

party L's value function is again scalable in k. In turn, this implies that the optimal investment is linear in capital stock.

Lemma 1 *The investment platform problem of party L can be written recursively, with a value function that takes the form* $V(k) = v + \log(k)/(1-\beta)$.

Using Lemma 1, we write party L's optimization as

$$\max_i G(D)\left(\log(\phi k - i) + \frac{\beta}{1-\beta}\log(\rho k + i)\right) + (1 - G(D))\left(\log(\phi k) + \frac{\beta}{1-\beta}\log(\rho k)\right), \tag{9}$$

where $D = \log((\phi k - i)/\phi k)$. Differentiating party L's objective with respect to i yields

$$G(D)\left(-\frac{1}{\phi k - i} + \frac{\beta}{(1-\beta)(\rho k + i)}\right) - \frac{g(D)}{\phi k - i}\left(\log\left(\frac{\phi k - i}{\phi k}\right) + \frac{\beta}{1-\beta}\ \log\left(\frac{\rho k + i}{\rho k}\right)\right).$$

Letting $i = \lambda k$, $D = \log\left(\frac{\phi k - i}{\phi k}\right)$ simplifies to $D = \log\left(\frac{\phi - \lambda}{\phi}\right)$, and our first order condition is[3]

$$G(D)\left(-\frac{1}{\phi - \lambda} + \frac{\beta}{(1-\beta)(\rho + \lambda)}\right) = g(D)\frac{1}{\phi - \lambda}\left(\log\left(\frac{\phi - \lambda}{\phi}\right) + \frac{\beta}{1-\beta}\left(\log\left(\frac{\rho + \lambda}{\rho}\right)\right)\right). \tag{10}$$

The left-hand side of (10) is the marginal effect of increasing savings today on the discounted stream of future consumption, absent electoral competition. The right-hand side of (10) is the marginal cost of increasing savings through raising the chances of Party R's winning, who will not save at all. It is the reduction in the probability of winning due to increased savings, $g(D)/(\phi - \lambda)$, times the difference in the payoffs from winning and investing λk, versus losing to party R who will not invest.

Party L seeks to maximizes social welfare just like the social planner, who does not face electoral competition. However, unlike the planner, party L can only implement their policy if elected, otherwise savings are chosen by party R at the worst possible level for voters. Thus, while the planner chooses $\lambda = \beta\phi - (1-\beta)\rho$ to make the left hand side zero, party L must reduce the saving rate to equate the left hand side with the marginal electoral costs of raising savings on the right hand side.

To characterize the equilibrium choices by party L, it is convenient to multiply both sides of (10) by $\phi - \lambda$, and rewrite the first-order condition as

$$G(D)\left(\beta\phi - (1-\beta)\rho - \lambda\right) = g(D)(\rho + \lambda)\left((1-\beta)\log\left(\frac{\phi - \lambda}{\phi}\right) + \beta\left(\log\left(\frac{\rho + \lambda}{\rho}\right)\right)\right). \tag{11}$$

[3] Observe that party L's objective function at the lower boundary of investment, $i = 0$, and the upper boundary of investment, $i = \phi k$, are the same. Moreover, the derivative of the objective at the lower boundary, $i = 0$, is positive if and only if $\beta(\rho + \phi) > \rho$, which is true by Assumption 1. Thus, we have an interior maximum.

Proposition 2

1. *Equilibrium investment is strictly less than the social optimum, i.e.,* $\lambda < \beta\phi - (1-\beta)\rho$.
2. *Let* $\lambda(\rho)$ *be the equilibrium level of investment given the depreciation rate* $1-\rho$. *Then if* ρ *goes to zero,*
   (a) *The winning probability* $G(D)$ *converges to* $G(0)$.
   (b) *The share of capital reinvested,* $\lambda(\rho)$, *converges to zero at the rate* $-G(0)\phi/(g(0)\log(\rho))$ *independent of the valence distribution, even though the social optimum converges to* $\beta\phi > 1$.
   (c) $\lambda(\rho)$ *is strictly increasing and convex in* ρ *for small* ρ.

That investment falls strictly below the social optimum follows because party L cares about winning in order to implement its policy, and hence equates the marginal benefit of increased investment with the marginal cost. At the social optimum, the marginal benefit is zero, while marginal cost is strictly positive. Hence, reducing investment is optimal.

The more crucial question is: how far below the social optimum does party L choose its investment? In some respects this is a quantitative question. However, clean characterizations obtain when the depreciation rate $1-\rho$ is large. One might expect that if party L has a very large ex-ante valence advantage, so that it is likely to win an election then party L would only marginally reduce its investment below the social optimum. After all, if the valence advantage is large, one might posit that $G(D)$ is large and $g(D)$ is small, and hence λ would be close to $\beta\phi - (1-\beta)\rho$.

This reasoning is flawed because the winning probability $G(D)$ is endogenous. In fact, there is a large marginal benefit of mimicking the demagogue's policy of $i = 0$ by lowering λ. As it turns out, the winning probability goes to $G(0)$, which is large when party L has a large ex-ante valence advantage, but it is strictly bounded away from one given the assumption that g is strictly positive on its support. As ρ becomes small, λ goes to zero and the difference between the socially optimal level of investment and the competitive level chosen by party L becomes maximal. In other words, even if the threat of a demagogue winning is remote, it can have large effects on the behavior of established parties and this effect is discontinuous. That is, if $G(D) = 1$ for all D then the socially optimal investment is made.[4] Otherwise, the distance between the socially optimal and the equilibrium λ goes to $\beta\phi$ as ρ approaches 0.

To see why λ must go to zero as ρ goes to zero, observe that $\log((\rho+\lambda)/\rho)$ would go to infinity if λ were bounded away from zero. That is, the marginal cost would become arbitrarily

[4] This obviously violates our assumption that g is strictly positive on its support.

large, making it optimal to lower λ. This implies that even though the social optimum would have investment of $\beta\phi > 1$, the socially-minded party L is so concerned about electoral competition that it drives investment down toward zero, albeit at a somewhat slow rate of $1/\log(\rho)$.

To establish the rate at which λ approaches zero we substitute $\lambda(\rho) = -G(0)\phi/(g(0)\log(\rho))$ into equation (11) and show that the left-hand side converges to the right-hand side.

Intuition for why this rate of convergence obtains when ρ is small comes from dropping terms that go to zero at a faster rate (e.g., ρ converges to zero faster than $\lambda(\rho)$). Thus, for small ρ, a solution of equation (11) is approximately equal to a solution of

$$-g(0)\lambda(\rho)\left(\log\left(\frac{\rho+\lambda(\rho)}{\rho}\right)\right)+G(0)\phi=0. \tag{12}$$

More work is still needed because (12) does not have a closed-form solution. This leads us to differentiate (12) with respect to ρ and then drop terms that go to zero at a faster rate, to get a simpler expression for the derivative. In particular, the derivative is

$$\lambda'(\rho)=\frac{\lambda^3(\rho)g(0)}{\rho(G(0)\phi\lambda(\rho)+G(0)\phi\rho+\lambda(\rho)^2g(0))}. \tag{13}$$

Both $\lambda(\rho)^2$ and ρ converge to zero at a faster rate than $\lambda(\rho)$. Thus, for small ρ we get

$$\lambda'(\rho)\approx\frac{\lambda^2(\rho)g(0)}{G(0)\phi\rho}. \tag{14}$$

Assuming that (14) holds with equality and solving the differential equation yields

$$\lambda(\rho)\approx\frac{G(0)\phi}{C-g(0)\log(\rho)}, \tag{15}$$

where C is a constant. Thus, we have an approximation of the solution to the original first-order condition, which shows that $\lambda(\rho)$ goes to zero at (the slow) rate of $-G(0)\phi/(g(0)\log(\rho))$. If also follows that $\lambda(\rho)$ is convex for small ρ.

The equilibrium competitive investment policy and the social optimal investment policy differ in other important ways. Observe that the socially optimal level of investment decreases linearly with ρ. This reflects that when the depreciation rate $1-\rho$ is smaller, less capital has to be re-invested. Proposition 2 establishes the surprising result that if ρ is small, then λ is *strictly increasing* in ρ—electoral competition causes party L to behave in the exact opposite direction of the social optimum. In fact, depending on parameters, $\lambda(\rho)$ can be increasing for all ρ, or it can decrease for ρ sufficiently close to one. This latter case occurs when the distribution G is shifted sufficiently far in favor of party L so that it is very likely to win—as a result, doing the 'right' thing begins to dominate electoral concerns when ρ is large.

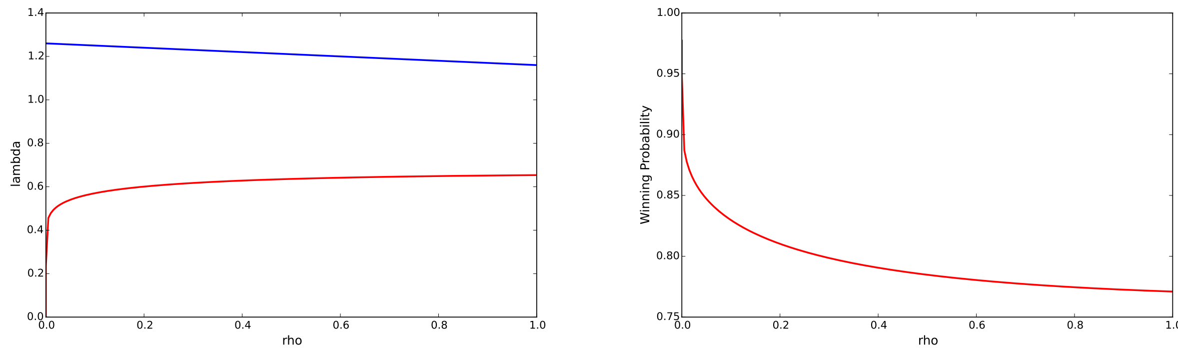

Figure 1: Investment (red), Social Optimal investment (blue), and L's winning probability for $\mu = -1$, $\sigma = 0.5$, $\phi = 1.4$, $\beta = 0.9$.

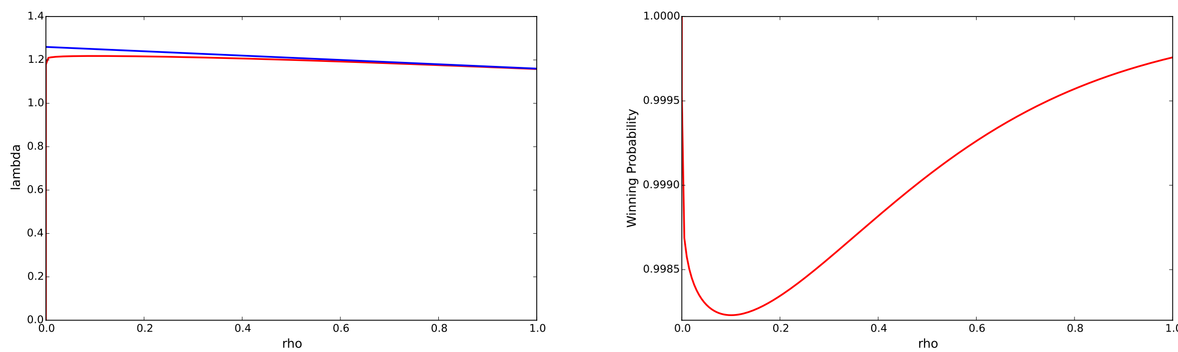

Figure 2: Investment (red), Social Optimal investment (blue), and L's winning probability for $\mu = -3.5$, $\sigma = 0.5$, $\phi = 1.4$, $\beta = 0.9$.

The two graphs have a normal distribution over party R's valence (dis)advantage with a standard deviation of 0.5 and different means. In both cases, party L has a significant valence advantage. In particular, for the parameterization in Figure 1, were L to perfectly mimic party R by offering $\lambda = 0$ then L would win with 97.7% probability. Nevertheless, λ is strictly increasing ρ and it is far less than the social optimum even when ρ is high. One measure of the tradeoff that party L faces is that its equilibrium probability of winning falls below 0.8 for $\rho > 0.4$. That is, to achieve the equilibrium benefits of saving, L reduces its ex-ante electoral advantage by over 20 percent. At the same time, from the perspective of an outsider, it looks as if L is pandering to voters, reducing investment by approximately 50% from the social optimum. In other words, pandering to short-sighted voters has large welfare costs.

To illustrate how λ can decrease with ρ for high levels of ρ, behaving similarly to the socially

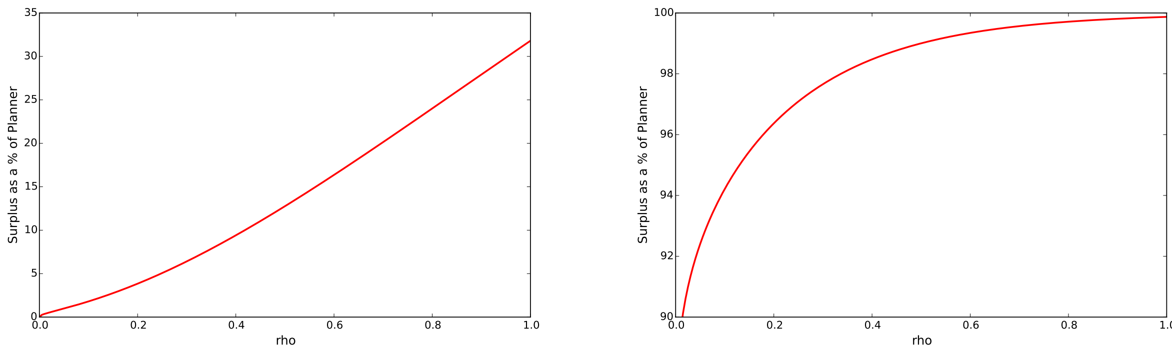

Figure 3: Welfare shortfall due to political competition, $\mu = -1.5$ left panel, and $\mu = -3.5$ right panel; $\sigma = 0.5$, $\phi = 1.4$, $\beta = 0.9$.

optimal investments, we maintain $\beta = 0.9$ and increase mean valence advantage of party L to μ to -3.5. At these values of μ and σ, the probability that candidate L loses if it perfectly mimics R is 1.28×10^{-12}, which is about 1,000 times less likely than winning the Powerball lottery. Figure 2 shows how the resulting investment rates vary with ρ. When ρ is small, λ is increasing as established in Proposition 2. However, as ρ increases, considerations of the marginal benefits in (11) start to dominate the marginal costs for larger ρ, causing party L to behave more like the social planner and reduce λ as ρ rises further. The non-monotonicity of λ is reflected in the lack of monotonicity of the winning probability. In particular, as λ decreases with ρ, party L's winning probability rises. Notably, even with such a large valence advantage, party L takes a 1/500 gamble of losing when $\rho = 0.1$ (roughly reversing the orders of magnitude in terms of winning the Powerball lottery), again highlighting the costs of doing the right policy when voters are short-sighted.

The figures illustrate that as $\mu \to \infty$, the equilibrium level of λ converges pointwise to the social optimum except at $\rho = 0$. Note that this convergence is not uniform. The graphs also indicate that this convergence is very slow.

The large difference between the optimal investment chosen by the planner compared to that by party L, even when L has a big valence advantage, translates into a large welfare loss. To compute this welfare loss, we calculate the level of constant consumption streams, c_P and c_L for the planner and party L, respectively, that would yield the same ex-ante expected utility. Figure 3 presents the ratio c_L/c_P in percent. The left panel, party L has a valence advantage of two standard deviations compared to the demagogue. Nevertheless, the presence of the demagogue results in a level of consumption that is always less than one third of that in a world without the demagogue. In the right-panel, party L's valence advantage is seven standard deviations. Nonetheless, even the now extremely remote possibility that the demagogue wins still meaningfully lowers welfare.

Proposition 2 implies that if ρ is small enough, then party L lowers λ below the capital replacement rate of $1-\rho$. Because party R never re-invests, it follows that the capital stock converges to zero in such a case. We say that the capital stock exhibits a death spiral if it converges to zero.

Corollary 1 *There exists a $\bar{\rho} > 0$ such that if $\rho \leq \bar{\rho}$, then the capital stock k_t converges to zero as $t \to \infty$.*

To prove Corollary 1 it suffices to establish that λ is close to zero when ρ is small. But death spirals are not limited to such extreme cases. In fact, death spirals can occur even when ρ is large, so that party L's investment rate exceeds the replacement rate of $1-\rho$. Indeed, we now show that the probability of a death spiral may go to one, even when the expected investment rate of $G(D)(\rho + \lambda(\rho)) + (1 - G(D))\rho$ (accounting for the probability that the demagogue wins) strictly exceeds the replacement rate $1-\rho$.

As a prelude to the characterization, we first provide necessary and sufficient conditions for a death spiral.

Lemma 2 *Let $\bar{k} > 0$ be the initial capital stock.*

1. *If $(\rho+\lambda)^{G(D)}\rho^{1-G(D)} < 1$, then for every $\varepsilon, \delta > 0$ there exists $\bar{t}$ such that the probability that $k_t < \delta\bar{k}$ exceeds $1-\varepsilon$ for all $t \geq \bar{t}$. That is, the probability of a death spiral goes to one.*

2. *If $(\rho+\lambda)^{G(D)}\rho^{1-G(D)} > 1$, then for every $\varepsilon > 0$ there exists $\bar{t}$ such that the probability that $k_t < \bar{k}$ is less than ε for all $t \geq \bar{t}$. That is, the probability of a death spiral goes to zero.*

Figure 4 plots the cutoff level of ρ below which death spirals are inevitable for different values of μ. Observe that if no party has a valence advantage, i.e., if $\mu = 0$, then death spirals occur unless ρ is close to one. Such a scenario could reflect a change in how the public perceives demagogues. If previous opprobrium by society toward demagogues declines, for example due to changes in political discourse or to changing attitudes about what is politically correct, then party L's valence advantage will decline. This contributes to death spirals in two ways. First, party R is more likely to win. But, what matters more, as our previous discussion explains and Figures 1 and 2 illustrate, is that party L begins to adopt policies that are closer to the demagogue's.

When party L's valence advantage rises, i.e., as μ falls, death spirals only emerge when ρ is smaller. Nonetheless, for intermediate levels of ρ, death spirals still arise when party L's

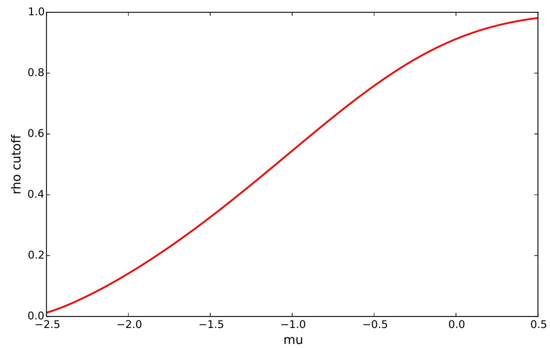

Figure 4: Cutoff value of ρ below which a death spiral occurs, $\sigma = 0.5$, $\phi = 1.4$, $\beta = 0.9$

valence advantage is large. To provide insights into when death spirals obtain, we identify the value of λ that maximizes long-run capital growth, solving

$$\max_{\lambda} (\rho + \lambda)^{G(D)} \rho^{1-G(D)}, \tag{16}$$

where, as usual, $D = \log((\phi - \lambda)/\phi)$. If the value of the objective of (16) exceeds one then Lemma 2 implies that no death spiral occurs. If it is less than one, then a death spiral occurs with probability 1.

The first-order condition of problem (16) is

$$G(D)(\phi - \lambda) - g(D)(\rho + \lambda) \log\left(\frac{\rho + \lambda}{\rho}\right) = 0. \tag{17}$$

To provide insights into the ways in which party L still acts as a long-run optimizer when it faces re-election concerns, we re-arrange party L's first-order condition for optimization, equation (11) to obtain

$$G(D)(\beta(\phi - \lambda) - (1 - \beta)(\rho + \lambda)) = g(D)(\rho + \lambda)\left((1 - \beta)\log\left(\frac{\phi - \lambda}{\phi}\right) + \beta \log\left(\frac{\rho + \lambda}{\rho}\right)\right),$$

which is equivalent to

$$\beta\left(G(D)(\phi - \lambda) - g(D)(\rho + \lambda)\log\left(\frac{\rho + \lambda}{\rho}\right)\right) + (1 - \beta)(\rho + \lambda)\,(G(D) - g(D)D) = 0. \tag{18}$$

When $\beta = 1$, equation (18) reduces to equation (17). Of course, party L's objective is not well-defined when $\beta = 1$, it means that when β is close to one, party L acts as if its primary objective is to avoid death spirals. More generally, party L's first-order condition for optimization can be re-written as a convex combination of the first-order condition of the problem (16) with weight β and a weight $(1 - \beta)$ on $(\rho + \lambda)(G(D) - g(D)D)$, which measures the distortion in party L's objective away from long-run survival.

Proposition 3 *The equilibrium investment rate λ increases in β. Let $\bar{\rho}$ be the minimum level of ρ such that death spirals do not emerge with probability one. Then there exists a $\hat{\rho} > \bar{\rho}$ such that party L's policy choice results in a death spiral even though there is a feasible policy that would ensure long-run survival.*

As β is increased toward 1, party L increasingly internalizes the future benefits of greater savings by increasing λ. At the value of λ that maximizes survival, the marginal benefit of increasing λ is less than the marginal cost, i.e., party L will choose a smaller λ. It follows that increases in β increase the set of depreciation rates for which the strategic party L invests enough to avoid a death spiral.

Democracies are not immune to demagogues — there is a lot of discussion on that, US constitution etc.

4.2 Log Utility with Non-myopic Populist

Markov Perfect Equilibria. First, we show that, restricting attention to linear Markov perfect equilibria, our results extend to the case of far-sighted populist. In particular, letting the state in each period be the initial capital stock, a Markovian strategy for player $P \in \{L, R\}$ is a mapping from the capital stock to an investment amount: $i_P(k) : [0, \infty) \to [0, \phi k]$. A linear Markovian strategy for player P is $i_P(k) = \lambda_P k$, for some $\lambda_P \in [0, \phi]$. A linear Markov perfect equilibrium is a linear strategy profile, $(\lambda_L k, \lambda_R k)$, of mutual best responses in every subgame.

Let $V_P(k; \lambda_{-P})$ be player P's value function when the initial capital stock is k and the other player uses the strategy $\lambda_{-P}\, k$. Then, (λ_L, λ_R) characterize a linear MPE if and only if they solve the following equations:

$$\begin{aligned} V_L(k;\lambda_R) = \max_{\lambda_L \in [0,\phi]} \quad & G\left(log\left(\frac{\phi-\lambda_L}{\phi-\lambda_R}\right)\right) \left(log((\phi-\lambda_L)k) + \beta\, V_L((\rho-\lambda_L)k;\lambda_R)\right) \\ + \quad & \left(1 - G\left(log\left(\frac{\phi-\lambda_L}{\phi-\lambda_R}\right)\right)\right) \left(log((\phi-\lambda_R)k) + \beta\, V_L((\rho-\lambda_R)k;\lambda_R)\right). \end{aligned}$$

$$V_R(k;\lambda_L) = \max_{\lambda_R \in [0,\phi]} G\left(log\left(\frac{\phi-\lambda_L}{\phi-\lambda_R}\right)\right) \left(\beta\, V_R((\rho-\lambda_L)k;\lambda_L)\right) + \left(1 - G\left(log\left(\frac{\phi-\lambda_L}{\phi-\lambda_R}\right)\right)\right) \left(r + \beta\, V_R((\rho-\lambda_R)k;\lambda_L)\right)$$

But expanding the populist's value function, we observe that it does not depend on the initial capital k:

$$v_R(\lambda_L) = V_R(k;\lambda_L) = \max_{\lambda_R \in [0,\phi]} \sum_{t=0}^{\infty} \beta^t \left(1 - G\left(log\left(\frac{\phi-\lambda_L}{\phi-\lambda_R}\right)\right)\right) r. \tag{19}$$

Therefore, the populist strategy is not to invest at all, i.e., to choose $\lambda_R = 0$. But from our analysis of the myopic case, we know that if the populist does not invest, investment strategies are linear. It remains to show that, if party L chooses a linear strategy, the the populist's best response is also linear. Using the scalability argument from before, we show that, in fact, if party L chooses a linear strategy, the the populist's best response is not to invest at all. Let $V_R(k)$ be part R's value function when the initial capital stock is k, and party L uses a linear strategy $\lambda_L k$.

$$V_R(\bar{k}) = \max_{\{i_t^R\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \left(1 - G\left(log\left(\frac{(\phi - \lambda_L)k_{t-1}}{\phi k_{t-1} - i_t^R}\right)\right)\right) r, \quad \text{(where } r \text{ is party } R\text{'s office rent)}$$

Let $\tilde{i}^R$ be the populist best response. Now, suppose the initial capital changes to $\alpha\bar{k}$, and consider a populist strategy in which the populist chooses $\hat{i}_t^R = \alpha\tilde{i}^R$. We have

$$V_R(\alpha\bar{k}) \geq \sum_{t=0}^{\infty} \left(1 - G\left(log\left(\frac{(\phi - \lambda_L)k_{t-1}}{\phi k_{t-1} - \tilde{i}_t^R}\right)\right)\right) r = V_R(\bar{k}). \tag{20}$$

Now, use $\frac{1}{\alpha}$ instead of α, and $\alpha\bar{k}$ instead of $\bar{k}$ to get $V_R(\bar{k}) \geq V_R(\alpha\bar{k})$. Thus, we have $V_R(\bar{k}) = V_R(\alpha\bar{k})$, and hence $V_R(k) = V_R(1)$. That is, if party L's uses a linear strategy, then the populist's payoff does not depend on the capital stock. Hence, the future capital stock is irrelevant for the populist, and hence he does not invest.

Thus, there is a unique linear MPE, which coincides with the behavior of a far-sighted idealist part and a myopic populist.[5]

$\sum_{t=0}^{\infty} \beta^t t$

The Limit of Finitely Many Periods. Next, we consider a finitely many period game, and show that our result with the myopic populist is reproduced in the limit when the number of periods goes to infinity.

The game begins in period $t = 0$ with initial capital $k_{-1} = \bar{k}$, and ends in period $T \geq 1$. We proceed by backward induction. In period T both players have a dominant strategy not to invest because. Thus,

$$i_T^R = i_T^L = 0, \;\; V_T(k_{T-1}) = log(\phi k_{T-1}),$$

where V_T is party L's (and R's) value function in period T. In period $t = T-1$, the populist does not invest because it hurts his current period payoff, and it does not change his future period payoffs. To analyze party L's optimization problem in period $t = T - 1$, let $V_{T-1}(k_{T-2})$ be its corresponding value function, and define $\lambda_{T-1} = \frac{i_{T-1}^L}{k_{T-2}}$. We have

$$\begin{aligned} V_{T-1}(k_{T-2}) = \max_{\lambda_{T-1}\in[0,\phi]} \quad & G\left(log\left(\frac{\phi - \lambda_{T-1}}{\phi}\right)\right) (log((\phi - \lambda_{T-1})k_{T-2}) + \beta V_T((\rho - \lambda_{T-1})k_{T-2})) \\ + \quad & \left(1 - G\left(log\left(\frac{\phi - \lambda_{T-1}}{\phi}\right)\right)\right) (log(\phi k_{T-2}) + \beta V_T(\rho k_{T-2}). \end{aligned}$$

[5] This result holds if we allow for linear strategies in which the coefficient depends on time, so that $i_t^P(k) = \lambda_t^P k$, where λ_t^P depends on t, but not on the history of play.

But recall that $V_T(k_{T-1}) = log(\phi k_{T-1})$, and hence we can factor out $log(k_{T-2})$ from the above maximization:

$$V_{T-1}(k_{T-2}) = (1+\beta)\, log(k_{T-2}) + \max_{\lambda_{T-1}\in[0,\phi]} \quad G\left(log\left(\frac{\phi-\lambda_{T-1}}{\phi}\right)\right)(log(\phi-\lambda_{T-1}) + \beta\, log(\rho-\lambda_{T-1})) + \left(1 - G\left(log\left(\frac{\phi-\lambda_{T-1}}{\phi}\right)\right)\right)(log(\phi)+\beta\, log(\rho)).$$

Thus, optimal λ_{T-1} depends on the parameters of the model, (ρ,ϕ,β), but on capital. Now, let

$$\tilde{V}_{T-1}(\beta,\phi,\rho) = \max_{\lambda_{T-1}\in[0,\phi]} \quad G\left(log\left(\frac{\phi-\lambda_{T-1}}{\phi}\right)\right)(log(\phi-\lambda_{T-1}) + \beta\, log(\rho-\lambda_{T-1})) + \left(1 - G\left(log\left(\frac{\phi-\lambda_{T-1}}{\phi}\right)\right)\right)(log(\phi)+\beta\, log(\rho)),$$

so that

$$V_{T-1}(k_{T-2}) = (1+\beta)\, log(k_{T-2}) + \tilde{V}_{T-1}(\beta,\phi,\rho). \tag{21}$$

Next, consider period $T-2$. Again, the populist chooses $i^R_{T-2} = 0$ because it maximizes his current period, and his future payoffs do not depend on capital stock—because party L's behavior, and hence R's winning probabilities do not depend on capital stock. Now, consider party L's optimization in period $T-2$. We repeat the above procedure.

$$V_{T-2}(k_{T-3}) = \max_{\lambda_{T-2}\in[0,\phi]} \quad G\left(log\left(\frac{\phi-\lambda_{T-2}}{\phi}\right)\right)(log((\phi-\lambda_{T-2})k_{T-3}) + \beta V_{T-1}((\rho+\lambda_{T-2})k_{T-3})) + \left(1 - G\left(log\left(\frac{\phi-\lambda_{T-2}}{\phi}\right)\right)\right)(log(\phi k_{T-3}) + \beta V_{T-1}(\rho k_{T-3}).$$

But recall that $V_{T-1}(x) = (1+\beta)\, log(x) + \tilde{V}_{T-1}(\beta,\phi,\rho)$, and hence we can factor out $log(k_{T-3})$ from the above maximization:

$$\begin{aligned} V_{T-2}(k_{T-3}) &= (1+\beta+\beta^2)\, log(k_{T-3}) \\ &+ \max_{\lambda_{T-2}\in[0,\phi]} G\left(log\left(\frac{\phi-\lambda_{T-2}}{\phi}\right)\right)(log(\phi-\lambda_{T-2}) + \beta(1+\beta)\, log(\rho+\lambda_{T-2}) + \beta\tilde{V}_{T-1}) \\ &+ \left(1 - G\left(log\left(\frac{\phi-\lambda_{T-2}}{\phi}\right)\right)\right)(log(\phi) + \beta(1+\beta)\, log(\rho) + \beta\tilde{V}_{T-1}). \end{aligned}$$

That is,

$$\begin{aligned} V_{T-2}(k_{T-3}) &= (1+\beta+\beta^2)\, log(k_{T-3}) + \beta\tilde{V}_{T-1} \\ &+ \max_{\lambda_{T-2}\in[0,\phi]} G\left(log\left(\frac{\phi-\lambda_{T-2}}{\phi}\right)\right)(log(\phi-\lambda_{T-2}) + \beta(1+\beta)\, log(\rho+\lambda_{T-2})) \\ &+ \left(1 - G\left(log\left(\frac{\phi-\lambda_{T-2}}{\phi}\right)\right)\right)(log(\phi) + \beta(1+\beta)\, log(\rho)). \end{aligned}$$

Thus, optimal λ_{T-2} depends on the parameters of the model, (ρ, ϕ, β), but on capital. Now, defining $\tilde{V}_{T-2}$ analogously as before we have

$$\tilde{V}_{T-2}(\beta, \phi, \rho) = \max_{\lambda_{T-2} \in [0, \phi]} \quad G\left(log\left(\frac{\phi - \lambda_{T-2}}{\phi}\right)\right)(log(\phi - \lambda_{T-2}) + \beta(1+\beta)\, log(\rho + \lambda_{T-2})) \\ + \left(1 - G\left(log\left(\frac{\phi - \lambda_{T-2}}{\phi}\right)\right)\right)(log(\phi) + \beta(1+\beta)\, log(\rho)).$$

Thus,

$$V_{T-2}(k_{T-3}) = (1 + \beta + \beta^2)\, log(k_{T-3}) + \tilde{V}_{T-2}(\beta, \phi, \rho) + \beta \tilde{V}_{T-1}(\beta, \phi, \rho). \tag{22}$$

Thus, one can show that

$$V_{T-n}(k_{T-(n+1)}) = (1 + \beta + \cdots + \beta^n)\, log(k_{T-(n+1)}) + \tilde{V}_{T-n} + \beta \tilde{V}_{T-(n-1)} + \cdots + \beta^{n-1} \tilde{V}_{T-1}. \tag{23}$$

Letting $n = T$, we have

$$V_0(k_{-1}; T) = \left(\sum_{t=0}^{T} \beta^t\right) log(k_{-1}) + \sum_{t=0}^{T-1} \beta^t \tilde{V}_t. \tag{24}$$

Thus, defining $v = \lim_{T\to\infty} \sum_{t=0}^{T-1} \beta^t \tilde{V}_t$ (**why does this limit exist? It corresponds to $V(1)$ in our initial problem. Why does $V(1)$ exist in our initial problem? We have lower and uppers bounds, and hence, we know it's finite, if it exists. But how do we know it converges?**), and recalling that $k_{-1} = \overline{k}$, we have

$$V(\overline{k}) = \lim_{T\to\infty} V_0(\overline{k}; T) = \frac{log(\overline{k})}{1-\beta} + v. \tag{25}$$

5 The Inevitability of Death Spirals

The previous section establishes that when voters have log preferences, the equilibrium probability with which party L wins does not vary with the level of the capital stock. As a result, party L's optimal policy is to propose a constant re-investment strategy. This means that death spirals either do or do not emerge, depending on parameters; and, in particular, whether they emerge does not depend on the existing capital stock. In practice, the macroeconomics literature provides extensive evidence that log preferences do not properly describe the level of risk aversion observed in consumer preferences, and that levels of relative risk aversion exceed one (estimates typically range from $s = 2$ to $s = 10$). We now show that when voters have CRRA utility with levels of relative risk aversion that exceed $s = 1$, then regardless of the other parameters describing the economy, given any initial capital stock k, the probability of a death spiral is strictly positive, bounded away from zero. We show that these death spirals become inevitable once capital levels fall too low [explain why bad economic shocks tend to drive more populist policies, cite Madison example in 1787, French municipalities, etc.]

Characterization of equilibrium behavior becomes far more challenging because the optimal investment as a share of capital is no longer constant. This reflects that for a given proposed investment rate policy λ, the utility that voters derive from the two policies now hinges on k, as $u((\phi - \lambda)k) - u(\phi k)$ is increasing in k. This means that, ceteris paribus, the larger is the capital stock, the more likely party R is to win re-election. The value function (and its derivative if it exists) that recursively describes party R's expected payoffs now depend in potentially complicated ways on the properties of the valence density. In particular, there is no reason for $G(u((\phi - \lambda)k) - u(\phi k))$ to be well-behaved. As a result, the standard arguments used to prove concavity or differentiability of the value function cannot be employed. In turn, this makes characterizing the optimal investment rate policy $\lambda(k) = \frac{i}{k}$ difficult.

The key step in our analysis is to identify a lower bound on the derivative of the value function. In turn, this delivers a lower bound on the marginal value of a positive investment rate. In particular, let $V_N(k)$ denote the value function if investment is zero in every period. We have $V_N(k) = \sum_{t=0}^{\infty} \beta^t \phi^{1-s} (\rho^{1-s})^t k^{1-s}/(1-s)$. Then, $k^{s-1} V_N(k)$ is finite if and only if $\beta\rho^{1-s} < 1$. In the analysis that follows, we maintain the assumption that $\beta\rho^{1-s} < 1$.

We establish the following bound.

Lemma 3 *Let $V(k)$ be the value function for party L. Then $V_N(k)$ is differentiable, and*

$$\liminf_{k' \to k} \frac{V(k) - V(k')}{k - k'} \geq V_N'(k). \tag{26}$$

The proof exploits the fact that were party L's probability of winning not to vary with k, then choosing a fixed investment rate $\lambda = i/k$ would be optimal. In turn, this would imply that $V(\alpha k) = \alpha^{1-s} V(k)$ for any k and $\alpha > 0$. But, for a fixed investment rate, the difference in utilities from the two proposed policies is decreasing in k, implying that the winning probability is increasing in k, and hence $\alpha^{1-s} V(k) \geq V(\alpha k)$ for any $\alpha \in (0, 1)$.

Similar arguments *cannot* be used to obtain *upper bounds* on the derivative of $V(k)$. In particular, even though $V_P > V(k)$, it need *not* be the case that $V_P'(k) > V'(k)$.[6]

Party L's optimization problem can be written recursively as follows:

$$V(k) = \max_{\lambda \in [0,\phi]} G(D)\Big(u((\phi - \lambda)k) + \beta V((\rho + \lambda)k)\Big) + (1 - G(D))\Big(u(\phi k) + \beta V(\rho k)\Big), \tag{27}$$

[6]If the derivative of V does not exist, we can replace it by the lim sup of the difference quotient as in Lemma 3.

where

$$D = u((\phi - \lambda)k) - u(\phi k).$$

Differentiating with respect to λ, and then using Lemma 3 to substitute in a lower bound on the marginal value of increased savings yields:

$$\begin{aligned}\frac{g(D)k^{1-s}}{(\phi-\lambda)^s}\Big(u((\phi-\lambda)k)+\beta V((\rho+\lambda)k)-u(\phi k)-\beta V(\rho k)\Big)\\ \geq G(D)\Big(-ku'((\phi-\lambda)k)+k\beta V_N'((\rho+\lambda)k)\Big).\end{aligned} \tag{28}$$

Substituting

$$V_N'(k) = \sum_{t=0}^{\infty}\beta^t(\rho^t)^{1-s}\phi^{1-s}k^{-s} = \frac{\phi^{1-s}k^{1-s}}{1-\beta\rho^{1-s}} \tag{29}$$

reveals that (28) is equivalent to

$$\begin{aligned}\frac{g(D)}{(\phi-\lambda)^s}\Big(u((\phi-\lambda)k)+\beta V((\rho+\lambda)k)-u(\phi k)-\beta V(\rho k)\Big)\\ \geq G(D)\left(-\frac{1}{(\phi-\lambda)^s}+\beta\frac{\phi^{1-s}}{1-\beta\rho^{1-s}}\right).\end{aligned} \tag{30}$$

Proposition 4 *For any capital level $k > 0$, the optimal investment rate for party L is strictly positive, i.e., $\lambda(k) > 0$.*

Proof. Suppose by way of contradiction that $\lambda = 0$. Then the right-hand side of (30) is strictly positive. In particular, cross multiplication yields the requirement that $\beta\phi > 1 - \beta\rho^{1-s}$, which holds given our assumption that $\beta\phi > 1$. The left-hand side is zero at $\lambda = 0$, a contradiction. ■

Proposition 5 *Suppose that the support of G is finite with $0 < G(0) < 1$. Then $\lim_{k\downarrow 0}\lambda(k) = 0$. Further, candidate L's winning probability satisfies $0 < G(D(k)) < G(0)$ for all $k > 0$.*

Proof. Suppose by way of contradiction that there exists a sequence k_n, $n \in \mathbb{N}$ with $\lim_{n\to\infty} k_n = 0$ and $\lambda(k_n) \geq \bar{\lambda} > 0$ for all $n \in \mathbb{N}$. Then $D(k_n) \to -\infty$, i.e., there exists $\bar{n}$ such that the probability that candidate L wins is zero for all $n \geq \bar{n}$. Thus, the payoff to candidate L is the same as from choosing $\lambda(k_n) = 0$. However, Proposition 4 established that the optimal $\lambda(k)$ is strictly positive, a contradiction.

Now suppose by way of contradiction that there exists a k such that $G(D(k)) = 0$. But then choosing $\lambda = 0$ provides the same payoff, contradicting the optimality of $\lambda(k) > 0$. Further, the optimality of $\lambda(k) > 0$ implies that $D(k) < 0$, and hence $G(D(k)) < G(0)$. ■

We next establish that even if party L has a large ex-ante valence advantage, and the initial capital stock is very large that the probability of a death spiral is strictly positive, bounded away from zero. In other words, as with log utility, the mere threat of competition from a demagogue can have large effects on party L, with severe consequences for the economy.

Proposition 6 *There exists a capital level $\bar{k}$ such that if $k \leq \bar{k}$ then a death spiral occurs with probability 1. Given any capital stock $k > \bar{k}$, the probability of dropping below $\bar{k}$ and entering a death spiral exceeds $(1 - G(0))^{n^*} > 0$, where $n^* = \lceil (\log(\bar{k}) - \log(k))/\log(\rho) \rceil$.*

Proof. The fact that $\lambda(k)$ converges to zero as k goes to zero implies that $\lambda(k) < 1 - \rho$ once k is sufficiently small. Thus, once k is sufficiently small, capital levels continue to fall even in the best case scenario in which party L wins.

Candidate L's winning probability can never exceed $G(0) < 1$. It follows immediately that for any k the probability of dropping below $\bar{k}$ is at least $(1 - G(0))^{n^*} > 0$, where $n^* = \lceil (\log(\bar{k}) - \log(k))/\log(\rho) \rceil$. ■

6 Investment in Good Times

We next establish how party L's investment policy choices vary with k when k is large. We show that the policies chosen are very different: rather than invest less than the replacement rate, as it does when capital is low, leading to a death spiral, party L now over-invests, choosing to save more than what it would absent electoral concerns.

Lemma 4 establishes that the amount of capital invested, λk, is uniformly bounded away from the maximum possible investment level, ϕk.

Lemma 4 *Suppose that the coefficient of relative risk aversion $s > 1$. Then there exists $\bar{\lambda} < \phi$ and $\bar{k} > 0$ such that $\lambda(k) < \bar{\lambda}$ for all $k \geq \bar{k}$.*

The formal proof is in the Appendix. Intuitively, it follows because if λk approaches ϕk then current consumption goes to zero, which implies that utility becomes arbitrarily negative. This, in turn, implies that the net benefit of investment is negative, a contradiction.

Lemma 4 implies that as k gets large, D_k converges uniformly to zero.

Our next step is to show that for large k, solution of Problem (27) are close to solutions of the following recursive optimization problem.

$$V_L(k) = \max_{\lambda \in [0,\phi]} G(0)\Big(u((\phi - \lambda)k) + \beta V_L((\rho + \lambda)k)\Big) + (1 - G(0))\Big(u(\phi k) + \beta V_L(\rho k)\Big). \tag{31}$$

In this optimization problem, party L's probability of winning does not depend on its platform choice—party L wins with probability $G(0)$, regardless of its action. In particular, if we replace $G(0)$ by 1 in (31) we get the social planner's problem, whose recursive form is given in (5). Thus, similar, to proposition 1 the value function takes the form $V_L(k) = k^{1-s}v_L/(1-s)$, where $v_L/(1-s) = V_L(1)$, i.e., the ex-ante expected utility of starting with one unit of capital. The investment share λ is again independent of k. Denote this share by λ^*.

To determine this share, note that the first-order condition of (31) is

$$(\phi - \lambda^*)^{-s}k^{1-s} = \beta v_L(\rho + \lambda^*)^{-s}k^{1-s}. \tag{32}$$

Solving this equation for λ^* yields

$$\lambda^* = \frac{(\beta v_L)^{1/s}\phi - \rho}{1 + (\beta v_L)^{1/s}}. \tag{33}$$

Lemma 5 shows that for large k, solutions of problem (27) are close to λ^*.

Lemma 5 *Let λ^* be the solution of problem* (31) *given by* (33)*. Let $\lambda(k)$ solve problem* (27)*. Then for every $\varepsilon > 0$ there exists $\hat{k}$ such that $|\lambda(k) - \lambda^*| < \varepsilon$, for all $k \geq \hat{k}$.*

In view of the lemma, it is sufficient to compare the planner's problem to (31). Proposition 1 shows that the value function for the planner's problem is $V_P(k) = k^{1-s}v_P/(1-s)$, where $v_P/(1-s) = V_P(1)$. Because ex-ante utility is maximized in the planner's problem, we get $V_P(1) > V_L(1)$. Because $1 - s < 0$, it follows that $v_P < v_L$. In general, if we have a value function of the form $k^{1-s}v/(1-s)$ then the investment share is

$$\lambda = \frac{(\beta v)^{1/s}\phi - \rho}{1 + (\beta v)^{1/s}}. \tag{34}$$

Note that λ as defined in (34) is strictly increasing in v. Because $v_P < v_L$ it follows that the investment share from problem (31), λ_L is strictly larger than the social planner's investment share, λ_P. Hence, lemma 5 implies that $\lambda(k) > \lambda_P$ for large k.

The same argument implies that that the solution of problem (31) is decreasing in $G(0)$. In other words, if the valence distribution becomes more favorable to the demagogue, then candidate L will respond by increasing the investment.

We now summarize these results.

Proposition 7 *If k is sufficiently large then:*

1. *Party L invests more than would a social planner who always gets his policy implemented, i.e., $\lambda(k) > \lambda_P$.*

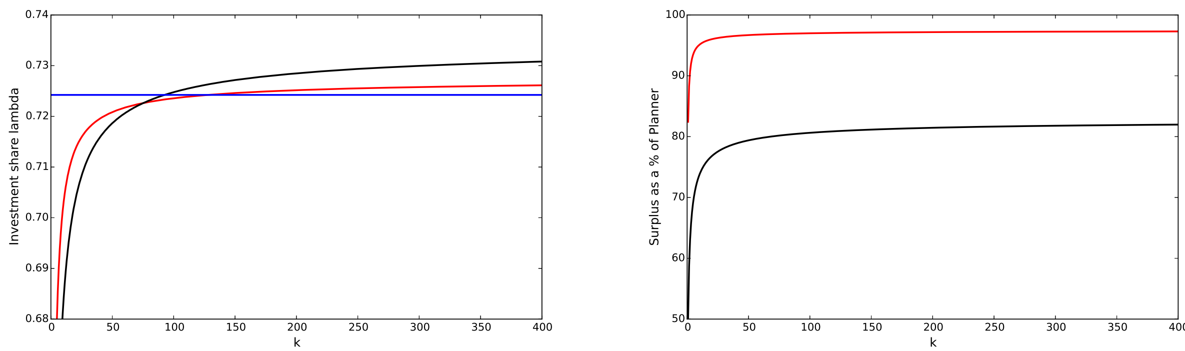

Figure 5: Investments and Welfare for $\mu = -1$ (red), $\mu = -0.5$ (black), and Social Optimal investment (blue), and investment for $\sigma = 0.5$, $\phi = 1.4$, $\beta = 0.9$, $\rho = 0.9$, and $s = 1.5$

2. *If party L's valence decreases (i.e., $G(0)$ decreases), then $\lambda(k)$ increases.*

As k grows very large, the probability that party L loses goes to $G(0)$, becoming insensitive to party L's investment choice. This leads party L to 'over-invest' to account for the fact that the investment is only made when it wins, i.e., to insure against capital being depreciated too much due to a string of wins by party R. In some sense, relative to a scenario with no political competition, candidate L pursues too much of a austerity policy in good times.

The left panel in figure 5 compares the social optimal investment to the investment chosen by candidate L. As proposition 7 indicates, candidate L's investment exceeds that of the planner when k is sufficiently large. The gap between these investment levels depends on $G(0)$.

7 Additional Notes:

Absent additional structure on G, there is no reason to expect that λ is monotone in k. In particular, k goes up, local incentives to save (marginal increased win probability) can do anything. This works for a distribution for which $g(D)/G(D)$ increases sufficiently for some D. This happens, for example, if g goes to zero on the interior of the distribution before increasing steeply. For similar reasons, we cannot expect concavity of λ in k. Moreover, while $V_S(k) > V(k)$, this does not imply that $\lambda_S(k) > \lambda(k)$ or that $V_S(k)$ varies more with k than $V(k)$. In fact, we will establish that the opposite holds whenever the level of capital is sufficiently high.

Of course, the probability increases if the capital stock is lower***careful... only establishes for the upper bound on the probability of a death spiral, and we can have $\lambda'(k) > 1$... (implying that the equilibrium probability of losing can increase in k).

8 Appendix

Proof of Proposition 1. We showed in the text that the value function is differentiable in k. We proceed by using Euler equations.

$$\begin{aligned} V_P(k_t) &= \max_{i_t \in [0, \phi k_{t-1}]} u(\phi k_t - i_{t+1}) + \beta V_P(k_{t+1}) \\ & \quad s.t.\ k_{t+1} = \rho k_t + i_{t+1}. \end{aligned}$$

Use the change of variable from i to $c_t = \phi k_t - i_{t+1}$ to rewrite the problem as

$$\begin{aligned} V_P(k_t) &= \max_{c_t \in [0, \phi k_t]} u(c_t) + \beta V_P(k_{t+1}) \\ & \quad s.t.\ k_{t+1} = (\rho + \phi) k_t - c_t. \end{aligned}$$

The associated first-order condition yields

$$u'(c_t) = \beta V_P'(k_{t+1}). \tag{35}$$

Similarly, for the next period, $k_{t+2} = (\rho + \phi) k_{t+1} - c_{t+1}$, we have

$$u'(c_{t+1}) = \beta V_P'(k_{t+2}). \tag{36}$$

By the Envelope Theorem,

$$V_P'(k_t) = \beta(\rho + \phi) V_P'(k_{t+1}). \tag{37}$$

Thus, for the next period,

$$V_P'(k_{t+1}) = \beta(\rho + \phi) V_P'(k_{t+2}). \tag{38}$$

Thus, from equations (35)–(38),

$$\frac{u'(c_{t+1})}{u'(c_t)} = \frac{V_P'(k_{t+2})}{V_P'(k_{t+1})} = \frac{1}{\beta(\rho + \phi)}. \tag{39}$$

The budget constraint for the next period is

$$k_{t+2} = (\rho + \phi) k_{t+1} - c_{t+1} \quad \Leftrightarrow \quad k_{t+1} = \frac{c_{t+1}}{\rho + \phi} + \frac{k_{t+2}}{\rho + \phi}.$$

Next, recursively apply the budget constraint to get

$$\begin{aligned} c_t &= (\rho + \phi) k_t - k_{t+1} \\ &= (\rho + \phi) k_t - \frac{c_{t+1}}{\rho + \phi} - \frac{k_{t+2}}{\rho + \phi} \\ &= (\rho + \phi) k_t - \frac{c_{t+1}}{\rho + \phi} - \frac{c_{t+2}}{(\rho + \phi)^2} - \frac{k_{t+3}}{(\rho + \phi)^2} \\ &\vdots \\ &= (\rho + \phi) k_t - \frac{c_{t+1}}{\rho + \phi} - \frac{c_{t+2}}{(\rho + \phi)^2} - \frac{c_{t+3}}{(\rho + \phi)^3} - \cdots . \end{aligned} \tag{40}$$

Thus, assuming existence, if (39) yields $c_{t+1} \propto c_t$, then it follows that $c_t \propto k_t$.

When $u(c) = \frac{c^{1-s}}{1-s}$, for $s \neq 1$, and $u(c) = log(c)$, for $s = 1$, then (39) yields

$$\left(\frac{c_t}{c_{t+1}}\right)^s = \frac{1}{\beta(\rho + \phi)} \Leftrightarrow \frac{c_{t+1}}{c_t} = [\beta(\rho + \phi)]^{1/s}. \tag{41}$$

Now, substituting from (41) into (40),

$$c_t = (\rho + \phi)k_t - \frac{[\beta(\rho + \phi)]^{1/s}}{\rho + \phi} c_t - \frac{([\beta(\rho + \phi)]^{1/s})^2}{(\rho + \phi)^2} c_t - \frac{([\beta(\rho + \phi)]^{1/s})^3}{(\rho + \phi)^3} c_t - \cdots .$$

Thus,

$$\begin{aligned}(\rho + \phi)k_t &= c_t + \left(\frac{[\beta(\rho + \phi)]^{1/s}}{\rho + \phi}\right) c_t + \left(\frac{[\beta(\rho + \phi)]^{1/s}}{\rho + \phi}\right)^2 c_t + \left(\frac{[\beta(\rho + \phi)]^{1/s}}{\rho + \phi}\right)^3 c_t + \cdots \\ &= \frac{c_t}{1 - \frac{[\beta(\rho+\phi)]^{1/s}}{\rho+\phi}}, \ \text{assuming } \beta(\rho + \phi)^{1-s} < 1, \text{ which is true by Assumption 1.}\end{aligned}$$

Cross-multiplying, yields that consumption is a linear function of k_t.

$$c_t = \left(\rho + \phi - [\beta(\rho + \phi)]^{1/s}\right) k_t.$$

Thus, investment is given by

$$i_{t+1} = \phi k_t - c_t = \left(\phi - \rho - \phi + [\beta(\rho + \phi)]^{1/s}\right) k_t = \left([\beta(\rho + \phi)]^{1/s} - \rho\right) k_t.$$

For the FOC to hold, i.e., for investment to be positive, we need $(\beta(\rho + \phi))^{1/s} > (\beta\phi)^{1/s} > 1 > \rho$, which holds by Assumption 1. ■

Proof of Lemma 1. Party L's payoff has an upper bound corresponding to the payoff to the social planner.

In state ω, let $i_t(\omega)$ be the optimal investment given an initial capital stock$\bar{k}$, and let $\tilde{i}_t(\omega)$ be the optimal investment given $\alpha\bar{k}$, where $\alpha > 0$. Let $k(\omega)$ and $\tilde{k}(\omega)$ be the associated capital stocks, and $D_t(\omega)$ and $\tilde{D}_t(\omega)$ be the associated valence cutoffs. Finally, let $\hat{i}_t(\omega) = \alpha i_t(\omega)$.

Note that the constraints of Problem (8) are satisfied by $\hat{i}_t(\omega)$ when the initial capital stock is $\alpha\bar{k}$. Moreover, the capital stock is given by $\hat{k}_t(\omega) = \alpha k_t(\omega)$. Let $\hat{D}_t(\omega)$ be the valence cutoff then $\hat{D}_t(\omega) = D_t(\omega)$. As a consequence, the expectation $E[\cdot]$ is the same under k_t, i_t starting at $\bar{k}$ and $\hat{k}_t, \hat{i}_t$, starting at $\alpha\bar{k}$.

Next, we can factor α out of the objective:

$$
\begin{aligned}
&E\left[\sum_{t=0}^{\infty}\beta^t\Big(W_t(\omega)\log(\phi\tilde{k}_{t-1}(\omega)-\tilde{i}_t(\omega))+(1-W_t(\omega))\log(\phi\tilde{k}_{t-1}(\omega))\Big);\{\tilde{i}_t(\omega),\tilde{k}_t(\omega)\}_{t\in\mathbb{N}}\right]\\
&\geq E\left[\sum_{t=0}^{\infty}\beta^t\Big(W_t(\omega)\log(\phi\alpha k_{t-1}(\omega)-\alpha i_t(\omega))+(1-W_t(\omega))\log(\phi\alpha k_{t-1}(\omega))\Big);\{\alpha i_t(\omega),\alpha k_t(\omega)\}_{t\in\mathbb{N}}\right]\\
&=\frac{\log(\alpha)}{1-\beta}+E\left[\sum_{t=0}^{\infty}\beta^t\Big(W_t(\omega)\log(\phi k_{t-1}(\omega)-i_t(\omega))+(1-W_t(\omega))\log(\phi k_{t-1}(\omega))\Big);\{i_t(\omega),k_t(\omega)\}_{t\in\mathbb{N}}\right].
\end{aligned}
\tag{42}
$$

Let $V(k)$ denote the expected discounted payoff to party L. Then this argument shows that

$$V(\alpha k)\geq\frac{\log(\alpha)}{1-\beta}+V(k). \tag{43}$$

Because k and α are arbitrary it follows that

$$V(k)=V\left(\frac{1}{\alpha}(\alpha k)\right)\geq\frac{\log\left(\frac{1}{\alpha}\right)}{1-\beta}+V(\alpha k). \tag{44}$$

Suppose that the inequality in (43) is strict. Then

$$V(k)\geq\frac{\log\left(\frac{1}{\alpha}\right)}{1-\beta}+V(\alpha k)>\frac{\log\left(\frac{1}{\alpha}\right)}{1-\beta}+\frac{\log(\alpha)}{1-\beta}+V(k)=V(k),$$

a contradiction. Similarly, we get a contradiction if (44) is strict. Thus, $V(k)=\log(k)/(1-\beta)+V(1)$, which concludes the proof. ■

Proof of Proposition 2. We first prove that $\lambda(\rho)$ is always strictly less than the socially optimal investment. Equation 6 of Proposition 1 show that the social optimum is $\lambda=\beta\phi-(1-\beta)\rho$. Inserting this into (11) implies that the left-hand side is zero. We next show that the right-hand side of (11) which represents the marginal cost of increasing λ, is strictly positive.

(11) implies that $\lambda>0$ in any social optimum. Thus, the dynamic payoff from investing $\lambda>0$ exceeds the payoff from investing nothing. Thus the cost to party L of losing the election is given by $\log((\phi-\lambda)/\phi)+\beta\log((\rho+\lambda)/\rho)>0$. Thus, marginal cost is strictly positive. As a consequence it is optimal to lower λ from the social optimum, i.e., to set $\lambda(\rho)<\beta\phi-(1-\beta)\rho$ for all $0<\rho\leq 1$.

We next show that $\lim_{\rho\downarrow 0}\lambda(\rho)=0$.

Let $\rho\downarrow 0$. Suppose by way of contradiction that λ remains bounded away from zero. Then there exists a sequence ρ_n, $n\in\mathbb{R}$ such that associated levels, λ_n converge to $\bar{\lambda}>0$. If $n\to\infty$

then the left-hand side of (11) converges to $G((\phi-\bar{\lambda})/\phi)\left(\beta\phi-\bar{\lambda}\right)$. In contrast, the right-hand side of (11) goes to ∞ because $\log((\rho_n+\lambda_n)/\rho_n)\to\infty$, while $\lim_{n\to\infty} g(D)(\rho+\lambda)=g((\phi-\bar{\lambda})/\phi)\bar{\lambda}$. Thus, the right-hand side of (11) exceeds the left-hand side for all sufficiently large n, a contradiction to the assumption that λ_n satisfies the first order condition for ρ_n. Thus, $\lambda(\rho)$ must converge to zero as $\rho\downarrow 0$.

Next, note that $\lim_{\rho\downarrow 0}\lambda(\rho)/\rho=\infty$.

The left-hand side of (11), i.e., the marginal benefit of saving converges to $G(0)\beta\phi$ as $\rho\downarrow 0$, which is strictly positive. Thus, the right-hand side of (11), i.e., the marginal cost, must also be non-zero in the limit. Again, note that D converges to 0 as $\rho\downarrow 0$. Next, because $\lim_{\rho\downarrow 0}\lambda(\rho)=0$ it follows that $\log((\phi-\lambda(\rho))/\phi)$ goes to zero. Thus, $(\rho+\lambda(\rho))\log((\rho+\lambda(\rho))/\rho)$ must be non-zero in the limit. Given that both ρ and $\lambda(\rho)$ go to zero, it follows that $\log((\rho+\lambda(\rho))/\rho)=\log(1+(\lambda(\rho)/\rho)$ goes to infinity. Thus, $\lambda(\rho)/\rho$ goes to infinity.

We can now conclude that $\lim_{\rho\downarrow 0}\rho\log((\rho+\lambda(\rho))/\rho)=0$.

Suppose by way of contradiction that there exists a sequence ρ_n such that $\lim_{\rho\downarrow 0}\rho\log((\rho_n+\lambda(\rho_n))/\rho_n)=a\neq 0$. Then this and the fact that $\lambda(\rho)/\rho$ becomes unbounded implies that $\lambda(\rho_n)\log((\rho_n+\lambda(\rho_n))/\rho_n)$ goes to infinity, a contradiction because the left-hand side of (11) is bounded, as shown above.

This can only be the case if $\lambda(\rho)/\rho$ goes to infinity if $\rho\downarrow 0$.

Taking the limit on both sides of (11) as $\rho\downarrow 0$ therefore yields

$$G(0)\beta\phi=g(0)\beta\lim_{\rho\downarrow 0}\lambda(\rho)\log\left(\frac{\rho+\lambda(\rho)}{\rho}\right). \tag{45}$$

Substituting $\lambda(\rho)=-G(0)\phi/(g(0)\log(\rho))$ into (45) yields

$$\begin{aligned} G(0)\phi/g(0) &= \lim_{\rho\to 0}\lambda(\rho)\log\left(\frac{\rho+\lambda(\rho)}{\rho}\right)=\lim_{\rho\to 0}\lambda(\rho)\log\left(\frac{\lambda(\rho)}{\rho}\right)\\ &= \lim_{\rho\to 0}\lambda(\rho)\log(\lambda(\rho))-\lim_{\rho\to 0}\lambda(\rho)\log(\rho)=-\lim_{\rho\to 0}\lambda(\rho)\log(\rho), \end{aligned} \tag{46}$$

where we have used the fact that $\lim_{x\to 0}x\log x=0$. Thus,

$$\lim_{\rho\to 0}\lambda(\rho)\left(-\frac{g(0)}{G(0)}\frac{\log(\rho)}{\phi}\right)=1. \tag{47}$$

This verifies that $\lambda(\rho)$ goes to zero at the indicated rate.

Because $-G(0)\phi/(g(0)\log(\rho))$ approximates $\lambda(\rho)$ for small ρ, it follows that $\lambda(\rho)$ is increasing and convex in ρ (for ρ small). ■

Proof of Lemma 2. We can assume that $\rho+\lambda>1$, else the result is trivial.

Let X denote the random variable that assumes the value 1 if L wins, and 0 if R wins. Party L's investment as a share of capital, λ, is given by (11), independent of k. Thus, if L wins in a period n then $k_n = (\rho + \lambda)k_{n-1}$. In contrast, if L loses, then the capital stock becomes $k_n = \rho k_{i-n}$.

Let u be the number of times that L wins and $t - u$ the number of times that R wins in t elections. Then the capital stock at the beginning of period t,

$$k_{t-1} = \bar{k}(\rho + \lambda)^u \rho^{t-u} = \bar{k}\left((\rho + \lambda)^{\frac{u}{t}} \rho^{1-\frac{u}{t}}\right)^t, \tag{48}$$

Let $p = G(D)$ be the probability that party L wins. Note that D is independent of k. Let X_t, $t \in \mathbb{N}$ be the stochastic process that assumes the value 1 if L wins, and 0, otherwise. Let $a > 0$. Then the weak law of large numbers implies that $\lim_{t\to\infty} \text{Prob}\left(\left\{(1/t)\sum_{n=0}^{t-1}\left|X_n - p\right| > a\right\}\right) = 0$. Hence for every $\delta > 0$, there exists $\bar{t}$ such $u/t < p + a$ for all $t \geq \bar{t}$ and $1 - u/t > 1 - p - a$. Thus, (48) implies

$$k_{t-1} \leq \bar{k}\left((\rho + \lambda)^{p+a} \rho^{1-p-a}\right)^t. \tag{49}$$

Thus, if $(\rho + \lambda)^{p+a}\rho^{1-p-a} < 1$ then (49) implies that k_{t-1} converges to zero. The result follows because a was arbitrary. ■

Proof of Proposition 3. Divide equation (18) by $\rho + \lambda$ to obtain:

$$\beta(G(D)\frac{\phi - \lambda}{\rho + \lambda} - g(D)\log(\frac{\rho + \lambda}{\rho})) + (1 - \beta)\Big(G(D) - g(D)\log(\frac{\phi - \lambda}{\phi})\Big) = 0.$$

We now show that the term with weight $1 - \beta$ is negative at the value where survival is maximized, i.e., where the term with weight β is zero. That is, at $G(D) = g(D)\log(\frac{\rho+\lambda}{\rho})\frac{\rho+\lambda}{\phi-\lambda}$, we show that

$$\log\left(\frac{\rho + \lambda}{\rho}\right)\frac{\rho + \lambda}{\phi - \lambda} > \log\left(\frac{\phi - \lambda}{\phi}\right).$$

Multiplying both sides by $\phi - \lambda$ yields

$$\log\left(\left(\frac{\rho + \lambda}{\rho}\right)^{\rho+\lambda}\right) > 0 > \log\left(\left(\frac{\phi - \lambda}{\phi}\right)^{\phi-\lambda}\right)$$

It follows that at the value of λ that maximizes survival, the marginal benefit of increasing λ is less than the marginal cost, i.e., party L will choose a smaller λ. Letting $x(\lambda)$ be the term multiplying β and $y(\lambda) < x(\lambda)$ be the term multiplying $1 - \beta$, we have $\frac{d\lambda}{d\beta} = -\frac{(x-y)}{\beta x' + (1-\beta)y'} > 0$ ■

Proof of Lemma 3. Without loss of generality, suppose that $k' < k$. Let $\alpha = k'/k$. Suppose that at capital level k the same proportional investment choices are made as at capital level k'. If the winning probabilities are unchanged than this would imply that $\alpha^{1-s}V(k) = V(\alpha k)$. However, increasing k increases D_k and thus, the winning probability. As a consequence, $\alpha^{1-s}V(k) \geq V(\alpha k)$. Thus,

$$\begin{aligned}\liminf_{k'\to k} \frac{V(k) - V(k')}{k - k'} &= \liminf_{\alpha\to 1} \frac{V(k) - V(\alpha k)}{(1-\alpha)k} \geq \liminf_{\alpha\to 1} \frac{V(k) - \alpha^{1-s}V(k)}{(1-\alpha)k} \\ &= \lim_{\alpha\to 1} \frac{V(k)}{k}\frac{1-\alpha^{1-s}}{1-\alpha} = \frac{(1-s)V(k)}{k}.\end{aligned} \tag{50}$$

Because $V(k) \geq V_N(k)$ equation (50) implies

$$\liminf_{k'\to k} \frac{V(k) - V(k')}{k - k'} \geq \frac{(1-s)V_N(k)}{k}. \tag{51}$$

Next, note that

$$k^{s-1}V_N(k) = \sum_{t=0}^{\infty} \beta^t \phi^{1-s} (\rho^{1-s})^t \tag{52}$$

converges uniformly if and only if $\beta\rho^{1-s} < 1$. As a consequence, $k^{s-1}V_N(k)$ is differentiable for $k > 0$. Thus, $V_N(k)$ is differentiable for $k > 0$. The derivative is given by

$$V_N'(k) = \sum_{t=0}^{\infty} \beta^t (\rho^t)^{1-s} \phi^{1-s} k^{-s} = \frac{(1-s)V_N(k)}{k}. \tag{53}$$

This, (50) and (53) imply the result. ■

Proof of Lemma 4. Suppose by way of contradiction that there exists a sequence $k_n \to \infty$ such that $\lim_{n\to\infty} \lambda(k_n) = \phi$. The net benefit of investment contingent on winning the election is

$$u((\phi-\lambda)k_n) - u(\phi k_n) + \beta\left(V((\rho+\lambda)k_n) - V(\rho k_n)\right) < u((\phi-\lambda)k_n) - u(\phi k_n) - \beta V_N(\rho k_n), \tag{54}$$

where the inequality follows because utility is negative for $s > 1$ and V_N is a lower bound on the continuation utility. Note that

$$\begin{aligned}&\lim_{n\to\infty} u((\phi-\lambda)k_n) - u(\phi k_n) - \beta V_N(\rho k_n) \\ &\quad = \lim_{n\to\infty} \frac{1}{(1-s)k_n^{s-1}} \left((\phi - \lambda(k_n))^{1-s} - \phi^{1-s} - \frac{\beta}{1-\beta\rho^{1-s}} \right) < 0,\end{aligned} \tag{55}$$

because $\lim_{n\to\infty} \lambda(k_n) = \phi$ implies $\lim_{n\to\infty}(\phi - \lambda(k_n))^{1-s} = -\infty$, and the other terms are bounded. However, this implies that the net benefit from investment given by (54) is strictly negative. By choosing $\lambda = 0$ candidate L can guarantee a higher payoff, contradicting the optimality of $\lambda(k_n)$.

■

Proof of Lemma 5. Let $W_L(k) = k^{s-1}V_L(k)$. Then $W(k)$ solves

$$\begin{aligned} W_L(k) &= \max_{\lambda\in[0,\phi]} G(0)\Big(k^{s-1}u((\phi-\lambda)k)+\beta W_L((\rho+\lambda)k)\Big)+(1-G(0))\Big(k^{s-1}u(\phi k)+\beta W_L(\rho k)\Big) \\ &= \max_{\lambda\in[0,\phi]} G(0)\left((\phi-\lambda)^{1-s}+\beta W_L((\rho+\lambda)k)\right)+(1-G(0))\left(\phi^{1-s}+\beta W_L(\rho k)\right). \end{aligned} \tag{56}$$

It immediately follows that $W_L(k) = k^{s-1}V_L(k) = v_L/(1-s)$, is independent of k. Further, the objective of (56) is strictly concave in λ.

Similarly, we define $W(k) = k^{s-1}V(k)$, where $V(k)$ is the value function for problem (27). Let $\lambda(k)$ be be the associated optimal investment share. Then $W(k)$ and $\lambda(k)$ solve

$$W(k) = \max_{\lambda\in[0,\phi]} G(D)\Big(k^{s-1}u((\phi-\lambda)k)+\beta W((\rho+\lambda)k)\Big)+(1-G(D))\Big(k^{s-1}u(\phi k)+\beta W(\rho k)\Big). \tag{57}$$

Next, assume by way of contradiction that $\lambda(k)$ does not converge to λ^* as $k\to\infty$. Then there exists a sequence $k_n\to\infty$ with $\lim_{n\to\infty}\lambda(k_n)\neq\lambda^*$. Without loss of generality, suppose that $\lambda(k_n)<\bar\lambda<\lambda^*$, for some $\bar\lambda$ (the analysis when $\bar\lambda>\lambda^*$ is analogous). Recall that λ^* solves problem (56) and that the objective is strictly concave in λ, and independent of k. From the strict concavity, choosing $\lambda(k_n)<\bar\lambda<\lambda^*$, reduces payoffs by an amount that is bounded away from zero: there exists a $\delta>0$ such that

$$\begin{aligned} &G(0)\left((\phi-\lambda(k_n))^{1-s}+\beta W_L((\rho+\lambda(k_n))k_n)\right)+(1-G(0))\left(\phi^{1-s}+\beta W_L(\rho k_n)\right)+\delta \\ &\quad< G(0)\left((\phi-\lambda^*)^{1-s}+\beta W_L((\rho+\lambda^*)k_n)\right)+(1-G(0))\left(\phi^{1-s}+\beta W_L(\rho k_n)\right). \end{aligned} \tag{58}$$

Recall that the social planner's value function takes the form $V_S(k) = v_s/(1-s)k^{1-s}$. Similarly, the value function if no investment takes place takes the form $V_N(k) = v_n/(1-s)k^{1-s}$. The value function of (27) is bounded from above and below by these two value functions. Therefore,

$$\frac{v_N}{1-s} = k^{s-1}V_N(k) \le W(k) \le k^{s-1}V_S(k) = \frac{v_S}{1-s}. \tag{59}$$

Thus, $W(k)$ is bounded.

Let $\varepsilon = (1+/(1-\beta)|v_s/(1-s)|) < \delta/2$. Then Lemma 4 and the argument in the text imply that there exists a $k_1>0$ such that $|G(D_k)-G(0)|<\varepsilon$ for all $k\ge k_1$.

Further, because $\beta<1$, there exists a $T\in\mathbb{N}$ such that $\left|\beta^T V_N(k_1)\right|<\varepsilon$. Let $k_2 = k_1/\rho^T$. This, and (59) imply that if we start in period $t=0$ with initial capital stock $k\ge k_2$ then the utility impact of actions in periods $\tau>T$ is less than ε. In addition, in periods $\tau\le T$ capital $k\ge k_1$. Thus, the fact that $|G(D_k)-G(0)|<\varepsilon$ and the boundedness of W established in (59) imply that changing the probabilities from $G(0)$ to $G(D_k)$ impacts utility by less than $\varepsilon|v_s/(1-s)|$ in each period, and hence by less than $\varepsilon\beta^T/(1-\beta)|v_s/(1-s)| < \varepsilon/(1-\beta)|v_s/(1-s)|$ in the first $T+1$ periods. Given the choice of ε it follows that $|W(k)-W_L(k)|<\delta$ for all $k\ge k_2$.

Finally, equation (58) implies that if we replace λ^* by $\lambda(k_n)$ then utility decreases by more than δ. Further, we have shown that replacing W_L by W changes payoffs by less than $\delta/2$. Thus,

$$\begin{aligned} &G(0)\left((\phi-\lambda(k_n))^{1-s}+\beta W((\rho+\lambda(k_n))k_n)\right)+(1-G(0))\left(\phi^{1-s}+\beta W(\rho k_n)\right) \\ &\quad< G(0)\left((\phi-\lambda^*)^{1-s}+\beta W((\rho+\lambda^*)k_n)\right)+(1-G(0))\left(\phi^{1-s}+\beta W(\rho k_n)\right), \end{aligned} \tag{60}$$

which contradicts that $\lambda(k_n)$ is the optimal investment share given k_n. This contradiction proves that $\lim_{k\to\infty}\lambda(k_n)=\lambda^*$. ■