

Policy Persistence and Drift in Organizations*

Germán Gieczewski[†]

May 2020

Abstract

This paper models the evolution of organizations that allow free entry and exit of members, such as cities and trade unions. In each period current members choose a policy for the organization. Policy changes attract newcomers and drive away dissatisfied members, altering the set of future policymakers. The resulting feedback effects take the organization down a “slippery slope” that converges to a myopically stable policy, even if the agents are forward-looking, but convergence becomes slower the more patient they are. The model yields a tractable characterization of the steady state and the transition dynamics. The analysis is also extended to situations in which the organization can exclude members, such as enfranchisement and immigration.

Keywords: dynamic policy choice, median voter, slippery slope, endogenous population, transition dynamics

1 Introduction

This paper studies the dynamic behavior of organizations that are member-owned—that is, whose members choose policies through a collective decision-making process—

*I am deeply grateful to Daron Acemoglu and Juuso Toikka for their continuous support and numerous suggestions that substantially improved the paper. I would also like to thank Avidit Acharya, Alessandro Bonatti, Glenn Ellison, Robert Gibbons, Matías Iaryczower, Mihai Manea, Robert Townsend, Alex Wolitzky, Muhamet Yildiz, three anonymous referees, and seminar participants at MIT, Columbia University, the University of Chicago, Stanford University, Yale University and the Warwick-Princeton-Utah Political Economy Conference for helpful comments.

[†]Department of Politics, Princeton University.

and allow for the free entry and exit of members. In this context, policy and membership decisions affect one another: different policies appeal to or drive away different prospective members, and different groups make different choices when in charge of the organization. As a result, the policy path may drift over time: an initial policy may attract a set of members wanting a different policy, which in turn attracts other agents, and so on. It may also exhibit path-dependence: two organizations with identical fundamentals but different initial policies may exhibit divergent behavior in the long run.

A prominent example where these issues arise is that of cities or localities. Conceptualize each city as an organization and its inhabitants as members. Cities allow people to move in and out freely, and their inhabitants vote for local authorities who implement policies, such as the level of property taxes, the funding of public schools and housing regulations. The interplay between policy changes and migration can lead to demographic and socioeconomic shifts interpreted as urban decay, revitalization, and gentrification (Marcuse 1986; Vigdor 2010).

The relationship between local taxes and migration has been studied since Tiebout (1956), which considers a population able to move across a collection of communities with fixed policies. Epple and Romer (1991) allow for redistributive policies and location decisions to respond to one another, but they study the problem in static equilibrium, i.e., under the assumption that any temporary imbalance between the people living in a community and the policies they want has already resolved itself.

A natural follow-up question to this literature is whether, in a dynamic setting, communities will converge to a static equilibrium quickly, slowly or not at all. The main reason convergence may fail to obtain is a fear of *slippery slopes*. Here is a simple example: suppose a community with a local tax rate $x_0 = 0.2$ attracts a population whose median voter, m_0 , prefers a tax rate $x_1 = 0.18$. Lowering the tax rate to x_1 would attract a different population whose median voter, m_1 , has bliss point $x_2 = 0.16$. In turn, the tax rate x_2 would beget a median voter wanting a tax rate $x_3 = 0.15$, and so on. If agents vote myopically, the tax rate will quickly move not to x_1 but to a much lower steady state, say $x_\infty = 0.1$. Foreseeing this, m_0 might prefer not to change the tax rate after all.

This paper shows that, in a dynamic model where both policies and membership are determined endogenously, communities *will*, in fact, converge to a steady state, and all steady states are myopically stable independently of the agents' discount

factor. However, dynamic concerns induce agents to make smaller policy changes in each period than their myopic preferences would dictate. In particular, when the median voter's bliss point is closer to the current policy than to the steady state, convergence is slow—that is, as agents become arbitrarily patient, policy changes in each period become arbitrarily small. Thus communities observed in the world at any given time may well fail to be in static equilibrium, and predicting their future behavior requires an understanding of their transition dynamics. The model also yields a tractable characterization of these dynamics, which allows us to describe the equilibrium speed of policy change in terms of the strength of the myopic incentives for change and the degree of expected disagreement with future pivotal agents.

The location of steady states is characterized as a function of the distribution of preferences. In general, policy drift leads organizations towards peaks of the distribution of policy bliss points, which favors centrism if said distribution is unimodal and symmetric. However, a pocket of agents concentrated at an extreme can also support a steady state. When there are multiple steady states, which one the organization converges to depends on its initial policy (i.e., there is path-dependence). Extreme steady states are more likely if agents' willingness to join is asymmetric (that is, extremists are more willing to join a moderate organization than vice versa). Relative to a setting with a fixed population, the location of steady states is more sensitive to the distribution of preferences: small changes in the distribution can result in arbitrarily large changes to the long-run policy.

This paper is connected to several strands of literature. First, as noted previously, it can be seen as a study of dynamic Tiebout competition. There is a large literature on the Tiebout hypothesis (see Cremer and Pestieau (2004) for a review), but most papers in it assume that policies and location decisions must be in static equilibrium, and hence are silent on the transition dynamics that this paper focuses on.

Second, the model can be applied to other organizations with open membership, such as trade unions, nonprofits, sports clubs and religious communities, and it is therefore relevant to existing work about such organizations. For instance, Grossman (1984) explains why increased international competition may not decrease wages in a unionized sector: layoffs selectively affect less senior workers, so the median voter within the union becomes more senior—hence more securely employed, and prone to making more aggressive wage demands. As in the Tiebout literature, Grossman (1984) assumes that policy and membership are always in static equilibrium, i.e., that

they adjust immediately after an external shock; this paper can be seen as providing a model of the transition dynamics.

Finally, the paper makes several contributions to a growing literature on dynamic political decision-making (Roberts 2015; Acemoglu, Egorov and Sonin 2015; Bai and Lagunoff 2011). Most papers in this literature study organizations which can strategically restrict the entry of newcomers, remove existing members, or deny them political power (relevant applications include enfranchisement and immigration). Despite the apparent substantive differences between this setting and mine, my results readily extend to this context. The reason is that both types of models are driven by the same tension, namely, that policies and decision-making power are coupled in a rigid manner, so agents cannot choose their ideal policy without relinquishing control over future decisions.

There are two main branches in this literature. The first one (Roberts 2015; Acemoglu, Egorov and Sonin 2008, 2012, 2015) assumes a fixed, finite policy space and obtains the result that, when agents are patient enough, convergence to a steady state is “fast”, and steady states may not be myopically stable. The set of steady states can be found by means of a recursive algorithm, but not described explicitly, and it is sensitive to the set of feasible policies. What I show is that, if a continuous policy space is assumed, these results are overturned: all steady states are myopically stable, and when agents are patient, there is slow convergence which can be characterized explicitly (in some cases, in closed form).

The second branch (Jack and Lagunoff, 2006; Bai and Lagunoff, 2011) considers continuous policy spaces and obtains some important results related to the ones in this paper; in particular, Bai and Lagunoff (2011) show that, in their model, the steady states of “smooth” equilibria are stable under the assumption of a fixed decision-maker (in our setting, this is equivalent to myopic stability). However, they do not provide a general characterization of which steady state the model will converge to, nor of the transition dynamics. Moreover, their analysis applies only to smooth equilibria, which do not exist generically. In contrast, I derive results that apply either to all equilibria or to classes of equilibria for which I can provide existence conditions.

On a technical note, the present paper is also the first in this literature to tractably analyze a setting that violates the single-crossing assumption on preferences—a necessary complication in a context with free entry and exit, stemming from the fact that agents unhappy with the chosen policy can cut their losses by leaving the organization.

The paper is structured as follows. Section 2 presents the model. Section 3 proves some fundamental properties of all equilibria and characterizes the organization's policy in the long run. Section 4 characterizes the transition dynamics. Section 5 adapts the results to a setting without free entry and exit. Section 6 discusses some implications of the results and revisits their relationship with the existing literature. Section 7 is a conclusion. All the proofs can be found in the Appendices.

2 The Model

There is an organization (henceforth, a club) existing in discrete time $t = 0, 1, \dots$ and a unit mass of agents distributed according to a continuous density f with support $[-1, 1]$. We refer to an agent's position α in the interval $[-1, 1]$ as her type. All agents are potential members of the club.

At each integer time $t \geq 1$, two events take place. First there is a *voting stage*, in which a set of *incumbent members* $I_{t-1} \subseteq [-1, 1]$ vote on a *policy* $x_t \in [-1, 1]$ to be implemented during the period $[t, t + 1)$. Immediately after, in the *membership stage*, all agents observe x_t and decide whether to be members during the upcoming period $[t, t + 1)$. Agents can freely enter and leave the club as many times as desired at no cost. The set of agents who choose to be members at time t constitutes I_t , the set of incumbent members at the $t + 1$ voting stage.¹ At $t = 0$ the game starts with a membership stage; the club's initial policy x_0 is exogenously given.

The essential feature of this setup is that membership affects both an agent's utility and her right to vote. Agents will decide whether to be in the club based on their private payoffs, since the impact of any individual agent's vote on future policies is nil, but aggregate membership decisions will influence future policies.

Preferences

An agent α has utility

$$U_\alpha((x_t)_t, I_\alpha) = \sum_{t=0}^{\infty} \delta^t I_{\alpha t} u_\alpha(x_t),$$

¹The assumption that agents vote the period *after* joining rules out equilibria in which agents who dislike the current policy might join because they expect the policy to immediately change to their liking. Equivalent results would be obtained by assuming that agents can enter or leave at any time $t \in \mathbb{R}_{\geq 0}$ but only gain voting rights after being members for a short time $\varepsilon \in (0, 1)$.

where $I_{\alpha t} = \mathbb{1}_{\{\alpha \in I_t\}}$ denotes whether α is a member at time t . In other words, the agent can obtain a payoff $u_\alpha(x_t)$ from being a member of the club, or a payoff of zero from remaining an outsider. $\delta \in (0, 1)$ is a common discount factor. We make the following assumptions on u .

A1 $u_\alpha(x) : [-1, 1]^2 \rightarrow \mathbb{R}$ is C^2 .

A2 There are $0 < M' < M$ such that $M' \leq \frac{\partial^2}{\partial \alpha \partial x} u_\alpha(x) \leq M$ for all α, x .

A3 $u_\alpha(\alpha) > 0$ for all $\alpha \in [-1, 1]$.

A4 For a fixed α_0 , $u_{\alpha_0}(x)$ is strictly concave in x with peak $x = \alpha_0$.

A5 For a fixed x_0 , $\frac{\partial u_\alpha(x_0)}{\partial \alpha} > 0$ if $\alpha < x_0$ and $\frac{\partial u_\alpha(x_0)}{\partial \alpha} < 0$ if $\alpha > x_0$.

The essence of assumptions A2-A5 is that agent α has bliss point α and wants to be in the club if the policy x_t is close enough to α ; higher agents prefer higher policies; and the set of agents desiring membership is always an interval. A useful example for building intuition is the quadratic case: $u_\alpha(x) = C - (\alpha - x)^2$, where $C > 0$. Finally, we impose the following tie-breaking rule.

A6 An agent α 's preferences at time t_0 are as defined by U_α when comparing any two paths $((x_t)_t, I_\alpha), ((\tilde{x}_t)_t, \tilde{I}_\alpha)$ with membership rules $I_\alpha, \tilde{I}_\alpha$ that are not both zero for all $t \geq t_0$. However, if $I_{\alpha t} = \tilde{I}_{\alpha t} = 0$ for all $t \geq t_0$, then α prefers $((x_t)_t, I_\alpha)$ to $((\tilde{x}_t)_t, \tilde{I}_\alpha)$ iff $u_\alpha(x_{t_0}) \geq u_\alpha(\tilde{x}_{t_0})$.

In other words, if an agent expects to permanently quit the organization immediately after the current voting stage, she breaks ties in favor of the path with the better current policy. This assumption prevents members who intend to quit from making arbitrary choices out of indifference.²

Solution Concept

We will use Markov Voting Equilibrium (MVE) (Roberts 2015; Acemoglu et al. 2015) as our solution concept. This amounts to imposing two simplifying assumptions on our equilibrium analysis. First, rather than explicitly modeling the voting process,

²This tie-breaking rule would be uniquely selected if the game were modified to add a small time gap between the voting and membership stages, so that outgoing members at time t would receive a residual payoff $\varepsilon u_\alpha(x_{t_0})$ from the policy x_{t_0} chosen right before they leave.

we assume that only Condorcet-winning policies can be chosen on the equilibrium path, as otherwise a majority could deviate to a different policy. Second, we focus on Markov strategies. That is, when votes are cast at time t , voters only condition on the set of incumbent members, I_{t-1} ; when entry and exit decisions are made, the only state variable is the chosen policy, x_t .³

Definition 1. Let $\mathcal{L}([-1, 1])$ be the Lebesgue σ -algebra on $[-1, 1]$. A Markov strategy profile (\tilde{s}, I) is given by a membership function $I : [-1, 1] \rightarrow \mathcal{L}([-1, 1])$, and a policy function $\tilde{s} : \mathcal{L}([-1, 1]) \rightarrow [-1, 1]$ such that $\tilde{s}(I) = \tilde{s}(I')$ whenever I and I' differ by a set of measure zero.

We denote by $s = \tilde{s} \circ I$ the *successor function*. A policy x induces a set of members $I(x)$, who will vote for a policy $\tilde{s}(I(x)) = s(x)$ in the next period. Hence, an initial policy y leads to a policy path $S(y) = (y, s(y), s^2(y), \dots)$.

Definition 2. An MVE is a Markov strategy profile (\tilde{s}, I) such that:

1. Given a policy x , $\alpha \in I(x)$ iff $u_\alpha(x) \geq 0$.
2. Given a set of voters I , the policy path $S(\tilde{s}(I))$ is a Condorcet winner among the available policy paths. That is, for each $y \neq \tilde{s}(I)$, a weak majority of I weakly prefers $S(\tilde{s}(I))$ to $S(y)$.⁴

From here on we describe equilibria in terms of I and s rather than I and \tilde{s} . This is without loss of detail, as the set of voters is always of the form $I(x)$ on the equilibrium path.⁵

We now provide some definitions that will be useful for our analysis. $x \in [-1, 1]$ is a *steady state* of a successor function s if $s(x) = x$. x is *stable* if there is a neighborhood $(a, b) \ni x$ such that, for all $y \in (a, b)$, $s^t(y) \xrightarrow[t \rightarrow \infty]{} x$. We refer to the largest such neighborhood as the *basin of attraction* of x .

We define the *median voter function* m as follows: for each policy x , $m(x)$ is the median member of the induced voter set $I(x)$, i.e., $\int_{-1}^{m(x)} f(y) \mathbb{1}_{\{\alpha \in I(x)\}} dy = \int_{m(x)}^1 f(y) \mathbb{1}_{\{\alpha \in I(x)\}} dy$.⁶ Finally, we will often be interested in whether an equilibrium

³Our solution concept is equivalent to pure-strategy Markov Perfect Equilibrium (MPE), if in each voting stage two office-motivated politicians engage in Downsian competition.

⁴Note that only one-shot deviations are considered: after a deviation to $y \neq s(x)$, it is expected that the MVE will be followed otherwise, i.e., the policy path will be $(y, s(y), \dots)$ instead of $(s(x), s^2(x), \dots)$. This is without loss of generality.

⁵If an agent deviates from her equilibrium membership decision, the resulting set of members will differ from $I(x)$ by a set of measure zero, so tomorrow's policy will be unchanged.

⁶ $m(x)$ is uniquely defined if $I(x)$ is an interval, as will turn out to be the case.

satisfies the Median Voter Theorem (MVT), i.e., whether the Condorcet-winning policy for a voter set $I(x)$ is also $m(x)$'s optimal choice. Formally, given a successor function s and a set $X \subseteq [-1, 1]$, we will say the MVT holds in X if, for each $x \in X$ and all $y \in [-1, 1]$, $m(x)$ weakly prefers $S(s(x))$ to $S(y)$.

Examples

As an illustration, we map the model to two concrete examples. The first one is Tiebout-style policy competition between cities. Assume that there is a universe of “normal” cities $c \in [-1, 1]$, and a “special” city c^* . Cities differ in two ways. First, each city has a *policy* $x_t(c) \in [-1, 1]$, denoting a certain level of taxation and public goods in city c at time t . For example, a higher x represents higher local taxes which finance better public schools and amenities. Second, c^* has an intrinsic attribute that makes it more desirable than normal cities (good weather, a strong economy, etc.). For simplicity, suppose that each normal city has a positive mass of immobile voters tied to it, and the median immobile voter in city c has bliss point c , so that $x_t(c) = c$ for all t, c . In addition, there is a unit mass of mobile agents in the model, whose bliss points are distributed according to a density f . $x_0(c^*)$ is exogenous.

We are interested in the policy path of c^* and the behavior of mobile agents. At each time t , each mobile agent α chooses a city to live in. Her flow payoff from choosing c is $u_\alpha(x, c) = C\mathbb{1}_{c=c^*} - (x_t(c) - \alpha)^2$, where $C > 0$ is the intrinsic value of c^* . Clearly her decision boils down to a binary choice: she should live either in c^* or in her most-preferred normal city, $c = \alpha$, yielding a flow payoff of zero. Living anywhere but in c^* is equivalent to leaving the club in the general model.

The second example we discuss is that of trade unions. Assume an economy with a unionized firm and a larger competitive (non-union) sector. Firms offer employment contracts (w, l) consisting of a wage w and a family leave policy l . The marginal productivity of all workers is normalized to 1, and a leave policy l signifies that the worker only works a fraction $1 - l$ of the time. In equilibrium, competitive firms are willing to offer any contract of the form $(1 - l, l)$; the competitive sector is assumed to be large enough that all such contracts are available. The union, through collective bargaining, extracts a wage $w_u > 1$ from the unionized firm, so its leadership can bargain for any contract of the form $(w_u - l, l)$, but the same contract must apply to all unionized workers. As in Grossman (1983), assume the union bargains on behalf

of its median voter.

Workers differ in their taste for family leave. A worker of type α has flow payoff $\tilde{u}_\alpha(w, l) = w + \alpha v(l)$, where $v(0) = 0$ and v is smooth, increasing and strictly concave in l . Workers can move freely between firms, including to the unionized firm; upon joining the latter, they automatically become union members.⁷

Of all the competitive firms, a worker α prefers to join one offering $l = l^*(\alpha)$, where $v'(l^*(\alpha)) = \frac{1}{\alpha}$. Let $u_\alpha(l) = \tilde{u}_\alpha(w_u - l, l) - \tilde{u}_\alpha(1 - l^*(\alpha), l^*(\alpha))$ be α 's net utility from joining the union sector when the union has bargained for a contract $(w_u - l, l)$. Up to a relabeling, u satisfies A1-5 and hence the model applies without changes.

3 Equilibrium Analysis

In this Section we prove some fundamental properties of all MVEs, which in particular allow us to characterize the club's policy in the long run. We start by solving for the equilibrium membership strategy, which is simple:

Lemma 1. *In any MVE, $I(y) = [y - d_y^-, y + d_y^+]$ is an interval, and $d_y^-, d_y^+ > 0$ are given by the condition $u_{y-d_y^-}(y) = u_{y+d_y^+}(y) = 0$.*

Since members can enter or leave at any time, it is optimal for α to join whenever the flow payoff of the current policy, $u_\alpha(x)$, is positive, and leave when it is negative; the Lemma then follows from Assumptions A3 and A5. An immediate corollary is that m is strictly increasing and C^1 . Additionally, since $I(x)$ is uniquely determined, we can describe MVEs solely in terms of successor functions.

Before characterizing s in general, it is instructive to consider two simple special cases. First, suppose that $I(y) = I$ is independent of y (for instance, $I(y) \equiv [-1, 1]$, i.e., everyone always prefers to be in the club). In this case, regardless of the current policy y , the Condorcet winner is the bliss point of the median member of I . Second, suppose that $\delta = 0$, i.e., agents are myopic. Given an initial policy y and set of members $I(y)$, the Condorcet winner is the bliss point of $m(y)$, and the policy path will be $(y, m(y), m^2(y), \dots)$, which converges to a myopically stable policy $m^*(y) = \lim_{k \rightarrow \infty} m^k(y)$. In both scenarios, the simplicity of the solution stems from the lack of tension between current payoffs and future control: in the former case there is no link between them, while in the latter case they are linked but voters do not care.

⁷This is a common arrangement is known as a “union shop”.

As a first step towards solving the general case we show that equilibrium paths are always monotonic.

Proposition 1. *In any MVE, for any y , if $s(y) \geq y$ then $s^k(y) \geq s^{k-1}(y)$ for all k , and if $s(y) \leq y$ then $s^k(y) \leq s^{k-1}(y)$ for all k .*

To see why this must be the case, imagine an equilibrium path (x_0, x_1, \dots) that increases up to x_k ($k > 0$) and decreases afterwards. Then $S(x_k)$ must be a Condorcet winner in $I(x_{k-1})$, and $S(x_{k+1})$ must be a Condorcet winner in $I(x_k)$. In particular, a majority in $I(x_{k-1})$ must prefer $S(x_k)$ to $S(x_{k+1})$ but a majority in $I(x_k)$ must prefer the opposite. This is impossible because $S(x_{k+1})$ has a lower average policy than $S(x_k)$, while the group $I(x_k)$ contains agents with higher bliss points than $I(x_{k-1})$.⁸ Our next result characterizes the long-run behavior of any MVE satisfying the MVT.

Proposition 2. *In any MVE s and for any y :*

- (i) *If $m(y) = y$ then $s(y) = y$.*
- (ii) *If $m(y) > y$ then $m^*(y) > s(y) \geq y$. Moreover, if the MVT holds in $[y, m^*(y)]$ then $s(y) > y$ and $s^k(y) \xrightarrow[k \rightarrow \infty]{} m^*(y)$.*
- (iii) *If $m(y) < y$ then $m^*(y) < s(y) \leq y$. Moreover, if the MVT holds in $[m^*(y), y]$ then $s(y) < y$ and $s^k(y) \xrightarrow[k \rightarrow \infty]{} m^*(y)$.*

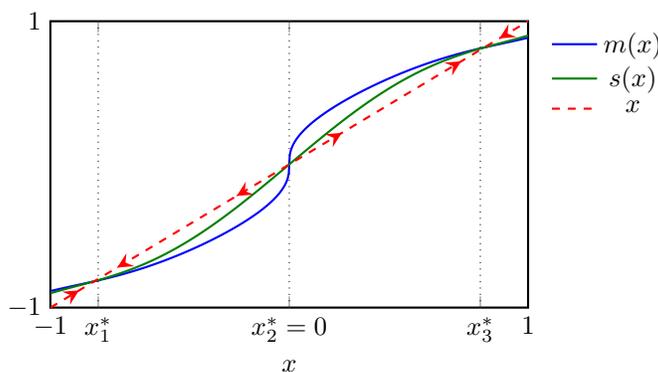


Figure 1: Convergence to steady states in MVE

⁸Analogous results are shown in Roberts (2015) and Acemoglu et al. (2015). The proof here is more involved because, owing to the infinite policy space, we have to rule out non-monotonic paths that never reach their supremum or infimum.

In other words, the steady states of any MVE s satisfying the MVT are simply the fixed points of the mapping $y \mapsto m(y)$. Moreover, stable (unstable) steady states of s are also stable (unstable) fixed points of m , and their basins of attraction coincide.

The intuition for why we should observe $s(x) \leq x$ if $m(x) < x$ and vice versa is straightforward: if $m(x) < x$ for x in an interval (x^*, x^{**}) , any policy in that interval attracts a set of voters whose median wants a lower policy. The main conclusion of Proposition 2 is that slippery slope concerns cannot create myopically unstable steady states, that is, $s(x) \neq x$ if there is a myopic incentive to change the policy. The logic behind the proof is as follows. Suppose $m(x) < x$, but $m(x)$ is afraid of further policy changes if she moves to any $y < x$. If $m(x)$ chooses a *slightly* better policy $y = x - \epsilon$, her flow payoff tomorrow will increase by roughly $\epsilon \left| \frac{\partial u}{\partial x} \right|$. In exchange, she will relinquish control over the continuation to a slightly different voter, $m(x - \epsilon)$. Because they have similar preferences (Assumption A2), $m(x - \epsilon)$'s optimal choice is also approximately optimal for $m(x)$. Hence the cost of losing control is small, that is, no higher than $M(m(x) - m(x - \epsilon)) \sum_t \delta^k |s^k(x) - s^k(s(x - \epsilon))|$. If $S(s(x - \epsilon))$ converges to $S(x)$ pointwise as $\epsilon \rightarrow 0$, this loss is of order $o(\epsilon)$, so $m(x)$ should deviate to $y = x - \epsilon$ for ϵ small enough. If not, it can be shown that $m(x)$ must be indifferent between $S(x)$ and $\lim_{\epsilon \rightarrow 0} S(s(x - \epsilon))$, and an analogous argument can then be made.

Figure 1 illustrates the equilibrium properties stated in Proposition 2 in an example with three steady states: x_1^* and x_3^* are stable, while x_2^* is unstable. This alternation of stable and unstable steady states occurs in general as long as m is well-behaved. Formally, in the rest of the paper we will assume the following:

B1 The equation $m(y) = y$ has finitely many solutions $x_1^* < x_2^* < \dots < x_N^*$. In addition, $m'(x_i^*) \neq 1$ for all i .

Corollary 1. *m has an odd number of fixed points. For odd i , $m'(x_i^*) < 1$ and x_i^* is a stable steady state of every MVE; for even i , $m'(x_i^*) > 1$ and x_i^* is unstable.*

Our last result in this Section guarantees that, in a sizable neighborhood of each stable steady state, every equilibrium satisfies the MVT, and hence the main conclusion of Proposition 2 applies. In addition, within the same neighborhood every equilibrium must be monotonic (a stronger property than path-monotonicity) and an equilibrium restricted to this neighborhood must exist.

Proposition 3. *Let x^* be such that $m(x^*) = x^*$ and $m'(x^*) < 1$; let $x^{***} < x^* < x^{**}$ be the unstable steady states adjacent to x^* . Then an MVE restricted to $I(x^*) \cap (x^{***}, x^{**})$ exists,⁹ and any MVE is weakly increasing and satisfies the MVT in $I(x^*) \cap (x^{***}, x^{**})$.*

The reason these results may fail to hold outside of $I(x^*) \cap (x^{***}, x^{**})$ is that they rely on pivotal voters not leaving the club on the equilibrium path; when pivotal voters quit the club at different times, the logic that voters with higher bliss points should like higher paths does not always apply.¹⁰

We finish this Section with two remarks. First, an alternative approach to solving for MVEs would be to study a game in which, given a policy x , the agent $m(x)$ is by assumption given direct control over tomorrow's policy.¹¹ In this closely related game, the main conclusion of Proposition 2 holds for all Markov equilibria. Second, as we will see next, under some conditions we will be able to guarantee the existence of MVE that are monotonic and satisfy the main conclusion of Proposition 2 everywhere.

4 Transition Dynamics

This Section analyzes the transition dynamics of the model in more detail. Without loss of generality, we restrict our analysis to the right side of the basin of attraction of a stable steady state, that is, an interval $[x^*, x^{**})$ such that $m(x^*) = x^*$, $m(x^{**}) = x^{**}$ and $m(y) < y$ for all $y \in (x^*, x^{**})$.¹²

We begin by noting that, under mild conditions, convergence to a steady state is far from instant, and becomes arbitrarily slow if agents are arbitrarily patient. Formally, say $m(x) \in (x^*, x^{**})$ is *reluctant* if $u_{m(x)}(x) > u_{m(x)}(x^*)$, i.e., $m(x)$ would rather stay at x than move instantly to x^* .¹³ If so, let $z(x)$ be the unique policy in $(x^*, m(x))$ for which $u_{m(x)}(x) = u_{m(x)}(z(x))$.

Proposition 4. *If $m(x)$ is reluctant, then, for all $y < z(x)$, $\exists K(y) > 0$ such that, for any δ and any MVE s of the game with discount factor δ , $\min\{t : s^t(x) \leq y\} \geq \frac{K(y)\delta}{1-\delta}$.*

⁹ s is a MVE restricted to $[a, b]$ if it is a MVE when the policy space is restricted to $[a, b]$.

¹⁰For instance, let $\alpha = 0.6$, $\tilde{\alpha} = 0.5$, and $S = (0.7, 0.1)$, $T = (0.65, 0)$ be two-period policy paths. If $\alpha, \tilde{\alpha}$ never leave under either path, $U_\alpha(S) - U_\alpha(T) > U_{\tilde{\alpha}}(S) - U_{\tilde{\alpha}}(T)$ by A2; but if $u_\alpha(0) < 0$ it is possible that $U_\alpha(S) - U_\alpha(T) < U_{\tilde{\alpha}}(S) - U_{\tilde{\alpha}}(T)$. By the same logic, the set of voters preferring S to T may not be an interval, so a winning coalition need not contain the median voter.

¹¹This approach has been taken in the literature, e.g., in Bai and Lagunoff (2011).

¹²For any $y \in (x^*, x^{**})$, $I(y)$ never wants to choose a policy outside of (x^*, x^{**}) , so $s|_{(x^*, x^{**})}$ can be studied in isolation. Results for a basin of attraction of the form $[x^*, 1]$ or $[x^{***}, x^*]$ are analogous.

¹³For instance, in the quadratic case, if $m'(x) > \frac{1}{2}$ for all x then every agent is reluctant.

The reason is simply that, if this condition were violated, $m(x)$ would rather stay at x forever.¹⁴ Thus, if there are reluctant agents, knowing the club's long-run policy is not enough to characterize the agents' utility even as $\delta \rightarrow 1$, unlike in related models (cf. Acemoglu et al. 2012); further analysis of the transition path is necessary.

In the remainder of this Section we propose a natural class of equilibria, which we call *1-equilibria* (henceforth 1Es), and study their transition dynamics. There are two reasons to focus on 1Es. The first is tractability: 1Es have a simple structure, and their transition dynamics can be explicitly characterized in the limit as agents become more patient. The second is robustness: 1Es are guaranteed to exist under some conditions we will provide, while other types of equilibria (including smooth equilibria) cannot be guaranteed to exist. We begin with a definition of 1Es and two related concepts.

Definition 3. Let s be a successor function on $[x^*, x^{**}]$. s is a *1-function* if there is a sequence $(x_n)_{n \in \mathbb{Z}}$ such that $x_{n+1} < x_n$ for all n , $x_n \xrightarrow{n \rightarrow -\infty} x^{**}$, $x_n \xrightarrow{n \rightarrow \infty} x^*$, and $s(x) = x_{n+1}$ if $x \in [x_n, x_{n-1})$. We call $(x_n)_n$ the *recognized sequence* of s .

s is a *1-equilibrium* (1E) if it is a 1-function and an MVE.

s is a *quasi-1-equilibrium* (Q1E) if it is a 1-function such that $(1-\delta)U_{m(x_n)}(S(x_{n+1})) = u_{m(x_n)}(x_{n+1})$ for all n .

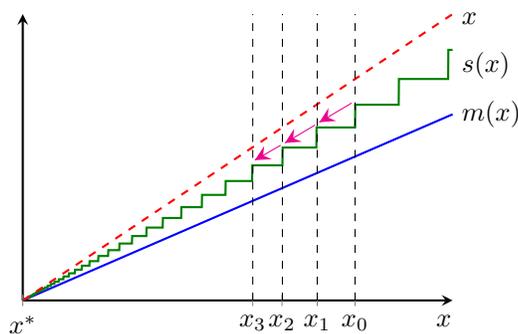


Figure 2: 1-equilibrium for $u_\alpha(x) = C - (\alpha - x)^2$, $m(x) = 0.7x$, $\delta = 0.7$

In a 1E, the chosen policies are always elements of the recognized sequence $(x_n)_n$. x_n today leads to x_{n+1} tomorrow; if the initial policy is not part of the recognized sequence, but is between x_n and x_{n-1} , then x_{n+1} is chosen, and the path follows the

¹⁴A partial converse holds: if $u_{m(x)}(x^*) > u_{m(x)}(x)$ for all x then, for all $y \in (x^*, x)$, $\min\{t : s^t(x) \leq y\} \leq \tilde{K}(y)$, with $\tilde{K}(y)$ independent of δ .

recognized sequence thereafter. An illustration is given in Figure 2. The notion of Q1E is useful to study because Q1Es are closely related to 1Es, but are easier to construct. The following Proposition summarizes the relationship between the two.

Proposition 5.

- (i) *Every 1E is also a Q1E.*
- (ii) *A Q1E exists. Moreover, for each $\underline{x} \in (x^*, x^{**})$, there is a Q1E $s_{\underline{x}}$ with $x_0 = \underline{x}$.*
- (iii) *In any Q1E, for all $x \in [x_n, x_{n-1})$ and for all k , a majority in $I(x)$ prefers $S(x_{n+1})$ to $S(x_k)$.*
- (iv) *Any Q1E such that $m(x_n) < x_{n+2}$ for all n is a 1E within $[x^*, m^{-1}(x^* + d_{x^*}^+)]$.*

To see why 1Es are also Q1Es, consider a 1E s with recognized sequence $(x_n)_n$. By construction, a majority in $I(x_n)$ prefers $S(x_{n+1})$ to $S(x_{n+2})$. But, for any x in a left-neighborhood of x_n , a majority of $I(x)$ prefers $S(x_{n+2})$ to $S(x_{n+1})$. Due to the fact that $S(x_{n+1}) = (x_{n+1}, S(x_{n+2}))$, a voter α prefers $S(x_{n+1})$ to $S(x_{n+2})$ iff $u_\alpha(x_{n+1}) \geq (1 - \delta)U_\alpha(S(x_{n+1}))$, and it can be shown that the set of voters with this preference is of the form $[\alpha^*, 1]$. Then we must have $\alpha^* = m(x_n)$. By continuity, $m(x_n)$ is indifferent between $S(x_{n+1})$ and $S(x_{n+2})$, i.e., $(1 - \delta)U_{m(x_n)}(S(x_{n+1})) = u_{m(x_n)}(x_{n+1})$.

Part (ii) of Proposition 5 guarantees that Q1Es always exist—in fact, there is a continuum of them. Part (iii) is a partial converse of (i): it shows that, in a Q1E, no deviations to other policies on the recognized sequence are possible. To see why $I(x_{n-1})$ will not deviate to $S(x_{n+2})$, for instance, note that $m(x_n)$ is indifferent between $S(x_{n+1})$ and $S(x_{n+2})$, so $m(x_{n-1})$ prefers $S(x_{n+1})$ to $S(x_{n+2})$, and is indifferent between $S(x_n)$ and $S(x_{n+1})$ by construction. Equivalently, then, the only reason a Q1E may fail to be a 1E is if a majority wants to deviate off the recognized sequence.

Intuitively, such deviations will not be desirable if the median voter, $m(x_n)$, is far to the left of the first few elements of the sequence following x_n . To see why, note that, if $m(x_n) \in [x_k, x_{k-1})$, deviating to a policy $y \in (x_l, x_{l-1})$ with $l < k$ would be even worse than deviating to x_l , while deviating to $y \in (x_l, x_{l-1})$ with $l > k$ would be worse than deviating to x_{l-1} . Hence the only deviations $m(x_n)$ might prefer are deviations to $y \in (x_k, x_{k-1})$. In particular, picking $y = m(x_n)$ is better than deviating to x_k . But such a deviation will still be unprofitable if $S(x_{n+1})$ is too strongly preferred to

$S(x_k)$. This is the idea behind part (iv); a more powerful result along these lines will be given in the next subsection.

Finally, note that, in a 1E, $m(x_n)$'s averaged per-period utility equals $u_{m(x_n)}(x_{n+1})$. Hence her net gain from not staying at x_n forever, $V(m(x_n)) := (1-\delta)U_{m(x_n)}(S(x_{n+1})) - u_{m(x_n)}(x_n)$, is approximately proportional to the equilibrium speed of policy change; specifically, it equals $(x_n - x_{n+1}) \left| \frac{\partial u_{m(x_n)}(y)}{\partial y} \right|$ for some $y \in [x_{n+1}, x_n]$.

Continuous Time Limit

We now characterize the limit of 1Es as the time gap between rounds of voting becomes arbitrarily small. This can be taken as an approximation of a setting in which voting happens periodically (e.g., at annual elections), but often enough relative to the agents' time horizon. The same results will also allow us to characterize the limit of 1Es with a fixed time gap between periods as $\delta \rightarrow 1$.

Denote $\delta = e^{-r}$. We will work with the following objects. First, for each $j \in \mathbb{N}$, consider a version of the game from Section 2 in which policy and membership decisions are made at every time $t \in \{0, \frac{1}{j}, \frac{2}{j}, \dots\}$ instead of at every integer time. We will call this the j -refined game, and denote a Q1E of this game by s_j . In addition, for each $t \in \mathbb{R}_{\geq 0}$, we denote by $s_j(x, t)$ the equilibrium policy at time t if the initial policy is x and s_j is played—that is, $s_j(x, t) = s_j^{\lfloor tj \rfloor}(x)$. Note that this game is, up to a relabeling, equivalent to the model in Section 2 with discount factor $\delta_j = e^{-\frac{r}{j}}$.

Finally, we define a continuous limit solution (CLS) as a function $s(x, t) : [0, +\infty) \times [x^*, x^{**}] \rightarrow [x^*, x^{**}]$ with the following properties: $s(x, t+t') \equiv s(s(x, t'), t)$; $s(x, 0) \equiv x$; s is weakly decreasing in t ; $s(x, t) \xrightarrow[t \rightarrow 0]{} x$ for all x ; and $\bar{U}_{m(x)}(S(x)) = u_{m(x)}(x)$ for all $x \in (x^*, x^{**})$, where $\bar{U}_\alpha(S(x)) = \int_0^\infty r e^{-rt} \max(u_\alpha(s(x, t)), 0) dt$.

The following Proposition relates the CLS to the Q1Es of the j -refined games.

Proposition 6. *Suppose that $m \in C^2$ and a CLS s exists. Then:*

- (i) s is the unique CLS, and s is C^1 as a function of t .
- (ii) For any sequence $(s_j)_j$, where s_j is a Q1E of the j -refined game, for all x and t , $s_j(x, t) \xrightarrow[j \rightarrow \infty]{} s(x, t)$.
- (iii) There is $\bar{\delta} < 1$ such that, for all $\delta > \bar{\delta}$, all Q1Es of the discrete-time game with discount factor δ are 1Es.

The intuition behind a CLS is the following. Fix x , and take a sequence $(s_j)_j$ of Q1Es of the j -refined games with $x_{j0} = x$. Recall that $U_{m(x_{jn})}(S_j(x_{j(n+1)})) = u_{m(x_{jn})}(x_{j(n+1)})$ for all j, n .¹⁵ Suppose that, as $j \rightarrow \infty$, the transition paths $S_j(x) = (s_j(x, t))_t$ converge pointwise to a continuous path $S(x)$. Then the differences $x_{j0} - x_{j1}$ go to zero, and in the limit, $U_{m(x)}(S(x)) = u_{m(x)}(x)$. This is why we require this condition of a CLS. Part (ii) of Proposition 6 is a converse to this argument: it shows that, when a CLS exists, the transition paths of all Q1Es must converge to it as $j \rightarrow \infty$.

Whether a CLS exists is a property of the primitives u and m ; it can be determined in isolation from our game. We can both verify whether a CLS exists and calculate it explicitly, as follows. Denote by $e(x) = -\frac{1}{\frac{\partial s(x,t)}{\partial t}|_{t=0}}$ the instantaneous delay of a CLS at x . If $U_{m(x)}(S(x)) = u_{m(x)}(x)$ for all x , then, differentiating with respect to x ,

$$m'(x) \frac{\partial U_{m(x)}(S(x))}{\partial \alpha} = m'(x) \frac{\partial u_{m(x)}(x)}{\partial \alpha} + \frac{\partial u_{m(x)}(x)}{\partial x}. \quad (1)$$

The key observation is that $\frac{\partial}{\partial x} \frac{\partial U_{m(x)}(S(x))}{\partial \alpha} = \left(\frac{\partial u_{m(x)}(x)}{\partial \alpha} - \frac{\partial U_{m(x)}(S(x))}{\partial \alpha} \right) re(x)$. Hence, differentiating Equation 1 yields an equation that pins down $e(x)$. After rearranging, we find

$$-\frac{\partial u}{\partial x} re(x) = 2m' \frac{\partial^2 u}{\partial \alpha \partial x} + \frac{\partial^2 u}{\partial x^2} + (m')^2 \left(\frac{\partial^2 u}{\partial \alpha^2} - \frac{\partial^2 U_{m(x)}(S(x))}{\partial \alpha^2} \right) - \frac{m''}{m'} \frac{\partial u}{\partial x}, \quad (2)$$

where u stands for $u_{m(x)}(x)$. This is an integral equation because $\frac{\partial^2 U}{\partial \alpha^2}$ is evaluated at $S(x)$, which depends on $e(y)$ for $y \in [x^*, x)$. But we can solve it forward, starting at x^* , to find $e(x)$ and, with it, the unique CLS. In fact, a CLS always exists unless the function $e(x)$ that solves Equation 2 turns negative. This is guaranteed not to happen if the forces pulling towards policy change are not too great:

Proposition 7. *Holding u constant, there exist $B, B', B'' > 0$ such that, if $x - m(x) \leq B$, $m'(x) \in [1 - B', 1 + B']$ and $m''(x) \geq -B''$ for all x , then a CLS exists.*

The quadratic case is useful for illustrative purposes. In that case, for $x \in I(x^*)$, Equation 2 reduces to

$$re(x) = \frac{2m'(x) - 1}{x - m(x)} + \frac{m''(x)}{m'(x)}, \quad (3)$$

¹⁵In this argument, we write $U_{m(x)}((x_t)_t) = (1 - \delta) \sum \delta^t I_{\alpha t} u_{\alpha}(x_t)$ to simplify notation.

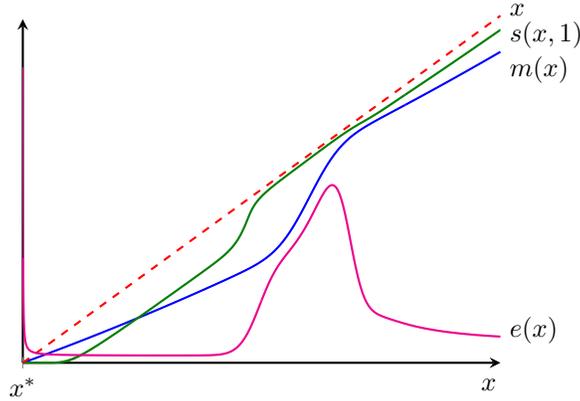


Figure 3: A continuous limit solution

so the speed of convergence can be calculated in closed form. Equation 3 reflects three forces at work in determining $e(x)$. First, the policy changes more slowly ($e(x)$ is higher) when the myopic incentive for policy changes, $x - m(x)$, is small. Second, the policy changes more slowly when $m'(x)$ is high. The reason is that changing the policy to $x - \epsilon$ entails yielding control to another agent $m(x - \epsilon)$. The higher m' is, the more costly this loss of control becomes. Third, the policy changes more slowly when $m''(x)$ is high. The reason is that, when m is convex, $m'(x)$ is higher than $m'(x - \epsilon)$; hence the agent $m(x)$ yields control to will not be as concerned about the behavior of her own successors, and so will make faster policy changes than $m(x)$ would like. These forces are illustrated in Figure 3.

Finally, part (iii) of Proposition 6 guarantees that, when there is a CLS, all Q1Es must be true 1Es when agents are patient enough. The reason is related to our discussion of Proposition 5: as the transition paths of Q1Es approach the CLS, they must feature smaller and smaller jumps between consecutive policies; in this scenario, deviations off the recognized sequence are never majority-preferred.

We conclude this Section with a few observations. First, the speed of policy change in a CLS is exactly inversely proportional to the agents' patience:

Remark 1. If $s(x, t)$ is a CLS for discount rate r , then $\tilde{s}(x, t) \equiv s(x, \frac{\tilde{r}}{r}t)$ is a CLS for discount rate \tilde{r} . The respective instantaneous delays $e(x)$, $\tilde{e}(x)$ satisfy $\tilde{e}(x) = \frac{r}{\tilde{r}}e(x)$.

The reason is that changing r is equivalent to a relabeling of the time variable. Second, when a CLS exists, Proposition 6 and Equation 2 together yield an asymptotic characterization of the transition path for all 1Es of the game from Section 2 when

agents are patient.¹⁶ Formally, let $e_1(x)$ be the solution to Equation 2 for $r = 1$. Then, for any $y < x$ and any collection of 1Es s_δ for $\delta \geq \bar{\delta}$,

$$(1 - \delta) \min\{t : s_\delta^t(x) \leq y\} \xrightarrow{\delta \rightarrow 1} \int_y^x e_1(z) dz. \quad (4)$$

This is, in effect, a more precise version of Proposition 4. Third, recall that, in a 1E, the net per-period gain $V(m(x_n))$ of a pivotal agent $m(x_n)$ from following the equilibrium path (relative to staying at x_n) is of the order $x_n - x_{n+1}$. Thus, if a CLS exists, for any x and any collection of 1Es s_δ , $V_\delta(m(x)) \rightarrow 0$ as $\delta \rightarrow 1$. In other words, the “rents” a pivotal agent gets from the best non-constant continuation evaporate as agents become more patient (or decisions are made more often). An intuition is that these rents are the result of agents being able to “lock in” their chosen policy for one period before losing control—hence they vanish as the periods shorten.

Fourth, it is not hard to find examples in which a CLS fails to exist.¹⁷ However, even when there is no CLS, a version of Proposition 6 holds: under some conditions, Q1Es can still be guaranteed to be 1Es for high δ , and all sequences of Q1E transition paths converge to a common limit, but this limit is no longer continuous. The details for this case are presented in Appendix B.

5 A Model of Political Power

We now discuss a variant of the model that overturns the assumption of free entry and exit. Consider a polity governed by an endogenous *ruling coalition*. At each time $t = 1, 2, \dots$ the ruling coalition chooses a policy x_t ; the initial policy x_0 is exogenous.

The model is the same as the one presented in Section 2, but with two differences. First, the policy x_t now directly determines not just the flow payoffs of all agents during the period $[t, t + 1)$, but also the ruling coalition at time $t + 1$. In other words, the mapping $x \mapsto I(x)$ is now taken as a primitive of the model. (We assume that $I(x)$ is still an interval $(x - d_x^-, x + d_x^+)$ for each x , with $x - d_x^-$, $x + d_x^+$ increasing and C^1 as functions of x .) Second, in this model, *all* agents are impacted by the policy,

¹⁶This is a consequence of Proposition 6 for a sequence $(\delta_j)_j$ of the form $\delta_j = e^{-\frac{x}{j}}$, but in fact the proof of the Proposition does not rely on the j 's being integers, only that $j \rightarrow \infty$.

¹⁷For example, if there is a non-reluctant agent, then a CLS cannot exist.

regardless of whether they are in the ruling coalition. In other words,

$$U_\alpha((x_t)_t, I_\alpha) = \sum_{t=0}^{\infty} \delta^t u_\alpha(x_t),$$

where u_α satisfies A1, A2 and A4. This setting is similar to the canonical model of “elite clubs” (Roberts, 2015), but with a continuous policy space. It can be framed as a model of enfranchisement (Jack and Lagunoff, 2006), institutional change (Acemoglu et al., 2012, 2015), or economic policymaking in a world where political influence is a function of wealth (Bai and Lagunoff, 2011). Clearly, slippery slope concerns apply here as well: a ruling coalition may want to expand the franchise (e.g., to lower unrest) but fear that the new voters will choose to expand it even further.

Our analysis of the main model extends to this case as follows. First, all of our previous Propositions continue to hold, with analogous proofs. Second, the conclusions of Propositions 3 and 5(iv) now hold for all $x \in [-1, 1]$, as opposed to only in a neighborhood of each stable steady state.¹⁸ In particular, an MVE exists; every MVE satisfies the MVT for all x ; and for every MVE $s^k(x) \xrightarrow[k \rightarrow \infty]{} m^*(x)$ for all x .

Although this version of the model represents a setting with different causal relationships between political power, membership, and flow payoffs, it is closely related to the model from Section 2: indeed, there is a mechanical equivalence between the components of both models, and the same tension between current payoffs and future control is present in both cases.

Other variants, allowing the model to fit new examples, are possible. For instance, the set of members can affect payoffs directly: $v_\alpha(x) = u_\alpha(x) + w_\alpha(I(x))$. So long as $v_\alpha(x)$ satisfies A1-4, all of our results apply. A natural example is immigration: if x_t is a country’s immigration policy and $I(x_t)$ is its set of citizens, x_t does not affect the payoffs of current citizens directly, but the entry of immigrants does.¹⁹

6 Discussion

This Section discusses some implications of our analysis.

¹⁸The proofs of Propositions 3 and 5(iv) assume that pivotal agents never stop receiving payoffs from the club’s policy. In the main model, this requires them not to leave the club; in this variant it does not matter if they are part of the ruling coalition.

¹⁹This example has been studied in the literature, although in an overlapping-generations framework (Ortega, 2005; Suwankiri, Razin and Sadka, 2016).

Myopic Stability of Steady States

Two central results of our analysis are that steady states are myopically stable (or, in the language of Roberts (2015), “extrinsic”),²⁰ and convergence to a steady state is slow when agents are patient. In contrast, in other papers in this literature (Roberts, 2015; Acemoglu et al., 2008, 2012, 2015), intrinsic steady states are possible, and the time it takes to converge to a steady state is uniformly bounded even as $\delta \rightarrow 1$.

These papers assume a fixed, finite policy space. Under this assumption, convergence is fast because there is a mechanical lower bound on the size of policy changes; as a result, intrinsic steady states *must* exist for δ high enough, if there are reluctant agents. What we show is that these results are overturned if arbitrarily small policy changes are allowed. For a fixed $\delta < 1$, our results also hold if the policy space is finite but fine enough; if we simultaneously take δ to 1 and make the policy space arbitrarily fine, whether intrinsic steady states exist depends on the order of limits.

The upshot is that, in practice, whether dynamic concerns can indefinitely stall policy changes depends not just on the agents’ foresight but also on institutional details that determine whether incremental changes are possible. For example, take a polity with a limited franchise considering a franchise extension on the basis of income. Suppose that each voter prefers a larger franchise than the smallest one she would be in (e.g., for all x , a voter at the x th income percentile wants to enfranchise everyone above the $(x - 5)$ th percentile). Then, if it is possible to enfranchise the top $y\%$ of the income distribution for any y , full democracy would eventually be reached through a series of small changes. However, if voting rights can only be extended based on a few criteria (e.g., only to men who can read, to landowners, to taxpayers, etc.), indefinite stalling is likely.²¹

Two important precursors to our analysis, Jack and Lagunoff (2006) and Bai and Lagunoff (2011), consider dynamic political decision-making with continuous policy spaces. In particular, Bai and Lagunoff (2011) show that in their model, steady states of “smooth” equilibria are also stable when the current decision-maker is assumed to retain power forever (in our setting, this is equivalent to myopic stability). However, their analysis uses a first-order approach, and so does not extend to other types of equilibria; moreover, smooth equilibria generically fail to exist, as the second-order

²⁰This is true for all 1Es; for all other MVEs in a neighborhood of each stable steady state; and globally for all MVEs in the model discussed in Section 5, which is closest to this literature.

²¹Jack and Lagunoff (2006) make the case that franchise extensions are typically gradual processes.

conditions are typically violated.²² We build on this result by showing that steady states must be myopically stable for all equilibria, including discontinuous ones.

Distribution of Steady States

Although f is the primitive of our model describing the distribution of preferences, our results are best stated in terms of m , the median voter function. Here, we briefly discuss the relationship between the two objects, focusing on how the shape of f affects the location of steady states.

In the quadratic case, or more generally whenever $I(x) \equiv (x-d, x+d)$ is symmetric around x , the distribution of steady states reflects the following intuition: if f is increasing within $I(x)$ then $m(x) > x$, and vice versa. Hence, stable steady states correspond roughly to maxima of the density function.

Remark 2. If $I(x) = (x-d, x+d)$ for all x , and x^* is a stable (unstable) steady state, then $I(x^*)$ contains a local maximum (minimum) of f .

In particular, if f is increasing (decreasing) everywhere, there is a unique steady state close to 1 (-1); if f is symmetric and single-peaked, 0 is the unique steady state. Thus, the organization always moves to the center if the distribution of preferences is bell-shaped. Yet, there are three scenarios in which the organization may converge to a policy more extreme than the bliss points of most voters.

First, even if most voters are near the center, a local maximum near an extreme may support a stable steady state, especially if d is low.²³

Remark 3. If $f'(x^*) = 0$ and $f''(x^*) < 0$, then there is $\bar{d} > 0$ such that for all $d < \bar{d}$, if $I(x) \equiv (x-d, x+d)$ for all x , $(x^* - d, x^* + d)$ contains a stable steady state.

Second, even if there is a unique steady state, its location will be unstable when f is close to uniform. For example, consider the densities $f_1(x) = \frac{1}{2} + \epsilon x$, $f_2(x) = \frac{1}{2} - \epsilon x$ and $f_3(x) = \frac{1+\epsilon}{2} - \epsilon|x|$, for a small $\epsilon > 0$. These are all close to each other ($\|f_i - f_j\|_\infty \leq 2\epsilon \forall i, j$) but f_1 has a unique steady state near -1 , f_2 has one near 1, and f_3 has one at 0. Hence, the long-run policy is potentially discontinuous in f , unlike in models of voting with a fixed population.

²²The existence properties of smooth equilibria are discussed in Appendix D.

²³Note that, when there are multiple steady states, which one the organization converges to depends on the initial policy, i.e., there is path-dependence. There is no guarantee that the organization will converge to a steady state that attracts the most members or maximizes aggregate welfare.

Third, the tendency towards policies preferred by a majority can be easily overturned when the voter sets $I(x)$ are not symmetric around x . In particular, if agents with extreme preferences are disproportionately more willing to join the organization, they can capture it despite being a minority, even locally.²⁴

For example, let the policy space be $[-1, 1]$, where -1 is the most moderate policy and 1 the most extreme, and assume $u_\alpha(x) = -|\alpha - x| + (1 + \alpha)$.²⁵ Then α wants to be a member whenever $x \in [-1, 2\alpha + 1]$, whence $I(x) = [\frac{-1+x}{2}, 1]$.

Suppose f is as follows: moderates constitute 60% of the population and have bliss points uniformly distributed in $[-1, -0.9]$; extremists, the remaining 40%, have bliss points uniformly distributed in $[-0.9, 1]$. It can then be shown that the unique steady state is $x^* = \frac{1}{3} > 0$. At the steady state policy, the set of members $I(\frac{1}{3}) = [-\frac{1}{3}, 1]$ is only 28% of the population, all of them extremists.

7 Conclusion

We conclude with a discussion of some issues that the model leaves out, possible extensions, and additional results presented in the Appendix.

Our model focuses on the behavior of a single organization, but organizations often compete—in particular, the usual assumption in Tiebout competition is that there are many districts. There are two ways of modeling competition. The first is to assume, as in the idealized Tiebout model, that districts are identical except for their policies. In such a model, the same dynamics we have studied would arise, but with the complication that policy changes in one district may lead to responses by other districts. At the same time, if there are enough districts, every agent would find a district near her bliss point, so the potential welfare impact of policy changes in individual districts would vanish as the number of districts grows.²⁶

An alternative approach is to assume that districts are imperfect substitutes. The example in Section 2 of a city with a competitive advantage over others is a version of this, and it can be generalized. Suppose that there are $k > 1$ special cities c_1^* ,

²⁴This asymmetry is likely in settings where too-extreme policies are perceived as reprehensible or criminal, but not the reverse (e.g., fringe political parties, violent protest movements, or advocacy groups whose causes can be perceived as racist or xenophobic).

²⁵The example is degenerate in that $\frac{\partial^2 u}{\partial \alpha \partial x}$ is only weakly positive and u is only weakly concave in x ; this is only for simplicity.

²⁶Tiebout (1956) conjectured that agents sorting into compatible districts would lead to an efficient outcome. This idea has been formalized (Wooders, 1989) as well as criticized (Bewley, 1981).

..., c_k^* and k groups of mobile agents, so that agent types are of the form (α, i) , for $i \in \{1, \dots, k\}$, and group i 's bliss points are distributed according to a density f_i . Assume that $u_{(\alpha, i)}(c) = C\mathbb{1}_{c=c_i^*} - (x_t(c) - \alpha)^2$, that is, agents in group i only value the city c_i^* more than normal cities. (For instance, agents in group i have immobile relatives in city i .) Our analysis goes through because each special city c_i^* competes only for agents in group i , and does so only with normal cities. If i 's value from $c_j^* \neq c_i^*$ is some intermediate $C' \in (0, C)$, the problem becomes more complicated, but the relevance of the forces we study does not vanish as $k \rightarrow \infty$. Thus, our model can be taken as an analysis of Tiebout competition in the presence of imperfect substitutability. A similar logic would apply if cities are ex ante identical but have ex post market power due to moving costs.²⁷

A related extension would allow for the endogenous creation of organizations. Our analysis suggests that an agent far from a steady state is less likely to create an organization, or more likely to create a non-democratic one.

The organizations we model are simple: all members have the same voting power, decisions are made by majority rule, and votes are cast independently. Appendix E.1 discusses how to extend our analysis to allow for supermajority rules or other electoral rules that make an agent other than the median pivotal.²⁸ However, there is much unexplored complexity regarding the internal structure of organizations. Agents may have endogenous voting power (seniority); or they may engage in collective behavior by voting as a bloc, joining an organization in large numbers in order to change its policy, or threatening to leave *en masse* to extract concessions. These behaviors are not likely in the context of Tiebout competition, but may be so in other applications.

Organizations may also set up various barriers to entry or membership restrictions (or, in the case of cities, there may be moving costs). In the paper, we consider two extreme cases: one with completely free entry and exit (Section 2) and one in which the organization can choose its set of members at will (Section 5). In Appendix E.2, we present an extension allowing for (exogenous) positive entry and exit costs. Because in equilibrium agents enter and leave the organization at most once, this does not change the analysis much. Modeling endogenous entry costs that can be chosen separately from the organization's main policy, on the other hand, is much

²⁷Moving costs and idiosyncratic preferences are suggested in Tiebout (1956) and Epple and Romer (1991), respectively, as forces preventing convergence to a Tiebout equilibrium in practice.

²⁸Even in democracies, higher-income agents may wield more political power (Benabou, 2000; Jack and Lagunoff, 2006).

harder, as the state space becomes multidimensional. However, the forces we study will still be present as long as the organization cannot perfectly control both its payoff-relevant policy and its membership (if it can, we are back in the world of a fixed decision-maker).

Our analysis abstracts away from history-dependent strategies by focusing on Markovian equilibria. In Appendix E.3, we show that Non-Markovian equilibria can support a large number of outcomes, but that several reasonable refinements select only Markovian equilibria. In particular, any equilibrium obtainable as a limit of discrete policy-space equilibria must be Markovian.

Finally, we do not explicitly model the organization's voting process. One way to interpret our results is that they will hold whenever the organization's collective decision-making process leads to Condorcet-winning policies being chosen. In Appendix E.4, we discuss possible microfoundations of this modeling assumption. In particular, the MVEs we study are Markov Perfect Equilibria of a game in which, in each round of voting, there are two short-lived, office-motivated candidates engaging in Downsian competition. However, not all political processes are so well-behaved. For instance, if the organization has more than two candidates running for office, or it is run by a deliberative decision-making process, then the Condorcet winner may not win.²⁹ In addition, leaders typically have some agency in practice. If they are long-lived and have heterogeneous appeal, or are policy-motivated, they may champion certain policies in an attempt to change the policy path, possibly permanently. Moreover, a politician may strategically push for policies that will attract a set of members predisposed to like her.³⁰

Appendices B, C and D contain technical results. Appendix B contains the proofs of Propositions 6 and 7 and characterizes the case in which no CLS exists. Appendix C shows that the limit solution described in Appendix B exists for all m satisfying a genericity condition. Appendix D discusses the existence properties of equilibria other than 1Es (in particular, smooth equilibria and k -equilibria, a generalization of 1Es); calculates explicit equilibria for the case of linear m and quadratic utility; and gives an example of an MVE that is non-monotonic outside of $I(x^*)$.

²⁹For example, Bouton and Gratton (2015) shows that Condorcet winners may lose in majority runoff elections with three candidates.

³⁰Glaeser and Shleifer (2005) discusses the case of Mayor Curley of Boston, who used wasteful policies in an effort drive out rich citizens of English descent, as he was most popular among the poor Irish population.

A Proofs

Lemma 2. *Let $S = (x_0, x_1, \dots)$, $T = (y_0, y_1, \dots)$ be two policy paths, and let $I(S) = \cup_{n=0}^{\infty} I(x_n)$, $I(T) = \cup_{n=0}^{\infty} I(y_n)$. Suppose that $\sup x_n < \inf y_n$; $I(S)$, $I(T)$ are intervals; and $I(S) \cap I(T) \neq \emptyset$. Then there is α_0 such that agents in $[-1, \alpha_0)$ strictly prefer S to T , and agents in $(\alpha_0, 1]$ strictly prefer T to S .*

Proof. Let $\underline{x} = \inf x_n$, $\bar{x} = \sup x_n$, $\underline{y} = \inf y_n$, $\bar{y} = \sup y_n$. By assumption, $\underline{x} \leq \bar{x} < \underline{y} \leq \bar{y}$. Note that all agents $\alpha < \underline{x}$ strictly prefer S to T by A4 and A6; likewise, all $\alpha > \bar{y}$ strictly prefer T to S .

Let $W(\alpha) = U_\alpha(T) - U_\alpha(S)$. Note that W is continuous and $W(\underline{x}) < 0 < W(\bar{y})$. Hence there is some $\alpha_0 \in [\underline{x}, \bar{y}]$ for which $W(\alpha_0) = 0$. For any $\alpha \in [\underline{x}, \bar{y}]$, let $I_S^-(\alpha) = \sum_{n=0}^{\infty} \delta^n \mathbb{1}_{x_n \leq \alpha, u_\alpha(x_n) > 0}$, $I_S^+(\alpha) = \sum_{n=0}^{\infty} \delta^n \mathbb{1}_{x_n > \alpha, u_\alpha(x_n) > 0}$. Define $I_T^-(\alpha)$, $I_T^+(\alpha)$ analogously.

If $\alpha \in [\bar{x}, \underline{y}]$, then $I_S^-(\alpha) \geq I_T^-(\alpha) = 0 = I_S^+(\alpha) \leq I_T^+(\alpha)$, and $I_S^-(\alpha) + I_T^+(\alpha) > 0$ by the assumption that $I(S) \cap I(T) \neq \emptyset$. Then $W'(\alpha) > 0$ by A5.

If $\alpha \in [\underline{y}, \bar{y}]$, then $I_S^+(\alpha) = 0$, and one of the following must be true. If $I_S^-(\alpha) \leq I_T^-(\alpha)$, then $W(\alpha) > 0$ by A4, and moreover $W(\alpha') > 0$ for all $\alpha' \in [\alpha, \bar{y}]$. If $I_S^-(\alpha) > I_T^-(\alpha)$, then $W'(\alpha) > 0$ by A2 and A5. Similarly, for each $\alpha \in [\underline{x}, \bar{x}]$, either $W'(\alpha) > 0$ or $W(\alpha') < 0$ for all $\alpha' \in [\underline{x}, \alpha]$.

In general, then, we can find thresholds z_0, z_1 such that $\underline{x} \leq z_0 \leq \bar{x} < \underline{y} \leq z_1 \leq \bar{y}$; $W'(\alpha) > 0$ for all $\alpha \in [z_0, z_1]$; $W(\alpha) < 0$ for all $\alpha \in [\underline{x}, z_0]$; and $W(\alpha) > 0$ for all $\alpha \in (z_1, \bar{y}]$. Hence W can vanish at most at one point, so α_0 is unique. \square

Lemma 3. *In any MVE s , for any y , $I(y) \cap I(s(y))$ has positive measure. Hence $I(S(y))$ is an interval for all y .*

Proof. Suppose WLOG $y < s(y)$. We argue that $y + d_y^+ > s(y) - d_{s(y)}^-$. Suppose this is false; then all voters in $I(y)$ get utility 0 from policy $s(y)$. If $E = \{\alpha \in I(y) : U_\alpha(S(s(y))) > 0\}$ is a strict majority of $I(y)$, this leads to a contradiction, as all of E would strictly prefer $S(s^2(y))$ to $S(s(y))$. Let $D = \{\alpha \in I(y) : U_\alpha(S(s(y))) \geq U_\alpha(S(y))\} \subseteq E$. Since $S(s(y))$ is a Condorcet winner in $I(y)$, D is a majority in $I(y)$. If $I(y) \subseteq D$ we are done. If not, $\exists \alpha_0 \in I(y) - D$. By the continuity of $U_\alpha(S(s(y)))$, $\exists \alpha_1$ such that $0 < U_{\alpha_1}(S(s(y))) < U_{\alpha_1}(S(y))$, and this inequality holds in some neighborhood $(\alpha_1 - \varepsilon, \alpha_1 + \varepsilon)$. Hence $E - D$ has positive measure and E is a strict majority of $I(y)$. \square

Corollary 2. *In any MVE, let $S = S(y)$ for some $y < x$ and $T = (x, x, \dots)$, with $\sup(S) \leq x$. Then there is $\alpha_0 \leq x$ such that agents in $[-1, \alpha_0)$ strictly prefer S to T , and agents in $(\alpha_0, 1]$ strictly prefer T to S .*

Proof. If $I(S) \cap I(x) \neq \emptyset$, this follows directly from Lemmas 2 and 3. If not, then all $\alpha \geq x - d_x^-$ strictly prefer T to S and all $\alpha \leq \sup I(S)$ strictly prefer S to T . Let α'_0 be such that $u_{\alpha'_0}(y) = u_{\alpha'_0}(x) < 0$. If $\alpha'_0 \in (\sup I(S), x - d_x^-)$ then take $\alpha_0 = \alpha'_0$. If $\alpha'_0 \leq \sup I(S)$ then take $\alpha_0 = I(S)$. If $\alpha'_0 \geq x - d_x^-$ take $\alpha_0 = x - d_x^-$. \square

Proof of Proposition 1. Suppose $S(y)$ is not monotonic. For this proof, denote $s_k = s^k(y)$, $S_k = S(s_k)$, $I_k = I(s_k)$, $\underline{y} = \inf(S(y))$ and $\bar{y} = \sup(S(y))$. For brevity, we will say α prefers a policy x to a path S if she prefers (x, x, \dots) to S . There are two cases:

Case 1: $S(y)$ attains \underline{y} or \bar{y} , i.e., $\exists k \in \mathbb{N}$ such that $s_k = \bar{y}$ or $s_k = \underline{y}$. Suppose WLOG the former. Then there is a $k \in \mathbb{N}$ such that $s_{k-1} < \bar{y}$, $s_k = \bar{y}$ and $s_{k+1} < \bar{y}$.³¹ Consider the decision made by voters in I_{k-1} and in I_k . Since S_k is the Condorcet winner in I_{k-1} , a majority of I_{k-1} prefer it to S_{k+1} . At the same time, S_{k+1} is Condorcet-winning in I_k , so a majority of I_k prefer S_{k+1} to S_k . Let $A = (s_{k-1} - d_{k-1}^-, s_k - d_k^-)$, $B = (s_k - d_k^-, s_{k-1} + d_{k-1}^+)$, $C = (s_{k-1} + d_{k-1}^+, s_k + d_k^-)$. Note that α prefers S_k to S_{k+1} iff he prefers s_k to S_{k+1} . Apply Corollary 2. If $\alpha_0 \in C$, all voters in $A \cup B$ strictly prefer S_{k+1} to S_k , a contradiction. If $\alpha_0 \in B$, all voters in A strictly prefer S_{k+1} to S_k and all voters in C strictly prefer S_k to S_{k+1} , a contradiction.

Case 2: $S(y)$ never attains its infimum nor its supremum. Then there must be a subsequence $(s_{k_i})_i$ such that $s_{k_i} \xrightarrow{i \rightarrow \infty} \bar{y}$. Construct a sub-subsequence $s_{k_{i_j}}$ such that $s_{k_{i_j}} \xrightarrow{j \rightarrow \infty} \bar{y}$ and $s_{k_{i_j}-1} \xrightarrow{j \rightarrow \infty} s_*^{-1}$ for some limit $s_*^{-1} \leq \bar{y}$. (We can do this because all the s_k are in $[-1, 1]$, which is compact.) Iterating this, construct a nested list of subsequences $((s_{k_{im}})_i)_m$ such that k_{im} is increasing in i for each m ; $K_m = \{k_{im} : i \geq 0\} \supseteq K_{m'}$ for $m < m'$; and, for each m , $s_{k_{im}+r} \xrightarrow{i \rightarrow \infty} s_*^r$ for any $r \in \{-m, \dots, m\}$, where $s_*^0 = \bar{y}$. Let $g_i = k_{ii}$. Then $(s_{g_i})_i$ is a subsequence of $(s_k)_k$ such that $s_{g_i+r} \xrightarrow{i \rightarrow \infty} s_*^r$ for any $r \in \mathbb{Z}$. We now consider four sub-cases.

Case 2.1: Suppose $s_*^r < \bar{y}$ for some $r < 0$ and for some $r' > 0$, and let $\underline{r} < 0 < \bar{r}$ be the numbers closest to 0 satisfying these conditions. Consider the decisions made by I_{g_i+r} and $I_{g_i+\bar{r}-1}$, for high i . In the limit, they imply that a majority in $I(s_*^{\underline{r}})$ prefers \bar{y} to $\tilde{S}(s_*^{\underline{r}})$, while a majority in $I(\bar{y})$ prefers $\tilde{S}(s_*^{\bar{r}})$ to \bar{y} (denoting $\tilde{S}(s_*^{\bar{r}}) = (s_*^{\bar{r}}, s_*^{\bar{r}+1}, \dots)$). As in Case 1, this contradicts Corollary 2.

³¹The same argument would apply if $s^k(y) = \dots = s^{k+m}(y) > s^{k-1}(y), s^{k+m+1}(y)$.

Case 2.2: Suppose $s_*^r < \bar{y}$ for some $r < 0$ but never for $r > 0$. Take \underline{r} maximal, so $s_*^{\underline{r}} < \bar{y}$ and $s_*^r = \bar{y} \forall r > \underline{r}$. Fix $0 < \nu < \bar{y} - s_*^{\underline{r}}$. For each i , let $r'(i)$ be such that $s_{g_i+r'(i)}$ is the first element of the sequence $(s_k)_k$ after $s_{g_i+\underline{r}}$ that is weakly smaller than $s_*^{\underline{r}} + \nu$. Construct a subsequence $(s_{g_{i_j}})_j$ such that $s_{g_{i_j}+r'(i_j)+l} \rightarrow s_{**}^l$ for $l \geq -1$ (in particular $s_{**}^{-1} \geq s_*^{\underline{r}} + \nu \geq s_{**}^0$). Now compare the decisions made by $I(s_{g_i+\underline{r}})$ and $I(s_{g_i+r'(i)-1})$. In the limit, they imply that a weak majority in $I(s_*^{\underline{r}})$ prefers \bar{y} to $\tilde{S}(s_{**}^0)$, while a weak majority in $I(s_{**}^{-1})$ prefers the opposite (here $\tilde{S}(s_{**}^0) = (s_{**}^0, s_{**}^1, \dots)$). This contradicts Corollary 2.

Case 2.3: Suppose $s_*^r < \bar{y}$ for some $r > 0$ but never for $r < 0$. Take \bar{r} minimal, so $s_*^{\bar{r}} < \bar{y}$ and $s_*^r = \bar{y} \forall r < \bar{r}$. Fix $0 < \nu < \bar{y} - s_*^{\bar{r}}$. Let $r'_\nu(i)$ be such that $s_{g_i+r'_\nu(i)}$ is the last element before s_{g_i} that is weakly smaller than $\bar{y} - \nu$. Clearly $r'_\nu(i) \xrightarrow{i \rightarrow \infty} -\infty$.

Consider the choice made by $I(s_{g_i+\bar{r}-1})$. In the limit, a majority in $I(\bar{y})$ prefers $(s_*^{\bar{r}}, s_*^{\bar{r}+1}, \dots)$ over \bar{y} . Apply Corollary 2. Clearly $\alpha_0 < \bar{y}$, so $m(\bar{y}) \leq \alpha_0 < \bar{y}$. As m is strictly increasing, $m^k(\bar{y})$ is strictly decreasing in k and converges to a limit \tilde{m} ; moreover, $m(y) < y$ for all $y \in (\tilde{m}, \bar{y}]$. Call $g_i+r'_\nu(i) = h_{i\nu}$ and let $s_*^\nu = \liminf_{i \rightarrow \infty} s_{h_{i\nu}}$. Let $s_{**} = \liminf_{\nu \rightarrow 0} s_*^\nu$. If $s_{**} < \bar{y}$, take a sequence of $\nu, h_{i\nu}$ such that $s_{h_{i\nu}} \rightarrow s_{**}$. By construction $s_{h_{i\nu}+\nu} \geq \bar{y} - \nu$ for $l = 1, \dots, L$ for L arbitrarily large as $\nu \rightarrow 0$, $h_{i\nu} \rightarrow \infty$. Then, in the limit, \bar{y} is a Condorcet winner in $I(s_{**})$; in particular, a majority prefers \bar{y} to $(s_*^{\bar{r}}, s_*^{\bar{r}+1}, \dots)$, which contradicts Corollary 2.

If $s_{**} = \bar{y}$ we must work away from the limit. Take $\varepsilon > 0$ such that $(\bar{y} - d_{\bar{y}}^- + \varepsilon, \bar{y} - \varepsilon) \subseteq I(\bar{y} - \nu)$ is a strict majority of $I(\bar{y} - \nu)$ for all $0 < \nu \leq \varepsilon$.³² Take a fixed $\nu' < \varepsilon$; a $\nu < \nu'$ such that $s_*^\nu \geq \bar{y} - \nu'$; and a subsequence s_{f_i} of $s_{h_{i\nu}}$ such that $s_{f_i} \rightarrow s_*^\nu$. Let M_i be the largest integer such that $s_{f_i+l} \in (\bar{y} - \nu, \bar{y})$ for $l = 1, \dots, M_i$ and K_i the set of $l \in 1, \dots, M_i$ such that $s_{f_i+l} \in (\bar{y} - \frac{\nu}{2}, \bar{y})$. Let $k_i = \min(K_i)$. By construction, $M_i, |K_i| \rightarrow \infty$. Then, for $\alpha \in (\bar{y} - d_{\bar{y}}^- + \varepsilon, \bar{y} - \varepsilon)$,

$$\begin{aligned} & \frac{1}{1-\delta} u_\alpha(s_{f_i}) - U_\alpha(S(s_{f_i+1})) = \sum_{t \in K_i} \delta^{t-1} (u_\alpha(s_{f_i}) - u_\alpha(s_{f_i+t})) + \\ & + \sum_{M_i \geq t \notin K_i} \delta^{t-1} (u_\alpha(s_{f_i}) - u_\alpha(s_{f_i+t})) + \sum_{t > M_i} \delta^{t-1} \left(u_\alpha(s_{f_i}) - \mathbf{1}_{\alpha \in I(s_{f_i+t})} u_\alpha(s_{f_i+t}) \right) \\ & \geq \delta^{k_i-1} (u_\alpha(s_{f_i}) - u_\alpha(s_{f_i+k_i})) + 0 - \frac{\delta^{M_i}}{1-\delta} C = \delta^{k_i-1} \left(u_\alpha(s_{f_i}) - u_\alpha(s_{f_i+k_i}) - \frac{\delta^{|K_i|}}{1-\delta} C \right) \end{aligned}$$

where $C = \max_\alpha u_\alpha(\alpha)$. Note that $u_\alpha(s_{f_i}) - u_\alpha(s_{f_i+k_i}) \geq u_{s_{f_i}}(s_{f_i}) - u_{s_{f_i}}(s_{f_i+k_i}) +$

³² ε exists because $m(\bar{y}) < \bar{y}$ and $y - d_y^-, y + d_y^+$ are continuous in y .

$M'(s_{f_i} - \alpha)(s_{f_i+k_i} - s_{f_i}) \geq M'(s_{f_i} - \bar{y} + \varepsilon)\frac{\nu'}{2}$, which converges to $M'(s_*^\nu - \bar{y} + \varepsilon)\frac{\nu'}{2} \geq M'\frac{\nu'}{2}(\varepsilon - \nu') > 0$ as $i \rightarrow \infty$. On the other hand, $\delta^{|K_i|} \rightarrow 0$ as $i \rightarrow \infty$. Hence all $\alpha \in (\bar{y} - d_{\bar{y}}^- + \varepsilon, \bar{y} - \varepsilon)$ prefer s_{f_i} to $S(s_{f_i+1})$ for high i , so $S(s_{f_i+1})$ is not a Condorcet winner in $I(s_{f_i})$, a contradiction.

Case 2.4: $s_*^r = \bar{y}$ for all r . In other words, the sequence spends arbitrarily long times near \underline{y} and \bar{y} (if not true for both boundaries, one of the former cases applies). We first prove the following claim: $m(y) = y$ for all $y \in [\underline{y}, \bar{y}]$.

Take any $y_0 \in (\underline{y}, \bar{y})$. Take a sequence $(h_i)_i$ such that, for each i , s_{h_i} is the last element of the sequence $(s_k)_k$ before s_{g_i} such that $s_k \leq y_0$. Intuitively, s_{h_i} is the last element of the sequence below y_0 before the sequence goes near \bar{y} for a long time. Take a diagonal subsequence (s_{k_i}) of (s_{h_i}) such that s_{k_i+l} has a limit s_{**}^l for all i . Clearly $s_{**}^0 \leq y_0$ and $s_{**}^l \geq y_0$ for all $l > 0$.

Consider the choice made by $I(s_{k_i})$. If $s_{**}^0 < y_0$, in the limit, a majority in $I(s_{**}^0)$ prefers $(s_{**}^1, s_{**}^2, \dots)$ over s_{**}^0 . Apply Corollary 2. Clearly $m(s_{**}^0) \geq \alpha_0 > s_{**}^0$, so $u_{m(s_{**}^0)}(s_{**}^0) \leq u_{m(s_{**}^0)}(y_0)$. If $s_{**}^0 = y_0$, then $m(y_0) < y_0$ leads to a contradiction by an analogous argument as in Case 2.3, so we must have $m(y_0) \geq y_0$. Conversely, considering sequences going near \underline{y} for arbitrarily long, we obtain that either $m(y_0) \leq y_0$ or there is $\tilde{s}_{**}^0 > y_0$ such that $m(\tilde{s}_{**}^0) < \tilde{s}_{**}^0$ and $u_{m(\tilde{s}_{**}^0)}(\tilde{s}_{**}^0) \leq u_{m(\tilde{s}_{**}^0)}(y_0)$.

For each $y \in (\underline{y}, \bar{y})$ such that $y \neq m(y)$, define $\hat{y} \neq y$ to be such that $u_{m(y)}(y) = u_{m(y)}(\hat{y})$. Take y_0 such that $|y_0 - \hat{y}_0|$ is maximal. WLOG $m(y_0) < y_0$, so there is $s_{**}^0 < y_0$ such that $m(s_{**}^0) > s_{**}^0$ and $u_{m(s_{**}^0)}(s_{**}^0) \leq u_{m(s_{**}^0)}(y_0)$. Since $m(y_0) > m(s_{**}^0)$, $u_{m(y_0)}(s_{**}^0) < u_{m(y_0)}(y_0)$, so $\hat{y}_0 > s_{**}^0$; but $\hat{s}_{**}^0 \geq y_0$. Hence $|s_{**}^0 - \hat{s}_{**}^0| > |\hat{y}_0 - y_0|$, a contradiction.

For the case where $m(y) = y$ for all $y \in [\underline{y}, \bar{y}]$, we use the following Lemma.

Lemma 4. *Let $S = (y, y, \dots)$, and let $T = (x_n)_n \neq S$. If x and x' both prefer T to S , and $x < y < x'$, then $x \notin I(x')$ or $x' \notin I(x)$.*

Proof. Suppose for that $x \in I(x')$ and $x' \in I(x)$. It is enough to check the case where T is contained in $[x, x']$: if not, define a path $(\tilde{x}_n)_n$ by $\tilde{x}_n = \min(\max(x_n, x), x')$. Then $(\tilde{x}_n)_n$ is contained in $[x, x']$ and is weakly better for both x and x' than T .

By assumption, both x and x' derive non-negative utility from all elements of T . Let $\bar{x} = (1 - \delta) \sum_n \delta^n x_n$, and $T'' = (\bar{x}, \bar{x}, \dots)$. If $T'' \neq T$, both x and x' strictly prefer T'' to T by Jensen's inequality and A4. Hence they both strictly prefer \bar{x} to y , a contradiction. If $T'' = T$, $\bar{x} \neq y$ and both agents prefer \bar{x} to y , a contradiction. \square

Take $\epsilon > 0$, $\nu > 0$ small and $y_0 = \underline{y} + \epsilon$. Construct s_{k_i} as before. It follows from previous arguments that $s_{**}^0 = y_0$. For all i , a majority in $I(s_{k_i})$ must prefer $S(s_{k_i+1})$ over s_{k_i} . Since s_{k_i} strictly prefers s_{k_i} over $S(s_{k_i+1})$ and $s_{k_i} = m(s_{k_i})$, this can only happen if there are voters both above and below s_{k_i} who prefer $S(s_{k_i+1})$. Let $y'_i < s_{k_i} < y''_i$ be the closest voters to s_{k_i} who weakly prefer $S(s_{k_i+1})$, and denote $y'_i - (s_{k_i} - d_{s_{k_i}}^-) = \eta'_i$, $y''_i - s_{k_i} = \eta''_i$. Note that $\eta'_i, \eta''_i \xrightarrow{i \rightarrow \infty} 0$.³³ In addition, $y''_i - d_{y''_i}^- > y'_i$; otherwise we obtain a contradiction as in Lemma 4. Let \tilde{y}_i be such that $\tilde{y}_i - d_{\tilde{y}_i}^- = y'_i$.

Given the path $T^i = S(s_{k_i+1})$ construct T'^i as follows. If $T_j^i \geq y''_i + \nu$, $T_j'^i = y''_i + \nu$. If $y''_i + \nu > T_j^i \geq \tilde{y}_i$, $T_j'^i = y''_i$. If $\tilde{y}_i > T_j^i \geq s_{k_i}$, $T_j'^i = z_i = \frac{\sum_{\tilde{y}_i > T_j^i \geq s_{k_i}} \delta^j T_j^i}{\sum_{\tilde{y}_i > T_j^i \geq s_{k_i}} \delta^j}$. If $s_{k_i} > T_j^i$, $T_j'^i = v_i = \frac{\sum_{s_{k_i} > T_j^i} \delta^j T_j^i}{\sum_{s_{k_i} > T_j^i} \delta^j}$. Then both y'_i and y''_i weakly prefer T'^i over T^i .

Moreover, T'^i is a linear combination of at most four policies; by an abuse of notation, $T'^i = \omega_1^i [y''_i] + \omega_2^i [y''_i + \nu] + \omega_3^i [z_i] + \omega_4^i [v_i]$ with $\sum_j \omega_j^i = 1$. In addition, since $S(s_{k_i+1})$ spends a long time near \bar{y} (hence above $y''_i + \nu$) before going back under s_{k_i} , $\frac{\omega_4^i}{\omega_2^i} \xrightarrow{i \rightarrow \infty} 0$.³⁴

Finally, take $0 < \omega_5^i \leq \omega_4^i$ such that $\omega_3^i z_i + \omega_5^i v_i = (\omega_3^i + \omega_5^i) s_{k_i}$,³⁵ and construct $T'''^i = \omega_1^i [y''_i] + \omega_2^i [y''_i + \nu] + (\omega_3^i + \omega_5^i) [s_{k_i}] + (\omega_4^i - \omega_5^i) [v_i]$, $T'''^i = w_1^i [y''_i] + w_2^i [y''_i + \nu] + w_3^i [v_i]$ where $w_1^i = \frac{\omega_1^i}{\omega_1^i + \omega_2^i + \omega_4^i - \omega_5^i}$, $w_2^i = \frac{\omega_2^i}{\omega_1^i + \omega_2^i + \omega_4^i - \omega_5^i}$, $w_3^i = \frac{\omega_4^i - \omega_5^i}{\omega_1^i + \omega_2^i + \omega_4^i - \omega_5^i}$ and $\frac{w_3^i}{w_2^i} \xrightarrow{i \rightarrow \infty} 0$. Then both y'_i and y''_i weakly prefer T'''^i over T^i and hence over s_{k_i} . Then, for some $C, c > 0$,

$$\begin{aligned} Cw_3^i &\geq w_3^i u_{y'_i}(v_i) \geq u_{y'_i}(s_{k_i}) = u_{y'_i}(s_{k_i}) - u_{y'_i - \eta'_i}(s_{k_i}) = \eta'_i \frac{\partial u_{\bar{\alpha}}(s_{k_i})}{\partial \alpha} \geq c\eta'_i \\ u_{y''_i}(s_{k_i}) &\leq w_1^i u_{y''_i}(y''_i) + w_2^i u_{y''_i}(y''_i + \nu) + w_3^i u_{y''_i}(v_i) \leq (w_1^i + w_2^i) u_{y''_i} \left(y''_i + \frac{w_2^i \nu}{w_1^i + w_2^i} \right) + w_3^i u_{y''_i}(s_{k_i}) \\ u_{y''_i}(y''_i - \eta''_i) &= u_{y''_i}(s_{k_i}) \leq u_{y''_i} \left(y''_i + \frac{w_2^i \nu}{w_1^i + w_2^i} \right) \leq u_{y''_i}(y''_i + w_2^i \nu) \end{aligned}$$

As A2 and A4 imply $u_\alpha(\alpha) - u_\alpha(\alpha - x) \in [\frac{M'}{2}x^2, \frac{M}{2}x^2]$, this means

$$\frac{M}{2}(\eta''_i)^2 \geq \frac{M'}{2}(w_2^i \nu)^2 \implies \frac{\eta''_i}{\eta'_i} \geq \sqrt{\frac{M'}{M} \frac{c}{C} \frac{w_2^i}{w_3^i} \nu}$$

Since (y'_i, y''_i) cannot be a strict majority in $I(s_{k_i})$, we must have $F(y''_i) - F(s_{k_i}) \leq$

³³ $\eta'_i \rightarrow 0$ by a similar argument to Case 2.3. Then, if η''_i did not converge to zero, (y'_i, y''_i) would be a strict majority in $I(s_{k_i})$ for large i , a contradiction.

³⁴For these arguments to work, we take ν, ϵ small enough that $y_0 + \nu < \bar{y}$ and $u_{y_0}(y) > 0$.

³⁵If this is not possible then y'_i could not have preferred T'^i over s_{k_i} , a contradiction.

$F(y'_i) - F(s_{k_i} - d_{s_{k_i}}^-)$ for all i . But this is impossible as $\frac{f(x)}{f(x')}$ is bounded and $\frac{\eta'_i}{\eta_i} \xrightarrow{i \rightarrow \infty} \infty$, a contradiction. \square

Proof of Proposition 2. Suppose $m(y) = y$ and $s(y) \neq y$; WLOG $s(y) < y$. A majority in $I(y)$ must prefer $S(s(y))$ to $S(y)$, i.e., they must prefer $S(s(y))$ to y . By Proposition 1, $s^k(y) \leq s(y)$ for all k . But then, for small enough $\epsilon > 0$, all agents in $(y - \epsilon, y + d_y^+)$ strictly prefer y to $S(s(y))$, a contradiction.

If $m(y) \neq y$, suppose WLOG that $m(y) < y$. If $s(y) > y$ then $s^k(y) \geq s(y) > y$ for all k , so all voters in $(y - d_y^-, y)$ (a strict majority in $I(y)$) strictly prefer y to $S(s(y))$, a contradiction. Hence $s(y) \leq y$. On the other hand, suppose $s(y) < m^*(y)$. Note that $m^*(y) < m(y)$; $m(m^*(y)) = m^*(y)$; $s^k(y) \leq s(y)$ for all k ; and choosing $m^*(y)$ leads to the policy path $(m^*(y), m^*(y), \dots)$ by the previous case. Then, for small enough $\epsilon > 0$, all voters in $(m(y) - \epsilon, y + d_y^+)$ prefer $S(m^*(y))$ over $S(s(y))$, a contradiction. Hence $s(y) \geq m^*(y)$. Next, suppose $s(y) = m^*(y)$ and consider $T = (m(y), s(m(y)), \dots)$. Since T is contained in $[m^*(y), m(y)]$ and $T_1 = m(y) > m^*(y)$, all voters in $(m(y) - \epsilon, y + d_y^+)$ for small $\epsilon > 0$ strictly prefer T over $S(s(y))$, a contradiction. Hence $s(y) > m^*(y)$.

We now show that, if the MVT holds on $[m^*(y), y]$, then $s(y) < y$. Suppose that $s(y) = y$. There must be ϵ_0 such that $s(y - \epsilon) < y - \epsilon$ for all $0 < \epsilon < \epsilon_0$ (otherwise, $m(y)$ would prefer the constant path $(y - \epsilon, y - \epsilon, \dots)$ to (y, y, \dots) for ϵ small enough).

Let $s_-(y) = \liminf_{\epsilon \rightarrow 0} s(y - \epsilon) \in [m^*(y), y]$. There are two cases: $s_-(y) = y$ and $s_-(y) < y$. If $s_-(y) = y$, then $s^k(x) \rightarrow y$ as $x \rightarrow y$ for all k . For all $x \in (y - \epsilon_0, y)$, $m(x)$ must prefer $S(s(x))$ to x . That is, denoting $W(x) = (1 - \delta)U_{m(x)}(S(s(x))) - u_{m(x)}(x)$, we must have $W(x) \geq 0$. Equivalently

$$(1 - \delta) \sum_{t=0}^{k_x} \delta^t u_{m(x)}(s^{t+1}(x)) - u_{m(x)}(x) \geq 0,$$

where $k_x = \max\{k : m(x) \in I(s^k(x))\} - 1$. (Note that $k_x \xrightarrow{x \rightarrow y} \infty$.) By the envelope theorem,

$$\begin{aligned} W'(x) &= (1 - \delta) \frac{\partial}{\partial \alpha} U_{m(x)}(S(s(x))) m'(x) - \frac{d}{dx} u_{m(x)}(x) = \\ &= \sum_{t=0}^{k_x} (1 - \delta) \delta^t \left(\frac{\partial}{\partial \alpha} u_{m(x)}(s^{t+1}(x)) - \frac{\partial u}{\partial \alpha} \right) m'(x) - (1 - \delta^{k_x+1}) \frac{\partial u}{\partial x} - \delta^{k_x+1} \frac{du}{dx} \end{aligned}$$

$$\begin{aligned} &\geq \sum_{t=0}^{k_x} \delta^t (-M(x - s^{t+1}(x))) m'(x) - (1 - \delta^{k_x+1}) \frac{\partial u}{\partial x} - \delta^{k_x+1} \frac{du}{dx} \\ &\xrightarrow{x \rightarrow y} -\frac{\partial}{\partial x} u_{m(y)}(y) > 0, \end{aligned}$$

where u stands for $u_{m(x)}(x)$ unless otherwise noted. Thus $W(x) \geq 0$ and $W'(x) > 0$ for all $x \in (y - \epsilon_1, y)$, whence $W(y) > 0$, which contradicts $s(y) = y$.

If $s_-(y) < y$, let $(y_n)_n$ be a sequence such that $y_n < y \forall n$, $y_n \rightarrow y$ and, for all t , $s^t(y_n)$ converges to a limit s_t as $n \rightarrow \infty$ (in particular $s_1 = s_-(y)$). By construction, $m(y)$ must prefer y to $S(s(y_n))$ for all n . We now aim to show that

$$\frac{U_{m(y)}(S(s(y_n))) - \frac{1}{1-\delta} u_{m(y)}(y)}{y - y_n} \xrightarrow{n \rightarrow \infty} 0.$$

$m(y_n)$ prefers $S(s(y_n))$ to y for all n . By continuity, $m(y)$ is indifferent between y and $(s_t)_t$. Moreover, $m(y_n)$ prefers $S(s(y_n))$ to all other $S(s(y_{n'}))$, hence to $(s_t)_t$. Thus

$$\begin{aligned} 0 &\geq U_{m(y)}(S(s(y_n))) - \frac{1}{1-\delta} u_{m(y)}(y) = U_{m(y)}(S(s(y_n))) - U_{m(y)}((s_t)_t) \geq \\ &\geq U_{m(y)}(S(s(y_n))) - U_{m(y)}((s_t)_t) + U_{m(y_n)}((s_t)_t) - U_{m(y_n)}(S(s(y_n))) = \sum_{t=0}^{\infty} \delta^t A_{tn}, \end{aligned}$$

where, denoting $v_\alpha(x) = \max(u_\alpha(x), 0)$,

$$A_{tn} = v_{m(y)}(s^{t+1}(y_n)) - v_{m(y)}(s_{t+1}) + v_{m(y_n)}(s_{t+1}) - v_{m(y_n)}(s^{t+1}(y_n)).$$

Let $B_{tn} = \frac{A_{tn}}{y - y_n}$. Then it is sufficient to show that B_{tn} is uniformly bounded (that is, $\exists \bar{B}$ such that $|B_{tn}| \leq \bar{B}$ for all t, n) and that, for all t , $\liminf_{n \rightarrow \infty} B_{tn} \geq 0$.

We first prove the boundedness. Using that $|\max(a, 0) - \max(b, 0)| \leq |a - b|$,

$$\begin{aligned} A_{tn} &\leq |A_{tn}| \leq |u_{m(y)}(s^{t+1}(y_n)) - u_{m(y)}(s^{t+1}(y_n))| + |u_{m(y_n)}(s_{t+1}) - u_{m(y)}(s_{t+1})| \leq \\ &\leq 2\bar{m}' \max_{\alpha, x} \left[\frac{\partial u_\alpha(x)}{\partial \alpha} \right] (y - y_n), \end{aligned}$$

where $\bar{m}' = m'(x)$. Next, we prove that $\liminf_{n \rightarrow \infty} B_{tn} \geq 0$. There are four cases. First, if $u_{m(y)}(s_{t+1}) > 0$, then there is n_0 such that for all $n \geq n_0$ $u_{m(y)}(s^{t+1}(y_n))$, $u_{m(y_n)}(s_{t+1})$, $u_{m(y_n)}(s^{t+1}(y_n)) > 0$. For all such n , by A2, there is $\tilde{M}_{tn} \in [M', M]$

such that $A_{tn} = \tilde{M}_{tn}(s^{t+1}(y_n) - s_{t+1})(m(y) - m(y_n))$, so $|B_{tn}| \xrightarrow{n \rightarrow \infty} 0$. Second, if $u_{m(y)}(s_{t+1}) < 0$, then for all large enough n $A_{tn} = 0$. Third, if $u_{m(y)}(s_{t+1}) = 0$ and $s^{t+1}(y_n) \geq s_{t+1}$, all the terms are positive and we use the same argument as in case 1. Fourth, if $u_{m(y)}(s_{t+1}) = 0$ and $s^{t+1}(y_n) < s_{t+1}$, then $u_{m(y_n)}(s_{t+1}) > u_{m(y_n)}(s^{t+1}(y_n))$ and $v_{m(y)}(s_{t+1}) = v_{m(y)}(s^{t+1}(y_n)) = 0$, so $B_{nt} \geq 0$.

Consider now the possibility of $m(y)$ choosing $S(y_n)$ instead. We can see that

$$\begin{aligned} (1 - \delta)U_{m(y)}(S(y_n)) - u_{m(y)}(y) &= \\ &= (1 - \delta)(u_{m(y)}(y_n) - u_{m(y)}(y)) + \delta((1 - \delta)U_{m(y)}(S(y_n)) - u_{m(y)}(y)) \\ &= (1 - \delta)(y_n - y) \frac{\partial u_{m(y)}(\tilde{y})}{\partial x} + o(y - y_n) \end{aligned}$$

for some $\tilde{y} \in [y_n, y]$. Since $\frac{\partial u_{m(y)}(\tilde{y})}{\partial x} \xrightarrow{n \rightarrow \infty} \frac{\partial u_{m(y)}(y)}{\partial x} < 0$, the above expression is positive for high n , so $s(y) = y$ is not optimal for $m(y)$, a contradiction.

Finally, we show that $s^k(y)$ converges to $m^*(y)$. Since $s^k(y) \in [m^*(y), y]$ for all y and the sequence $(s^k(y))_k$ is decreasing, it has a limit $s^* \in [m^*(y), y]$. Suppose $s^* > m^*(y)$. Then $m(s^*) < s^*$, so there is k_0 such that $m(s^k(y)) < s^*$ for all $k \geq k_0$. For such k , $m(s^k(y))$ would strictly prefer $S(s^{k+2}(y))$ to $S(s^{k+1}(y))$, a contradiction. \square

Proof of Corollary 1. Let $x_i^* < x_{i+1}^*$ be consecutive fixed points of m . Since m is continuous, either $m(y) > y \forall y \in (x_i^*, x_{i+1}^*)$ or $m(y) < y$ for all such y . The first case implies $m'(x_i^*) \geq 1$ and $m'(x_{i+1}^*) \leq 1$, and vice versa; since $m'(x_j^*) \neq 1$, these inequalities are strict, which implies that the intervals must alternate.

A fixed point of m is stable if $m'(x^*) < 1$ and unstable if $m'(x^*) > 1$ (see e.g. Elaydi (2005), Chapter 1.5). Since $m(-1) > -1$ and $m(1) < 1$, x_1^* and x_n^* are both stable, and stable and unstable fixed points alternate in between. \square

Proof of Proposition 3. We first prove the monotonicity. Fix $\epsilon > 0$ small. Let $x < y \in [x^*, x^* + \epsilon]$, where $m(x^*) = x^*$, $m(x^{**}) = x^{**}$, $m(y) < y$ for all $y \in (x^*, x^{**})$ and $\epsilon < x^{**} - x^*$. Call $s^k(x) = x_k$, $s^k(y) = y_k$ and suppose $x_1 > y_1$. Then $S(x_1)$ is preferred to $S(y_1)$ by a majority in $I(x)$, and the opposite happens in $I(y)$. Since all agents in $I(y) - I(x) = (x + d_x^+, y + d_y^+]$ prefer $S(x_1)$ due to A6, there must also be $z_0 \in I(x) - I(y) = [x - d_x^-, y - d_y^-)$ that prefers $S(x_1)$ (in fact there must be enough of them, but we only need one).

Let l be such that $x_i > y_i$ for $i = 1, 2, \dots, l$ but not for $i = l+1$. If $x_{l+1} = y_{l+1}$ (and hence $S(x_{l+1}) = S(y_{l+1})$) we have a contradiction, as any $z_0 \in I(x) - I(y)$ would prefer

$S(y_1)$ to $S(x_1)$ pointwise. By a similar argument, there must be $z_l \in I(y_l) - I(x_l)$ that prefers $S(y_{l+1})$ to $S(x_{l+1})$. If $x_i < y_i$ for all $i \geq l + 1$ this also yields a contradiction, as any such z_l would prefer $S(x_{l+1})$ over $S(y_{l+1})$ pointwise. More generally, if the ordering between x_l and y_l only changes a finite number of times, we can obtain a contradiction by looking at the last change. The only case left to consider is if there are arbitrarily high i 's and j 's for which $x_i > y_i$ and $x_j < y_j$. If so, note that

$$0 \leq U_{z_0}(S(x_1)) - U_{z_0}(S(y_1)) = \sum_{t=0}^{l-1} \delta^t (u_{z_0}(x_{1+t}) - u_{z_0}(y_{1+t})) + \sum_{t \geq l} \delta^t (u_{z_0}(x_{1+t}) - u_{z_0}(y_{1+t}))$$

$$- \frac{\partial u_{z_0}(x)}{\partial x} \Big|_{x^*} \sum_{t=0}^{l-1} \delta^t (x_{1+t} - y_{1+t}) \leq \sum_{t=0}^{l-1} \delta^t (u_{z_0}(y_{1+t}) - u_{z_0}(x_{1+t})) \leq \sum_{t \geq l} \delta^t (u_{z_0}(x_{1+t}) - u_{z_0}(y_{1+t}))$$

(Note that z_0 gets positive utility from all policies in $S(x_1)$ and $S(y_1)$.) Then

$$0 \leq U_{z_l}(S(y_{l+1})) - U_{z_l}(S(x_{l+1})) = \sum_{t \geq l} \delta^{t-l} (u_{z_l}(y_{1+t}) - u_{z_l}(x_{1+t}))$$

$$- \frac{\partial u_{z_0}(x)}{\partial x} \Big|_{x^*} \sum_{t=0}^{l-1} \delta^t (x_{1+t} - y_{1+t}) \leq \sum_{t \geq l} \delta^t (u_{z_0}(x_{1+t}) - u_{z_l}(x_{1+t}) - u_{z_0}(y_{1+t}) + u_{z_l}(y_{1+t}))$$

$$- \frac{\partial u_{z_0}(x)}{\partial x} \Big|_{x^*} \sum_{t=0}^{l-1} \delta^t (x_{1+t} - y_{1+t}) \leq \sum_{t \geq l} \delta^t M(z_0 - z_l) |x_{1+t} - y_{1+t}|$$

$$- \frac{\partial u_{z_0}(x)}{\partial x} \Big|_{x^*} \max_{0 \leq t \leq l-1} \{|x_{1+t} - y_{1+t}|\} \leq \frac{1}{1-\delta} M(z_0 - z_l) \sup_{t \geq l} \{|x_{1+t} - y_{1+t}|\}$$

$$\left(- \frac{\partial u_{z_0}(x)}{\partial x} \Big|_{x^*} \right) \frac{1-\delta}{M(z_0 - z_l)} \max_{0 \leq t \leq l-1} \{|x_{1+t} - y_{1+t}|\} \leq \sup_{t \geq l} \{|x_{1+t} - y_{1+t}|\}$$

Since $z_0 - z_l \leq \epsilon - d_{x^*+\epsilon}^- + d_{x^*}^- \xrightarrow{\epsilon \rightarrow 0} 0$, by taking ϵ small enough, we can guarantee

$$D \max_{0 \leq t \leq l-1} \{|x_{1+t} - y_{1+t}|\} \leq \sup_{t \geq l} \{|x_{1+t} - y_{1+t}|\}$$

for a fixed $D > 2$ (we can take D arbitrarily large). Take $t_0 = \arg \max_{0 \leq t \leq l-1} \{|x_{1+t} - y_{1+t}|\}$ and t_1 the smallest $t \geq l$ for which $|x_{1+t_1} - y_{1+t_1}| \geq 2|x_{1+t_0} - y_{1+t_0}|$. We can apply the same argument to obtain t_2 such that $|x_{1+t_2} - y_{1+t_2}| \geq 2|x_{1+t_1} - y_{1+t_1}|$, and so on for t_3 , etc. Then, for large enough j , $|x_{1+t_j} - y_{1+t_j}| > x^{**} - x^*$, a contradiction.

This argument proves (i) for an interval $[x^*, x^* + \epsilon)$. Now let

$$\hat{x} = \inf\{\tilde{x} : s \text{ is not monotonic on } [x^*, \tilde{x}]\} \geq x^* + \epsilon.$$

Suppose WLOG that $x^{**} > m^{-1}(x^* + d_{x^*}^+)$. We will now show that $\hat{x} \geq m^{-1}(x^* + d_{x^*}^+) > x^* + d_{x^*}^+$. Suppose $\hat{x} < m^{-1}(x^* + d_{x^*}^+)$.

By construction, for any $\varepsilon > 0$, there must be pairs $x^\varepsilon, y^\varepsilon$ such that $x^\varepsilon < y^\varepsilon$, $s(x^\varepsilon) > s(y^\varepsilon)$ and $x^\varepsilon, y^\varepsilon \in (\hat{x} - \varepsilon, \hat{x} + \varepsilon)$. There are two cases. If there are arbitrarily small ε for which $s^k(x^\varepsilon), s^k(y^\varepsilon) \geq \hat{x}$ for all k , and moreover $\lim_{k \rightarrow \infty} s^k(x^\varepsilon) = \lim_{k \rightarrow \infty} s^k(y^\varepsilon) \geq \hat{x}$, then we obtain a contradiction by repeating our previous argument. Else we are able to pick x^ε and y^ε so that, in the addition to the above conditions, $s^k(x^\varepsilon) > s^k(y^\varepsilon)$ for all $k \geq 1$.³⁶ Label $s^k(x^\varepsilon) = x_k$, $s^k(y^\varepsilon) = y_k$.

A majority in $I(y_0)$ must prefer $S(y_1)$ to $S(x_1)$, and a majority in $I(x_0)$ must prefer $S(x_1)$ to $S(y_1)$. Then some $z_0 \in I(x_0) - I(y_0)$ prefers $S(x_1)$ to $S(y_1)$. A2 implies that all $z \in [z_0, x^* + d_{x^*}^+]$ strictly prefer $S(x_1)$ to $S(y_1)$. For ϵ small enough, $m(I(y_0)) < x^* + d_{x^*}^+$, so a strict majority in $I(y_0)$ prefers $S(x_1)$ to $S(y_1)$, a contradiction.

Next, we prove that the MVT must hold. Let $y \in I(x^*) \cap [x^*, x^{**})$ and suppose $m(y)$ strictly prefers $S(y')$ to $S(s(y))$, where $y' < s(y)$. Since s is increasing in $I(x^*) \cap [x^*, x^{**})$, $s^k(y') \leq s^k(s(y))$ for all k , so by A2 all voters $x < m(y)$ prefer $S(y')$ to $S(s(y))$. Some voters $x > m(y)$ close to $m(y)$ also prefer $S(y')$ by continuity. Hence $S(s(y))$ is not a Condorcet winner in $I(y)$, a contradiction. Next, suppose $s(y) < y' \leq y$. Then all voters in $[m(y), x^* + d_{x^*}^+]$ prefer $S(y')$ by A2, and some voters $x < m(y)$ prefer $S(y')$ by continuity. On the other hand, voters $x \in (x^* + d_{x^*}^+, y + d_y^+]$ prefer $S(y')$ to $S(s(y))$ because $x \geq y$ and $s^k(y') \geq s^k(s(y))$ for all k . Hence $s(y)$ is not a Condorcet winner in $I(y)$, a contradiction.

For the existence, it is enough to prove existence of MVE for the model in Section 5.³⁷ For this, we refine an incomplete argument given in Acemoglu et al. (2015) (Theorem B3). Briefly, for any finite policy space $X \subseteq [-1, 1]$, an MVE can be found by backward induction, and its monotonicity can be proved as above. Take a sequence of finite spaces $X_1 \subseteq X_2 \subseteq \dots$ such that $\cup_{i \in \mathbb{N}} X_i$ is dense in $[-1, 1]$, and take an MVE

³⁶If eventually $s^k(x^\varepsilon) < \hat{x}$, or $s^k(x^\varepsilon)$ and $s^k(y^\varepsilon)$ converge to different limits, the inequality $s^k(x^\varepsilon) \leq s^k(y^\varepsilon)$ can only flip finitely many times as k grows, and we can look at the last time it flips.

³⁷Any such MVE is also an MVE of the main model within $I(x^*) \cap [x^*, x^{**})$. The reason is that any such MVE is monotonic and satisfies the MVT everywhere, i.e., $\forall y, y', m(y)$ prefers $S(s(y))$ to $S(y')$. $m(y)$ then also prefers $S(s(y))$ to all $S(y')$ in the main model if $y \in I(x^*)$. By the monotonicity, and combining A2 and A6, the sets of voters in $I(y)$ preferring $S(s(y))$ to $S(y')$ or vice versa are both intervals, so a majority prefers $S(s(y))$ iff $m(y)$ does.

$\tilde{s}_i : X_i \rightarrow X_i$ for each. Let s_i be a monotonic extension of \tilde{s}_i to $[-1, 1]$. By a diagonal argument, abusing notation, find a subsequence $(s_j)_j$ such that $(s_j(x))_j$ converges at every element of $\cup_{i \in \mathbb{N}} X_i$. $(s_j(x))_j$ must in fact converge at all but countably many points, so we can find a subsequence that converges for all x . Denote the limit by \tilde{s} . This is the construction from Acemoglu et al. (2015). But \tilde{s} **need not** be an MVE, as there is no guarantee that $\tilde{S}(\tilde{s}(y)) = \lim_{j \rightarrow \infty} S_j(s_j(y))$ if \tilde{s} is not continuous.

Say $S = (x_t)_t$ is an *optimal path* for y if there is a sequence $y_j \rightarrow y$ such that $S_j(s_j(y_j)) \rightarrow S$ (i.e., $s_j^t(y_j) \xrightarrow{j \rightarrow \infty} x_t \forall t$). Denote by $\mathcal{S}(y)$ the set of y 's optimal paths. Then it can be shown that the elements of $\mathcal{S}(y)$ are ordered for each y ; if $S \in \mathcal{S}(y)$ and $S' \in \mathcal{S}(y')$, with $y > y'$, then $S \geq S'$; and, if $S_j \in \mathcal{S}(y_j) \forall j$, $y_j \rightarrow y$, and $S_j \rightarrow S$, then $S \in \mathcal{S}(y)$. In addition, for any $(x_0, x_1, \dots) \in \mathcal{S}(y)$, $(x_1, \dots) \in \mathcal{S}(x_0)$. Moreover, (x_1, \dots) must be the maximal element of $\mathcal{S}(x_0)$, that is, it must be $\lim_{y_j \searrow x_0} \min(\mathcal{S}(y_j))$. Indeed, $m(x_0)$ is indifferent between all elements of $\mathcal{S}(x_0)$; by A2, $m(y)$ strictly prefers the maximal one. Then, if (x_1, \dots) is not the maximal element, $m(y)$ would deviate to $x_0 + \epsilon$ for $\epsilon > 0$.

Define then s by $s(y) = \inf_{y' > y} \tilde{s}(y')$. We can show by induction on t that $S(s(y)) = \max(\mathcal{S}(y))$, and from there that s constitutes an MVE. \square

Proof of Proposition 4. Let $t_0 = \min\{t : s^t(x) \leq y\}$. By Proposition 1, $s^t(x) \leq s^{t_0}(x) \leq y$ for all $t \geq t_0$. A majority of $I(x)$ must prefer $S(s(x))$ to $S(x)$. Then, by Corollary 2, $m(x)$ must have this preference:

$$\begin{aligned} \frac{u_{m(x)}(x)}{1 - \delta} &\leq U_{m(x)}(S(s(x))) \leq \frac{1 - \delta^{t_0}}{1 - \delta} u_{m(x)}(m(x)) + \frac{\delta^{t_0}}{1 - \delta} \max(u_{m(x)}(y), 0) \\ \delta^{t_0} &\leq \frac{u_{m(x)}(m(x)) - u_{m(x)}(x)}{u_{m(x)}(m(x)) - \max(u_{m(x)}(y), 0)} \end{aligned}$$

$$t_0 \frac{1 - \delta}{\delta} \geq t_0 \ln \left(\frac{1}{\delta} \right) \geq \ln \left(\frac{u_{m(x)}(m(x)) - \max(u_{m(x)}(y), 0)}{u_{m(x)}(m(x)) - u_{m(x)}(x)} \right) =: K(y). \quad \square$$

Proof of Proposition 5. Parts (i) and (iii) follow from the arguments given in the text plus Lemma 2.

For part (ii), we first construct a sequence of approximate Q1Es as follows. For each $i = 1, 2, \dots$, let $\epsilon(i) = \frac{1}{i}$ and take y_1, y_2 such that $x^* < y_1 < y_2 < x^* + \epsilon(i)$ and such that, moreover, $u_{m(y_2)}(y_2) < u_{m(y_2)}(y_1)$. Define $\tilde{x}_{ik} = y_1$ for all $k > 0$ and

$\tilde{x}_{i0} = y_2$. Then, for $k = -1, -2, \dots$ define \tilde{x}_{ik} such that $m(\tilde{x}_{ik})$ is indifferent between the policy $\tilde{x}_{i(k+1)}$ and the path $(\tilde{x}_{i(k+2)}, \tilde{x}_{i(k+3)}, \dots)$. (We can show by induction that \tilde{x}_{ik} is uniquely defined and strictly decreasing in k for all $k < 0$, by Corollary 2.) Let \tilde{s}_i denote the associated successor function, i.e., $\tilde{s}_i(y) = \tilde{x}_{i(k+1)}$ for all $y \in [\tilde{x}_{ik}, \tilde{x}_{i(k-1)})$.

We now make some useful observations. First, \tilde{s}_i satisfies all the conditions to be a Q1E for $k < 0$. Indeed, $m(\tilde{x}_{ik})$ is indifferent between $S(\tilde{x}_{i(k+1)})$ and $S(\tilde{x}_{i(k+2)})$ by construction; by A2 and Corollary 2, she prefers these policy paths to any other $S(\tilde{x}_{ik'})$. Second, it can be shown by induction that $\tilde{x}_{ik}(y_1, y_2)$ is a continuous function for all $k < 0$. Third, $\tilde{x}_{ik} \leq m^{-1}(\tilde{x}_{i(k+1)})$ for all $k < 0$; in particular, $\tilde{x}_{ik} \leq m^k(\tilde{x}_{i0}) = m^k(y_2)$. Fourth, $\tilde{x}_{ik} \xrightarrow[k \rightarrow -\infty]{} x^{**}$.³⁸

Next, we argue that y_1, y_2 can be chosen so that some element of the sequence $(\tilde{x}_{ik})_k$ equals \underline{x} . For an arbitrary initial choice of y_1, y_2 satisfying the requirements above, let k_0 be such that $\underline{x} > \tilde{x}_{ik_0}$. Now lower (y_1, y_2) continuously towards x^* while satisfying the conditions that $x^* < y_1 < y_2 < x^* + \epsilon(i)$ and $u_{m(y_2)}(y_2) < u_{m(y_2)}(y_1)$. Then $\tilde{x}_{ik_0}(y_1, y_2) \leq m^k(y_2) \xrightarrow[y_2 \rightarrow x^*]{} x^*$, so there are intermediate values of y_1, y_2 for which $\tilde{x}_{ik_0}(y_1, y_2) = \underline{x}$. Denote $y_{i1} = y_1, y_{i2} = y_2, x_{ik} = \tilde{x}_{i(k+k_0)}(y_{i1}, y_{i2}), s_i = \tilde{s}_i(y_{i1}, y_{i2})$. Note that $x_{i0} = \underline{x}$ for all i .

We now construct a true Q1E s by taking the limit of a subsequence of s_i . We use a diagonal argument: $x_{i0} \rightarrow x_0 = \underline{x}$ by construction. For all i, x_{i1} is contained in $[x^*, \underline{x}]$, so we can take a convergent subsequence such that $x_{i_1,1} \rightarrow x_1$. Next, we take a subsequence such that the $x_{i_1,2}$ also converge, etc. By an abuse of notation, let x_{jk} denote the result of this argument, so that $x_{jk} \rightarrow x_k$ for all k .

The indifference conditions that made the s_i Q1Es under m_i make s a Q1E under m by continuity. To guarantee that s is a proper Q1E, we must also show that $x_k > x_{k+1}$ for all k ; $x_k \xrightarrow[k \rightarrow +\infty]{} x^*$; and $x_k \xrightarrow[k \rightarrow -\infty]{} x^{**}$.

For all these claims it is enough to show that there cannot be two sequences $x_{ik(i)}, x_{ik'(i)}$ such that $k(i) < k'(i)$ for all i but $\lim_{i \rightarrow \infty} x_{ik(i)} = \lim_{i \rightarrow \infty} x_{ik'(i)} \in (x^*, x^{**})$. In turn, it is enough to show that this cannot happen for $k'(i) = k(i) + 1$. Suppose it does, and relabel the sequences as follows: $y_{il} = x_{i(l+k(i))}$. (If necessary, take a subsequence such that $y_{il} \rightarrow y_l$ for all l .) Then we just have to show that $y_0 > y_1$. Clearly $y_0 \geq y_1$ as $y_{i0} > y_{i1}$ for all i , so suppose $y_0 = y_1$. If $y_2 > y_1$, then $m(y_0) = m(y_1)$ must be indifferent between $y_1, S(y_2)$ and y_2 , which implies $y_2 < m(y_1) < y_1$. But then $m(y_{1i})$

³⁸If $\tilde{x}_{ik} \xrightarrow[k \rightarrow -\infty]{} y < x^{**}$, we obtain a contradiction by the same argument as in Proposition 2.

would strictly prefer y_{i2} to $S(y_{i2})$ for high enough i , a contradiction. Hence $y_1 = y_2$, and by the same argument $y_2 = y_3 = y_4 = \dots$

This will lead to a contradiction by a similar argument as in Proposition 2. Let $V(y) = (1 - \delta)U_{m(y)}(S(s(y))) - u_{m(y)}(y)$ as in that proof. The fact that $y_1 = y_0$ implies that $V_i(y_{i0}) \rightarrow 0$. Now take an arbitrary sequence $(g(i))_i \subseteq \mathbb{N}$, and denote $y_{ig(i)} = y_{i0} - \epsilon_i$. Then, by the argument in Proposition 2,

$$\begin{aligned} V_i(y_{i0}) &\geq (1 - \delta)\epsilon_i \left(-\frac{\partial}{\partial x} u_{m(y_{i0})}(\tilde{y}) \right) + \delta (V(y_{i0}) - M\epsilon_i(E(S(y_{i1})) - E(S(y_{i(g(i)+1)}))) \\ V_i(y_{i0}) &\geq \epsilon_i \left(-\frac{\partial}{\partial x} u_{m(y_{i0})}(\tilde{y}_i) - \frac{\delta}{1 - \delta} M(E(S(y_{i1})) - E(S(y_{i(g(i)+1)}))) \right) \end{aligned}$$

for some $\tilde{y}_i \in (y_{ig(i)}, y_{i0})$, where $E(S(y)) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t s^t(y)$. Given some $0 < \epsilon' < \epsilon$ and $i \in \mathbb{N}$, we say that $g(i) \in \mathbb{N}$ is ϵ', ϵ -valid if $\epsilon_i \in (\epsilon', \epsilon)$ and $E(S(y_{i1})) - E(S(y_{i(g(i)+1)})) \leq \frac{1-\delta}{2\delta} \frac{\frac{\partial}{\partial x} u_{m(y_{i0})}(y_{i0})}{M}$. Clearly, if there are $0 < \epsilon' < \epsilon$ with ϵ small enough, for which we can find valid $g(i)$ for arbitrarily high i , we obtain a contradiction, as $\liminf V_i(y_{i0}) > 0$. If not, then there must be a fixed $\epsilon > 0$ and a sequence $\epsilon'_i \rightarrow 0$ for which there are no ϵ'_i, ϵ -valid values of $g(i)$ for any $i \geq i_0$. If there are arbitrarily high values of i for which $(y_{ik})_i \cap (y_{i0} - \epsilon, y_{i0} - \epsilon'_i)$ is empty, then let $y_{ih(i)}$ be the last element to the right of this gap, i.e., $y_{i(h(i)+1)} < y_{i0} - \epsilon, y_{i0} - \epsilon'_i < y_{ih(i)}$ and relabel the sequence so that $z_{i0} = y_{ih(i)}$. Then $z_1 < z_0 = y_0 = z_{-1}$, which leads to a contradiction by our previous argument. If there are arbitrarily high values of i for which there is $g(i)$ such that $\epsilon_i \in (\epsilon'_i, \epsilon)$, but $E(S(y_{i1})) - E(S(y_{i(g(i)+1)})) > C$ for a fixed C , this implies that there are fixed C' and k_0 such that $y_{i0} - y_{i(g(i)+k_0)} > C'$, and hence $V_i(y_{i(g(i)+k)}) \geq C''$ for some $0 < k \leq k_0$. Note that k_0, C, C' and C'' are fixed even as we take ϵ to 0, which implies that $\frac{\partial V_i(y)}{\partial y}$ must become arbitrarily large and negative as $i \rightarrow \infty$, a contradiction.

Next, we show that s is a 1E in $[x^*, x^* + d_{x^*}^-]$ iff $m(x_n) < x_{n+2}$ for all n .

Note that $m(x_n) < x_{n+1}$ always holds (otherwise $m(x_n)$ would strictly prefer x_{n+1} to $S(x_{n+2})$). If $m(x_n) > x_{n+2}$, $m(x_n)$ prefers $m(x_n)$ to x_{n+2} ; hence he prefers $S(m(x_n))$ to $S(x_{n+2})$, and hence to $S(x_{n+1})$. This implies that $S(x_n)$ cannot be a Condorcet winner in $I(x_n)$, as the MVT must hold in this interval by Proposition 3, and thus s is not a 1E.

Conversely, suppose that, for some $x \in [x_{n+1}, x_n)$, $I(x)$ prefers $S(y)$ to $S(x_{n+2})$ for some $y \in [x_k, x_{k-1})$. If $k \leq n + 2$, this is impossible as all agents in $[x - d_x^-, m(x) + \epsilon]$

would strictly prefer $S(x_{n+2})$ to $S(y)$. Suppose then that $k \geq n + 3$. By the MVT, $m(x)$ prefers $S(y)$ to $S(x_{n+2})$. Suppose $m(x) \in [x_b, x_{b-1})$; we will argue that $b = k$. If $b < k$, $m(x)$ prefers $S(x_{n+2})$ to $S(x_b)$ to $S(y)$, a contradiction. If $b > k$, $m(x)$ prefers $S(x_{n+2})$ to $S(x_k)$ to $S(y)$, a contradiction.

Next, note that, if indeed $m(x)$ prefers $S(y)$ to $S(x_{n+2})$, she then prefers $S(y)$ to $S(x_{k-1})$, and so do all agents z such that $y - d_y^- < z < m(x)$ by A2. Hence a majority in $I(x_{k-2})$ should prefer $S(y)$ to $S(x_{k-1})$. By the above argument, since $y \in [x_k, x_{k-1})$ it must be that $m(x_{k-2}) \in [x_k, x_{k-1})$, a contradiction. \square

Proof of Remark 1. This follows from $e^{-rt}u_\alpha(s(x, t)) = e^{-\tilde{r}\frac{rt}{\tilde{r}}}u_\alpha\left(s\left(x, \frac{\tilde{r}}{r}\frac{rt}{\tilde{r}}\right)\right)$. \square

Proof of Remark 2. If there are $x_1 < x_2 < x_3 \in (x - d, x + d)$ such that $f(x_1), f(x_3) < f(x_2)$, then there is a local maximum of f in $(x_1, x_3) \subseteq (x - d, x + d)$. Hence, if there is no local maximum, there must be $x^* \in (x - d, x + d)$ such that f is decreasing in $(x - d, x^*]$ and increasing in $[x^*, x + d)$. Suppose WLOG that $f(x - d) \leq f(x + d)$. By definition, $F(m(x)) - F(x - d) = \frac{F(x + d) - F(x - d)}{2}$; this implies $f(x)m'(x) = \frac{f(x + d) + f(x - d)}{2}$, given that $m(x) = x$. Since x is a stable steady state, $m'(x) < 1$, so $f(x) > \frac{f(x + d) + f(x - d)}{2} \geq f(x - d)$. Hence $x > x^*$. But then $f|_{(x - d, x)} \leq f(x) \leq f|_{(x, x + d)}$, where the first inequality is sometimes strict. Hence $F(x + d) - F(x) > F(x) - F(x - d)$, which contradicts $m(x) = x$. The other case is analogous. \square

Proof of Remark 3. There must be \hat{d} such that f is strictly increasing in $[x^* - \hat{d}, x^*]$ and strictly decreasing in $[x^*, x^* + \hat{d}]$. Take $\bar{d} = \frac{\hat{d}}{2}$. Then, for all $d < \bar{d}$, f is strictly increasing in $[x^* - 2d, x^*]$ and strictly decreasing in $[x^*, x^* + 2d]$, so $m(x^* - d) > x^* - d$ and $m(x^* + d) < x^* + d$. Hence there is a stable steady state in $[x^* - d, x^* + d]$. \square

Additional proofs and robustness checks are found in the online Appendices B-E.

References

- Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin**, “Coalition Formation in Non-Democracies,” *The Review of Economic Studies*, 2008, 75 (4), 987–1009.
- , – , and – , “Dynamics and Stability of Constitutions, Coalitions, and Clubs,” *American Economic Review*, 2012, 102 (4), 1446–1476.
- , – , and – , “Political Economy in a Changing World,” *Journal of Political Economy*, 2015, 123 (5), 1038–1086.
- Bai, Jinhui H and Roger Lagunoff**, “On the Faustian Dynamics of Policy and Political Power,” *The Review of Economic Studies*, 2011, 78 (1), 17–48.
- Benabou, Roland**, “Unequal Societies: Income Distribution and The Social Contract,” *American Economic Review*, 2000, 90 (1), 96–129.
- Bewley, Truman F**, “A Critique of Tiebout’s Theory of Local Public Expenditures,” *Econometrica: Journal of the Econometric Society*, 1981, pp. 713–740.
- Bouton, Laurent and Gabriele Gratton**, “Majority Runoff Elections: Strategic Voting and Duverger’s Hypothesis,” *Theoretical Economics*, 2015, 10 (2), 283–314.
- Cremer, Helmuth and Pierre Pestieau**, “Factor Mobility and Redistribution,” in “Handbook of Regional and Urban Economics,” Vol. 4, Elsevier, 2004, pp. 2529–2560.
- Elaydi, Saber**, *An Introduction to Difference Equations*, Springer, 2005.
- Epple, Dennis and Thomas Romer**, “Mobility and Redistribution,” *Journal of Political Economy*, 1991, pp. 828–858.
- Glaeser, Edward L and Andrei Shleifer**, “The Curley effect: The Economics of Shaping the Electorate,” *Journal of Law, Economics, and Organization*, 2005, 21 (1), 1–19.
- Grossman, Gene M**, “Union Wages, Temporary Layoffs, and Seniority,” *The American Economic Review*, 1983, 73 (3), 277–290.

- , “International Competition and the Unionized Sector,” *Canadian Journal of Economics*, 1984, 17 (3), 541–556.
- Hunt, Brian R, Tim Sauer, and James A Yorke**, “Prevalence: a translation-invariant ”almost every” on infinite-dimensional spaces,” *Bulletin of the American mathematical society*, 1992, 27 (2), 217–238.
- Jack, William and Roger Lagunoff**, “Dynamic Enfranchisement,” *Journal of Public Economics*, 2006, 90 (4-5), 551–572.
- Marcuse, Peter**, “Abandonment, Gentrification, and Displacement: The Linkages in New York City,” in “Gentrification of the City,” Routledge, 1986, pp. 169–193.
- Ortega, Francesc**, “Immigration Quotas and Skill Upgrading,” *Journal of Public Economics*, 2005, 89 (9-10), 1841–1863.
- Roberts, Kevin**, “Dynamic Voting in Clubs,” *Research in Economics*, 2015, 69 (3), 320–335.
- Suwankiri, Benjarong, Assaf Razin, and Efraim Sadka**, “The Welfare State and Migration: a Dynamic Analysis of Political Coalitions,” *Research in Economics*, 2016, 70 (1), 122–142.
- Tiebout, Charles M**, “A Pure Theory of Local Expenditures,” *Journal of Political Economy*, 1956, pp. 416–424.
- Vigdor, Jacob L**, “Is Urban Decay Bad? Is Urban Revitalization Bad Too?,” *Journal of Urban Economics*, 2010, 68 (3), 277–289.
- Wooders, Myrna Holtz**, “A Tiebout Theorem,” *Mathematical Social Sciences*, 1989, 18 (1), 33–55.

B Proofs (Continuous Time Limit) (For Online Publication)

In this Section we assume m is C^2 . We first define some useful objects. Say $s(x, t)$ is a policy mapping if $s(x, t + t') \equiv s(s(x, t'), t)$; $s(x, 0) \equiv x$; and s is weakly decreasing in t . Given a policy mapping $s(x, t)$, denote $V(\alpha) = U_\alpha(S(m^{-1}(\alpha))) - u_\alpha(m^{-1}(\alpha))$ and $W(x) = U_{m(x)}(S(x)) - u_{m(x)}(x)$, where $S(x) = (s(x, t))_t$. Given a policy path S , denote $V_\alpha(S) = U_\alpha(S) - u_\alpha(m^{-1}(\alpha))$.

Remark 4. If $s(x, t)$ is C^1 and decreasing in t , there are functions $d(x, y) : [x^*, x^{**}]^2 \rightarrow \mathbb{R}$, $e(z) : [x^*, x^{**}] \rightarrow \mathbb{R}_+$ such that $s(x, d(x, y)) = y$ and $d(x, y) = \int_y^x e(z) dz$.

$d(x, y)$ measures the time it takes the policy path to get from x to y , if $x > y$ (if $x < y$ then $d(x, y) = -d(y, x)$). This time can be expressed as an integral of the instantaneous delay $e(z)$ at each policy z .

We first show that, if a CLS exists, it solves Equation 2. We restate it here:

$$e(x) = \frac{1}{r} \left(-\frac{\frac{\partial^2 u}{\partial x^2}}{\frac{\partial u}{\partial x}} - 2\frac{\frac{\partial^2 u}{\partial \alpha \partial x}}{\frac{\partial u}{\partial x}} m'(x) - \frac{\frac{\partial^2 u}{\partial \alpha^2}}{\frac{\partial u}{\partial x}} m'(x)^2 + \frac{\frac{\partial^2 U_{m(x)}(S(x))}{\partial \alpha^2}}{\frac{\partial u}{\partial x}} m'(x)^2 + \frac{m''(x)}{m'(x)} \right). \quad (*)$$

Lemma 5. *Let $x_0 \in (x^*, x^{**})$. If a policy mapping $s(x, t)$ is such that $W(x) = 0$ for all x in a neighborhood of x_0 , then $d(x_0, x)$ is differentiable with respect to its second argument at (x_0, x_0) , and $e(x) = -\frac{\partial d(x_0, x)}{\partial x}$ is given by Equation (*).*

Proof. First, assume a C^1 policy mapping $s(x, t)$. Denote $n(\alpha) = m^{-1}(\alpha)$ and $\alpha_0 = m(x_0)$. By the envelope theorem,

$$\begin{aligned} V'(\alpha_0) &= \frac{\partial U_\alpha(S(x))}{\partial \alpha} \Big|_{\alpha_0, x_0} - \frac{du_\alpha(n(\alpha))}{d\alpha} \Big|_{\alpha_0} \\ V''(\alpha_0) &= \frac{\partial^2 U_\alpha(S(x))}{\partial \alpha^2} \Big|_{\alpha_0, x_0} + n'(\alpha_0) re(x_0) \left[\frac{\partial u_\alpha(x)}{\partial \alpha} \Big|_{\alpha_0, x_0} - \frac{\partial U_\alpha(S(x))}{\partial \alpha} \Big|_{\alpha_0, x_0} \right] \\ &\quad - \frac{\partial^2 u}{\partial \alpha^2} - 2n' \frac{\partial^2 u}{\partial \alpha \partial x} - n'^2 \frac{\partial^2 u}{\partial x^2} - n'' \frac{\partial u}{\partial x}. \end{aligned}$$

We can use the fact that $V(\alpha) \equiv 0$ in a neighborhood of α_0 , and hence $V' \equiv V'' \equiv 0$,

to determine $e(x)$:

$$\begin{aligned}
0 &= V'(\alpha_0) = \frac{\partial U_\alpha(S(x))}{\partial \alpha} \Big|_{\alpha_0, x_0} - \frac{\partial u_\alpha(x)}{\partial \alpha} \Big|_{\alpha_0, x_0} - n'(\alpha) \frac{\partial u_\alpha(x)}{\partial x} \Big|_{\alpha_0, x_0} \\
\implies re(x) \left[-\frac{1}{m'(x)} \frac{\partial u_{m(x)}(x)}{\partial x} \right] &= re(x) \left[\frac{\partial u_{m(x)}(x)}{\partial \alpha} - \frac{\partial U_{m(x)}(S(x))}{\partial \alpha} \right] = \\
&= -m'(x) \frac{\partial^2 U_{m(x)}(S(x))}{\partial \alpha^2} + m'(x) \frac{d^2 u_\alpha(n(\alpha))}{d\alpha^2} \Big|_{\alpha_0} = \\
&= -m'(x) \frac{\partial^2 U_{m(x)}(S(x))}{\partial \alpha^2} + \frac{1}{m'(x)} \frac{\partial^2}{\partial x^2} u + 2 \frac{\partial^2}{\partial \alpha \partial x} u + m'(x) \frac{\partial^2}{\partial \alpha^2} u - \frac{m''(x)}{m'(x)^2} \frac{\partial}{\partial x} u.
\end{aligned}$$

Now we show that d must be differentiable at (x_0, x_0) . Let $(z_n)_n$ be a sequence such that $z_n \rightarrow x_0$. WLOG assume $z_n < x_0$ for all n . Note that

$$\begin{aligned}
\frac{\partial U_{m(x_0)}(S(x_0))}{\partial \alpha} \Big|_{m(x_0)} - \frac{\partial U_{m(x_0)}(S(z_n))}{\partial \alpha} &= \int_0^\infty re^{-rt} \left[\frac{\partial u_{m(x_0)}(s(x_0, t))}{\partial \alpha} - \frac{\partial u_{m(x_0)}(s(z_n, t))}{\partial \alpha} \right] dt \\
&= \int_0^{d(x_0, z_n)} re^{-rt} \frac{\partial}{\partial \alpha} u_{m(x_0)}(s(x_0, t)) dt - (1 - re^{-rd(x_0, z_n)}) \frac{\partial U_{m(x_0)}(S(z_n))}{\partial \alpha} \\
&= (1 - re^{-rd(x_0, z_n)}) \left(\frac{\partial u_{m(x_0)}(\tilde{x})}{\partial \alpha} - \frac{\partial U_{m(x_0)}(S(z_n))}{\partial \alpha} \right)
\end{aligned}$$

for some $\tilde{x} \in (z_n, x_0)$. Then

$$\begin{aligned}
0 &= V'(m(x_0)) - V'(m(z_n)) = \\
&= \frac{\partial U_\alpha(S(x))}{\partial \alpha} \Big|_{m(x_0), x_0} - \frac{du_\alpha(n(\alpha))}{d\alpha} \Big|_{m(x_0)} - \left(\frac{\partial U_\alpha(S(x))}{\partial \alpha} \Big|_{m(z_n), z_n} - \frac{du_\alpha(n(\alpha))}{d\alpha} \Big|_{m(z_n)} \right) \\
&= \frac{\partial U_\alpha(S(x))}{\partial \alpha} \Big|_{m(x_0), x_0} - \frac{\partial U_\alpha(S(x))}{\partial \alpha} \Big|_{m(x_0), z_n} + \frac{\partial U_\alpha(S(x))}{\partial \alpha} \Big|_{m(x_0), z_n} - \frac{\partial U_\alpha(S(x))}{\partial \alpha} \Big|_{m(z_n), z_n} - \\
&\quad - (m(x_0) - m(z_n)) \frac{d^2 u_\alpha(n(\alpha))}{d\alpha^2} \Big|_{\tilde{\alpha}} \\
&= (1 - e^{-rd(x_0, z_n)}) \left(\frac{\partial u_{m(x_0)}(\tilde{x})}{\partial \alpha} - \frac{\partial U_{m(x_0)}(S(z_n))}{\partial \alpha} \right) + (m(x_0) - m(z_n)) \frac{\partial^2 U_\alpha(S(x))}{\partial \alpha^2} \Big|_{\hat{\alpha}, z_n} - \\
&\quad - (m(x_0) - m(z_n)) \frac{d^2 u_\alpha(n(\alpha))}{d\alpha^2} \Big|_{\tilde{\alpha}} \\
0 &= \frac{\partial^2 U_\alpha(S(x))}{\partial \alpha^2} \Big|_{\hat{\alpha}, z_n} + \frac{1 - e^{-rd(x_0, z_n)}}{m(x_0) - m(z_n)} \left(\frac{\partial u_{m(x_0)}(\tilde{x})}{\partial \alpha} - \frac{\partial U_{m(x_0)}(S(z_n))}{\partial \alpha} \right) - \frac{d^2 u_\alpha(n(\alpha))}{d\alpha^2} \Big|_{\tilde{\alpha}}.
\end{aligned}$$

This implies that $\lim_{n \rightarrow \infty} \frac{1 - e^{-rd(x_0, z_n)}}{m(x_0) - m(z_n)} = n'(m(x_0))re(x_0)$, which in turn implies that $\lim_{n \rightarrow \infty} \frac{d(x_0, z_n)}{x_0 - z_n} = e(x_0)$, as we wanted. \square

Lemma 6. Equation (*) has a unique solution, in the following sense: for any $x_1 > x_0 \geq x^*$ and given a candidate path $S(x_0)$, there is at most one way to choose $e : (x_0, x_1) \rightarrow \mathbb{R}_{\geq 0}$ so that Equation (*) holds for all $x \in (x_0, x_1)$.

Proof. Let $\tilde{g}(x) = re(x) \left[-\frac{1}{m'(x)} \frac{\partial u_{m(x)}(x)}{\partial x} \right]$ and $g(x) = \max(\tilde{g}(x), 0)$. The issue is that Equation (*) is an integral equation, since $\frac{\partial^2 U_{m(x)}(S(x))}{\partial \alpha \partial x}$ is an integral that depends on $S(x)$, which depends on $g(x')$ for $x' < x$. We prove the result for the case $x_0 = x^* < x_1$ but other cases are analogous.

Given $x_1 \in (x^*, x^{**})$, let $\mathcal{C}_{x_1} = \{h : [x^*, x_1] \rightarrow \mathbb{R}_{\geq 0} \text{ continuous}\}$ with the norm $\|h\|_\infty$, and define $T_{x_1} : \mathcal{C}_{x_1} \rightarrow \mathcal{C}_{x_1}$ as follows:

$$T_{x_1}(g)(x) = \max \left(-m'(x) \frac{\partial^2 U_{m(x)}(S(x))}{\partial \alpha^2} + m'(x) \frac{d^2 u_{m(x)}(x)}{dm(x)^2}, 0 \right)$$

for $x \in [x^*, x_1]$. Let $g_1, g_2 \in \mathcal{C}_{x_1}$ with $\|g_1 - g_2\| \leq K$, and, for each $x_2 \in (x^*, x_1)$, define g_{x_2} by: $g_{x_2}(x) = g_1(x)$ if $x \leq x_2$ and $g_{x_2}(x) = g_2(x)$ otherwise. Then

$$\begin{aligned} |T_{x_1}(g_1)(x) - T_{x_1}(g_2)(x)| &\leq m'(x) \left| -\frac{\partial^2 U_{m(x)}(S_{g_1}(x))}{\partial \alpha^2} + \frac{\partial^2 U_{m(x)}(S_{g_2}(x))}{\partial \alpha^2} \right| = \\ &= m'(x) \left| \int_{x^*}^x \frac{\partial}{\partial x_2} \left(\frac{\partial^2 U_{m(x)}(S_{g_{x_2}}(x))}{\partial \alpha^2} \right) dx_2 \right| = \\ &= m'(x) \left| \int_{x^*}^x r e^{-rd(x, x_2)} \frac{m'(x_2)}{r \frac{\partial u_{m(x)}(x)}{\partial x}} (g_1(x_2) - g_2(x_2)) \left[\frac{\partial^2 U_{m(x)}(x_2)}{\partial \alpha^2} - \frac{\partial^2 U_{m(x)}(S_{g_{x_2}}(x))}{\partial \alpha^2} \right] dx_2 \right| \leq \\ &\leq \int_{x^*}^x \frac{\bar{m}'^2 KL(x_2 - x^*)}{M'(x_2 - m(x_2))} dx_2 \leq \int_{x^*}^x \frac{\bar{m}'^2 KL(x_2 - x^*)}{M'(x_2 - x^*)(1 - m'(\tilde{x}(x_2)))} dx_2 \leq KC(x - x^*) \end{aligned}$$

for some constant $C > 0$.³⁹ If x_1 is close enough to x^* , $C(x - x^*) < 1$ and hence T_{x_1} is a contraction. Thus g (and hence e) is uniquely determined in a neighborhood of x^* . By repeating the same argument we can extend the solution uniquely on any interval (x^*, x) where $e(x') > 0$ for all $x \in (x^*, x)$. \square

Proof of Proposition 6. For (i), the uniqueness is proven by Lemmas 5 and 6; that s is C^1 follows from the fact that the RHS of Equation (*) is continuous. (ii) will be proven as part of Proposition 8.

³⁹ L is a Lipschitz constant for $\frac{\partial^2 u}{\partial \alpha^2}$, and $\bar{m}' = \sup_x m'(x)$. The argument still goes through if we only require $\frac{\partial^2 u}{\partial \alpha^2}$ to be Hölder continuous for some positive exponent.

For (iii), suppose not. Then there is a sequence $(\delta_n)_n$ with $\delta_n \rightarrow 1$ and a sequence of Q1Es s_n for each δ_n , such that s_n is not a 1E for all n . By part (ii), we know that $s_n(x, t) \rightarrow s(x, t)$ for all x, t . Suppose that, for each n , there is x_k^n for which $S_n(x_{k+1}^n)$ is not a Condorcet winner in $I(x_k^n)$ because a strict majority strictly prefers $S_n(y_n)$, and assume $x_k^n \rightarrow x$ and $y_n \rightarrow y$. Note that $y_n \leq x_{k+1}^n$, as otherwise all agents to the left of $m(x_k^n)$ and some to the right would strictly prefer $S_n(x_k^n)$ over $S_n(y_n)$; and thus $y \leq x$. If $x \in (x^*, x^{**})$ and $y < x$, this leads to a contradiction as $U_\alpha(S_n(y_n)) \rightarrow U_\alpha(S(y))$ and $U_\alpha(S_n(x_k^n)) \rightarrow U_\alpha(S(x))$ for all α , and $U_\alpha(S(x)) > U_\alpha(S(y))$ for all $\alpha \in (m(x) - \epsilon, x + d_x^+)$ for some $\epsilon > 0$. If $y = x$, suppose $y_n \in (x_{k+l_n}^n, x_{k+l_n-1}^n)$, where $l_n \geq 2$. It is clear that $U_\alpha(S_n(x_{k+1}^n)) > U_\alpha(S_n(x_{k+l_n-1}^n)) > U_\alpha(S_n(y_n))$ for all $\alpha \in (x_{k+l_n-1}^n, x_k^n + d_{x_k^n}^+)$ and $U_\alpha(S_n(x_{k+1}^n)) > U_\alpha(S_n(x_{k+l_n}^n)) > U_\alpha(S_n(y_n))$ for all $\alpha \in (m(x), x_{k+l_n}^n)$, so it must be that some $\alpha_n \in (x_{k+l_n-1}^n, x_{k+l_n}^n)$ prefers $S_n(y_n)$ to $S_n(x_{k+1}^n)$. But then

$$\begin{aligned}
0 &\geq \frac{1 - \delta}{1 - \delta^{l_n-1}} [U_{\alpha_n}(S_n(x_{k+1}^n)) - U_{\alpha_n}(S_n(y_n))] = \\
&= \frac{1 - \delta}{1 - \delta^{l_n-1}} \left[\sum_{t=0}^{l_n-1} \delta^t u_{\alpha_n}(x_{k+1+t}^n) - u_{\alpha_n}(y_n) - (1 - \delta^{l_n-1}) \sum_{t=1}^{\infty} \delta^t u_{\alpha_n}(x_{k+l_n+t}^n) \right] = \\
&= \frac{1 - \delta}{1 - \delta^{l_n-1}} \left[u_{\alpha_n}(x_{k+1}^n) - u_{\alpha_n}(y_n) + \sum_{t=1}^{l_n-1} \delta^t u_{\alpha_n}(x_{k+1+t}^n) - (1 - \delta^{l_n-1}) \delta U_{\alpha_n}(S(x_{k+l_n+1}^n)) \right] \\
&\xrightarrow{n \rightarrow \infty} 0 + \delta [u_x(x) - (1 - \delta)U_x(S(x))] > 0,
\end{aligned}$$

a contradiction.

An analogous proof can be written if $x = x^*$ after a normalization argument. Briefly, if $x = x^*$, assume WLOG that $x^* = 0$ to simplify notation, and denote $T_n(y) = x_k^n y$ and $U_\alpha^n(y) = U_{\alpha x_k^n}(\alpha x_k^n) - \frac{1}{(x_k^n)^2} (U_{\alpha x_k^n}(\alpha x_k^n) - U_{\alpha x_k^n}(y x_k^n))$. In the normalized version of the problem, x_k^n maps to $y_k^n = 1 > 0$ and we can apply the above arguments. The case $x = x^{**}$ is similar. \square

Proof of Proposition 7. WLOG assume $r = 1$. Suppose that there is a CLS with $e(x) \geq A$ for all $x \leq x_0$. Take $D > 0$ fixed, and let $L > 0$ be such that, for all $\alpha, x, x' \in [-1, 1]$, $\left| \frac{\partial u_\alpha^2(x)}{\partial \alpha^2} - \frac{\partial u_\alpha^2(x')}{\partial \alpha^2} \right|, \left| \frac{\partial u_\alpha^2(x)}{\partial \alpha \partial x} - \frac{\partial u_\alpha^2(x')}{\partial \alpha \partial x} \right|, \left| \frac{\partial u_\alpha^2(x)}{\partial x^2} - \frac{\partial u_\alpha^2(x')}{\partial x^2} \right| \leq L|x - x'| + D$. (For any D , such L exists because u is C^2 .) Note then that $\frac{\partial^2 u_y(y)}{\partial x^2} \equiv -\frac{\partial^2 u_y(y)}{\partial \alpha \partial x} \in [M', M]$; $\left| \frac{\partial^2 u_{m(y)}(y)}{\partial \alpha \partial x} - \frac{\partial^2 u_{m(y)}(m(y))}{\partial \alpha \partial x} \right|, \left| \frac{\partial^2 u_{m(y)}(y)}{\partial x^2} - \frac{\partial^2 u_{m(y)}(m(y))}{\partial x^2} \right| \leq L(y - m(y)) + D$; and

$\left| \frac{\partial u_{m(y)}(y)}{\partial x} \right| \in [M'(y - m(y)), M(y - m(y))]$. In addition, $|x - s(x, t)| \leq \frac{t}{A}$ for all t , so $\left| \frac{\partial^2 \max(u_{m(x)}(s(x, t)), 0)}{\partial \alpha^2} - \frac{\partial^2 u_{m(x)}(x)}{\partial \alpha^2} \right| \leq \frac{\tilde{L}t}{A} + D$, where $\tilde{L} = \max\left(L, \frac{\max|\frac{\partial^2 u}{\partial \alpha^2}|}{\min d_y}\right)$. In turn, this means that

$$\left| \frac{\partial^2 U_{m(x)}(S(x))}{\partial \alpha^2} - \frac{\partial^2 u_{m(x)}(x)}{\partial \alpha^2} \right| \leq \int_0^\infty e^{-t} \left(\frac{\tilde{L}t}{A} + D \right) dt = \frac{\tilde{L}}{A} + D.$$

Putting this all together, by Equation (*),

$$e(x_0) \geq \frac{(2m'(x_0) - 1)M' - (2m'(x_0) + 1)(LB + D)}{MB} - \frac{(1 + B')^2 \tilde{L}}{AMB} - \frac{(1 + B')^2 D}{MB} - \frac{B''}{1 - B'} \stackrel{?}{\geq} A.$$

We now choose $A = \sqrt{\frac{(1+B')^2 \tilde{L}}{MB}}$. Then it is enough if

$$\frac{(1 - 2B')M' - (3 + 2B')(LB + D)}{MB} - \frac{(1 + B')^2 D}{MB} - \frac{B''}{1 - B'} \geq 2\sqrt{\frac{(1 + B')^2 \tilde{L}}{MB}}.$$

Choose any $B' < \frac{1}{2}$ and any B'' . Choose D such that $(1 - 2B')M' > (3 + 2B')D + (1 + B')^2 D$. Then this condition holds for B small enough. We can show with a similar argument that, under these parameter conditions, $e(x) \xrightarrow{x \rightarrow x^*} \infty$, so the unique solution to Equation (*) must satisfy $e(x) \geq A \forall x$ by an argument similar to Lemma 6. \square

We now define a (not necessarily continuous) limit solution (LS) as a policy mapping s such that

- (i) $x = \arg \max_{y \in [-1, 1]} U_{m(x)}(S(y)) \forall x \in [x^*, x^{**})$.
- (ii) If there is $c > 0$ s.t. $W(x_0 - \epsilon) = 0 \forall \epsilon \in [0, c]$, then $d(x_0^+, x_0) = 0$.
- (iii) If $W(x_0) = 0$ and $W(x') > 0$ for all x' in a left-neighborhood of x_0 , then $d(x_0^+, x_0^-)$ satisfies⁴⁰

$$e^{\frac{rd(x_0^+, x_0^-)}{2}} = 1 + m'(x_0) \frac{\frac{\partial V_{m(x_0)}(S(x_0^-))}{\partial \alpha}}{\frac{\partial u_{m(x_0)}(x_0)}{\partial x}}.$$

⁴⁰We denote $f(x^-) = \lim_{t \nearrow x} f(x)$ and $f(x^+) = \lim_{t \searrow x} f(x)$.

This definition is backward-engineered so that the transition path generated by an LS will be the limit of Q1E transition paths as $\delta \rightarrow 1$. (i) requires that an LS has to be as if agents could choose their preferred continuation. (ii) says that, if in a left-neighborhood of x_0 pivotal agents are indifferent between the LS transition path and a constant path, then the policy path cannot stop at x_0 for a positive length of time. The significance of (iii) is that non-CLS transition paths will have intervals in which $W(x) > 0$ and, as a result, the policy moves quickly (in the limit, instantaneously) through such intervals. At a point x_0 where W hits 0 again, the transition path has to slow down dramatically (in the limit, stop for some time $d(x_0^+, x_0^-)$) in order to increase the average policy of the path so that W' can have a kink and $W'(x_0^+)$ can be non-negative. But Q1Es have “inertia”: when they slow down, they do so for long enough that the average policy increases enough that $W'(x_0^+) > 0$. Property (iii) requires the correct value of $d(x_0^+, x_0^-)$ to match the behavior of Q1E transition paths around such points.

The following properties, defined jointly for the parameters u , m and a policy mapping s , will help us to ascertain the properties of an LS.

B2.1 $m'(x^*) > \frac{1}{2}$. (This implies that $e(x) > 0$ for x in a neighborhood of x^* .)⁴¹

B2.1' $V_{m(x)}(S(x)) \equiv 0$ and there is $K > 0$ s.t. $e(x) \geq K$ for all $x \in (x^*, x^{**})$.

B2.2 u , m are C^3 , and there is no point $x \in [x^*, x^{**})$ for which $V_{m(x)}(S(x)) = \frac{\partial V_{m(x)}(S(x))}{\partial \alpha} = \frac{\partial^2 V_{m(x)}(S(x))}{\partial \alpha^2} = \frac{\partial^3 V_{m(x)}(S(x))}{\partial \alpha^3} = 0$.

B2.3 For all $x \in (x^*, x^{**})$ such that $W(x) = 0$ and $W(x') > 0$ for $x' < x$ arbitrarily close to x , $W'(x^-) < 0$. We refer to such points x as *vertex points*.

B2.1 and B2.1' are not generic, but hold in an open set. B2.2 and B2.3 are generic conditions under the assumption that u , m are C^3 ; this is shown in Appendix C. On an intuitive level, B2.3 requires that, at points where an interval of fast policy change ends (that is, where W hits 0), the derivative of W does not happen to also equal zero; B2.2 is a similar but weaker condition involving higher-order derivatives of W .

We will now build towards a characterization of LS.

⁴¹As seen in Lemma 5, $\frac{1}{m'(x^*)} \frac{\partial^2}{\partial x^2} u|_{x^*, x^*} + 2 \frac{\partial^2}{\partial \alpha \partial x} u|_{x^*, x^*} > 0$ is enough to guarantee that $e(x) > 0$ for x close to x^* —in fact, the condition guarantees $e(x) \geq \frac{C}{x-x^*}$ for some $C > 0$. In addition, Assumptions A1, A4 imply that $\frac{\partial^2}{\partial x^2} u|_{x^*, x^*} = -\frac{\partial^2}{\partial \alpha \partial x} u|_{x^*, x^*}$.

Lemma 7. *If any LS $s(x, t)$ satisfying Conditions B2.2 and B2.3 is such that $W(x) > 0$ for some $x \in (x^*, x^{**})$, then there are sequences $(y_l)_{l \in \mathbb{N}_{\geq 0}}$, $(e_l)_{l \in \mathbb{N}_{\geq 1}}$ such that $(y_l)_l$ is strictly increasing in l ; $W(y_l) = 0$ for all l ; $W(y) > 0$ for all $y \in (y_l, y_{l+1})$ for any l ; $d(y_{l+1}^-, y_l^+) = 0$; $d(y_l^+, y_l^-) = e_l$; $W'(y_l^+) > 0 > W'(y_l^-)$ for all $l \geq 1$;*

$$e^{\frac{re_l}{2}} = 1 + m'(y_l) \frac{\frac{\partial V_{m(y_l)}(S(y_l^-))}{\partial \alpha}}{\frac{\partial u_{m(y_l)}(y_l)}{\partial x}}$$

for all l ; $y_l \rightarrow x^{**}$ as $l \rightarrow \infty$; and $W(z) = 0$ for $z < y_0$.

Proof. Let (a, b) be the largest interval containing x such that $W(y) > 0$ for all $y \in (a, b)$, and denote $a = y_0$, $b = y_1$. That $d(x, x') = 0$ for all $x' < x \in (a, b)$ follows from the following argument. Take $x \in (a, b)$. Since $W(x) > 0$, there is $\tilde{x} \in (m(x), x)$ such that $u_{m(x)}(\tilde{x}) = U_{m(x)}(S(x)) > u_{m(x)}(x)$. Then $d(x, \tilde{x}) = 0$, as otherwise we would have $U_{m(x)}(S(\tilde{x})) > U_{m(x)}(S(x))$, contradicting the definition of a LS. Now suppose $d(x, x') > 0$ for some $x' < x \in (a, b)$. Construct a decreasing sequence $x = \tilde{x}_0 > \tilde{x}_1 > \tilde{x}_2 > \dots$ such that, for all n , $u_{m(\tilde{x}_n)}(\tilde{x}_{n+1}) = U_{m(\tilde{x}_n)}(S(\tilde{x}_n))$ and $d(\tilde{x}_n, \tilde{x}_{n+1}) = 0$ per the above argument. Let $\tilde{x}_\infty = \lim \tilde{x}_n$. If $\tilde{x}_\infty < x'$ we have a contradiction and the proof is done. If not, it follows by continuity that $u_{m(\tilde{x}_\infty)}(\tilde{x}_\infty) = U_{m(\tilde{x}_\infty)}(S(\tilde{x}_\infty))$, i.e., $W(\tilde{x}_\infty) = 0$, a contradiction.

Let $d(b^+, b^-) = e_1$. That e_1 is as required follows from the definition of LS. Note that Condition B2.3 implies that $e_1 > 0$. In addition, $W'(b^+) > 0$. To see this, in general let $\epsilon_l = \frac{\partial}{\partial \alpha} V_{m_l}(S(y_l^-))$ and $\epsilon'_l = \frac{\partial}{\partial \alpha} V_{m_l}(S(y_l^+))$, and suppose $\epsilon_l < 0$ as per Condition B2.3. Then

$$\begin{aligned} \epsilon'_l &= \frac{\partial V_{m_l}(S(y_l^+))}{\partial \alpha} = e^{-re_l} \left(\frac{\partial V_{m_l}(S(y_l^-))}{\partial \alpha} \right) + (1 - e^{-re_l}) \left(-\frac{1}{m'(x)} \frac{\partial u_{m(x)}(x)}{\partial x} \right) \\ &= \frac{\frac{\partial V_{m_l}(S(y_l^-))}{\partial \alpha} \frac{1}{m'(x)} \frac{\partial u_{m(x)}(x)}{\partial x}}{-\frac{\partial V_{m_l}(S(y_l^-))}{\partial \alpha} - \frac{1}{m'(x)} \frac{\partial u_{m(x)}(x)}{\partial x}} = -\frac{\partial V_{m_l}(S(y_l^-))}{\partial \alpha} \frac{1}{1 + \frac{\frac{\partial V_{m_l}(S(y_l^-))}{\partial \alpha}}{\frac{1}{m'(x)} \frac{\partial u_{m(x)}(x)}{\partial x}}} = \\ &= -\epsilon_l \frac{1}{1 + \frac{\epsilon_l}{\frac{1}{m'(x)} \frac{\partial u_{m(x)}(x)}{\partial x}}} > 0. \end{aligned}$$

This implies that there is $y_2 > y_1$ such that $W(y) > 0$ for $y \in (y_1, y_2)$, with $W(y_2) = 0$ and $W'(y_2^-) < 0$, and so on.

Next we argue that $y_l \rightarrow x^{**}$ as $l \rightarrow +\infty$. Suppose instead that $y_l \rightarrow y^* < x^{**}$, and let $m_l = m(y_l)$, $m^* = m(y^*)$. Since V is continuous, $V_{m(y^*)}(S(y^*)) = 0$. In addition, $\frac{\partial}{\partial \alpha} V_{m^*}(S(y^{*-}))$ must equal zero.⁴²

Suppose, then, that $\frac{\partial^2}{\partial \alpha^2} V_{m^*}(S(y^{*-})) \neq 0$. If this is positive, we have $V_{m(x)}(S(x)) > 0$ for all $x < y^*$ in a neighborhood of y^* , a contradiction.

If it is negative, we will obtain a contradiction by showing that $(\epsilon_l)_l$ cannot go fast enough to 0 for $(y_l)_l$ to converge. Note that $\epsilon_l < 0 < \epsilon'_l$ and $\epsilon_l + \epsilon'_l \in \mathcal{O}(\epsilon_l^2)$ since $\epsilon'_l = -\epsilon_l \frac{1}{1 + \frac{-\epsilon_l}{-\frac{1}{m'(x)} \frac{\partial u_{m(x)}(x)}{\partial x}}}$, as shown above, and $-\frac{1}{m'(x)} \frac{\partial u_{m(x)}(x)}{\partial x}$ is bounded away from 0 in a neighborhood of y^* . Next, we argue that $\epsilon_{l+1} = \epsilon_l + \mathcal{O}(\epsilon_l^2)$.

Let $N(\alpha) = \frac{\partial V_\alpha(S(m^{-1}(\alpha)))}{\partial \alpha}$ and $M(\alpha) = \frac{\partial^2 V_\alpha(S(m^{-1}(\alpha)))}{\partial \alpha^2}$. We claim that M is left-continuous at m^* —indeed, for this to not be the case we would require $\sum_l \epsilon_l = +\infty$, which implies M is not bounded in a neighborhood of y^* , a contradiction. Thus, since $M(m^*) < 0$, $M(\alpha) < 0$ for all $\alpha < m^*$ in a neighborhood of m^* .

Let $\overline{M}_l = \max_{\alpha \in (\alpha_l, \alpha_{l+1})} -M(\alpha)$, $\underline{M}_l = \min_{\alpha \in (\alpha_l, \alpha_{l+1})} -M(\alpha)$. Note that $\overline{M}_l - \underline{M}_l \leq L(m_{l+1} - m_l)$ for some fixed constant L , i.e., $\overline{M}_l - \underline{M}_l \in \mathcal{O}(m_{l+1} - m_l)$.⁴³ Since $V_{m_l} = V_{m_{l+1}} = 0$,

$$0 = \int_{m_l}^{m_{l+1}} N(\alpha) = N(m_l^+) (m_{l+1} - m_l) + \int_{m_l}^{m_{l+1}} M(\alpha) (m_{l+1} - \alpha)$$

, where $N(m_l^+) = \epsilon'_l$. This implies

$$\begin{aligned} \underline{M}_l \frac{(m_{l+1} - m_l)^2}{2} &\leq \epsilon'_l (m_{l+1} - m_l) \leq \overline{M}_l \frac{(m_{l+1} - m_l)^2}{2} \\ \underline{M}_l \frac{m_{l+1} - m_l}{2} &\leq \epsilon'_l \leq \overline{M}_l \frac{m_{l+1} - m_l}{2} \end{aligned}$$

Now $\epsilon_{l+1} = \epsilon'_l + \int_{m_l}^{m_{l+1}} M(\alpha) = \epsilon'_l - (m_{l+1} - m_l) \tilde{M}$, for some $\tilde{M} \in (\underline{M}_l, \overline{M}_l)$. From the above, $\epsilon'_l = \tilde{M}_l \frac{m_{l+1} - m_l}{2} + \mathcal{O}((m_{l+1} - m_l)^2)$. Then $\epsilon_{l+1} = -\epsilon'_l + \mathcal{O}((m_{l+1} - m_l)^2)$. In addition, it follows that $\mathcal{O}(\epsilon'_l) = \mathcal{O}(m_{l+1} - m_l)$. Since $\epsilon'_l = -\epsilon_l + \mathcal{O}(\epsilon_l^2)$, we have that $\epsilon_{l+1} = \epsilon_l + \mathcal{O}(\epsilon_l^2)$, i.e., $(\epsilon_l)_l$ at most decays (or grows) at the rate of a harmonic series,

⁴²Indeed, if this derivative is negative, it follows that $V_{m(x)}(S(x)) \geq V_{m(x)}(S(y^{*-})) > 0$ for all $x < y^*$ in a neighborhood of y^* , contradicting that $y_l \rightarrow y^*$. If it is positive, then $V_{m(x)}(S(y^{*-})) \leq -c(y^* - x)$ for x in such a neighborhood and some $c > 0$. From the fact that $V_{m(x)}(S(x)) \geq 0$ and $0 = V_{m^*}(S(y^*)) \geq V_{m(x)}(S(y^*))$ it then follows that $E(S(y^*)) - E(S(x)) \geq c' > 0$ for all $x < y^*$, which is impossible.

⁴³This follows from the assumption that u is C^3 .

whence $\sum_l \epsilon_l = \infty$. Since $\epsilon_l \in \mathcal{O}(m_{l+1} - m_l)$, we have $\sum_l (m_{l+1} - m_l) = \infty$ as well, which contradicts $y_l \rightarrow y^*$.

Finally, suppose that $P = \frac{\partial^3}{\partial \alpha^3} V_{m^*}(S(y^{*-})) \neq 0$. If it is negative, we again have $V(\alpha) > 0$ for α in a left-neighborhood of m^* , a contradiction, so it must be positive; and, as before, $P(\alpha) = \frac{\partial^3}{\partial \alpha^3} V_\alpha(S(m^{-1}(\alpha)))$ must be left-continuous at m^* , i.e., it must be close to P for α close to m^* . Note that

$$\begin{aligned} 0 &= \int_{m_l}^{m_{l+1}} N(\alpha) = \epsilon'_l (m_{l+1} - m_l) + \int_{m_l}^{m_{l+1}} M(\alpha) (m_{l+1} - \alpha) \\ &= \epsilon'_l (m_{l+1} - m_l) + \frac{(m_{l+1} - m_l)^2}{2} M(\tilde{\alpha}_l) \end{aligned}$$

for some $\tilde{\alpha}_l \in (m_l, m_{l+1})$. This implies

$$\begin{aligned} \epsilon'_l &= -M(\tilde{\alpha}_l) \frac{m_{l+1} - m_l}{2}, \quad \epsilon'_{l+1} = -M(\tilde{\alpha}_{l+1}) \frac{m_{l+2} - m_{l+1}}{2} \\ \epsilon_{l+1} &= \epsilon'_l + \int_{m_l}^{m_{l+1}} M(\alpha) = \epsilon'_l + (m_{l+1} - m_l) M(\tilde{\alpha}_l) \\ \implies \epsilon_{l+1} &= (m_{l+1} - m_l) \frac{2M(\tilde{\alpha}_l) - M(\tilde{\alpha}_l)}{2}, \end{aligned}$$

where $\tilde{\alpha}_l, \tilde{\alpha}_{l+1} \in (m_l, m_{l+1})$. To finish the proof we will need to be more specific about the positions of these values in the interval (m_l, m_{l+1}) . Due to the left-continuity of $P(\alpha)$, $M(\alpha)$ is roughly linear in each interval (m_l, m_{l+1}) . This, coupled with the above, implies that $\tilde{\alpha}_l = \frac{2m_l + m_{l+1}}{3} + o(m_{l+1} - m_l)$ and $\tilde{\alpha}_{l+1} = \frac{m_l + m_{l+1}}{2} + o(m_{l+1} - m_l)$. In addition, $M(m_l^+) - M(m_l^-) \in \mathcal{O}(\epsilon_l) \in o(m_l - m_{l-1})$. Then

$$\begin{aligned} &M(\tilde{\alpha}_{l+1}) - 2M(\tilde{\alpha}_l) + M(\tilde{\alpha}_l) = M(\tilde{\alpha}_{l+1}) - M(\tilde{\alpha}_l) + M(\tilde{\alpha}_l) - M(\tilde{\alpha}_l) = \\ &= \mathcal{O}(\epsilon_{l+1}) + (P + o(l)) \left(\frac{2m_{l+1} + m_{l+2}}{3} - \frac{m_l + m_{l+1}}{2} + o(m_{l+2} - m_l) \right) + \\ &\quad + (P + o(l)) \left(\frac{2m_l + m_{l+1}}{3} - \frac{m_l + m_{l+1}}{2} + o(m_{l+2} - m_l) \right) \\ &= P \left(\frac{m_{l+2}}{3} - \frac{m_l}{3} \right) + o(m_{l+2} - m_l) > 0, \text{ so that} \\ \frac{-\epsilon_{l+1}}{-M(\tilde{\alpha}_{l+1})} &= \frac{m_{l+1} - m_l}{2} \frac{-2M(\tilde{\alpha}_l) + M(\tilde{\alpha}_l)}{-M(\tilde{\alpha}_{l+1})} = \\ &= \frac{m_{l+1} - m_l}{2} \left(1 + \frac{M(\tilde{\alpha}_{l+1}) - 2M(\tilde{\alpha}_l) + M(\tilde{\alpha}_l)}{-M(\tilde{\alpha}_{l+1})} \right) \geq \frac{m_{l+1} - m_l}{2} \end{aligned}$$

$$\begin{aligned} \implies \frac{m_{l+2} - m_{l+1}}{2} &= \frac{\epsilon'_{l+1}}{-M(\tilde{\alpha}_{l+1})} = \frac{-\epsilon_{l+1} + \mathcal{O}(\epsilon_{l+1}^2)}{-M(\tilde{\alpha}_{l+1})} = \frac{-\epsilon_{l+1}}{-M(\tilde{\alpha}_{l+1})} (1 + \mathcal{O}(m_{l+1} - m_l)) \geq \\ &\geq \frac{m_{l+1} - m_l}{2} + \mathcal{O}((m_{l+1} - m_l)^2), \end{aligned}$$

which as before implies $m_l \rightarrow \infty$, a contradiction. As a result, we can conclude that if $y_l \rightarrow y^* < x^{**}$ then m violates Condition B2.2 at y^* , which is what we wanted.

As for what happens to the left of a , if $W'(a^+) = 0$, we are done. If $W'(a^+) > 0$ instead, then by the same arguments as before $W'(a^-) < 0$ and there must be $y_{-1} < y_0$ s.t. $W(y) > 0$ for $y \in (y_{-1}, y_0)$ and $W(y_{-1}) = 0$, etc. If this sequence of intervals to the left of a is finite, re-index the sequence $(y_l)_l$ appropriately and we are done. If it is infinite, we would have an infinite decreasing sequence $y_0 > y_{-1} > \dots$ such that $y_l \rightarrow y_*$ as $l \rightarrow -\infty$. If $y_* > x^*$, we obtain a contradiction analogously to our previous arguments. If $y_* = x^*$ we still obtain a contradiction by a slightly different argument—near x^* and x^{**} it is not true that $\epsilon'_l = -\epsilon_l + \mathcal{O}(\epsilon_l^2)$, as $\frac{\partial u_m(x)}{\partial x}$ approaches zero, but it is still true that $\epsilon'_l \leq -\epsilon_l$, so it is possible that ϵ_l is shrinking fast enough for $(y_l)_l$ to converge as $l \rightarrow +\infty$, but not as $l \rightarrow -\infty$.

Finally, the fact that $W(x) = 0$ for $x < y_0$ follows from the fact that, if this were false, there would be a sequence $(\tilde{y}_l)_l$ with $\tilde{y}_l < x < y_0$ and $\tilde{y}_l \rightarrow x^{**}$ as $l \rightarrow +\infty$, which contradicts $W'(y_0^+) = 0$. \square

We can now construct a *canonical LS*, s_* , as follows. Under Condition B2.1, construct a smooth LS based on Lemma 5 for a maximal interval (x^*, x_0) where this is possible—either (x^*, x^{**}) if $e(x) > 0$ everywhere, or else up to a point x_0 where $e(x_0) = 0$. In the latter case, to the right of x_0 , Condition B2.2 guarantees that $W'''(x_0^+) > 0$, so $W(x) > 0$ in a right-neighborhood of x_0 . We can then construct the solution based on sequences $(y_l)_l$, $(e_l)_l$ as described above, with Condition B2.3 guaranteeing that $e_l > 0$ and $y_{l+1} > y_l$ for all l .

Lemma 8. *If s_* satisfies Conditions B2.1, B2.2 and B2.3, then it is the unique LS.*

Proof. Let \hat{s} be another another LS. Suppose that $W(x) > 0$ for some x , so s_* features a sequence $(y_l)_l$ as in Lemma 7. Let $\hat{y}_0 = \inf\{y \in (x^*, x^{**}) : \hat{W}(y) > 0\}$. Note that s_* and \hat{s} must be identical for x between x^* and $\min(y_0, \hat{y}_0)$ by Lemma 6.

If $y_0 < \hat{y}_0$, it follows that $\hat{V}_{m(y)}(\hat{S}(y)) = 0$ for y in a right-neighborhood of y_0 , but at the same time $\hat{V}_{m(y)}(\hat{S}(y)) \geq \hat{V}_{m(y)}(\hat{S}(y_0)) = V_{m(y)}(S(y_0)) > 0$, a contradiction.

If $\hat{y}_0 < y_0$, there are two cases. First, suppose that $\hat{W}(y) > 0$ for all y in a right-neighborhood of \hat{y}_0 . Then we can apply the previous argument at \hat{y}_0 . Second, suppose $W(y) > 0$ and $W(y) = 0$ are both obtained for $y > \hat{y}_0$ arbitrarily close to \hat{y}_0 . Then there must be an infinite collection of intervals $(a_n, b_n)_{n \in \mathbb{Z}_{\leq 0}}$ such that $b_n > a_n \geq b_{n-1}$ for all n ; $a_n \xrightarrow{n \rightarrow -\infty} \hat{y}_0$; $W(y) > 0$ for all $y \in (a_n, b_n)$ and $W(y) = 0$ for all $y = a_n$ or $y = b_n$. This case leads to a contradiction by arguments developed in Lemmas 5 and 7. Briefly, for $y > \hat{y}_0$ close enough to \hat{y}_0 , $S(y)$ and $\hat{S}(y)$ are similar; $V''(m(y)) = 0$; and $e(y) \geq C > 0$, so $\hat{V}''(\alpha) \leq \tilde{C} < 0$ for any α such that $m(\alpha) \in (a_n, b_n)$. This implies that \hat{s} in fact satisfies Conditions B2.2 and B2.3, which contradicts $a_n \xrightarrow{n \rightarrow -\infty} \hat{y}_0$ by Lemma 7.

Hence $y_0 = \hat{y}_0$. Then s_* and \tilde{s} must be identical for $x > y_0$ because their behavior is uniquely pinned down by Lemma 7. Finally, note that, in the case where $e(x) > 0 \forall x \in (x^*, x^{**})$, the same proof goes through and we do not need to assume that s_* satisfies B2.2 (in particular, we do not need to assume u or m are C^3). \square

The following Proposition summarizes our results and extends Proposition 6 to the case without a CLS.

Proposition 8. *Let $\tilde{e}(x)$ be the solution to Equation (*). Then, if $\tilde{e}(x) \geq 0$ for all $x \in [x^*, x^{**}]$, there is a CLS s_* given by $e \equiv \tilde{e}$. Moreover, this is the unique LS.*

Otherwise, assume that the canonical LS s_ satisfies Conditions B2.1, B2.2 and B2.3. Then it is the unique LS, and it is given by $e(x) = \tilde{e}(x)$ for x up to some \tilde{x} , and by two sequences $(y_l)_{l \in \mathbb{N}_{\geq 0}}$, $(e_l)_{l \in \mathbb{N}_{\geq 1}}$ such that: $(y_l)_l$ is increasing, $y_0 = \tilde{x}$ and $y_l \xrightarrow{l \rightarrow \infty} x^{**}$; $d(y_l^-, y_{l+1}^+) = 0$ and $d(y_l^+, y_l^-) = e_l$ for all $l \geq 1$; and*

$$U_{m(y_{l+1})}(S(y_l^+)) = u_{m(y_{l+1})}(y_{l+1})$$

$$e^{\frac{re_l}{2}} = 1 + m'(y_l) \frac{\frac{\partial V_{m(y_l)}(S(y_l^-))}{\partial \alpha}}{\frac{\partial u_{m(y_l)}(y_l)}{\partial x}}$$

For any sequence $(s_j)_j$, where s_j is a Q1E of the j -refined game, and for any fixed x , $s_j(x, t) \xrightarrow{j \rightarrow \infty} s_(x, t)$ a.s. (more precisely, $s_j(x, t) \xrightarrow{j \rightarrow \infty} s_*(x, t) \forall t$ where $s_*(x, t)$ is continuous in t).*

In addition, if $m(y_l) < y_{l-1}$ for all l or $\tilde{e}(x) \geq 0 \forall x$, there is $\bar{\delta} < 1$ such that all Q1Es for discount factor δ are 1Es within $[x^, m^{-1}(x^* + d_{x^*}^+)]$.*

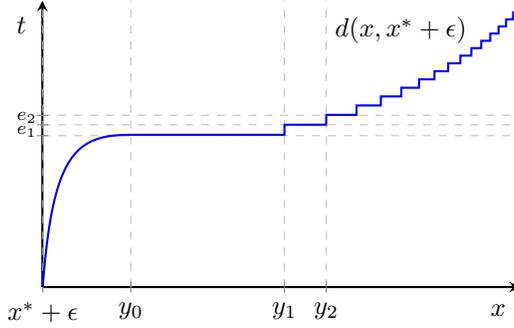


Figure 4: Discontinuous limit solution

Proof of Proposition 8. The last claim is a corollary of Proposition 5, and the characterization of s_* follows from Lemmas 7 and 8. It remains to show that all sequences of Q1Es of the j -refined games $(s_j)_j$ converge to s_* a.e., i.e., $s_j(x, t) \xrightarrow{j \rightarrow \infty} s_*(x, t) \forall x, t$ where $s_*(x, t)$ is continuous. Take a fixed x_0 and let $p(x) = d(x, x_0)$. Then we have to show $p_j(x) \xrightarrow{j \rightarrow \infty} p(x) \forall x$ where p is continuous.

Suppose not, so there is a sequence $(s_j)_j$ and an x_1 for which p is continuous at x_1 but $p_j(x_1) \not\rightarrow p(x_1)$. Take a subsequence $(s_l)_l$ such that p_l converges pointwise to some \hat{p} , and label the associated policy mapping \hat{s} .⁴⁴ $\hat{p} \neq p$ as, in particular, $\hat{p}(x_1) \neq p(x_1)$.

We will now prove the result simply by proving that \hat{s} is a LS.

(i): Let $x \in (x^*, x^{**})$. For each j and x' , $U_{m(x)}(S_j(s_j(x))) \geq U_{m(x)}(S_j(x'))$. If $s_j(x) \xrightarrow{j \rightarrow \infty} x$ we obtain $U_{m(x)}(\hat{S}(x)) \geq U_{m(x)}(\hat{S}(x'))$ by taking the limit. If not, and $s_{j'}(x) \xrightarrow{j' \rightarrow \infty} \tilde{x} < x$ for some subsequence, then $U_{m(x)}(\hat{S}(\tilde{x})) \geq U_{m(x)}(\hat{S}(x'))$. But this also implies $\hat{p}(x) - \hat{p}(\tilde{x}) = 0$, so $U_{m(x)}(\hat{S}(x)) = U_{m(x)}(\hat{S}(\tilde{x})) \geq U_{m(x)}(\hat{S}(x'))$.

In turn, the fact that \hat{s} satisfies (i) means that Lemmas 5 and 6 apply to it.

(ii): Suppose \hat{s} violates this condition at some $a \in (x^*, x^{**})$, i.e., $\hat{p}(a^+) - \hat{p}(a^-) = e^* > 0$. By an argument similar to Lemma 5, we have $\hat{W}'(a^+) > 0$ and hence \hat{p} is constant on some interval (a, b) .

Take $\epsilon > 0$ small, and let $(x_{j_n})_n$ be the recognized sequence of s_j for each j . By construction $(x_{j_n})_n$, must have $j\epsilon^* + j\hat{d}(a^-, a - \epsilon) + o(j)$ elements in $(a - \epsilon, a + \epsilon)$, and $j\hat{d}(a - \epsilon, a - 2\epsilon) + o(j)$ elements in $(a - 2\epsilon, a - \epsilon)$. In particular, given $\eta > 0$, for high

⁴⁴Use a diagonal argument to find a subsequence $(s_{l'})_{l'}$ such that $(p_{l'})_{l'}$ converges at all rational points. This guarantees convergence at all points except points of discontinuity of $\limsup_{l' \rightarrow \infty} p_{l'}$, which are countable because the function in question is increasing. Use another diagonal argument to get $(s_l)_l$ such that p_l also converges at all discontinuities of $\limsup_{l \rightarrow \infty} p_l$.

enough j there must be an element $x_{jt} \in (a - 2\epsilon, a - \epsilon)$ such that $x_{jt} - x_{j(t+1)} \geq \frac{1}{j\bar{e}(1+\eta)}$ for $\bar{e} = \max_{x \in (a-2\epsilon, a-\epsilon)} e(x)$. Let $x_{jt'}$ be the right-most element of $(x_{jn})_n$ contained in $(a - \epsilon, a + \epsilon)$. The above implies $t' - t \geq je^* + j\hat{d}(a^-, a - \epsilon) + o(j)$.

Now, denoting $x_{jn} = x_n$, $m(x_{jn}) = m_n$ and $S_j(x_{jn}) = S(x_n)$, and exploiting the indifference conditions $U_{m_{n-1}}(S(x_n)) = U_{m_{n-1}}(S(x_{n+1})) = u_{m_{n-1}}(x_n)$ and $U_{m_n}(S(x_{n+1})) = U_{m_n}(S(x_{n+2})) = u_{m_n}(x_{n+1})$,

$$\begin{aligned} V_{m_{n-1}}(S(x_{n+1})) - V_{m_n}(S(x_{n+1})) &= U_{m_{n-1}}(S(x_{n+1})) - U_{m_n}(S(x_{n+1})) - u_{m_{n-1}}(x_{n-1}) + u_{m_n}(m_n) \\ &= -(u_{m_{n-1}}(x_{n-1}) - u_{m_{n-1}}(x_n)) + (u_{m_n}(m_n) - u_{m_n}(x_{n+1})) \\ m'(\hat{x}_n)(x_{n-1} - x_n) \frac{\partial}{\partial \alpha} V_{\tilde{\alpha}_n}(S(x_{n+1})) &= -(x_{n-1} - x_n) \frac{\partial}{\partial x} u_{m_{n-1}}(\tilde{x}_n) + (x_n - x_{n+1}) \frac{\partial}{\partial x} u_{m_n}(\tilde{x}_{n+1}) \\ &\quad - \frac{\partial}{\partial x} u_{m_n}(\tilde{x}_{n+1}) \\ (x_{n-1} - x_n) &= (x_n - x_{n+1}) \frac{-\frac{\partial}{\partial x} u_{m_n}(\tilde{x}_{n+1})}{-m'(\hat{x}_n) \frac{\partial}{\partial \alpha} V_{\tilde{\alpha}_n}(S(x_{n+1})) - \frac{\partial}{\partial x} u_{m_{n-1}}(\tilde{x}_n)} \end{aligned}$$

for some $\tilde{\alpha}_n \in (m_n, m_{n-1})$, $\tilde{x}_n, \hat{x}_n \in (x_n, x_{n-1})$, $\tilde{x}_{n+1} \in (x_{n+1}, x_n)$. In addition

$$\begin{aligned} \frac{\partial}{\partial \alpha} V_\alpha(S(x_n)) &= e^{-\frac{r}{j}} \frac{\partial}{\partial \alpha} V_\alpha(S(x_{n+1})) + (1 - e^{-\frac{r}{j}}) \left(\frac{\partial u_\alpha(x_n)}{\partial \alpha} - \frac{\partial u_\alpha(m^{-1}(\alpha))}{\partial \alpha} - \frac{\frac{\partial u_\alpha(m^{-1}(\alpha))}{\partial x}}{m'(m^{-1}(\alpha))} \right) \\ \frac{\partial}{\partial \alpha} V_\alpha(S(x_n)) &= e^{-\frac{rk}{j}} \frac{\partial}{\partial \alpha} V_\alpha(S(x_{n+k})) + (1 - e^{-\frac{rk}{j}}) \left(\mathcal{O}(\epsilon) - \frac{\frac{\partial}{\partial x} u_\alpha(m^{-1}(\alpha))}{m'(m^{-1}(\alpha))} \right) \end{aligned}$$

for $\alpha \in (m(a - 2\epsilon), m(a + \epsilon))$, $x_n, x_{n+k} \in (a - 2\epsilon, a + \epsilon)$.

Then, for $n \in \{t, \dots, t'\}$,

$$\begin{aligned} x_{j(n-1)} - x_{jn} &\geq (x_{jn} - x_{j(n+1)}) \frac{-\frac{\partial}{\partial x} u_{m(a)}(a) - K\epsilon}{-\tilde{m}' \frac{\partial}{\partial \alpha} V_{\tilde{\alpha}_n}(S_j(x_{j(n+1)})) - \frac{\partial}{\partial x} u_{m(a)}(a) + K'\epsilon} \geq \\ &\geq (x_{jn} - x_{j(n+1)}) \frac{-\frac{\partial}{\partial x} u_{m(a)}(a) - K\epsilon}{-e^{-\frac{r(t-n)}{j}} \tilde{m}' \frac{\partial}{\partial \alpha} V_{\tilde{\alpha}_n}(S_j(x_{j(t+1)})) - e^{-\frac{r(t-n)}{j}} \frac{\partial}{\partial x} u_{m(a)}(a) + K''\epsilon} \\ &\geq (x_{jn} - x_{j(n+1)}) \frac{-\frac{\partial}{\partial x} u_{m(a)}(a) - K\epsilon}{G(j) - e^{-\frac{r(t-n)}{j}} \frac{\partial}{\partial x} u_{m(a)}(a) + K''' \epsilon} = \\ &= (x_{jn} - x_{j(n+1)}) \frac{1 - K'''' \epsilon}{e^{-\frac{r(t-n)}{j}} + \tilde{G}(j) + K'''' \epsilon} \end{aligned}$$

for some function $G(j)$ such that $G(j) \xrightarrow{j \rightarrow \infty} 0$, as

$$\left| \frac{\partial}{\partial \alpha} V_{\tilde{\alpha}_n}(S_j(x_{j(t+1)})) \right| \xrightarrow{j \rightarrow \infty} \left| \frac{\partial}{\partial \alpha} V_{\alpha^*}(\hat{S}(x^*)) \right| \leq \left| \frac{\partial}{\partial \alpha} V_{m(x^*)}(\hat{S}(x^*)) \right| + C\epsilon = C\epsilon$$

for some $\alpha^* \in [m(a - 2\epsilon), m(a + \epsilon)]$, $x^* \in [a - 2\epsilon, a - \epsilon]$ and $C > 0$.⁴⁵ Then

$$\begin{aligned} x_{j(n-1)} - x_{jn} &\geq (x_{jt} - x_{j(t+1)}) \prod_{k=0}^{t-n} \frac{1 - K'\epsilon}{e^{-\frac{rk}{j}} + \tilde{G}(j) + K'''\epsilon} \\ \implies x_{jt'} - x_{j(t'+1)} &\geq \frac{1}{je^*(1+\eta)} \prod_{k=0}^{je^*-1} \frac{1 - K'\epsilon}{e^{-\frac{rk}{j}} + \tilde{G}(j) + K'''\epsilon}. \end{aligned}$$

If we take ϵ small enough that $1 - K'\epsilon > 0$, the right-hand side grows to infinity as $j \rightarrow \infty$. In particular, for j high enough, $x_{jt'} - x_{j(t'+1)} > 3\epsilon$, a contradiction.

(iii): This follows from a calculation analogous to the one used for (ii). Briefly, if (iii) is violated at x_0 and $\hat{p}(x_0^+) - \hat{p}(x_0^-)$ is higher than the value required by (iii), then $x_{jt'} - x_{j(t'+1)} \xrightarrow{j \rightarrow \infty} \infty$, a contradiction. If $\hat{p}(x_0^+) - \hat{p}(x_0^-)$ is lower than the value required by (iii), then it can be shown that $j(x_{jn} - x_{j(n+1)}) \xrightarrow{j \rightarrow \infty} 0$ for all n such that $x_{j(n+1)} \geq a$, which implies that the number of elements of $(x_{jn})_j$ in $(a - \epsilon, a + \epsilon)$ grows faster than j , a contradiction. □

C Genericity of Conditions on m (For Online Publication)

In this Section we show that the conditions B1, B2.2 and B2.3 imposed on the function m are “generic”.⁴⁶ We employ two different notions of genericity. On the one hand, we show that these conditions hold on an open and dense set (or, at least, a residual set) within the function space with a natural metric. In addition, we show that some of these conditions hold on a *prevalent* set, a notion introduced in Hunt, Sauer and Yorke (1992) that generalizes the measure-theoretic notion of “almost everywhere” to

⁴⁵If necessary, take a convergent subsequence so that $(\tilde{\alpha}_{n_j})_j$ and $x_{j(t_j+1)}$ converge for this argument.

⁴⁶We say m satisfies conditions B2.1, B2.2, B2.3 if there is an LS for this m that satisfies them—equivalently, if the canonical LS for this m satisfies them.

infinite-dimensional spaces where an analog of the Lebesgue measure is not available.

Claim 1. *Consider the set of functions*

$$X_1 = \{m : [-1, 1] \longrightarrow [-1, 1] : m \in C^1, m \text{ weakly increasing} \}$$

with the norm $\|m\| = \max(\|m\|_\infty, \|m'\|_\infty)$. The subset $Y_1 \subseteq X_1$ satisfying B1 is open and dense,⁴⁷ and also prevalent in the sense of Hunt et al. (1992).

Proof. We first show Y_1 is open. Let $m_0 \in Y_1$; $x_1^* < \dots < x_N^*$ be the fixed points of m_0 ; $\alpha_i = m_0'(x_i^*)$ for $i = 1, \dots, N$; $\epsilon > 0$ and $\nu > 0$ such that $|m_0'(y) - 1| \geq \nu$ for $y \in I_i = (x_i^* - \epsilon, x_i^* + \epsilon)$ for any i ; $\nu > 0$ such that $|m(y) - y| \geq \nu$ for $y \notin I_i$ for any i ; $\eta = \min(\epsilon, \nu, \nu)$; and $m_1 \in B(m_0, \eta)$. Then $m_1(y) = y$ implies $|m_0(y) - y| < \eta \leq \nu$, so $y \in I_i$ for some i , so $|m_1'(y) - 1| \geq |m_0'(y) - 1| - |m_1'(y) - m_0'(y)| > \nu - \nu = 0$. This shows that $m_1'(y) \neq 1$ at any fixed point y of m_1 . Moreover, by construction either $m_0'(y) > 1$ for all $y \in I_i$ and $m_1'(y) > 1$ for all $y \in I_i$ as well, or the reverse inequalities hold, whence m_1 can have at most one fixed point in I_i for each i , and the set of fixed points is finite.

Next, we show Y_1 is dense. Let $m_0 \in X_1$ and $\epsilon > 0$. We want to show that there is $m_1 \in B(m_0, \epsilon) \cap Y_1$. Since m_0' is continuous in $[-1, 1]$, it is uniformly continuous, so we can take $\nu > 0$ such that if $|y - y'| < \nu$ then $|m_0'(y) - m_0'(y')| < \frac{\epsilon}{4}$. Partition $[-1, 1]$ into intervals I_1, I_2, \dots, I_J as follows: $I_j = [y_{j-1}, y_j)$, where $y_j = -1 + j\nu$, for $j < J$, and $I_J = [y_{J-1}, 1]$. For each j , if $m_0'(y_{j-1}) \geq 1$, let $m_2'(y) = m_0'(y) + \frac{\epsilon}{4}$ for all $y \in I_j$ (which implies $m_2'(y) > 1$ for $y \in I_j$); otherwise let $m_2'(y) = m_0'(y) - \frac{\epsilon}{4}$ for all $y \in I_j$ (so $m_2'(y) < 1$ for $y \in I_j$), and then define m_2 by integrating m_2' , with $m_2(-1) = m_0(-1)$. By construction m_2 has at most one fixed point in each interval I_j and $m_2' \neq 1$ at such points. Moreover, $\|m_2' - m_0'\| \leq \frac{\epsilon}{4}$ and $\|m_2 - m_0\| \leq 2\frac{\epsilon}{4} = \frac{\epsilon}{2}$. If $m_2(y_j) \neq y_j$ for all y_j , we can construct a “smoothed-out” version of m_2' , which we’ll call m_1' , that is in $B(m_0, \epsilon) \cap Y_1$. If $m_2(y_j) = y_j$ for some j , and $m_2'(y) > 1$ for $y \in I_j \cup I_{j+1}$ or $m_2'(y) < 1$ for $y \in I_j \cup I_{j+1}$, this is not a problem. If $m_2(y_j) = y_j$ for some j and $m_2'(y) > 1$ for $y \in I_j$, $m_2'(y) < 1$ for $y \in I_{j+1}$, we can construct a smooth m_1' such that $m_1'(y_j) = 1$, $m_1'(y) > 1$ to the left of y_j and < 1 to the right, and $m_1(y_j) < m_2(y_j) = y_j$. The remaining case is analogous.

For the last claim, note that, if a C^1 function m defined on a compact interval has $m' \neq 1$ at all its fixed points, it automatically has a finite number of them. Consider

⁴⁷The statement is also true within the space of C^3 functions, taken with the appropriate norm.

the translation $X_1 - v$, where v is the identity function. Then $m \in Y_1$ iff $m - v$ has no points where $(m - v)(y) = (m - v)'(y) = 0$. Finally, the fact that $Y_1 - v$ is prevalent in $X_1 - v$ follows from Proposition 3 in Hunt et al. (1992). \square

Claim 2. *B2.1 holds in an open set Y_2 within*

$$X_2 = \{m : [x^*, x^{**}] \longrightarrow [x^*, x^{**}] : m \in C^2, m \text{ weakly increasing,} \\ m(x^*) = x^*, m(x^{**}) = x^{**}, m(x) < x \forall x \in (x^*, x^{**})\}$$

taken with the norm $\|m\| = \max(\|m\|_\infty, \|m'\|_\infty, \|m''\|_\infty)$.⁴⁸

Proof. Trivial. \square

Claim 3. *B2.1' holds in an open set Y_4 within X_2 .*

Proof. This amounts to showing that e is continuous in m'' , and it follows from an argument similar to the proof of the uniqueness of e from Lemma 6. \square

Claim 4. *Assume u is C^3 . Let*

$$X_3 = \{m \in Y_2 : m \in C^3, m(x^*) = x^*, m(x^{**}) = x^{**}, m(x) < x \forall x \in (x^*, x^{**}), \\ m \text{ strictly increasing, } m \text{ satisfies B2.1}\}$$

taken with the norm $\|m\| = \max(\|m\|_\infty, \|m'\|_\infty, \|m''\|_\infty, \|m'''\|_\infty)$. For each $y \in (x^*, x^{**})$, the set $Y_3(y) \subseteq X_3$ of functions m for which B2.2 and B2.3 hold in $[x^*, y]$ is open and dense.

Proof. We proceed in two steps. First, we show that the set $Y_5(y) \subseteq X_3$ for which B2.2 holds in $[x^*, y]$ is open and dense. Second, we show that the set $Y_3(y)$ is open and dense within $Y_5(y)$.

To show that $Y_5(y)$ is open, take $m \in Y_5(y)$ and suppose there is a sequence $(m_n)_n$ such that $m_n \notin Y_5(y)$ for all n but $m_n \rightarrow m$. For each m_n we can construct a LS s_n (possibly not unique) by finding a convergent sequence of discrete-time equilibria $(s_{nj})_j$ for $\delta = e^{-\frac{r}{j}}$ with $j \rightarrow \infty$, as in Proposition 8. Using a diagonal argument, we can find a convergent subsequence of $(s_n)_n$, which by continuity must converge to a LS for m , \hat{s} . WLOG assume $(s_n)_n \rightarrow \hat{s}$. We will need the following Lemma:

⁴⁸Again, this is also true within the space of C^3 functions.

Lemma 9. *If s_* satisfies Conditions B2.1 and B2.2, then it is the unique LS. Moreover, s_* has a finite number of vertex points in $[x^*, y]$ for any $y < x^{**}$.*

Proof. Briefly, if s_* has an infinite number of vertex points in $[x^*, y]$, they must accumulate at some $y^* \in (x^*, y]$, which must satisfy $V_{m(y^*)}(S(y^*)) = \frac{\partial V_{m(y^*)}(S(y^*))}{\partial \alpha} = 0$. If $\frac{\partial^2 V_{m(y^*)}(S(y^*))}{\partial \alpha^2} > 0$ we obtain $V > 0$ in a neighborhood of y^* , a contradiction. If $\frac{\partial^2 V_{m(y^*)}(S(y^*))}{\partial \alpha^2} < 0$ this guarantees Condition B2.3 in a neighborhood of y^* , which means the vertex points near y^* must be part of a single sequence, contradicting Lemma 7.

Suppose that there are infinitely many vertex points on a left-neighborhood of y^* (the other case is analogous). Similar arguments apply if $\frac{\partial^3 V_{m(y^*)}(S(y^*))}{\partial \alpha^3} < 0$ or $\frac{\partial^3 V_{m(y^*)}(S(y^*))}{\partial \alpha^3} > 0$, respectively.

As for the uniqueness of s_* , the proof in Lemma 8 can be extended to this case. \square

From this we conclude that $\hat{s} = s_*$. Letting W_n be the value function for s_n , we then have $W_n \rightarrow W$. It can be shown in addition that, at every y that is not a vertex point of s_* , $W'_n(y) \rightarrow W'(y)$, $W''_n(y) \rightarrow W''(y)$ and $W'''_n(y) \rightarrow W'''(y)$, by using Lemmas 6 and 9.

Next, we show that $Y_5(y)$ is dense. Take $m \in X_3$ and $\epsilon > 0$. Consider \hat{m} given by: $\hat{m}(x^*) = x^*$, $\hat{m}'(x^*) = m'(x^*)$, $\hat{m}''(x^*) = m''(x^*)$ and $\hat{m}'''(x) = m'''(x) + \eta(x)$, where $|\eta(x)| \leq \epsilon$ will be defined as 0 except where we specify otherwise. We will argue that, by picking η correctly, we can find a $\hat{m} \in Y_5(y)$ that is close to m .

Apply the following algorithm. Take $\nu > 0$ small and $N > 0$ large. Let $\eta_0 \equiv 0$ and $m_0 \equiv m$. Let

$$x_0 = \inf \left\{ x \in (x^*, y] : \max \left(\left| V_{m_0(x)}(S_0(x)) \right|, \left| \frac{\partial}{\partial \alpha} V_{m_0(x)}(S_0(x)) \right|, \left| \frac{\partial^2}{\partial \alpha^2} V_{m_0(x)}(S_0(x)) \right|, \left| \frac{\partial^3}{\partial \alpha^3} V_{m_0(x)}(S_0(x)) \right| \right) \leq \frac{\epsilon}{N} \right\},$$

where $S_0(x)$ is a the policy path starting at x for a LS given median voter function m_0 .⁴⁹ Let $\alpha_0 = m(x_0)$. Define $\eta_1(x) = -\epsilon$ for $x \in [x_0, x'_0)$ and $\eta_1(x) = 0$ for all other x , with x'_0 taken so that $m_1(x'_0) = m_1(x_0) + \nu$. Next, let x_1 be the infimum of $x \in (x_0 + \nu, y]$ for which $|V_{m_1(x)}(S_1(x))|$, $|\frac{\partial}{\partial \alpha} V_{m_1(x)}(S_1(x))|$, $|\frac{\partial^2}{\partial \alpha^2} V_{m_1(x)}(S_1(x))|$ and

⁴⁹Note that, for ϵ small enough, $x_0 > x^*$ since $e(x) \geq C > 0$ in a neighborhood of x^* , which implies $\frac{\partial^2}{\partial \alpha^2} V_{m_0(x)}(S_0(x)) \geq C' > 0$.

$|\frac{\partial^3}{\partial \alpha^3} V_{m_1(x)}(S_1(x))| \leq \frac{\epsilon}{N}$, and define $\alpha_1 = m(x_1)$ and $\eta_2(x) = -\epsilon$ for $x \in [x_1, x'_1]$ and $\eta_2(x) = \eta_1(x)$ for all other x , with x'_1 taken so that $m_2(x'_1) = m_2(x_1) + \nu$. Define $x_k, \alpha_k, \eta_{k+1}, m_{k+1}$ for $k = 2, 3, \dots$ in the same fashion until $x_K = +\infty$ for some K .⁵⁰ Let $\tilde{m} = m_K$.

We argue that, if ν and N are taken to be small and large enough, respectively, \tilde{m} satisfies Condition B2.2. To explain why, we will need the following

Remark 5. A function $f : [a, b] \rightarrow \mathbb{R}$ is uniformly continuous iff there is an increasing function $h : [0, +\infty) \rightarrow [0, +\infty)$ such that $h(0) = 0$, h is continuous at 0, and $|f(x) - f(y)| \leq h(|x - y|)$ for all $x, y \in [a, b]$. We say a function h satisfying these properties is a *bounding function*.

Now note that, for any k and any $x < x'$ such that $m(x) = a, m(x') = a'$ satisfy $a, a' \in [\alpha_k, \alpha_k + \nu]$, we have

$$\begin{aligned} & \left| \frac{\partial^3}{\partial \alpha^3} V_a(\tilde{S}(x)) - \frac{\partial^3}{\partial \alpha^3} V_{a'}(\tilde{S}(x')) \right| \leq \left| \frac{\partial^3}{\partial \alpha^3} U_a(\tilde{S}(x)) - \frac{\partial^3}{\partial \alpha^3} U_a(\tilde{S}(x')) \right| \\ & + \left| \frac{\partial^3}{\partial \alpha^3} U_a(\tilde{S}(x')) - \frac{\partial^3}{\partial \alpha^3} U_{a'}(\tilde{S}(x')) \right| + \left| -\frac{d^3}{d\alpha^3} u_a(\tilde{m}^{-1}(a)) + \frac{d^3}{d\alpha^3} u_{a'}(\tilde{m}^{-1}(a')) \right| \\ & \leq \left| \frac{\partial^3}{\partial \alpha^3} U_a(\tilde{S}(x)) - \frac{\partial^3}{\partial \alpha^3} U_a(\tilde{S}(x')) \right| + h_1(a' - a) + h_2(a' - a) \\ & \leq K \frac{\epsilon}{N} + K'(a' - a) + h_1(a' - a) + h_2(a' - a) \leq K \frac{\epsilon}{N} + h_3(a' - a) \end{aligned}$$

where h_1, h_2, h_3 are bounding functions, $K, K' > 0$, and K, h_1, h_2, h_3 are independent of ν and N .

The bound $|\frac{d^3}{d\alpha^3} u_a(\tilde{m}^{-1}(a)) + \frac{d^3}{d\alpha^3} u_{a'}(\tilde{m}^{-1}(a'))| \leq h_2(|a - a'|)$ uses the uniform continuity of $\frac{d^3}{d\alpha^3} u_{\tilde{a}}(\tilde{m}^{-1}(\tilde{a}))$, which follows from the fact that u is C^3 and \tilde{m} is C^3 on $[a, a'] \subseteq [\alpha_k, \alpha_k + \nu]$. The bound $|\frac{\partial^3}{\partial \alpha^3} U_a(\tilde{S}(x')) - \frac{\partial^3}{\partial \alpha^3} U_{a'}(\tilde{S}(x'))| \leq h_1(|a - a'|)$ uses the uniform continuity of $\frac{\partial^3}{\partial \alpha^3} U$, and the fact that the mapping $x \mapsto \max(x, 0)$ is Lipschitz. The first bound is the trickiest, and is based on the idea that, if $\frac{\partial V}{\partial \alpha}$ and $\frac{\partial^2 V}{\partial \alpha^2}$ are low, then $\tilde{s}(x, t)$ changes relatively quickly as a function of t , so $\tilde{S}(x)$ and

⁵⁰This must happen for a finite K , as $\alpha_k - \alpha_{k-1} \geq \nu > 0$ for all k .

$\tilde{S}(x')$ are similar. Formally:

$$\begin{aligned}
& \left| \frac{\partial^3}{\partial \alpha^3} U_a(\tilde{S}(x)) - \frac{\partial^3}{\partial \alpha^3} U_a(\tilde{S}(x')) \right| \leq \left(1 - e^{-r\tilde{d}(x',x)}\right) 2 \max_{a,y} \frac{\partial^3}{\partial \alpha^3} u_\alpha(y) \\
1 - e^{-r\tilde{d}(x',x)} & \leq r\tilde{d}(x',x) = r \left(\sum_{\substack{\tilde{y}_l \in (x,x') \\ \text{vertex point}}} e_l + \int_x^{x'} e(y) dy \right) \leq r \frac{2}{r} \sum_{\tilde{y}_l \in (x,x') \text{ v.p.}} (e^{\frac{r e_l}{2}} - 1) + r\bar{e}(x' - x) \leq \\
& \leq 2 \sum_{\tilde{y}_l \in (x,x') \text{ v.p.}} \frac{|\max \tilde{m}'|}{\min_{x \in [x_0, y]} \left| \frac{\partial u_{m(x)}(x)}{\partial x} \right|} \left| \frac{\partial}{\partial \alpha} V_{a_l}(\tilde{S}(y_l^-)) \right| + r\bar{e}(x' - x) \leq \\
& \leq K'' \left(\left| \frac{\partial}{\partial \alpha} V_a(\tilde{S}(x)) \right| + \int_a^{a'} \left| \frac{\partial^2}{\partial \alpha^2} V_{\tilde{a}}(\tilde{S}(n(\tilde{a}))) \right| d\tilde{a} \right) + r\bar{e}(x' - x) \leq K'' \frac{\epsilon}{N} + K'''(a' - a).
\end{aligned}$$

Pick ν and N so that $K \frac{\epsilon}{N} + h_3(\nu) \leq \frac{\epsilon}{2}$ and $N \geq 4$. Now \tilde{m} satisfies Condition B2.2 because, for $a \in [\alpha_k, \alpha_k + \nu)$,

$$\left| \frac{\partial^3}{\partial \alpha^3} V_a(\tilde{S}(x)) \right| \geq \left| \frac{\partial^3}{\partial \alpha^3} V_{\alpha_k}(\tilde{S}(x_k)) \right| - \left| \frac{\partial^3}{\partial \alpha^3} V_{\alpha_k}(\tilde{S}(x_k)) - \frac{\partial^3}{\partial \alpha^3} V_a(\tilde{S}(x)) \right| \geq \epsilon - \frac{\epsilon}{N} - \frac{\epsilon}{2} \geq \frac{\epsilon}{4} > 0.$$

On the other hand, if $a \notin [\alpha_k, \alpha_k + \nu)$ for any k , then $|\frac{\partial^i}{\partial \alpha^i} V| > \frac{\epsilon}{N}$ for some $i = 0, 1, 2, 3$ by construction.

The only remaining issue is that \tilde{m} is not C^3 because η_K is not continuous at α_k and $\alpha_k + \nu$ for $k = 0, 1, \dots, K - 1$. However, it is easy to construct a continuous η close to η_K that fixes this problem.⁵¹

Next, we argue that $Y_3(y)$ is open. As shown in Proposition 8, Conditions B2.1, B2.2 and B2.3 taken together imply that the equilibrium path $s(x, t)$ will be given by either a smooth path with $e(x) > 0$ for all $x \in (x^*, y]$ or a smooth path up to some y_0 followed by a finite sequence of jumps and stops with stops at y_1, y_2, \dots, y_l . It is enough to show that e' is continuous in m''' , which follows from the arguments in Proposition 8, and that y_i is continuous in m''' for $i = 1, 2, \dots, l$, which is elementary (in fact, y_i is continuous in m').

Finally we argue that $Y_3(y)$ is dense. Take $m \in X_3$ and $\epsilon > 0$. Because $Y_5(y)$

⁵¹WLOG, take $a = \alpha_k$. If $V_a(S(x^-)) > 0$, it is easy to perturb η_K to make it continuous at a without violating Condition B2.2. If not, but $V_{\tilde{a}}(S(\tilde{x}^-)) > 0$ for $\tilde{a} < a$ arbitrarily close to a , we can perturb η_K at one such \tilde{a} instead. If $V_{\tilde{a}}(S(\tilde{x}^-)) = 0$ for all $\tilde{a} < a$ close to a , but $\frac{\partial^2}{\partial \alpha^2} V$ is nonzero close to a , we can do the same argument. If $\frac{\partial^2}{\partial \alpha^2} V$ is also zero in an interval to the left of a , then Condition B2.2 would be violated for $\tilde{a} < a$, a contradiction.

is dense, there is $\hat{m} \in B(m, \epsilon) \cap Y_5(y)$. Because $Y_5(y)$ is open, there is $\epsilon' > 0$ such that $B(\hat{m}, \epsilon') \subseteq B(m, \epsilon) \cap Y_5(y)$. Next, we claim that there is $\hat{\hat{m}} \in B(\hat{m}, \epsilon') \cap Y_3(y)$, which completes the proof. This can be shown by construction. If $\hat{e}(x) > 0$ for all $x \in [x^*, y]$, we are done. If not, \hat{s} induces a policy path that is continuous up to some y_0 and then features a sequence of jumps and stops with stops at y_1, y_2, \dots, y_l . (By the arguments in Proposition 8, this sequence cannot be infinite.) If Condition B2.3 holds at y_1, \dots, y_l , we are done. If not, suppose WLOG that it first fails at y_l . \hat{m} can be perturbed near y_l to obtain $\hat{m}_2 \in B(\hat{m}, \epsilon')$ that satisfies Condition B2.3 at y_1, \dots, y_l . Similarly, if \hat{m}_2 first fails Condition B2.3 at some $y_{l'} > y_l$, we can construct $\hat{m}_3 \in B(\hat{m}, \epsilon')$, a perturbation of \hat{m}_2 near $y_{l'}$, that satisfies Condition B2.3 up to $y_{l'}$. If this process stops in a finite number of steps, we are done. If not, let \hat{m}_∞ be the pointwise limit of $(\hat{m}_k)_k$. \hat{m}_∞ must feature an infinite sequence of vertex points $y_1 < y_2 < \dots$ with $y_l \xrightarrow{l \rightarrow +\infty} y^* \leq y$, but, as $\hat{m}_\infty \in Y_5(y)$, \hat{m}_∞ satisfies Condition B2.2, leading to a contradiction. \square

Corollary 3. *The set $Y_3 \subseteq X_3$ of functions for which B2.2 and B2.3 hold in $[x^*, x^{**}]$ is a residual set.*

Claim 5. *In the case of quadratic utility, the set of functions m for which B2.2 holds for $x \in (x^*, x^* + d)$ is prevalent.*

Proof. The result follows from Theorem 3 in Hunt et al. (1992). Following their notation, take $M = \{y \in \mathbb{R}^4 : \frac{2y_3-1}{y_1-y_2} + \frac{y_4}{y_3} = 0\}$ and $Z = \{y \in \mathbb{R}^5 : (y_1, y_2, y_3, y_4) \in M \text{ and } \frac{2y_4(y_1-y_2)-2y_3(1-y_3)}{(y_1-y_2)^2} + \frac{y_5y_3-y_4^2}{y_3^2} = 0\}$. We need to check that M is a manifold of codimension 1, and that the projection $\pi : M \rightarrow \mathbb{R}$ given by $y \mapsto y_1$ is a submersion; both follow from the Implicit Function Theorem. Finally, we need to check that Z is a zero set in $M \times \mathbb{R}$, which can also be shown using the Implicit Function Theorem. Theorem 3 from Hunt et al. (1992) then implies that the set of functions m for which there is an x such that

$$\begin{aligned} e(x) &= \frac{2m'(x) - 1}{x - m(x)} + \frac{m''(x)}{m'(x)} = 0 = \\ &= \frac{2m''(x)(x - m(x)) - 2m'(x)(1 - m'(x))}{(x - m(x))^2} + \frac{m'''(x)m'(x) - m''(x)^2}{m'(x)^2} = e'(x) \end{aligned}$$

is shy, i.e., its complement is prevalent. \square

I conjecture that B2.2 and B2.3 hold in a prevalent set even for general utility functions, but this is hard to prove.

D Other Equilibria (For Online Publication)

The discrete time model in Section 2 may admit MVEs other than 1-equilibria. We discuss two possible types here: k -equilibria (k E), which are composed of k interleaved sequences, and continuous equilibria. For brevity, we present our analysis for the model in Section 5.

Definition 4. Let s be a MVE on $[x^*, x^{**}]$. s is a k -equilibrium if there is a sequence $(x_n)_{n \in \mathbb{Z}}$ such that $x_{n+1} < x_n$ for all n , $x_n \xrightarrow{n \rightarrow -\infty} x^{**}$, $x_n \xrightarrow{n \rightarrow \infty} x^*$, and $s(x) = x_{n+k}$ if $x \in [x_n, x_{n-1})$. s is a continuous equilibrium if it is an MVE and continuous.

Figure 5 shows a 2E (5b) compared to a 1E (5a). Although k Es and continuous equilibria do not exhaust the set of possible equilibria, studying them sheds light on the general behavior of non-1Es. Our main conclusion will be that the existence of these equilibria is not robust in any sense analogous to what is shown in Propositions 5 and 6 for 1Es. This is why the paper does not focus on them.

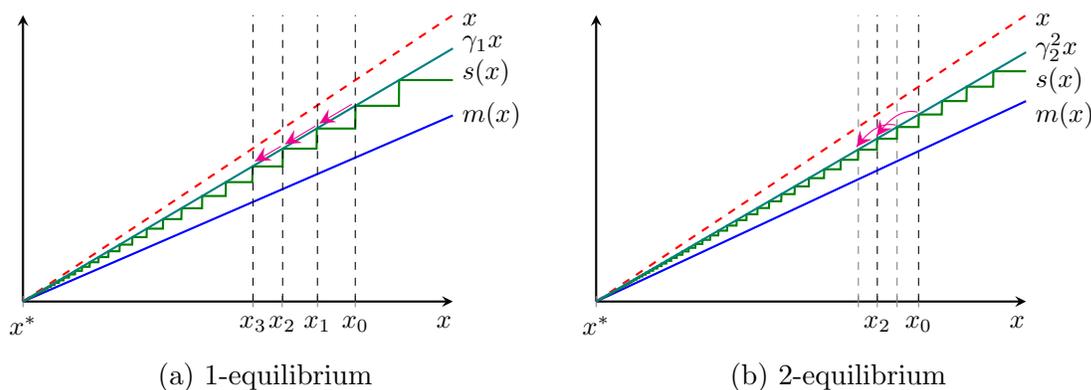


Figure 5: Equilibria for $u_\alpha(x) = C - (\alpha - x)^2$, $m(x) = 0.7x$, $\delta = 0.7$

We first note that, when m is linear⁵² and u is quadratic, we can explicitly find k Es for all k , as well as a continuous equilibrium.

⁵²We can construct densities f such that $m(x) = ax$ for $x \in [-d, d]$. For example, for a continuous f symmetric around $x = 0$, take $f(y) = 1 - \frac{1-a}{d}y$ for $y \in [0, d]$ and $f(y) = a + (1-a)(2a^2 + 1) - \frac{(1-a)(2a^2 + 1)}{d}y$ thereafter.

Proposition 9. Let $u_\alpha(x) = C - (\alpha - x)^2$ and $m(x) = ax$ for $x \in [-d, d]$. Assume $\delta \geq \frac{2}{3}$ and $a \in (\frac{1}{2}, 1)$. Then, for each k and $\underline{x} < d$, there is a kE s_k restricted to $[-d, d]$ such that $x_0 = \underline{x}$, given by $x_n = \gamma_k^n \underline{x}$, where $\gamma_k \in (0, 1)$. There is also a continuous equilibrium s_∞ given by $s_\infty(x) = \gamma_\infty x$. γ_k^k is decreasing in k and $\gamma_k^k \rightarrow \gamma_\infty$.

Proof of Proposition 9. Given $k \geq 1$, assume a kE of the form $s(x_n) = \gamma_k^k x_n$. Since $s(x_n) = x_{n+k}$ but $s(x_n - \epsilon) = x_{n+k+1}$, $m(x_n)$ must be indifferent between x_{n+k} and x_{n+k+1} . This implies

$$\begin{aligned} -\sum \delta^t (ax_n - x_{n+(t+1)k})^2 &= -\sum \delta^t (ax_n - x_{n+(t+1)k+1})^2 \\ \sum \delta^t (a - \gamma^{(t+1)k})^2 &= \sum \delta^t (a - \gamma^{(t+1)k+1})^2 \\ \frac{\gamma^k(1+\gamma)}{1-\delta\gamma^{2k}} &= \frac{2a}{1-\delta\gamma^k} \end{aligned}$$

We now argue that the expression $\frac{\gamma^k - \delta\gamma^{2k}}{1 - \delta\gamma^{2k}}$ is increasing in γ . It is equivalent to show that $\frac{x - \delta x^2}{1 - \delta x^2}$ is increasing in x for $x \in [0, 1]$. This is true because $x - \delta x^2 \leq 1 - \delta x^2$ but $1 - 2\delta x > -2\delta x$, so the log-derivative of $\frac{x - \delta x^2}{1 - \delta x^2}$ is positive. Hence, $\frac{(\gamma^k - \delta\gamma^{2k})(1+\gamma)}{1 - \delta\gamma^{2k}}$ is also increasing in γ , and equals $2a$ for a unique value of γ which we denote $\gamma_k \in (0, 1)$.

Denote $A_k(x) = \frac{(x - \delta x^2)(1 + x^{\frac{1}{k}})}{1 - \delta x^2}$. Then $A_k(\gamma_k^k) = 2a$ for all k , and $A_k(x)$ is increasing in x and k . It follows that γ_k^k is decreasing in k .

We now show that the constructed s_k supports an MVE. By the same argument given in Proposition 5 for 1Es, since $m(x_n)$ is indifferent between $S(x_{n+k})$ and $S(x_{n+k+1})$, all $\alpha > m(x_n)$ strictly prefer $S(x_{n+k})$ to $S(x_{n+k+1})$, and $\alpha < m(x_n)$ strictly prefer $S(x_{n+k+1})$. Hence, $m(x_n)$ prefers $S(x_{n+k})$ to $S(x_r)$ for all $r \neq n + k$.

Next, we show that $m(x_n)$ prefers x_{n+k} to other policies $x \notin (x_n)_n$. We do this in two steps. First, we argue that $\gamma^{k+1} > a$, which implies $x_{n+k+1} > m(x_n)$. Second, we note that this yields our result by the same argument as in Proposition 5. For the first part, note that $\gamma^{k+1} > a$ iff

$$\begin{aligned} (\gamma^k + \gamma^{k+1})(1 - \delta\gamma^k) &< 2\gamma^{k+1}(1 - \delta\gamma^{2k}) \\ \iff (1 - \gamma) &< \delta(\gamma^k(1 - \gamma^{k+1}) + \gamma^{k+1}(1 - \gamma^k)) \\ \iff 1 &< \delta(\gamma^k + 2\gamma^{k+1} + \dots + 2\gamma^{2k}) \end{aligned}$$

Note that A_k is decreasing in δ , so γ is increasing in δ and a . Hence the worst case is $\delta = \frac{2}{3}$, $a = \frac{1}{2}$. Now suppose $k = 1$. Then the required inequality is $1 < \frac{2}{3}(\gamma + 2\gamma^2)$,

which holds if $\gamma \geq \frac{2}{3}$, so it is enough to verify $1 > A_1(\frac{2}{3}) = \frac{50}{57}$. If $k \geq 2$, then it is enough to satisfy $1 < \frac{2}{3}(\gamma^k + 4\gamma^{2k})$, which is true if $\gamma^k \geq \frac{1}{2}$. We then check that $1 > A_k(\frac{1}{2})$. Because A_k is increasing in k , it is enough to check $1 > \lim_{k \rightarrow \infty} A_k(\frac{1}{2}) = \frac{4}{5}$.

Finally, we construct a continuous equilibrium. In general, s must solve

$$\begin{aligned} s(x) &= \arg \max_y \sum_{t=0}^{\infty} \delta^t (C - (m(x) - s^t(y))^2) \\ \implies 0 &= \sum_{t=0}^{\infty} \delta^t \left(-2(m(x) - s^t(y)) \prod_{i=0}^{t-1} s'(s^i(y)) \right) \end{aligned}$$

if s is smooth. Since $m(x) = ax$, we look for a solution of the form $s_{\infty}(x) = \gamma x$:

$$\sum_{t=0}^{\infty} \delta^t \left((a - \gamma^{t+1}) \prod_{i=0}^{t-1} \gamma \right) = \sum_{t=0}^{\infty} \delta^t ((a - \gamma^{t+1})\gamma^t) = 0,$$

whence $\frac{a}{1-\delta\gamma} = \frac{\gamma}{1-\delta\gamma^2}$. By similar arguments as before, this equation has a unique solution $\gamma_{\infty} \in (0, 1)$ and $\gamma_k^k \rightarrow \gamma_{\infty}$ because $A_k(x) \xrightarrow[k \rightarrow \infty]{} \frac{x-\delta x^2}{1-\delta x^2}$. Finally, $\frac{\partial U_{m(x)}(S(y))}{\partial y} |_{y=y_0} > 0 (<)$ for $y_0 < s(x)$ ($>$) follows from combining Assumption A2 with the fact $\frac{\partial U_{m(x)}(S(y))}{\partial y} |_{y=s(x)} = 0 \forall x$. Hence $y = s(x)$ maximizes $U_{m(x)}(S(y))$. \square

In the general case, however, k Es for $k > 1$ and continuous equilibria may not exist. The issue is the following. Suppose that a k E s_k exists in a right-neighborhood of a stable steady state, $[x^*, x^* + \epsilon)$ (even this is not guaranteed in general). s_k can then be extended at least to $[x^*, x^* + d_{x^*}^+]$ (Lemma 10), but its extension may fail to be a k E. Similarly, the unique extension of a continuous equilibrium may have discontinuities; whether this happens depends on arbitrarily small details of m .

Here is an intuition. Assume $u_{\alpha}(x) = C - (\alpha - x)^2$, $\delta > \frac{2}{3}$, $a > 0.5$, and $\tilde{m}(x) = ax + \frac{a}{4} \max(c - |x - x'|, 0)$, where $c > 0$ is small. We are in the linear case, except \tilde{m} has a small ‘‘bump’’ around x' . Let s be a 2E for $m(x) = ax$ such that $x_0 = x'$, and let \tilde{s} be a 2E for \tilde{m} such that $\tilde{x}_n = x_n$ for $n > 0$. As $m(x_0)$ and $\tilde{m}(\tilde{x}_0)$ must both be indifferent between $S(x_2)$ and $S(x_3)$, $m(x_0) = \tilde{m}(\tilde{x}_0)$, but $\tilde{m}(x_0) > m(x_0)$, so $\tilde{x}_0 < x_0$. Meanwhile $\tilde{x}_{-1} = x_{-1}$. But $m(\tilde{x}_{-2})$, being indifferent between $\tilde{S}(\tilde{x}_0)$ and $S(x_1)$, must be lower than $m(x_{-2})$ because \tilde{x}_0 being lower makes the former path more attractive than $S(x_0)$, so \tilde{x}_{-2} is lower. On the other hand $\tilde{x}_{-3} > x_{-3}$ because it is defined by indifference between $S(x_{-1})$ and $\tilde{S}(\tilde{x}_0)$ (more attractive than $S(x_0)$). Continuing in

this fashion, the subsequence $(\tilde{x}_0, \tilde{x}_{-2}, \tilde{x}_{-4}, \dots)$ is lower than (x_0, x_{-2}, \dots) , and the opposite is true for the odd elements. Eventually $\tilde{x}_{2l} < \tilde{x}_{2l+1}$ for some l , i.e., the even subsequence becomes so attractive that a voter $m(\tilde{x}_{2l+1})$, though indifferent between $\tilde{S}(\tilde{x}_{2l+3})$ and $\tilde{S}(\tilde{x}_{2l+4})$, instead prefers $\tilde{S}(\tilde{x}_{2l+2})$ to both, so no one votes for x_{2l+3} .

k Es for $k > 1$ are unstable for this reason. We now give a local argument to this effect. Let $E_n = (1 - \delta) \sum_{t=0}^{\infty} \delta^t x_{n+tk}$, $W_n = (1 - \delta) \sum_{t=0}^{\infty} \delta^t x_{n+tk}^2$, and characterize a k E recursively as follows, using the indifference of $m(x_n)$ between S_{n+k} and S_{n+k+1} , i.e., $-(m(x_n) - E_{n+k})^2 - W_{n+k} = -(m(x_n) - E_{n+k+1})^2 - W_{n+k+1}$:

$$E_n = (1 - \delta)x_n + \delta E_{n+k} = (1 - \delta)m^{-1} \left(\frac{1}{2} \frac{W_{n+k} - W_{n+k+1}}{E_{n+k} - E_{n+k+1}} \right) + \delta E_{n+k}$$

$$W_n = (1 - \delta)x_n^2 + \delta W_{n+k} = (E_n - \delta E_{n+k})x_n + \delta W_{n+k}$$

Taking $Y_n = (E_n, \dots, E_{n+k+1}, W_{n+1}, \dots, W_{n+k+1})$ as the state variable of the recursion, its linearization around an equilibrium is given by $Y_n = M_n Y_{n+1}$, where

$$M_n = \begin{pmatrix} 0 & \dots & 0 & A & B & 0 & \dots & 0 & C & D \\ 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 2x & 0 & \dots & 0 & -2\delta x & 0 & 0 & \dots & 0 & \delta \\ 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix} \begin{matrix} \left. \vphantom{\begin{matrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 2x \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{matrix}} \right\} 1 \\ \left. \vphantom{\begin{matrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 2x \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{matrix}} \right\} k \\ \left. \vphantom{\begin{matrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 2x \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{matrix}} \right\} 1 \\ \left. \vphantom{\begin{matrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 2x \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{matrix}} \right\} k-1 \end{matrix}$$

where $x = x_{n+1}$; $B = \frac{\partial E_n}{\partial E_{n+k+1}}$, $D = \frac{\partial E_n}{\partial W_{n+k+1}}$ and so on. Note that

$$\det(M_n) = -\delta B - 2\delta x_{n+1} D = \delta(1 - \delta) \frac{x_{n+1} - m(x_n)}{m'(x_n)(E_{n+k} - E_{n+k+1})}$$

$$\begin{aligned}
\det(M_n \dots M_{n-k+1}) &\geq \delta^k (1 - \delta)^k \left[\min_{0 \leq l \leq k-1} \left(\frac{x_{n-l+1} - m(x_{n-l})}{m'(x_{n-l})} \right) \right]^k \frac{1}{\prod_{l=0}^{k-1} (E_{n-l+k} - E_{n-l+k+1})} \\
&\geq \delta^k (1 - \delta)^k \left[\min_{0 \leq l \leq k-1} \left(\frac{x_{n-l+1} - m(x_{n-l})}{m'(x_{n-l})} \right) \right]^k \frac{k^k}{(E_{n+1} - E_{n+k+1})^k} \\
&= \delta^k \left[\min_{0 \leq l \leq k-1} \left(\frac{x_{n-l+1} - m(x_{n-l})}{m'(x_{n-l})} \right) \right]^k \frac{k^k}{(x_{n+1} - E_{n+k+1})^k}
\end{aligned}$$

where we have used the AM-GM inequality. Now, if δ is close to 1 and the equilibrium is close to a CLS in the sense of Proposition 6, then $\frac{x-m(x)}{m'(x)(x-E(S(s(x))))} \approx 1$ (see Appendix B) and $\det(M_n \dots M_{n-k+1}) \approx k^k$. (In particular, these statements hold with equality for the linear case we have solved above.) Hence there must be an eigenvalue of absolute value at least $k^{\frac{k}{2k+1}} > 1$. Thus any deviation from an equilibrium resulting from a local perturbation of m which adds a nonzero component to a generalized eigenvector of this eigenvalue (in the Jordan form decomposition of the matrix) will grow exponentially.

Similarly, in the example given above of a linear m with a bump, a continuous equilibrium constructed up to $x' - c$ would have a discontinuity at $s^{-1}(x')$. This insight extends more generally. For brevity, we show this for smooth (C^∞) equilibria.

Proposition 10. *Assume $u_a(x) = C - (\alpha - x)^2$. Let $s : [x^*, x^{**}] \rightarrow [x^*, x^{**}]$ be a smooth equilibrium for a given m . Let $x_0 \in (x^*, x^{**})$. A perturbation \tilde{m} of m is an increasing function $\tilde{m} = m + \rho\kappa$ where $\kappa : [x^*, x^{**}] \rightarrow [x^*, x^{**}]$ has support $(x_0 - \epsilon, x_0 + \epsilon)$. For each \tilde{m} , let \tilde{s} be an equilibrium under \tilde{m} such that $\tilde{s}|_{[x^*, x_0 - \epsilon]} = s|_{[x^*, x_0 - \epsilon]}$.*

Suppose m is C^∞ . Then, if κ is C^k but its $(k+1)$ th derivative has a discontinuity in $(x_0 - \epsilon, x_0 + \epsilon)$, \tilde{s} has a discontinuity in $[x^, s^{-k-1}(x_0 + \epsilon)]$ for all $\rho \neq 0$.*

Proof. Denote $E(y) = E(S(y))$, $W(y) = W(S(y))$, and $\frac{1}{2} \frac{\partial W(y)}{\partial E(y)} = L^{-1}(y)$. Then

$$s(x) = \arg \max_y -m(x)^2 + 2m(x)E(y) - W(y) \Rightarrow m(x) = \frac{1}{2} \frac{\partial W(y)}{\partial E(y)} \Big|_{s(x)}$$

$$s(x) = \left(\frac{1}{2} \frac{\partial W(y)}{\partial E(y)} \right)^{-1} (m(x)) = L(m(x))$$

$$(W, E)(x) = ((1 - \delta)x^2 + \delta W(L(m(x))), (1 - \delta)x + \delta E(L(m(x))))$$

In particular, $W(E)$ must be a strictly convex function so that s is surjective, and it must have no kinks, i.e., s must be strictly increasing (if s is locally constant at x ,

it will be discontinuous at $s^{-1}(x)$ as long as $s(y) > m(y)$ in this area), so L^{-1} and L are strictly increasing and well-defined.

Now suppose E is C^{l+1} around $s(x)$ but s has a $(l+1)$ -kink at $s(x)$, i.e., it is C^{l+1} in $(s(x) - \eta, s(x)) \cup (s(x), s(x) + \eta)$ but only C^l in $(s(x) - \eta, s(x) + \eta)$. Then s' has a l -kink at $s(x)$. Since

$$\begin{aligned} \frac{\partial W(y)}{\partial E(y)} &= \frac{\frac{\partial W(y)}{\partial y}}{\frac{\partial E(y)}{\partial y}} = \frac{2(1-\delta)y + \delta W'(s(y))s'(y)}{(1-\delta) + \delta E'(s(y))s'(y)} \\ &= \frac{W'(s(y))}{E'(s(y))} + \frac{(1-\delta)(2y - \frac{W'(s(y))}{E'(s(y))})}{1-\delta + \delta E'(s(y))s'(y)} = 2m(y) + \frac{(1-\delta)(2y - 2m(y))}{1-\delta + \delta E'(s(y))s'(y)}, \end{aligned}$$

L^{-1} has a l -kink at $s(x)$; L has a l -kink at $m(x)$; and s has a l -kink at x .

Then, if κ has a $(k+1)$ -kink at x , then \tilde{m} has a $(k+1)$ -kink at x ; \tilde{s} has a $(k+1)$ -kink at x ; \tilde{s} has a k -kink at $\tilde{s}^{-1}(x)$; ... and \tilde{s} has a discontinuity at $\tilde{s}^{-k-1}(x)$. \square

A similar result holds even for smooth perturbations.

Proposition 11. *Assume $u_a(x) = C - (\alpha - x)^2$. Let $s : [x^*, x^{**}] \rightarrow [x^*, x^{**}]$ be a smooth equilibrium for a given m with $m' \geq A > \frac{1}{2}$ everywhere. Let $z \in (x^*, x^{**})$ and $z' = s(z)$. Given a smooth function $\kappa : [x^*, x^{**}] \rightarrow [x^*, x^{**}]$ with support contained in $[z', z]$ and $\rho \in \mathbb{R}$, let $m_\rho = m + \rho\kappa$, and let s_ρ be the unique equilibrium under m_ρ that equals s within $[x^*, z']$, if it exists. Then there exists κ such that s_ρ is discontinuous for all $\rho \neq 0$.*

Proof. Pick $K \in \mathbb{N}$ and define a set of $2K$ sequences (y_{ni}) ($n \in \mathbb{Z}$, $i \in \{1, \dots, 2K\}$) as follows. $y_{01} = z'$, $y_{0(2K)} = z$, $(y_{0i})_i$ is an arithmetic sequence, and, for all n , i , $y_{(n+1)i} = s(y_{ni})$.

Let $\tilde{\kappa}$ be a non-negative C^∞ function with support $[0, 1]$ such that $\tilde{\kappa}(\frac{1}{2}) = 1$. Define $\kappa(x) = \sum_{i \leq 2K \text{ odd}} \tilde{\kappa}\left(\frac{x - y_{0i}}{y_{0(i+1)} - y_{0i}}\right)$, so that κ has a copy of $\tilde{\kappa}$ "squeezed" into each interval $[y_{0i}, y_{0(i+1)}]$ for i odd. Any such κ will work.

We write $W(E)$ to denote $W(E^{-1}(E))$, W_ρ to denote W with perturbation $\rho\kappa$, etc. Assume $\rho > 0$. It is easy to show that $s_\rho(x) \equiv s(x)$ for $x \in [y_{ni}, y_{n(i+1)}]$ for all n and even i , and $s_\rho(x) > s(x)$ for $x \in [y_{0i}, y_{0(i+1)}]$ for odd i (because $s_\rho(x) = s(y)$ whenever $m_\rho(x) = m(y)$). In addition, and $W_\rho(E) < W(E)$ for $x \in [E(y_{0i}), E(y_{0(i+1)})]$ for odd i , since $\frac{\partial W}{\partial E}(y) = 2m(y)$ but, for a path $(x_t)_t$, decreasing x_0 by ϵ decreases $E((x_t)_t)$ by $(1-\delta)\epsilon$ and $W((x_t)_t)$ by $(1-\delta)\epsilon(2x_0 - \epsilon)$, that is, $\frac{\Delta W}{\Delta E} \approx 2x_0 > 2m(x_0)$.

For odd i , let $\Delta W_{0i} = \max_{E \in [E(y_{0i}), E(y_{0(i+1)})]} [W(E) - W_\rho(E)]$, and denote by \hat{E}_{0i} the argmax. Then there must be a point $\hat{E}_{0i} \in [E(y_{0i}), E(y_{0(i+1)})]$ for which $\frac{\partial W}{\partial E}(\hat{E}_{0i}) - \frac{\partial W_\rho}{\partial E}(\hat{E}_{0i}) \geq \frac{\Delta W_{0i}}{E(y_{0(i+1)}) - E(y_{0i})}$.

Let \acute{y} be such that $2m(\acute{y}) = \frac{\partial W}{\partial E}(\hat{E}_{0i})$ and \grave{y} be such that $2m(\grave{y}) = \frac{\partial W_\rho}{\partial E}(\hat{E}_{0i})$. Then $\acute{y} - \grave{y} \geq \frac{1}{2\bar{m}'} \frac{\Delta W_{0i}}{E(y_{0(i+1)}) - E(y_{0i})}$. Then, denoting $\dot{E} = E_\rho(\grave{y})$ and $\acute{E} = E_\rho(\acute{y})$, $W_\rho(\dot{E}) = (1 - \delta)\grave{y}^2 + \delta W_\rho(\hat{E}_{0i}) \leq (1 - \delta)\acute{y}^2 + \delta W(\hat{E}_{0i})$, and $W(\dot{E}) \geq W(\acute{E}) - (\acute{E} - \dot{E}) \frac{\partial W}{\partial E}(\acute{E}) = (1 - \delta)\acute{y}^2 + \delta W(\hat{E}_{0i}) - (\acute{E} - \dot{E})2m(\acute{y})$. Also note that $\grave{y} > s(\acute{y}) > m(\acute{y})$ and $y - E(s(y)) \geq 2(y - m(y))$ for all y . Hence

$$\begin{aligned} \Delta W_{1i} &\geq W(\dot{E}) - W_\rho(\dot{E}) \geq (1 - \delta)(\acute{y} - \grave{y})(\acute{y} + \grave{y} - 2m(\acute{y})) \\ &\geq \frac{1 - \delta}{2\bar{m}'} \left[\frac{\acute{y} - m(\acute{y})}{E(y_{0(i+1)}) - E(y_{0i})} \right] \Delta W_{0i} \\ \prod_i \Delta W_{1i} &\geq \left[\frac{1 - \delta}{2\bar{m}'} \right]^K \prod_i \left[\frac{\acute{y}_i - m(\acute{y}_i)}{E(y_{0(i+1)}) - E(y_{0i})} \right] \prod_i \Delta W_{0i} \\ &\geq \left[\frac{(1 - \delta)K}{2\bar{m}'} \right]^K \prod_i \left[\frac{\acute{y}_i - m(\acute{y}_i)}{E(y_{0(2K)}) - E(y_{01})} \right] \prod_i \Delta W_{0i} \\ &\geq \left[\frac{K}{4\bar{m}'} \right]^K \prod_i \left[\frac{\acute{y}_i - m(\acute{y}_i)}{y_{0(2K)} - m(y_{0(2K)})} \right] \prod_i \Delta W_{0i} \geq \left[\frac{K}{8\bar{m}'} \right]^K \prod_i \Delta W_{0i} \end{aligned}$$

If we choose $K > 8\bar{m}'$, iterating this argument, we find that there must be n, i (possibly functions of ρ) for which $\Delta W_{ni} > (x^{**} - x^*)^2$, a contradiction. \square

It follows that, if $m|_{[x^*, z]}$ admits a smooth equilibrium s , then the set of extensions of m to $[x^*, x^{**})$ that admit a smooth extension of s to $[x^*, x^{**})$ is shy. I conjecture that $\forall \epsilon > 0$, the set of m 's admitting a smooth equilibrium on $[x^*, x^* + \epsilon]$ is also shy.

Non-Monotonic Equilibria

Proposition 3 shows that MVEs must be monotonic in a neighborhood of a stable steady state, and 1Es are monotonic everywhere. However, non-monotonic MVEs may exist; we provide an example here. Assume that $u_\alpha(x) = C - (\alpha - x)^2$ and let $d = \sqrt{C} = d_x^- = d_x^+$ for all x . In addition, suppose that $m(x) = x - \rho d$ for all $x \in \mathbb{R}$, where $\rho \in [\frac{1}{2}, 1)$ is a parameter.⁵³ For simplicity we will take the MVT as a primitive,

⁵³ m can be obtained as a median voter function if $f(x) = ke^{-\hat{\rho}x}$ for an appropriately chosen $\hat{\rho} > 0$. This is a degenerate example in the sense that there is an infinite mass of voters distributed on the

i.e., we will analyze the game in which $m(x)$ picks $s(x)$.

Assume $\delta = 0$. Then $s(x) = m(x)$, so $S(s(x)) = (x - \rho d, x - 2\rho d, x - 3\rho d, \dots)$. Crucially $x - \rho d, x - 2\rho d \in (m(x) - d, m(x) + d)$ but $x - 3\rho d \notin (m(x) - d, m(x) + d)$.

Now, suppose that δ is small but positive. Assume a successor function s_1 of the form $s_1(x) = x - \rho d + \eta_1$ with η_1 small, such that $s_1^2(x) > m(x) - d$ but $s_1^3(x) < m(x) - d$. For s_1 to be an equilibrium, η_1 must maximize

$$\begin{aligned} u_{m(x)}(x - \rho d + \eta) + \delta u_{m(x)}(s_1(x - \rho d + \eta)) &= C(1 + \delta) - \eta^2 - \delta(\rho d - \eta_1 - \eta)^2 \\ \implies \eta_1 &= \frac{\delta}{1 + 2\delta} \rho d. \end{aligned}$$

Note that, since $\eta_1 > 0$, this calculation will be invalid for ρ close enough to $\frac{1}{2}$, as in fact we will then have $s_1^3(x) > m(x) - d$.

Next, assume a successor function $s_2(x) = x - \rho d + \eta_2$ with η_2 such that $s_2^3(x) > m(x) - d$ but $s_2^4(x) < m(x) - d$. For s_2 to be an equilibrium, we must have

$$\begin{aligned} \eta_2 &= \arg \max_{\eta'} u_{m(x)}(x - \rho d + \eta') + \delta u_{m(x)}(s_2(x - \rho d + \eta')) + \delta^2 u_{m(x)}(s_2^2(x - \rho d + \eta')) \\ \implies \eta_2 &= \frac{\delta + 2\delta^2}{1 + 2\delta + 3\delta^2} \rho d. \end{aligned}$$

Now suppose that s_2 is being played, and x considers deviating to some η_3 such that $s_2^2(x - \rho d + \eta_3) < m(x) - d$. The (locally) optimal η_3 must satisfy

$$-2\eta_3 - 2\delta(\eta_3 + \eta_2 - \rho d) = 0 \implies \eta_3 = \frac{\delta}{1 + \delta}(\rho d - \eta_2) = \frac{\delta}{1 + \delta} \frac{1 + \delta + \delta^2}{1 + 2\delta + 3\delta^2} \rho d.$$

Note that $\eta_2 > \eta_1 > \eta_3 > 0$. In particular, since $\eta_2 > \eta_3$, we can choose ρ, δ so that $x - 3\rho d + 3\eta_2 > m(x) - d > x - 3\rho d + \eta_3 + 2\eta_2$. Furthermore, we can choose them so that $U_{m(x)}(S_2(x - \rho d + \eta_3)) = U_{m(x)}(S_2(x - \rho d + \eta_2))$.⁵⁴ If we choose our parameters this way then $m(x)$ is indifferent about deviating (this is true for all x). Now construct a successor function s as follows: $s(x) = s_2(x)$ for all $x < x_0$; $s(x_0) = x - \rho d + \eta_3$; and $s(x)$ for $x > x_0$ is defined by backward induction. This is a non-monotonic equilibrium by construction. While this example relies on indifference,

real line, as opposed to a unit mass with support $[-1, 1]$, but it allows for a simpler construction.

⁵⁴This follows from a continuity argument: when $s_2^2(x - \rho d + \eta_3) = m(x) - d$, $U_{m(x)}(S_2(x - \rho d + \eta_3)) < U_{m(x)}(S_2(x - \rho d + \eta_2))$, whereas when $s_2^2(x - \rho d + \eta_2) = m(x) - d$ then $U_{m(x)}(S_2(x - \rho d + \eta_3)) > U_{m(x)}(S_2(x - \rho d + \eta_2))$, so we can choose intermediate values of ρ, δ for which $U_{m(x)}(S_2(x - \rho d + \eta_3)) = U_{m(x)}(S_2(x - \rho d + \eta_2))$.

we can adjust m to make it strict.

E Additional Results (For Online Publication)

E.1 Supermajority Requirements and Other Decision Rules

We assume as part of our solution concept that chosen policies are Condorcet winners. The analysis readily extends to other decision rules. We briefly discuss two.

First, suppose that, given a policy x and a set of members $I(x)$, an unmodeled political process gives some agent $n(x)$ the right to choose tomorrow's policy. (For example, the function $n(x)$ might reflect the notion that the policy x affects the relative power of different agents within the organization.) All the results extend to this case, substituting $n(x)$ for $m(x)$, even if $n(x)$ is not a median voter function.

Second, consider an organization with a bias towards inaction in which policy changes require a supermajority $\rho > \frac{1}{2}$. Define $m_\rho(x)$ as the ρ th percentile-member of $I(x)$, and assume that a policy $y > x$ can only be chosen over the current policy x if $m_{1-\rho}(x)$ votes for it, while a change to $y < x$ is only possible if m_ρ votes for it. It follows that, in intervals where $x > m_\rho(x)$, the game is equivalent to the main model with $n(x) = m_\rho(x)$; in intervals where $x < m_{1-\rho}(x)$, it is equivalent to setting $n(x) = m_{1-\rho}(x)$; and in intervals where $m_{1-\rho}(x) < x < m_\rho(x)$, no policy changes are possible. In other words, steady states are now intervals rather than points, and we will observe lower policy drift, but the gist of the results is unchanged.

E.2 Positive Entry and Exit Costs

The main model assumes free entry and exit. This assumption adds a lot of tractability, but is rarely exactly true in a descriptive sense. Here, we demonstrate that our results are also relevant in a setting with positive entry and exit costs. We begin by considering a variant of the game with only entry costs: every time an outsider chooses to join the club, she must pay a cost $c > 0$.

A full extension of the results to this case is difficult because the introduction of entry costs adds intertemporal concerns to entry and exit decisions: agents considering entry now need to think about what the policy will be several periods from now, whether they will want to leave later, etc. Relatedly, agents with identical preferences may behave differently depending on their current status: if the policy x is stable over

time, and an agent α 's flow payoff from membership, $u_\alpha(x)$, is positive but very small, then α would choose to remain in the club if she is already a member but not bother entering otherwise. As a result, the club's current policy is no longer the only payoff-relevant state variable; I_t is now an (infinite-dimensional) state variable as well.

In spite of this, the main thrust of the paper's results—namely, that the club should converge to a myopically stable policy—still carries over in this model if we impose some reasonable simplifications. Concretely, we will assume the following:

- (i) As in Section 4, we restrict the analysis to $[x^*, x^{**}]$, the right side of the basin of attraction of a stable steady state x^* .
- (ii) Assume $x_{t+1} \leq x_t$ for all $x \in [x^*, x^{**}]$, i.e., the policy cannot move to the right.
- (iii) Assume an initial $x_0 \in [x^*, x^{**}]$ such that $u_{x_0-d_{x_0}}(x^*) \geq (1-\delta)c$. In other words, x_0 is close enough to x^* that all agents who might consider joining the club as the policy moves left from x_0 will strictly prefer not to quit later.⁵⁵
- (iv) Assume an initial set of members $I_0 = I(x_0)$.
- (v) Voting behavior at time t can condition on I_t only up to a set of measure zero, i.e., if I_1, I_2 differ by a set of Lebesgue measure zero then $s(I_1) = s(I_2)$.

An MVE of this game is given by mappings $s(I)$ and $I(x, I')$ satisfying the above conditions such that I reflects optimal entry and exit decisions given current policy x , an existing set of members I' and the expected continuation; and $S(s(I))$ is a Condorcet winner among all $S(y)$ for the set of voters I .

It turns out that the set of equilibria of this game corresponds exactly to the set of equilibria of a game with free entry and exit but modified utility functions. Let $G(u, c)$ denote the game just described, with cost of entry c and utility functions $u_\alpha(x)$. Let $G(v, 0)$ denote the game with free entry and exit and utility functions $v_\alpha(x)$ given by $v_\alpha = u_\alpha$ for $\alpha \geq x_0 - d_{x_0}$ and $v_\alpha = u_\alpha - (1 - \delta)c$ for $\alpha < x_0 - d_{x_0}$.

Proposition 12. *In any MVE (s, I) of $G(u, c)$, $I(x, I') = I_v(x)$ for all (x, I') on the equilibrium path.*

For any MVE (s, I) of $G(u, c)$, there is an MVE \tilde{s} of $G(v, 0)$ given by $\tilde{s}(x) = s(I_v(x))$. Conversely, for any MVE \tilde{s} of $G(v, 0)$, there is an MVE (s, \tilde{I}) of $G(u, c)$ given by $s(I) = \tilde{s}(m_v(I))$ and $\tilde{I}(x, I') = I_v(x)$.

⁵⁵Effectively, this means we find equilibria restricted to $[x^*, x_0]$.

Proof. For the first claim, suppose $(x_t, I_t)_t$ is an equilibrium path, and assume that $I_t = I_v(x_t)$. We aim to show that $I_{t+1} = I_v(x_{t+1})$.

There are four types of agents to consider. First, suppose $\alpha \notin I_t$ and $\alpha > x_t$, i.e., α is an outsider with a policy preference to the right of x_t . Then $\alpha \notin I_v(x_t)$, i.e., $u_\alpha(x_t) < 0$, so $u_\alpha(x_{t+1}) \leq u_\alpha(x_t) < 0$, whence $\alpha \notin I_v(x_{t+1})$. Moreover, since $u_\alpha(x_s) < 0$ for all $s \geq t$, α should not join the club at time $t + 1$, i.e., $\alpha \notin I_{t+1}$.

Second, suppose $\alpha \in I_t$ and $\alpha \geq x_0 - d_{x_0}$. Then α is an incumbent member at time $t + 1$, and will choose to remain a member iff $u_\alpha(x_{t+1}) \geq 0$, which is also the condition that determines whether $\alpha \in I_v(x_{t+1})$ as $u_\alpha = v_\alpha$ for this agent.

Third, suppose $\alpha \in I_t$ and $\alpha < x_0 - d_{x_0}$. Since $\alpha \in I_t$, α is an incumbent member at time $t + 1$. Since $\alpha < x_0 - d_{x_0}$, we have $u_\alpha(x_s) \geq (1 - \delta)c > 0$ for all $s \geq t$. This means both that α will choose to remain a member forever (in particular, at time $t + 1$) and that $\alpha \in I_v(x_{t+1})$.

Fourth, suppose $\alpha \notin I_t$ and $\alpha < x_t$, i.e., α is an outsider with a policy preference to the left of x_t . Then α should join at time $t + 1$ iff $u_\alpha(x_{t+1}) \geq (1 - \delta)c$, i.e., iff $\alpha \in I_v(x_{t+1})$.⁵⁶ Since $G(u, c)$ has the same membership behavior as $G(v, 0)$, the two games are equivalent, so the sets of equilibrium successor functions are the same. \square

As for the case of positive exit costs, it can be shown that, with a positive exit cost $c' > 0$, the game $G(u, c, c')$ is still equivalent to a game with free entry and exit, except that now the relevant utility functions are $v_\alpha = u_\alpha + (1 - \delta)c'$ for $\alpha \geq x_0 - d_{x_0}$ and $v_\alpha = u_\alpha - (1 - \delta)c$ for $\alpha < x_0 - d_{x_0}$.

We can apply Propositions 2 and 3 to $G(u, c, c')$ to determine the organization's long-run policy in this setting. Let $m_v(y) = m(I_v(y))$, and note that $m_v(y) \geq m_u(y)$ for all $y \in [x^*, x_0]$. Let $y^*(c, c', \delta)$ be the highest $y \in [x^*, x_0]$ for which $m_v(y^*) = y^*$. Then it follows that $x_t \rightarrow y^*$ for any equilibrium path $(x_t, I_t)_t$.

A few interesting observations can be made. First, $y^* > x^*$: the existence of entry and exit costs affects the long-run policy, as some marginal agents near $x^* - d_{x^*}^-$ never enter, or some agents near $x^* + d_{x^*}^+$ never exit. Second, y^* is a function of δ , unlike x^* ; this is because the entry and exit decisions of marginal agents now involve intertemporal trade-offs. Third, it can be shown that $y^*(c, c', \delta) \rightarrow x^*$ as $c, c' \rightarrow 0$, or as $\delta \rightarrow 1$ if we take c, c' as fixed. That is, small entry and exit costs only have a

⁵⁶In all of these arguments, α expects the equilibrium path not to change as a function of her behavior, because her joining or leaving the club amounts to a measure zero change to I_{t+1} .

small effect on the organization’s long-run policy, and even sizable costs matter less and less as agents become more patient, as they only have to be paid once.

E.3 Non-Markov Equilibria

The restriction to Markov equilibria may appear restrictive: allowing strategies to condition only on the current policy prevents agents from doling out history-dependent rewards and punishments in ways that might be plausible in some applications. This Section discusses the set of non-Markov equilibria of the game.

We make two main points. First, if non-Markov equilibria are allowed, many outcomes are possible; under some conditions, an “anything goes” result is obtained. Thus, no strong predictions can be made if we take SPE as our solution concept. Second, there are several substantively plausible perturbations of the game which rule out all non-Markov equilibria. Taken together, the results suggest that Markov equilibria are the most sensible to study in this setting.

For simplicity, we restrict our analysis in the following ways. First, as in Section 4, we restrict our analysis to $[x^*, x^{**})$, the right side of the basin of attraction of a stable steady state x^* . Second, we write the results for the framework given in Section 5, which avoids some non-essential technical issues related to entry and exit. Third, we assume the MVT as a primitive, i.e., we study a game in which, given a current policy x , $m(x)$ directly chooses tomorrow’s policy.⁵⁷ Recall our definition of reluctant agents from Section 4, and say $m(x)$ is *very reluctant* if she is reluctant and $(1 - \delta)u_{m(x)}(x') + \delta u_{m(x)}(z(x')) \leq u_{m(x)}(x)$ for all $x' \in (z(x), x)$.

Proposition 13. *If every $x \in [x^*, x^{**})$ is very reluctant, then, for every weakly decreasing path $(y_t)_{t \in \mathbb{N}_0} \subseteq [x^*, x^{**})$ such that $U_{m(y_t)}((y_{t+1}, y_{t+2}, \dots)) \geq \frac{u_{m(y_t)}(y_t)}{1-\delta}$ for all t , there is an SPE with policy path $(y_t)_t$.*

Proof. We construct a suitable successor function $s(x, T)$, where T is a payoff-irrelevant function of the history. T can take on the values 0, 1 or 2. s is defined as follows:

- (i) $s(y_t, 0) = y_{t+1}$ for all t , and $s(x, 0) = x$ for all $x \notin (y_t)_t$;
- (ii) $s(x, 1) = x$;
- (iii) $s(x, 2) = z(x)$.

⁵⁷In the model from Section 5, the MVT always holds, so this assumption is only for brevity.

Define $T_0 = 0$, and T_τ for $\tau > 0$ according to the following mapping H :

- (i) If $T = 0$ and $x = y_t$, $x' = y_{t+1}$ for some t , then $H(x, T, x') = 0$;
- (ii) else, if $x' \notin (z(x), x)$, $H(x, T, x') = 1$;
- (iii) else $H(x, T, x') = 2$.

In other words, in state 0, the policy follows the intended equilibrium path, $(y_t)_t$, and T_τ remains equal to zero. In state 1, the policy path is constant and T_τ remains equal to 1. In state 2, the current decisionmaker, $m(x)$, chooses the lowest policy that she weakly prefers to x , and the state then changes to 1. Deviations to myopically attractive policies (policies that $m(x)$ strictly prefers to x) are punished by switching to state 2, while deviations to myopically unattractive policies are punished by switching to state 1.

We can verify that this is an SPE. If $(x_\tau, T_\tau) = (y_t, 0)$ for some t , then $m(x_\tau)$'s equilibrium continuation utility is $U_{m(y_t)}(S(y_{t+1})) \geq \frac{u_{m(y_t)}(y_t)}{1-\delta}$. If she deviates to $x' \notin (z(x_\tau), x_\tau)$, then $T_{\tau+1} = 1$ and the continuation is given by $x_{\tau'} = x'$ for all $\tau' > \tau$, yielding utility $\frac{u_{m(y_t)}(x')}{1-\delta} \leq \frac{u_{m(y_t)}(y_t)}{1-\delta}$. If she deviates to $x' \in (z(x_\tau), x_\tau)$, then $T_{\tau+1} = 2$ and the continuation is $x_{\tau+1} = x'$, $x_{\tau'} = z(x')$ for all $\tau' > \tau + 1$, yielding utility $u_{m(y_t)}(x') + \frac{\delta}{1-\delta} u_{m(y_t)}(z(x')) \leq \frac{u_{m(y_t)}(y_t)}{1-\delta}$.

If $(x_\tau, T_\tau) = (x, T)$ with $T = 1$ or $T = 2$, then $m(x)$'s equilibrium continuation utility is $u_{m(x)}(x)$. If she deviates to $x' \notin (z(x), x)$, she gets utility $\frac{u_{m(x)}(x')}{1-\delta} \leq \frac{u_{m(x)}(x)}{1-\delta}$. If she deviates to $x' \in (z(x), x)$, she gets $u_{m(x)}(x') + \frac{\delta}{1-\delta} u_{m(x)}(z(x')) \leq \frac{u_{m(x)}(x)}{1-\delta}$. \square

The condition $U_{m(y_t)}((y_{t+1}, y_{t+2}, \dots)) \geq u_{m(y_t)}(y_t)$ is clearly necessary—otherwise, $m(y_t)$ would deviate to staying at y_t forever. What this result shows is that, aside from this common-sense restriction, anything goes.⁵⁸ In particular, for each $x \in [x^*, x^{**}]$ there is an SPE with policy path constantly equal to x , so any policy can become an intrinsic steady state in the right SPE.

However, there are several arguments for focusing on Markov equilibria.

- (i) Non-Markovian behavior must be supported by non-Markovian behavior arbitrarily close to x^* . Moreover, assuming any particular form of Markov behavior in a neighborhood of x^* collapses the set of equilibria to a single (Markov) equilibrium:

⁵⁸The condition that all x be very reluctant is a joint condition on u and δ , which is not hard to satisfy. For example, in the quadratic-linear case given by $u_\alpha(x) = C - (\alpha - x)^2$ and $m(x) = ax$, it holds if $2a\delta \geq 1$, i.e., if $a > \frac{1}{2}$ and δ is high enough.

Lemma 10. *Let $\epsilon > 0$ and $\tilde{s} : [x^*, x^* + \epsilon] \rightarrow [x^*, x^* + \epsilon]$ such that $\tilde{s}(x) < x$ for all x . Let s, s' be two SPEs on $[x^*, x^{**}]$ such that $s(x, h) = s'(x, h) = \tilde{s}(x)$ for all $x \in [x^*, x^* + \epsilon]$ and h , and assume that s and s' obey the following tie-breaking rule: if the set $\arg \max_{y \leq x} U_{m(x)} S(y, h)$ has multiple elements, then $s(y)$ is the highest element of the set. Then $s \equiv s'$ and $s(x, h)$ is independent of h .*

Proof of Lemma 10. The intuition behind this result is a simple unraveling argument: suppose two equilibria coincide up to some point $x^* + \epsilon$. Then, for y slightly above $x^* + \epsilon$, $I(y)$ will be choosing between successors in $[x^*, x^* + \epsilon]$, which have the same continuation in both equilibria, so the same choice will be made. Formally, let

$$A = \{x \in [x^*, x^{**}] : \exists \hat{s} \text{ s.t. } \forall h, \forall y \in [x^*, x], s(y, h) = s'(y, h) = \hat{s}(y)\},$$

and $x_0 = \sup(A)$. By assumption, $x_0 \geq x^* + \epsilon$. Suppose $x_0 < x^{**}$.

There are two cases. First, suppose $x_0 \notin A$. Then the same proof as in Proposition 2 shows that $u_{m(x_0)}(x_0) < \max_{y \in [x^*, x_0]} U_{m(x_0)} \hat{S}(y)$, so $s(x_0, h), s'(x_0, h) < x_0 \forall h$. The tiebreaking rule then implies $s(x_0, h) = s'(x_0, h) = \max(\arg \max_{y \in [x^*, x_0]} U_{m(x_0)} \hat{S}(y)) \forall h$, whence \hat{s} can be extended to x_0 , a contradiction.

Second, suppose $x_0 \in A$. Then there is a sequence $(x_n)_n$ such that $x_n \rightarrow x_0$ and $x_n > x_0 \forall n$ such that, for each n , $s(x_n, h_n) > x_0$ for some history h_n , as otherwise the tiebreaking rule would guarantee that $s(x_n, h) \equiv s'(x_n, h)$ are independent of h .

Note that, for each n , $m(x_n)$ always has the option of jumping to any policy $z \in [x^*, x_0]$, and that the continuation would be the history-independent path $\hat{S}(z)$; hence, the optimality of s requires that $U_{m(x_n)}(S(s(x_n, h_n), h_n)) \geq U_{m(x_n)}(\hat{S}(z))$.

For each n , label the continuation path starting at $s(x_n, h_n)$ as $S_n = (s_0^n, s_1^n, \dots)$, where $s_0^n = s(x_n, h_n)$. Let $s_{k_n}^n$ be the first policy in this path that is in $[x^*, x_0]$. Note that $(s_{k_n}^n, s_{k_n+1}^n, \dots) = \hat{S}(s_{k_n}^n)$ is history-independent, and $m(x_n)$ always has the option of jumping directly to $s_{k_n}^n$, so

$$U_{m(x_n)}(\hat{S}(s_{k_n}^n)) \leq U_{m(x_n)}(S_n) \leq \frac{1}{1 - \delta^{k_n+1}} \sum_{t=0}^{k_n-1} \delta^t u_{m(x_n)}(s_t^n) \leq \frac{u_{m(x_n)}(x_0)}{1 - \delta}.$$

As $m(x_n)$ can also choose any $z \in [x^*, x_0]$, it must be that

$$\max_{z \in [x^*, x_0]} U_{m(x_n)}(\hat{S}(z)) \leq U_{m(x_n)}(S_n) \leq u_{m(x_n)}(x_0).$$

By continuity, $U_{m(x_0)}(\hat{S}(\hat{s}(x_0))) \leq u_{m(x_0)}(x_0)$, which contradicts Proposition 2. \square

(ii) Consider a discrete approximation of the problem in which the policy space is restricted to a finite set X . Then, for a generic choice of $X = \{x_1, \dots, x_N\} \subseteq [-1, 1]$, there is a unique subgame perfect equilibrium, which is Markov.⁵⁹ Hence, if we are interested in equilibria that can be obtained as limits of discrete-policy space equilibria,⁶⁰ we need only to study Markov equilibria.

(iii) Consider a variant of the game with a finite number of periods $t = 0, 1, \dots, T$. For each T the game has a unique equilibrium s_T , which is Markov in (x, t) . A limit of such equilibria as we take $T \rightarrow \infty$ may not be Markov in x exclusively, but Propositions 2 and 3 can still be extended to this case, so intrinsic steady states are also ruled out under this refinement.

Finally, it is worth noting that the MVEs we construct in the main text, which converge to a myopically stable policy, are strictly preferred by all pivotal decision-makers to an SPE in which the policy never changes. In other words, the fall down the slippery slope is *desired* by agents. More precisely, assume an initial policy x_0 , and consider an MVE s such that $s^t(x_0) \xrightarrow{t \rightarrow \infty} x^*$, and an SPE \tilde{s} where the policy remains constant at x_0 . By Lemma 2, there is α_0 such that all agents to the left of α_0 strictly prefer the continuation under s ; since $m(x_0)$ has this preference, $\alpha_0 \geq m(x_0)$. Consequently, all agents in $[x^*, m(x_0)]$ —in particular, all agents who will be pivotal on the equilibrium path—have the same preference. Hence, by focusing on Markov equilibria, we are not ruling out preferable equilibria that just require mutually desirable coordination on the part of the players to arise.

E.4 Explicit Voting Protocols

Our solution concept assumes that, for a policy y to be chosen by a set of voters $I(x)$, $S(y)$ must be a Condorcet winner. This assumption is agnostic about the actual voting process taking place. Here, we discuss two natural microfoundations.

⁵⁹This can be shown by proving Proposition 1 in the discrete case, and then applying backward induction (Acemoglu et al., 2015). The equilibrium is unique so long as there are no indifferences.

⁶⁰Formally, denoting s_X to be the equilibrium for policy space s_X , an equilibrium s is a limit of discrete-policy space equilibria if there is a sequence $(X_n)_n$ such that $\max_{y \in [-1, 1]} d(X_n, y) \xrightarrow{n \rightarrow \infty} 0$, i.e., the sets X_n become arbitrarily fine, and $s_{X_n}(x_n) \rightarrow s(x) \forall (x_n)_n$ s.t. $x_n \in X_n \forall n$ and $x_n \rightarrow x$.

The first is Downsian competition. Suppose that, at each voting stage, there are two politicians A_t, B_t who simultaneously propose policies x_{At}, x_{Bt} ; voters observe the two proposals and then vote for the candidate proposing the policy whose continuation path they prefer. Assume either that the politicians are short-lived (they are replaced every period) or that they play Markov strategies, and they are office-motivated, i.e., they obtain $R > 0$ from winning and zero from losing. An equilibrium of the voting stage is given by policy proposal strategies $x_A(I), x_B(I)$ and voting strategies $v_\alpha(x_A, x_B)$ such that: for each candidate i , offering $x_i(I)$ maximizes i 's winning probability given a set of voters I , and $v_\alpha(x_A, x_B) = i$ if $U_\alpha(S(x_i)) > U_\alpha(S(x_{-i}))$, where $S(x_i)$ is the equilibrium continuation starting at x_i . Then:

Remark 6. Given an MVE s from the main model, we can construct an MPE of the Downsian model as follows: $x_A(I(x)) = x_B(I(x)) = s(x)$ and $v_\alpha(x_A, x_B) = \mathbb{1}_{\{U_\alpha(S(x_A)) > U_\alpha(S(x_B))\}}$ for all x, x_A, x_B .

Conversely, for any MPE of the Downsian model in pure strategies, $x_i(I)$ must be a Condorcet winner for every i, I . Moreover, if $x_A(I) = x_B(I)$ for all I , then $s(x) = x_i(I(x))$ constitutes an MVE of the main model.

In other words, requiring $S(s(x))$ to be a Condorcet winner among $I(x)$ is equivalent to assuming Downsian competition at the voting stage, except that we implicitly rule out situations where there is no Condorcet winner (in which case the Downsian model might still have equilibria with mixed proposal strategies), and we rule out mixed policy choices by voters ($s(x)$ is assumed to be deterministic).

Another possible microfoundation is a sequential proposal protocol similar to that used in Acemoglu et al. (2008, 2012). Suppose that, at each voting stage, there is a period of continuous time $[0, 1]$ (this is measured on a different scale from the time that passes between periods, and no discounting accrues during the voting stage). At each instant $y \in [0, 1]$, policy y is proposed to the organization and each voter α casts a vote $v_\alpha(y, I) \in \{0, 1\}$. If a majority votes 1 when y is under consideration, the voting stage ends and y is chosen; if no policy receives a majority then the policy next period is equal to the current policy. Voters are strategic and voting strategies are pure and weakly undominated. Any MVE s from the main model can also be implemented as an equilibrium of this model.