Voter polarization and extremism[*]

Jon Eguia[†]   Tai-Wei Hu[‡]

February 23, 2022

–Preliminary–

Abstract

We present a theory of endogenous policy preferences and electoral competition with boundedly rational voters who find it costly to recall detailed information. Voters are otherwise fully rational, and they strategically choose how much memory to devote to processing political information. We find that even if all voters start with a common prior such that given this prior they all prefer a moderate policy over either a left or a right alternative, and even if they only observe common signals that in the limit would make a perfectly rational observer certain that the moderate policy is indeed best for everyone, voters with costly memory capacity will eventually prefer extreme policies, and the electorate polarizes: some voters support the left policy, and some support the right policy. Two fully rational political parties polarize as well, one pandering to the voters on the left, and the other pandering to the voters on the right.

Key words: Polarization, extremism, rational inattention, bounded memory, electoral competition.

[†]Michigan State University, email: eguia@msu.edu

[‡]University of Bristol, email: taiwei.hu@bristol.ac.uk

1 Introduction

We study how the cost of processing political information for an individual voter relates to the polarization of the electorate. We find conditions under which voters would embrace a policy consensus if they were fully informed or fully uninformed, but their constrained optimal processing of the available information leads them to polarize into two opposed extreme camps. To explain this phenomenon, we microfound voters' behavior with explicit formulations of voters' motives to vote and of the costs they bear to process political information.

Consider an electorate facing a set of different policies, and a list of candidates running for office. Voters are endowed with preferences over certain economic and social outcomes (their individual wealth, society's wealth and inequality, pollution, etc.) as a primitive. A voter's preferences over policies or over her voting options (i.e. over each candidate on the ballot, and abstention) are not a primitive. Voters face uncertainty about how their vote affects the election and the implemented policy, and uncertainty about how this policy affects the downstream socio-economic outcomes. Each voter's preferences over policies are endogenously derived from the voter's primitive preferences over outcomes, and from her information about how the implemented policy will affect the outcomes over which her primitive preferences are defined.

Given their endogenous preferences over policies, voters derive their preferences over their voting options by combining an outcome-oriented motivation with an "expressive" motivation. The outcome-oriented motivation depends on the effect of their vote on the chosen policy; in a large election, this motivation vanishes as the probability that an individual vote has any effect becomes negligible (Ledyard 1984). The expressive motivation to vote for a given candidate is that voters enjoy supporting good causes: by supporting a cause or a party, a voter *becomes* a supporter of this party or cause (Schuessler 2000), and voters enjoy identifying as members of a group that champions good policies.[1]

Voters who vote for a candidate because they enjoy the act of supporting a good policy need to know (or at least they need to believe) that the policy their chosen

---

[1] *"Prosperity has many parents; adversity only one"* (Tacitus 2014 [94 AD], page 53, in the original language: *"Prospera omnes sibi vindicant; adversa uni imputatur."*).

candidate would implement is good.[2] In an uncertain environment, these voters need to process information to determine which are the best policies, and thus which candidates (if any) are worth supporting by voting for them.

In the information age, voters are flooded with political information, freely available across multiple media platforms. If processing information were costless, a rational voter would use all this freely available information and Bayes rule to update and formulate a precise posterior belief about the mapping from policies to the outcomes of interest, and would vote to maximize her expected utility according to that posterior belief. Alas, processing information is costly. Voters need to weigh this cost against the benefit of being better informed (Downs 1957; Davis, Hinich and Ordeshook 1970). Strategic voters exposed to an over-abundance of political news, and with limited memory capacity to correctly process and store all this information need a simpler, (constrained) optimal rule to determine how to process information, which important pieces of information to keep in mind, and which ones to discard and to forget about.

We construct a theory of political participation under the following two premises. First, voters enjoy supporting a policy in direct proportion to their expected utility if this policy were implemented. Further, each voter enjoys voting for candidates who support the policy the voter thinks is best, while she dislikes voting for candidates who support policies that are very different from the one she thinks best. These expressive preferences over voters' own actions may be weighed arbitrarily little relative to the weight on standard preferences over outcomes, but as long as their importance is not zero, they will influence voters' behavior. Second, voters decide how much cost to incur assimilating and processing political information by weighing how much this information helps them determine which policy is best, against the costs of processing this information.

We formalize the cost of processing information by relaxing the assumption of free perfect recall: it is costly for voters to keep track of all the information they have observed, and to remember precise details to obtain a precise posterior. Instead,

---

[2]This motivation to vote contingent on "getting it right" is unlike, say, an intrinsic motivation to vote regardless of who or what one votes for (Riker and Ordershook 1968) or the motivation to vote for a candidate whose identity is exogenously given as part of the voter's type (Coate and Conlin 2004; Feddersen and Sandroni 2005).

we postulate that voters have a limited set of "memory states" that aggregate past information. A fully rational agent with perfect recall would be one with unlimited memory states, who can record any minutely different history of informative signals (or any sufficient statistic such as a posterior belief recorded to any degree of precision) into its own memory state. A more realistic voter has only a limited memory capacity to deal with all the information, so she must lump sufficiently similar information histories into the same memory state with a blurred belief about the state of affairs. Since it is costly to increase that memory capacity, voters trade-off this cost of processing information, with the benefit of a more precise understanding about how alternative policies would affect outcomes of interest, and thus deriving more satisfaction from supporting the alternative they think would have the best effect.

We use finite-automata to model this endogenously imperfect information processing as in Wilson (2014). An automaton consists of a collection of finitely many memory states, and a transition rule taking the process from a given memory state to another depending on the information received in that period. Limited memory capacity implies that agents' beliefs are categorized discretely instead of being represented by precise posteriors. This then departs from the rational inattention literature (Maćkowiak, Matějka and Wiederholt 2021), because, under rational inattention, agents strategically choose which information to attend to, and then update in a fully rational way with a precise posterior belief. In contrast, under our formulation agents strategically choose both their memory capacity and, given this capacity, how best to process all the information they observe.

Given the number of memory states an agent chooses, each memory state represents the agent's state of mind or her set of thoughts about the uncertain state of the world relevant for her voting behavior. Each memory state corresponds to a category of partial histories and a qualitative "belief". This categorization is analogous to the one made by an agent who evaluates sovereign default risk solely based on credit ratings "A", "B" , "C", etc., and who then forgets all the detailed information that fed into the rating, including all information that make some countries rated "B" less likely to default than others with the same rating. A distinct feature of this updating process is the discreteness in information processing; there is no "feather that breaks the camel's back" because feathers (small bits of information) do not induce a sufficient update to

change from one memory state to another, and are thereafter forgotten; rather, it takes quite a substantial bit of information to trigger a transition across memory states.[3]

Agents who use different updating processes may process the same news very differently. To avoid infinite costs, each agent will choose only a finite number of distinct memory states, accommodating only finitely many different views of the world. The optimal number of memory states and the optimal rule to transition across memory states are determined endogenously by the preferences over outcomes, and by the cost structure. We assume homogeneous costs of memory across voters but we allow heterogeneity in preferences, and this heterogeneity generates endogenously the difference in the constrained optimal updating processes.

Consider the following environment. There is a set of three policies: a moderate policy and two extreme ones, one on each side (left and right) of the ideological divide. Each policy matches one state of the world, and is Pareto superior to the other two in this state: every voter strictly prefers the socioeconomic outcome if the state-matching policy is implemented than the outcome if any other policy is implemented. Signals that the state is "moderate" (which we take it to be the normal, expected state) are abundant in every state and hence are commonplace (say a day of below average street violence), while the signals that shift preferences toward extreme policies (say a shocking case of either coordinated police violence against peaceful demonstrators, or coordinated violence by armed rioters against peaceful bystanders and police) will make big news as they are rare but very informative.

Our main result is that in this environment, voters with costly memory capacity polarize once a signal about the extreme state realizes. Even if all voters start with a common prior about the state of the world and under this prior all voters prefer the moderate policy, even if voters only observe common signals, and even if these common signals are such that any voter with perfect recall would formulate a posterior that the moderate policy is indeed best for everyone, given their limited memory, all voters end up favoring extreme policy alternatives, and diverging in their preferences: some

---

[3]Rationally inattentive agents may also optimally ignore small signals, updating only after highly informative ones (Kominers, Mu and Peyshakovich 2018), and may optimally choose discrete actions, even if under full attention their optimal action as a function of beliefs would span a continuum. Under our model it is the belief that is categorized, which has implications for dynamic learning.

prefer the left policy, while others prefer the right policy, both away from moderation. Chasing their voters, two office-motivated parties polarize as well, and the implemented policy becomes extreme.

The underlying mechanism that drives our result is that, under the constrained optimal rule, voters ignore the commonplace signals for moderation after seeing the big news indicating an extreme state of the world. These updating process will eventually drive voters away from moderation, and toward the political extremes. However, even under the common signals, voters with a common prior still polarize at opposite extremes. Why? Say the common prior is a belief that the state is likely "normal", and under such a normal state, all agents prefer a "moderate" policy. Once the news of the day reveal a commonly-observed signal that the state of the world is extraordinary (which is rare but will eventually happen), all voters agree that the state is indeed more likely to be an extraordinary one that calls for an extreme policy solution. But having concluded that the state is not normal and that a moderate policy would be unsuitable, voters may disagree about which extreme policy may be appropriate: some may support left policies, while others support right policies. The policy disagreement stems from the heterogeneity in the relative distaste over policy mistakes in one or the other direction. Nobody supports moderation, and society polarizes.

In this theory, polarization is micro-founded by the individual decisions of each voter to simplify her information environment, by coarsening the partition of possible beliefs under consideration. Polarization is, at heart, an aggregate phenomenon that can be decomposed as a large number of independent (and disparate) decisions to become extreme made by each individual voter in isolation. Polarization, in this account, is not elite driven, and it is not driven by the electorate's network interactions, nor by biased media that reinforces the beliefs of like-minded voters in their own informational bubble. Rather, we show that a large society of Robinson Crusoes, each isolated in their own island, all endowed with a common prior at the time of arrival to their own island, and observing common signals in the night sky each night, would also polarize.

In what follows, we first discuss the related literature. Thereafter, in Section 2 we present a model on information processing and preference formation for voters who face memory costs; in Section 3 we show how an electorate composed of such voters polarizes; and in Section 4 we show that party platforms and implemented policies

will polarize as well in any equilibrium of a stylized electoral competition theory. We discuss our theoretical results in light of new empirical evidence on polarization in Section 5. All proofs are relegated to an Appendix.

Related Literature

Our modeling of how voters process information relates in its motivation and substance to theories of rationally inattentive voters (Prato and Wolton 2016; Matějka and Tabellini 2021). In these models of rational inattention, typically voters can choose from a menu of costly signal-generating processes such that more informative processes are costlier. In contrast, in our model all signals are commonly observed, and the differences arise in what voters do with the signals; technically, our model draws from decision-theoretic work on finite memory (Cover and Hellman 1970 and Wilson 2014)[4] and on updating of diffuse beliefs defined by a set of priors (Chambers and Hayashi 2010).

Our theoretical finding that voters' costly recall leads to extreme beliefs and to policy divergence contributes to a literature documenting political polarization (Abramowitz and Saunders 2008; McCarty, Poole and Rosenthal 2016; Gentzkow 2016), studying its consequences (Gordon and Landa 2017; Buisseret and van Weelden 2021), suggesting ways to mitigate it (Axelrod, Daymude and Forrest 2021), or explaining some of its causes. Among the latter, Glaeser, Ponzetto and Shapiro (2005), Serra (2010), Bol, Matakos, Troumpounis and Xefteris (2018), Tolvanen, Tremewan and Wagner (2021), and McMurray (2021) focus on candidates' polarization. With regard to voter polarization, it can arise if voters choose to follow different sources of information (Nimark and Sundaresan 2019; Che and Mierendorff 2019; Perego and Yuksel 2022); or if they pay disproportionate attention to the issues they care more about (Yuksel, forthcoming) or to the issues in which the candidates' proposals differ most (Nunnari and Zapal 2020); if they share news with their connections (Bowen, Dmitriev and Galperti, 2021); or, even under commonly-observed signals, if voters face ambiguity and are averse to it (Baliga, Hanany and Klibanoff 2013).

Perhaps closest to our work in their linking of voters' memory constraints to polar-

---

[4]The automata approach to model imperfect recall has been recently evaluated by Oprea (2020), and is supported by experimental evidence in Banovetz and Oprea (2020).

ization are two theories in which voters observe common information, but they process it in a boundedly rational way that leads to polarization. Fryer, Harms and Jackson (2019), assume that voters coarsen the space of signals about the state of the world. In their model, voters reinterpret each uninformative signal as an informative one that conforms with their prior. Voters then update this prior as if the signal had been truly informative; this self-confirming miss-processing of signals, together with heterogeneous priors, leads to polarization. In fact, if some signals are equivocal rather than uninformative, upon observing equivocal signals, fully rational agents with different priors about the meaning of equivocal signals also polarize toward their priors (Benoit and Dubra 2019). In either case, if agents shared a common prior, they would not polarize.

Levy and Razin (2021), like us, present a dynamic theory of electoral competition and polarization with two parties and three possible policies (a moderate one, and an extreme one to each side), in which the driver of polarization is voters' limited temporal memory: voters remember all information for a fixed time, and after this lapse of time they forget.[5] Polarization is candidate-driven, and arises because parties are policy motivated and the median's preference uncertain (as in Wittman 1983, or Calvert 1985). However, as parties polarize, the extreme policies they implement reveal more information about the state of the world, allowing voters to infer which is the right policy, and forcing candidates to converge to it; once voters forget their history, they are indeed bound to repeat it, and parties are able to polarize again. Policy polarization is thus cyclical, while voters' beliefs never polarize, as all voters share a common update of the environment. We complement their account with a theory of voter polarization.

2   The Model

Consider a large democratic society, represented by a set I of voters, with unit mass. Each voter $i \in I$ is faced with a choice over three policy alternatives in each of infinitely many periods. Let $\mathcal{A} \equiv \{a^L, a^M, a^R\}$ denote the set of alternatives, where a^L denotes

---

[5]In contrast, our voters' memory constraint is one of capacity, like computers' memory: our voters can only carry limited information in their memory from period to period.

a "left" alternative, a^M a "moderate alternative", and a^R a "right" alternative. Let $\Theta \equiv \{L, M, R\}$ denote the set of possible states of the world (where L, M, and R again respectively denote Left, Moderate, and Right), and let $\theta \in \Theta$ denote a state of the world. All voters share a common prior probability distribution $\mathbf{P}_0$ over Θ about the state of the world. We envision an environment in which the moderate state M is the most likely, and states L or R represent an extraordinary event or shock. Formally, we assume that the common prior among the agents is such that

$$\mathbf{P}_0(L) = \mathbf{P}_0(R) = p_0 < \frac{1}{4}. \tag{1}$$

In each period t $\in \{1, 2, ..., \infty\}$ each voter i chooses which policy alternative to support. Let $a_t^i \in \mathcal{A}$ denote the alternative that voter i supports in period t (more generally we denote individual agent labels as superscripts, and period labels as subscripts). Let a_t^I denote the policy alternative collectively chosen by society in period t; in Section 4 we model how this collective choice is made through party competition in democratic elections.

Each voter i cares about the policy outcome a_t^I and about the policy a_t^i that she supports, in each period. Let $\lambda \in (0, 1)$ denote the relative weight assigned to the policy outcome, and $1 - \lambda$ the weight assigned to the expressive component of her political preferences, so that in each period t, each voter i derives instantaneous utility

$$\lambda u(a_t^I, \theta, b^i) + (1 - \lambda)u(a_t^i, \theta, b^i), \tag{2}$$

where b^i is voter i's type as described below.[6] We assume that voters' intertemporal patience is captured by a discount factor $\delta \in (0, 1)$ across periods, so that the total utility for voter i for an infinite sequence of individual and collective choices is

$$\lambda \sum_{t=1}^{\infty} (\delta)^t u(a_t^I, \theta, b^i) + (1 - \lambda) \sum_{t=1}^{\infty} (\delta)^t u(a_t^i, \theta, b^i), \tag{3}$$

where the first term is the utility from the sequence of policy outcomes, and the second

---

[6]For an overview of citizen's motivations for voting, see Brennan and Lomasky (1993) or a survey by Hamlin and Jennings (2018); and Glazer (1987) for an early theory of elections under expressive voting.

9

term is the expressive utility from the sequence of individual choices to express support for an alternative.

We assume that in each state $\theta \in \Theta$, every voter derives highest period utility from the policy alternative a^θ that matches the state, so we refer to alternative a^θ as the "correct" alternative in state θ. Each cell in the left matrix in Table 1 shows the utility function $u(a, \theta, b^i)$ as a function of the action a in each column, and of the state of the world θ in each row, with $c \in [0, 1]$, and type $b^i \in (-\bar{b}, \bar{b})$ for each voter i, for some $\bar{b} \in (-1, 1)$. The correct policies are on the diagonal of the matrix. Notice as well that in each state, utilities are single-peaked with respect to the standard left-to-right order, and that, given Assumption 1 on the prior, states Left and Right are ex-ante much less likely than the Moderate state. We thus refer to policy alternatives a^L and a^R, and to states L and R as "extreme."

We assume that the distribution over voter types has a full support over $(-\bar{b}, \bar{b})$. Type b^i captures a mild asymmetry, bias or lean on voter i's preferences over policy alternatives, as follows: Subject to a Moderate state of the world, voters have a common-value symmetric preferences over policies, with ideal alternative a^M. Voters also have a common preference order over alternatives in either of the two extreme states of the world, and they all agree that alternative a^M is ex-ante the best given their common prior $\mathbf{P}_0$ over the state of the world. Voters only vary on how much they gain from the correct extreme alternative in each extreme state.

We say that a voter "leans left" if she has a stronger preference for the left action in the Left state, than for the right action in the Right state; and that she "leans right" if she has a stronger preference for the right action in the Right state, than for the left action in the Left state. Type b^i is then a measure of this lean, with $b^i < 0$ implying that voter i leans left, and $b^i > 0$ that she leans right. But such leanings are mild, in that they only arise if the state is extreme; under the Moderate state, voters have common values, and their sole concern in this state is to choose the moderate alternative.

The state of the world, however, is not observable. Instead, in each period $t \in \mathbb{N}$, voters observe a common signal s_t drawn from the set $\mathcal{S} \equiv \{\ell, m, r\}$ independently in each period. Conditional on the state of the world $\theta \in \{L, M, R\}$, signal $s \in \{\ell, m, r\}$ is drawn with probability μ_s^θ. Each cell of the right matrix in Table 1 denotes this

$\theta \backslash a$	a^L	a^M	a^R
L	$1 - b^i$	0	$-c$
M	0	1	0
R	$-c$	0	$1 + b^i$

$\theta \backslash s$	ℓ	m	r
L	$\mu - \epsilon$	$1 - \mu$	ϵ
M	ϵ	$1 - 2\epsilon$	ϵ
R	ϵ	$1 - \mu$	$\mu - \epsilon$

Table 1: Left: payoff matrix; right: signal structure

probability μ_s^θ, as a function of the state θ in each row and the signal s in each column, with $\mu < 1$ and $\epsilon \in (0, \mu/2)$.

In each period $t \in \mathbb{N}$, voters observe the common signal s_t and their own actions a_t^i. Later on when we introduce the election game, voters will also observe the platforms chosen by political parties. In each period t, each voter i must choose an alternative, a choice that can in principle depend on all relevant past information, which we denote by $h_t^i \equiv (((s_\tau, a_\tau^i))_{\tau=1}^{t-1}, s_t)$. We refer to h_t^i as a "partial history for voter i". A partial history h_t^i for voter i includes the history of signals $(s_1, ..., s_t)$ up to period t, and the sequence $(a_1^i, ..., a_{t-1}^i)$ of voter i's own actions up to period $t - 1$. Let H_t denote the set of all possible pairs of $((s_\tau)_{\tau=1}^t, (a_\tau)_{\tau=1}^{t-1})$, the first sequence representing a history of signals up to period t, and the second one representing a sequence of actions by an arbitrary agent up to period $t - 1$.

A decision rule for agent i is a function $D^i : \bigcup_{t=1}^\infty H_t \to \mathcal{A}$. The decision rule maps each possible partial history for agent i to the set of alternatives. However, since all agents (voters and political parties) share the same sequence of signals and that is the only relevant source of information for the state of the world, under perfect recall all agents share the same posterior belief about Θ and that is the only relevant state variable, updated only according to the current common signal realization.

As a result, the unconstrained optimal rule can be fully characterized by the posterior, p over Θ: for each voter i, the optimal decision rule is to choose alternative a that maximizes $\sum_{\theta \in \Theta} p(\theta) u(a, \theta, b^i)$. The posterior is computed according to Bayes rule. We use $\Delta(\Theta)$ to denote the set of posteriors, and we use $\Delta^\theta(b^i)$ to denote the set of posteriors under which a^θ is an optimal action for agent i with type b^i.

Finite automata and implementation

Storing and remembering precise, detailed information is costly for voters. Boundedly rational voters that optimize their choices must take into account this cost in making their decisions. The unconstrained optimal rule described above requires costless processing of an arbitrarily long sequence of signals and actions to compute a precise posterior belief, in order to remain optimal. Rational agents for which such memory capacity and information processing is costly—and these include any voter in any real-world application—will seek to find a constrained optimal rule that is less costly to use. We assume that our agents summarize their past memory using finitely many "memory states," and update their memory only using the most relevant information, namely, the commonly observed signal. More precisely, we formulate this process as a finite automaton, as described below.

We use finite automata to model this cost of processing infinite sequences of information. A **stochastic finite-state automaton (SFSA)** consists of a list $\langle Q, q_0, \tau, d, \rangle$, where Q is a finite set of *memory states*, $q_0 \in Q$ is the initial memory state, $\tau : Q \times \mathcal{S} \to \Delta(Q)$ is the *transition rule*, and $d : Q \to \Delta(\mathcal{A})$ is the *decision rule*. We use $\tau(q, s; q')$ to denote the transition probability from memory state q to memory state q' when receiving signal s. Using results from Kalai and Solan (2004), with no loss of generality we can restrict attention to deterministic action rules. Then the decision rule is $d : Q \to \mathcal{A}$ and $d(q)$ denotes the action taken when current memory state is q.[7] If the transition rule is also deterministic, we say the finite automaton is deterministic (abbreviated as DFSA), and we use $\tau(q, s) = q'$ to denote the transition rule.

Let $\mathcal{Q}$ denote the set of all such stochastic finite-state automata.

A voter using one of these finite automata no longer needs to keep track of an arbitrarily long sequence of signals and actions. Rather, the voter only needs to remember her transition rule τ and her decision rule d, to keep track of the memory state $q \in Q$, and to observe the latest signal s. With just that, she can transition to a new memory

---

[7]We could conceive of an automaton with a transition rule from $Q \times \mathcal{A} \times \mathcal{S}$ to Q, according to which a voter's choice of an alternative together with the observed signal jointly drive the transition to a new memory state. However, since a voter's own choice does not convey any information to the voter about the state of the world, we simplify the class of automata under consideration to be ones that only transition to a new memory state based on the signals observed by the voter, and not based on the choices she makes.

state according to her transition function, the memory state she is at and the signal she observes; and she can take a decision over alternatives according to her decision rule. If the set of memory states Q is small, this is a simple enough exercise. More complex automata, with more memory states, require more memory. We assume that using a finite automaton with $|Q|$ memory states, costs $\kappa \cdot |Q|$, for some $\kappa \in \mathbb{R}_{++}$.

Each voter i chooses her automaton to maximize her discounted, total expected utility, by solving the optimization problem

$$\max_{\langle Q, q_0, \tau, d \rangle \in \mathcal{Q}} \left(\mathbf{E} \left[\lambda \sum_{t=1}^{\infty} (\delta)^t u(a_t^I, \theta, b^i) + (1 - \lambda) \sum_{t=1}^{\infty} (\delta)^t u(d(q_t), \theta, b^i) \right] - \kappa \cdot |Q| \right). \quad (4)$$

In a large society, the probability that an individual agent's choice affects the collective choice is negligible —with a unit mass of agents, each agent is infinitesimal, and this probability is exactly zero—and therefore the first summation in the expectation drops out of each voter's optimization problem (Brennan and Hamlin 1998), which simplifies to finding the automaton that maximizes ex-ante expected expressive utility, net of costs of running the automaton. Formally, the optimal automaton for voter i is one that solves

$$\max_{\langle Q, \tau, d, q^o \rangle \in \mathcal{Q}} \left(\mathbf{E} \left[(1 - \lambda) \sum_{t=1}^{\infty} (\delta)^t u(d(q_t), \theta, b^i) \right] - \kappa \cdot |Q| \right). \quad (5)$$

In Section 3 we find the optimal automaton and we describe the resulting voters' individual decisions. In Section 4 we study how the democratic process shapes the collective choice a_t^I and total welfare, in light of these voters' decisions.

3   Voter Polarization

We say that a voter i "becomes extreme" in period t if $a_{t-1}^i = a^M$, and for any sequence of signals $(s_\tau)_{\tau=t}^{\infty}$ and any future period $\tau > t$, $a_\tau^i \neq a^M$. That is, we use the term only in a strong, irreversible sense that the voter has abandoned moderation once and for all, never to return regardless of any further information revealed to her.

We define "voter polarization" as the phenomenon by which a positive mass of

voters supports a^L and a positive mass support a^R. Given that ex-ante —before the first signal is revealed—all voters support a^M; that they all agree on the best alternative for any state of the world; and that they share the same prior and observe a common sequence of signals that (asymptotically) reveals the true state of the world, it might seem that optimizing agents could not polarize for long.

Indeed, rational agents with unlimited memory and perfect capacity of information processing, a common prior and a common set of observed signals will agree on their posterior over the state of the world, and once this posterior becomes sufficiently close to degenerate, given their near common-value preferences, they will also agree on which alternative to support. Whereas, we shall show that agents with limited memory capacity have divergent posteriors and become extreme and polarized, despite their common priors and common signals.

We start by considering an environment in which voters with limited memory *do not* polarize: the special case in which extreme signals (ℓ or r) reveal the state of the world (formally, extreme signals are fully informative if ϵ in Table 1 is zero). In this case, the unconstrained optimal rule is straightforward: the posterior on M increases as long as all signals are m, and hence alternative a^M continues to seem optimal as long as $s^\tau = m$ for any $\tau \in \{1, ..., t\}$. In contrast, a single ℓ-signal, if $\epsilon = 0$, drives the posterior of L to one and reveals that a^L is the correct alternative. Symmetric argument holds for signal r and action a^R as well.

Such a simple rule is accessible for a voter with very little memory, using any SFSA in the following class.

Definition 3.1. *A SFSA is a **"no-return 3-state automaton"** if it has three memory states $\{q^L, q^M, q^R\}$, initial state q^M, decision rule $d(q^\theta) = a^\theta$ and transition rule probabilities given by Table 2, where each cell (q_t, s_t) lists the probability of transitioning, respectively, to q^L, to q^M, and to q^R, and with free parameters α and β in $[0, 1]$.*

Let FA_3 denote the class of no-return 3-state automata, and for any α and β in $[0, 1]$, let $FA_3(\alpha, \beta) \in FA_3$ denote the specific no-return 3-state automaton with transition parameters α and β, as depicted in Figure 1.

Any no-return 3-state automaton starts out in the moderate memory state q^M and stays there, choosing the moderate alternative a^M, as long as all signals it observes are

14

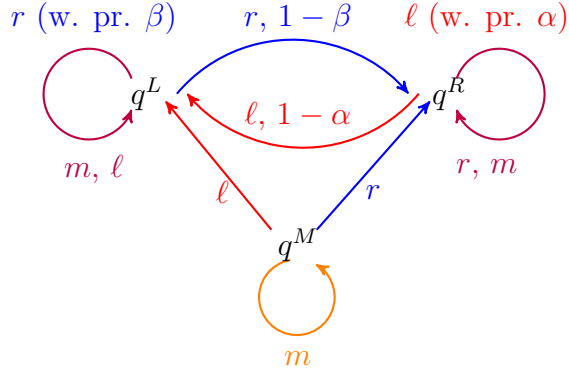| $q_t \backslash s_t$	ℓ	m	r
q^L	$(1,0,0)$	$(1,0,0)$	$(\beta,0,1-\beta)$
q^M	$(1,0,0)$	$(0,1,0)$	$(0,0,1)$
q^R	$(1-\alpha,0,\alpha)$	$(0,0,1)$	$(0,0,1)$

Table 2: Transition probabilities under automata in FA_3

Figure 1: No return 3-state automaton $FA_3(\alpha,\beta)$

moderate (m). But as soon as it observes an extreme (ℓ or r) signal, it transition to the corresponding extreme memory state (q^L or q^R), and it chooses the corresponding extreme alternative (a^L or a^R). Once it arrives at an extreme memory state, a no-return 3-state automaton never returns to moderation; neither in its memory state, nor in the alternative it chooses (hence its name). In the special case in which extreme signals are perfectly informative ($\epsilon = 0$), once an extreme signal reveals the state, any no-return 3-state automaton $FA_3(\alpha,\beta)$ stays at the correct extreme memory state choosing the correct alternative in every future period. Thus, if ($\epsilon = 0$), any no-return 3-state automaton executes the unconstrained optimal decision rule (Lemma 6.1 in the Appendix).

Of course, the case with fully revealing extreme signals is a knife-edge case; if ϵ is positive, a perfectly rational agent with unlimited memory continues to update her belief, even after a very informative signal ℓ or r. Indeed, under state of the world M, although the agent would occasionally receive the strong signals ℓ or r indicating L or R, she would also receive many more m-signals and, in the long run, she would conclude the state of the world is M. Similarly, under a extreme state of the world,

say, $\theta = L$, extreme ℓ-signals would be frequent enough that voters would eventually conclude that state of the world is in all likelihood L.

We reach our case of interest: $\epsilon > 0$ but small, so that extreme signals are rare and hence very (but not perfectly) informative when they arise. In this environment, any no-return 3-state automaton $FA_3(\alpha, \beta)$ in class FA_3 features some attractive qualities for voters who find memory capacity costly ($\kappa > 0$). First, any 3-state automata (including those in class FA_3) is cheap, as the voter only needs to keep tract of which of the three states she is in. Second, $FA_3(\alpha, \beta)$ follows the optimal decision rule (namely, to choose the moderate action) as long as all signals are moderate (which is likely to be for a long while if ϵ is small). Third, if the first extreme signal happens soon enough, $FA_3(\alpha, \beta)$ again follows the unconstrained optimal rule in following the signal to the corresponding extreme memory state and choice of alternative. So far, so good.

The no return 3-state automata only make two kinds of mistakes, relative to the unlimited-memory unconstrained optimum. First, they disregard the small evidence provided by moderate signals. Moderate signals are quite likely in every state, so one such signal does not shift a perfectly-computed posterior much. The no return 3-state automata regard the very little information contained in a moderate signal as negligible, and do not budge in any way upon observing it. But even if one moderate signal does not mean much, an abundance of them does. So no return 3-state automata err in not returning to the moderate memory state after observing a sufficiently long history of signals in which moderate signals are overwhelmingly preponderant. The key to the appeal of no return 3-state automata is that it ϵ is small, it takes a long time to accumulate the large number of moderate signals necessary to compensate for a single extreme signal even if we were to compute the posterior with unlimited memory. So by not returning to moderation when they should, the no return 3-state automata depart from the unconstrained optimal decision rule only far into the future; and an impatient voter finds choices consigned to a distant future to be of little relevance, and not worth incurring a higher cost of memory capacity.

The second problem for a no return 3-state automaton is that, since it does not keep track of how many ℓ or r signals it has observed, it finds itself at a bit of a quandary when it is at an extreme memory state (say q^R) and it observes the opposite extreme signal (say ℓ). Should it ignore the signal, or should it switch memory states? It turns

out that randomization in such situation may be useful, and the values of α and β are relevant to pin down the optimal automaton, which is $FA_3(\alpha, \beta)$ for some α, β in $[0, 1]$.

Proposition 3.1. *(**Voter extremism**) For any cost κ sufficiently small, there exists $\bar{\epsilon} > 0$ such that for all $\epsilon \in (0, \bar{\epsilon})$, each voter optimally follows a no return 3-state automaton, and thus with probability converging to one, she becomes extreme.*

The proof of Proposition 3.1 can be found in the Appendix. Proposition 3.1 tells us that eventually, all voters become extreme. It does not tell us whether they polarize. For that, we turn to the next proposition: even though all voters start with the same prior, that they all observe the same signals, that all automata they each follow are all in the class FA_3, and that they all become extreme, it turns out that they polarize, with different voters arriving at different extremes, each according to their ex-ante lean.

Proposition 3.2. *(**Voter polarization**) For any $\bar{b}$ sufficiently close to one and c sufficiently small, there exist a threshold lean b^*, a range of memory costs (κ) and probabilities of an incorrect extreme signal (ϵ), such that under all states of the world, any optimal behavior by voters is such that with probability converging to one, voters polarize as follows:*

1. *a positive mass of agents with a sufficiently left lean (namely, $\{i \in I : b^i \in (\bar{b}, -b^*)\}$) support a^L;*

2. *agents with a small lean ($\{i \in I : b^i \in (-b^*, b^*)\}$) can swing between supporting a^L and a^R over time, but if the state is moderate, then those who lean left ($b^i < 0$) are more likely to support a^L and those who lean right ($b^i > 0$) are more likely to support a^R; and*

3. *a positive mass of agents with sufficiently right lean (namely, $\{i \in I : b^i \in (b^*, \bar{b})\}$) support a^R.*

We provide the proof, and the exact expression for the threshold b^* in the Appendix. Proposition 3.2 shows that preference heterogeneity affects how agents with limited memory capacity process information. Voters who, in any state of the world, agree on a commonly preferred alternative, who all have a common prior, and who all observe

the same signals, nevertheless end up polarized, supporting opposite extremes in the policy spectrum even if the state of the world is moderate.

The constrained optimal processing of the common signals under limited memory pulls voters' posterior beliefs and policy choices toward the alternative aligned with their political lean. Put differently: voters with a large enough lean act as if they believed the state of the world they prefer among the two extreme ones had come to pass. They polarize in this manner not due to wishful thinking, but rather, as the result of optimizing their choices given the cost of processing infinite sequences of information, with limited memory capacity.

The intuition for this polarizing behavior is as follows: voters start out with the common prior that the correct policy is the moderate one, and they all support the moderate policy as long as they do not see any signal that goes against this prior. If extreme signals are sufficiently rare under the moderate state of the world, when voters see one, they all perceive it as very informative, and they all agree to treat it as if it were correct, and to support the extreme policy that corresponds to this extreme signal (Proposition 3.1). As long as subsequent signals are either additional realizations of the same extreme signal, or hardly informative moderate signals that all voters ignore, all voters continue to agree and to support the extreme alternative congruent with the extreme signals they've seen.

Disagreement, and polarization, arises when voters first receive a contradictory extreme signal; that is, either the first extreme signal was ℓ and now they see an r, or the first extreme signal was r and now they see an ℓ. This is a surprise, and it generates greater uncertainty as to whether the state is L or R (voters all know that the state could also be M, but observing an additional extreme signal does not make this event seem more likely).

Voters who follow an automaton optimally behave as if they follow beliefs that are categorized: each memory state represents such a category, and only big news would trigger a transition from one to another.[8] In the context of our model, the moderate signal does not trigger nay transition but only extreme signal do. However, upon receiving an extreme signal opposite to the memory state they are at, voters who

---

[8]These results are formalized in the Appendix, Lemmas 6.2 and 6.3, which generalize similar findings in Wilson (2014) to three states of the world.

follow a no return 3-state automaton, behave as if they reached a category of beliefs according to which the Moderate state is excluded, but with high uncertainty as to which extreme it is.

Agents with such a category of beliefs would part ways according to their lean: given uncertainty as to which extreme policy is correct, those who lean left will choose the left alternative, and those who lean right will choose the right alternative. Those whose lean is close to zero will mix, with the exact mix compensating their lean just so as to make them indifferent between the two alternatives.[9]

Note that we assumed common signals and common prior as that makes our polarization result strongest. Polarization, of course, does not hinge on such commonality of priors and signals; polarization arises as well if we introduce heterogeneity in prior beliefs and if we allow agent to observe private signals.

4 Candidate Polarization and Policy Extremism

We now close our theory of polarization and extremism by introducing a stylized model of political competition and policy implementation, in which in each period, two parties P^1 and P^2 announce policy platforms, voters vote, and the party that obtains most votes wins and implements its announced platform.

Players. We consider an electoral competition game played by the continuum of voters introduced in Section 2, and by two political parties P^1 and P^2.

Let F_b denote the cumulative distribution function of voter types over $(-\bar{b}, \bar{b})$. We had assumed that this distribution of types has full support; we now also assume that it has a density function f_b that is continuous and symmetric around zero.

We treat each party P^1 and P^2 as a fully rational unitary actor that follows Bayes rule with perfect recall to update beliefs. Parties' sole strategic decisions are to choose policy platforms in each period, as a function of the observed game history. For each $j \in \{1, 2\}$, let $a_t^j \in \mathcal{A}$ denote the platform chosen by Party P^j in period t, where for

---

[9]The higher the probability that a voter sticks to the right memory state q^R after she observes signal ℓ, the less that being in the right state q^R implies to the voter that there were more r than ℓ signals in the now forgotten past (and thus, that the state is R). For a voter who leans toward the right, indifference is thus attained by staying in q^R with higher probability. We compute the optimal transition probabilities $\alpha(b^i)$ and $\beta(b^i)$ in the proof in the Appendix.

any $a \in \mathcal{A}$, committing to policy platform $a_t^j = a$ implies that if P^j wins the period t election, then action a is the alternative a_t^I collectively chosen by society through the political process in period t.

Party differentiation. We assume that parties are endogenously differentiated in their policy-specific quality, as in Hirsch and Shotts (2015) or Eguia and Giovannoni (2017).

If both parties propose the same platform, i.e. if $a_t^1 = a_t^2 = a \in \mathcal{A}$, then one of the two parties is perceived by all voters to be the "leader" on this policy alternative, while the other party is perceived to be the "follower." We may say that the leader "owns the issue" (Petrocik 1996) in the sense that voters perceive that on this policy alternative, the leader has greater policy-specific competence (or "valence") than the follower.

We assume that leader status on a given policy is acquired through experience, and that once acquired, it is sticky, as long as the party continues to choose the same policy.

For each policy $a \in \mathcal{A}$, and for each period $t \in \mathcal{N}$, the party that has proposed platform a for the longest number of consecutive periods up to and including t is the leader on a (so if parties propose different platforms, then each of them is a leader on its chosen platform). If a common platform $a_t^1 = a_t^2 = a$ represents a change of policy from the the previous period's platform for both parties, i.e., if $a_{t-1}^1 \neq a \neq a_{t-1}^2$, we assume that "leader" status is randomly determined at the moment platforms are announced, and this status carries thereafter in future periods for as long as the party that is a leader on a given policy continues to propose this same policy.

Timing and information. The timing in each period t is as follows:
-First, the parties simultaneously commit to their individual platforms, $a_t^i \in \mathcal{A}$, $i = 1, 2$.
-Second, if both parties announce the same platform, and if leader status is undetermined by the past history of play, Nature randomly chooses one of the two parties to be a leader on this common platform (otherwise Nature plays no role at this step). Either way, voters observe the pair of platforms (a_t^1, a_t^2), and if $a_t^1 = a_t^2 = a$ they also observe which of the two parties is the leader on a.
-Third, the common signal $s_t \in \mathcal{S}$ about the state of the world is commonly observed.
-Fourth, each voter i chooses which policy alternative $a_t^i \in \mathcal{A}$ to support.

-Fifth, each voter i chooses one of three voting alternatives: vote for P^1, vote for P^2, or abstain.

-Sixth, the party that obtains a greater share of votes implements its platform $a_t^{w_t}$, with ties broken randomly.

We assume that parties observe the pair of platforms, each party's leader status on its chosen policy, the common signal about the state of the world, the total mass of votes for each party, and the winning party in each period.

In contrast, in each period, each voter observes only the pair of platforms, each party's leader status on its chosen policy, the common signal about the state of the world, and her own private choices of which policy alternative to support and how to vote.

"We model the voters' cost of increasing their memory capacity to process information as a choice of an automaton with costly memory states (as described in sections 2 and 3). Thus, the observed partial history up to period t enters voter i's decision-making in period t only partially and indirectly through its effect on the memory state q_t^i of the finite automaton $\langle Q^i, q_0^i, \tau^i, d^i \rangle$ that voter i uses to guide her decisions. In particular, the finite automaton employed by voter i enters period t at a given memory state q_{t-1}^i; it observes the public signal s_t released at the third step; and between the third and the fourth step in the timing of the strategic environment above, it transitions to a new memory state q_t^i according to τ^i (which can be stochastic), and it produces a recommended alternative to support, $d^i(q_t^i) \in \mathcal{A}$.

Parties' motivations. Parties are office motivated, obtaining a period payoff of 1 if they win, 0 otherwise. We assume that lexicographically, and as long as it does not reduce their probability of winning, parties strictly prefer to obtain a greater vote. We assume that parties are fully impatient, with discount factor $\delta_P = 0$ across periods.[10]

Thus, each party's decision problem in each period is to choose the policy platform that maximizes the probability of winning the current election, and, if there are multiple

---

[10]We have in mind that while "parties" are infinitely lived, in each election the "party" is run by a candidate who only runs once, and thus optimizes solely for the one-shot period game. Unlike infinitely lived voters, who can be interpreted as families who care about their future generations, we argue that party candidates are best understood as individuals interested in their own office prospects, rather than on their party's future generations. As we discuss below, the equilibria with party polarization is robust if parties are patient.

solutions, then to choose among them the policy platform that maximizes vote share.

Voters' motivations. We assume that in each period, each individual voter solves two individual choice problems sequentially.

In each period $t \in \mathbb{N}$, each voter i first chooses which alternative $a_t^i \in \mathcal{A}$ to support. The voter makes this choice following the optimal automaton that solves the optimization problem (5) in Section 2.[11] The solution is as detailed in Proposition 3.1 in Section 3, trading off the desire to choose to support the right alternative, with the cost of recalling information.

Agent i following optimal automaton $\langle Q^i, q_0^i, \tau^i, d^i, \rangle$ chooses to support action $a_t^i = d^i(q_t^i) \in \mathcal{A}$, where q_t^i is the memory state i's automaton reaches in period t, and d^i is the automaton's decision rule.

Once voter i has identified which alternative in $a_t^i \in \mathcal{A}$ she supports, voter i faces a second choice problem; namely, whether to vote for Party 1, to vote for Party 2, or to abstain. Let $v_t^i \in \{P^1, P^2, \emptyset\}$ denote the voting decision of agent i in period t, with $v_t^i = \emptyset$ representing abstention.

We assume that voters obtain an expressive payoff from voting, additive in each period to the utility expression (3), so that the expression of agent i's overall utility in the democratic environment with an election in each period is

$$\lambda \underbrace{\sum_{t=1}^{\infty}(\delta)^t u(a_t^I, \theta, b^i)}_{\text{Instrumental utility}} + (1-\lambda) \underbrace{\sum_{t=1}^{\infty}(\delta)^t u(a_t^i, \theta, b^i)}_{\text{Support expressive utility}} + \underbrace{\sum_{t=1}^{\infty}(\delta)^t u_v(v_t^i)}_{\text{Vote expressive u.}}, \qquad (6)$$

where u_v is the period expressive utility derived from voting.

Since a voter's individual vote cannot have any influence over the election outcome, nor over future play (it cannot even be individually observed by other agents), the instrumental utility component drops out of the summation, and each voter's behavior is driven exclusively by the expressive payoffs.

---

[11]Notice that there is a restriction here. Namely, voters' automata only process information about the state of the world directly obtained through the sequence of signals $(s_t)_{t=1}^{\infty}$, but they do not recognize the potentially informative indirect signaling content of the parties' announced equilibrium platforms. This restriction is supported by evidence that agents overweigh their own experience and their own private signals, over the information that can be inferred from the behavior of other agents (Kogan 2008; Kaustia and Knüpfer 2008)

Expression (6) decouples the act of supporting a policy alternative, from the act of voting. A citizen can support any policy alternative by advocating for it in conversation, in writing, or in civic activism, deriving an expressive payoff from any of these activities. A citizen can only vote by casting a ballot for one of the two competing parties, and it is this specific act that delivers the additional expressive utility term $u_v(v_t^i)$. The expressive utility from the choice of alternative to support is the one maximized, net of the cost of memory, by the optimal automaton in the optimization problem (5). The expressive utility from voting is determined by the voter by her voting choice in each period.

We normalize the expressive utility from abstaining to zero, so $u_v(\emptyset) = 0$ for any agent i, for any period t. We assume that the expressive utility from voting depends on whether the vote aligns or not with the policy alternative that the voter has determined is best, according to her optimal automaton. Namely, if voter i votes sincerely for a party that commits to the alternative chosen by voter i's optimal automaton, then voter i obtains a positive expressive payoff; whereas, if voter i votes for a party committed to an alternative that is not the one chosen by voter i's optimal automaton, then voter i incurs a disutility from such vote.

Formally, there exists a parameter $\bar{u}_v > 0$ such that, for each $j \in \{1, 2\}$, for each voter i and for each period t, if $d^i(q_t^i)$ is the alternative chosen in period t by the optimal automaton chosen by voter i to solve her optimization problem (5), then

$$
u_v(v_t^i) = \begin{cases}
\bar{u}_v & \text{if } v_t^i = P^j, \ a_t^j = d^i(q_t^i) \text{ and } P^j \text{ is the leader at } a_t^j; \\
\bar{u}_v/2 & \text{if } v_t^i = P^j, \ a_t^j = d^i(q_t^i) \text{ and } P^j \text{ is not the leader at } a_t^j; \\
< 0 & \text{if } v_t^i = P^j \text{ and } a_t^j \neq d^i(q_t^i); \text{ and} \\
0 & \text{if } v_t^i = \emptyset.
\end{cases}
\tag{7}
$$

We can think of $d^i(q_t^i)$ as the optimal automaton's recommendation, so that voter i has agency over the choice of alternative $a_t^i \in \mathcal{A}$ given this recommendation. If voters follow their optimal automaton, they support the alternative chosen by their automaton (that is, if $a_t^i = d^i(q_t^i)$ for every voter i and period t). If so, agents derive expressive utility from voting for the alternative they support.

23

Whereas, if voter i deviates and chooses to support an alternative $a \neq d^i(q_t^i)$ that is not the one recommended by the voter's optimal automaton, then the voter enjoys a positive expressive payoff of voting if she votes for a party that commits to alternative $d^i(q_t^i)$, not for voting for a party that commits to a.[12]

An agent's vote has no effect over the agent's current period instrumental payoff, no effect over the current period expressive payoff from the choice of an alternative to support, and no effect on future play. Therefore, the agent's voting problem in each period t reduces to the static optimization problem

$$\max_{v_t^i \in \{P^1, P^2, \emptyset\}} u_v(v_t^i). \tag{8}$$

Equilibrium concept. In our model, an equilibrium is a profile in which each voter chooses an optimal automaton and takes actions aligned with her automaton's recommendations, and in which parties' actions are sequentially rational given voters' behavior and given beliefs updated by Bayes rule. We next formally define this concept.

At the beginning of the game, each voter i chooses a stochastic finite state automaton $\langle Q^i, q_0^i, \tau^i, d^i \rangle \in \mathcal{Q}$, which in period t would transit to memory state q_t^i and produce a recommendation $d^i(q_t^i)$ for which alternative to support, letting agent i make the final choices. A voter i has observed the following at the time she chooses an action in period t: the party platforms and leader status, the public signals, her automaton's memory states and recommended action, and her own chosen action and voting decision in every period up to $t-1$, plus the party platforms and leader status, the public signals and her automaton's memory states and recommended action in period t. Let voter i "support function" refer to a mapping from the set of all these observables for any period t, to the set of actions $\mathcal{A}$. Similarly, the set of all observables at the time voter i chooses her vote in period t includes all of the above, plus her own choice of an action to support in the current period. Let voter i's "voting function" refer to a mapping from the set of all such observables for any period t, to the set of voting

---

[12]We interpret alternative $d^i(q_t^i)$ as the one voter i thinks is best, while alternative a_t^i is the one voter i claims to support in public. Vote v_t^i is cast in a secret ballot, where the positive expressive utility from voting comes from voting one's conscience sincerely for what one thinks best.

options $\{P^1, P^2, \emptyset\}$.

For each party, a pure strategy is a standard object: a mapping from information sets to the set of actions $\mathcal{A}$, and a mixed strategy is a probability distribution over pure strategies.

Definition 4.1. *An **equilibrium** is a support function and a voting function for each voter, and a mixed strategy profile for candidates that satisfy the following.*

1. ***(Voters optimize)*** *There exists a continuous mapping $\phi : (-\bar{b}, \bar{b}) \to \mathcal{Q}$ such that for each voter i, $\phi(b^i) \equiv \langle Q^i, q_0^i, \tau^i, d^i, \rangle \in \mathcal{Q}$ solves optimization problem (5) and is such that for any period t,*

   (a) ***(Sincere support)*** *The action a_t^i that i chooses in period t is $d^i(q_t^i)$ for any realization of all the observables observed by i up to her choice of which action to support in period t.*

   (b) ***(Sincere voting)*** *The vote v_t^i is a solution to*

$$\max_{v_t^i \in \{P^1, P^2, \emptyset\}} u_v(v_t^i).$$

   for any realization of observables observed by i up to her period t vote.

2. ***(Parties optimize)*** *The parties' strategy profile is a sequentially rational profile given the voters' action function and voting function, and given that parties update beliefs according to Bayes rule.*

The intuition behind this formal notion of equilibrium is as follows. Voters want to learn which alternative is best, but it is costly for them to keep track of all the informative signals in detail, so they resort to a cost-efficient automata. An optimal automaton makes the best possible recommendation to maximize the expressive utility from supporting an action, based on the available signals and on the memory constraints induced by the cost of memory capacity. Equilibrium condition $1(a)$ requires voters to follow this recommendation: each voter supports the alternative that is recommended by an automaton that is optimal for her, given her type.

Equilibrium condition $1(b)$ requires each voter to vote optimally, given what she thinks best. Equilibrium condition 2 is that parties best respond at every information

set, given standard Bayesian-updating beliefs. Equivalently, taking voters' optimal behavior as given, parties play a Weak Perfect Bayes Nash equilibrium (Mas Colell, Whinston and Green 1995) of the 2-player electoral competition game induced by voters' behavior.

Condition 1(a) includes a technical requirement: the optimal automata followed by the agents must vary continuously in agents' types. We require this to guarantee that the share of agents of a given type that vote for a given party as a function of type is integrable over the range of possible types. This integrability over the range of types allows us to compute the share of the total population that supports each party.[13]

Results.

We show existence of an equilibrium in Lemma 6.5 in the Appendix. The intuition is as follows. First, each voter's problem has a solution and the solution set is continuous in the voter's type, so there exists a solution that is continuous across types. Second, parties are fully impatient, so they play the election game each period as if it were an independent game; and taking the voters' behavior as given, such a two-player period game played by the two parties is a finite game.

Once voter i has determined which is the best alternative to support according to her optimal automaton, her optimal voting behavior is determined by her expressive payoff of voting: voter i votes for the party that is a leader on the policy that i's optimal automaton recommends; and if neither party chooses this policy alternative, then voter i abstains.

Policy divergence among the two parties follows from our assumption of candidate differentiation between leader and follower: once identified as a "follower", no party will continue to mimic the leader's platform.[14]

What also follows from voter behavior, but perhaps less transparently, is that parties polarize at the extremes in all states of the world; this is the main result in this section.

Proposition 4.1. *(Platform Polarization)* *For any preference parameters $\bar{b}$ sufficiently close to one and c sufficiently small, there exist a range of memory costs (κ)*

---

[13]It would suffice to require instead that ϕ feature at most finitely many discontinuities.

[14]For more sophisticated and richer theories of elections with differentiated candidates, see Krasa and Polborn (2010, 2012, 2014).

and probabilities of an incorrect extreme signal (ϵ), such that under all states of the world, in all equilibria, with probability converging to one in t, in every period $t' \geq t$ parties polarize: one party chooses platform a^L and the other chooses a^R.

Voters' polarization to the extremes (Proposition 3.2) drives candidates' polarization to the same extremes. Expressive voters are uncompromising in the sense that they only vote for a candidate who embraces a policy position that they sincerely support. Candidates, thus, must react as in the quote attributed to French Minister Ledru-Rollin *"there go the people, and I must follow them, for I am their leader."*

In the Moderate state of the world (M), this candidate behavior constitutes pandering. In the state of the world $\theta = M$, over time parties accumulate sufficient moderate signals to be arbitrarily close to certain that the moderate policy a^M is best for every citizen, as it is in fact the case. If voters were fully rational with no memory capacity constraints (i.e. $\kappa = 0$), they too would learn that moderate policies are best, and at least one party would offer moderation and would get elected. However, voters with limited memory capacity are swayed more by a few strong extreme signals than by the very many forgettable –and indeed forgotten– moderate ones, so they remain poorly informed. Even though parties know a moderate policy would be best, they cater to voters by choosing extreme policy platforms.[15] Implemented policies are extreme, and they are volatile, switching back and forth between stretches of periods of extreme right and stretches of extreme left policy, their duration until the next switch in public sentiment determined by the noise in the publicly observed signals. Aggregate welfare thus suffers.

On the other hand, if the state of the world θ is extreme –and if the distribution of types over $(-\bar{b}, \bar{b})$ is sufficiently close to symmetric around zero– then in most periods the majority of voters supports the correct extreme policy, and votes for the party that leads on this extreme policy.

Summarizing our results, in an environment in which extreme signals are rare, and moderate signals are abundant, each individual moderate signal conveys very little information. With costly memory capacity, such weak signals are best ignored, and thus

---

[15]The rationale for such pandering was perhaps best articulated by then Prime Minister of Luxembourg, Jean-Paul Juncker: *"We all know what to do, but we don't know how to get re-elected once we have done it"*, quoted in The Economist in "The quest for prosperity", March 15, 2007.

an extreme signal cannot be overcome by any number of moderate signals. Voters who follow constrained optimal decision rules with limited memory capacity then become extreme regardless of the state of the world. Given conflicting extreme signals, voters polarize over the two extremes according to their relative gains and losses from each extreme alternative in each extreme state. Candidates know better, but chasing votes, they too locate at the extremes, where the voters are found. Each party consolidates a reputation on one particular extreme policy. Party positions thus become stable, at opposite extremes. Electoral majorities and implemented policies become volatile, switching back and forth between the two extremes.

Party polarization at opposite extremes also holds in equilibrium if we relax the assumption that parties are fully impatient, and we assume instead that parties are forward looking with some patience. Suppose parties P^1 and P^2 were patient, with respective discount factors $\delta_1 \in (0,1)$ and $\delta_2 \in (0,1)$. Suppose that party P^1 is a leader on platform a^L at some period t, where t is a period after the first extreme signal, and that party P^1 chooses policy a^L in every subsequent period. Regardless of parties' patience, after the first extreme signal, the center of the ideological distribution of voters hollows out irreversibly (Proposition 3.2). Voters (who are sincere in the sense of Definition 4.1) thereafter never vote for a moderate party. Given our modeling of differentiated candidates with of the endogenous policy-specific advantages, a party that is a late adopter of a policy position cannot successfully challenge a party that is an established leader on this position. Combining these two factors, Party P^2 would obtain no votes in any period in which it offered policy a^L or a^M. Therefore, regardless of its patience, Party P^2 best responds to an invariant sequence of left-extreme platforms by choosing its own extreme sequence of platforms $p_\tau^2 = a^R$ for any subsequent period $\tau > t$, for any continuation history.

The party polarization result is not driven by parties' impatience; rather, it is driven by voters' polarization, and by the parties' desire to pander to the polarized voters.

5 Discussion

Processing and recalling information is costly. Rational voters with costly recall who observe a sequence of signals about the state of the world must optimally trade-off

	L	M	R
2004	28%	57%	15%
2017	42%	37%	21%

Table 3: Polarization of the US electorate, 2004-2017.

the benefits of becoming better informed, with the costs of processing these signals. In an environment in which extreme signals are rare and very informative when they do occur, and moderate signals are very abundant and thus not very informative, it is optimal for an impatient voter with costly recall to pay attention only to extreme signals. This constrained optimal behavior introduces a bias, in that eventually, all voters become extreme in all states of the world.

We show that even if all agents share a common prior, all signals are publicly observed, and all voters share a common preference order over alternatives in any state of the world, voters still polarize at opposite extremes in the presence of conflicting signals. Voters polarize according to their relative gains and losses from choosing the wrong alternative in each extreme state (Proposition 3.2).

Fully rational office-motivated parties with perfect recall pander to the electorate and polarize as well: policy platforms diverge and become stable, with one party at each extreme (Proposition 4.1).

This theory of voter-driven polarization, in which both the electorate and the two main parties become polarized, with one party and a mass of voters at each of two opposite extremes, is consistent with the following stylized empirical evidence on the joint polarization of the US electorate and of the perceived positions of the two main political parties in the US, in the past two decades.

First, longitudinal survey data collected by the Pew Center shows that the US electorate polarized from 2004 to 2017.[16] The Pew surveys asked ten ideological questions, and placed respondents on a left/right scale from -10 to $+10$ based on their answers. To fit our simpler model, we partition the scale into three equal length intervals: L for values from -10 to -4 or below, M for intermediate values from -3 to $+3$, and R for values from $+4$ to $+10$. Table 3 shows the sharp change in the distribution of

---

[16]We use the dataset from the following Pew Center Political Typology surveys of 2004 (December 1-16, 2004, n=2,000) and 2017 Phase A (June 8-18, 2017, n=2,504).

voters in each of these three categories from 2004 to 2017. While respondents in the intermediate category constitute a large majority of the 2004 electorate, the size of this group shrinks sharply in little more than a decade, while at the same time the masses of voters on the left and on the right both enlarge substantially. Put differently, while in 2004 for every four voters located in the middle, there were only three at the extremes, by 2017 there were seven voters at the extremes, for every four in the middle.[17]

Figure 2 illustrates this polarization graphically. In the background, and with faded colors, it shows the electorate's distribution over L ("Left"), M, and R ("right") in 2004. At the front, in bright colors, the analogous distribution in 2017. We report additional descriptive evidence using the latest five Pew Center longitudinal surveys on this common set of questions in Table 5 in the Appendix. That expanded table shows a monotonic pattern of polarization from 2004 to 2011 to 2014 and then to 2017, which stabilizes from 2017 to 2019.

The American electorate's polarization was accompanied by the (perceived) polarization of its two main parties. Longitudinal data from the American National Election Studies (ANES) surveys shows the evolution of the position of the Democratic and Republican parties on a 0 to 10 ideological scale, at each Presidential election year since 2004, as perceived by respondents. As shown by Figure 3, the gap between the two parties has steadily widened since 2008, more than doubling by 2020, and if from 2004 to 2012 both parties were perceived to locate within the middle third of the scale, in 2020 the parties are perceived to be located one at each of the two extreme thirds of the scale.

The simultaneous voters' polarization and party polarization is qualitatively consistent with our theory. We say that our theory "is consistent" with these empirical findings, rather than saying that these findings "are as predicted" by our theory, because our awareness of voters' and parties' polarization (albeit not of its exact quantitative values), precedes and motivates our development of the theory. We provide a micro-founded explanation for the observed polarization that inspired our study.

---

[17]Slightly stronger evidence of polarization arises if we follow instead Pew's partition of respondents into five categories, with the middle category corresponding only to values from -2 to $+2$ (Pew Research Center, 2017, page 11), or if we consider the Pew Center's subsample of citizens who coded by Pew as "politically engaged", i.e. citizens who are registered to vote, follow government and public affairs most of the time and say they vote "always" or "nearly always'."
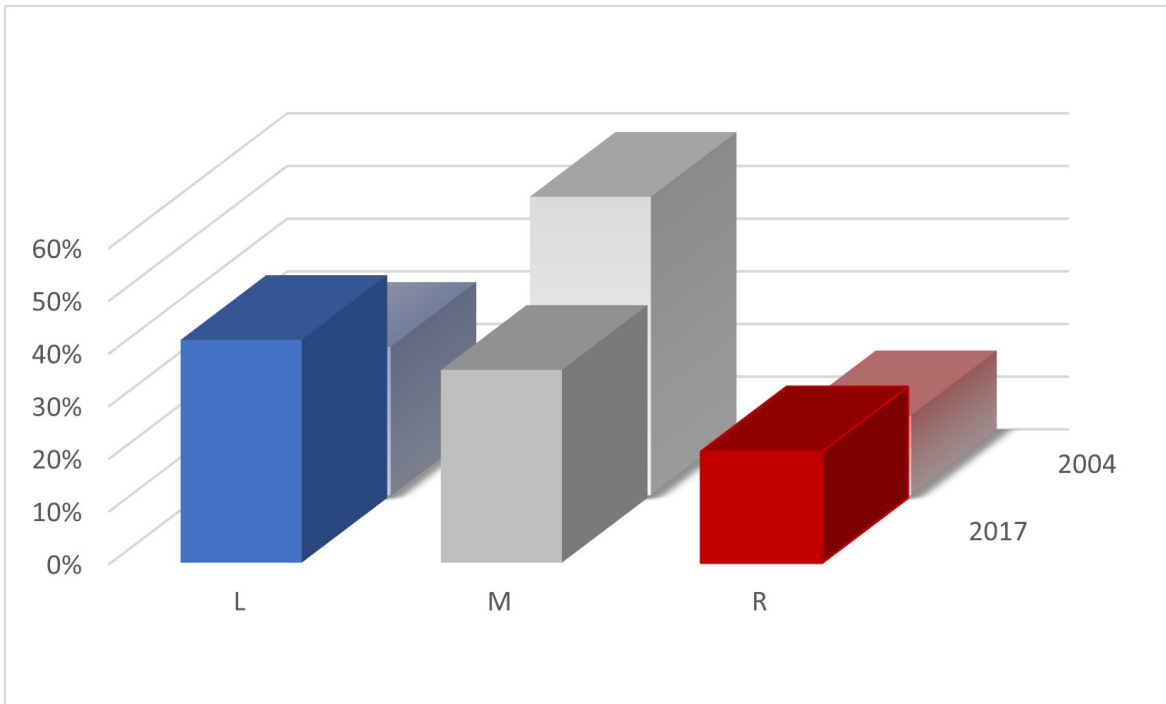
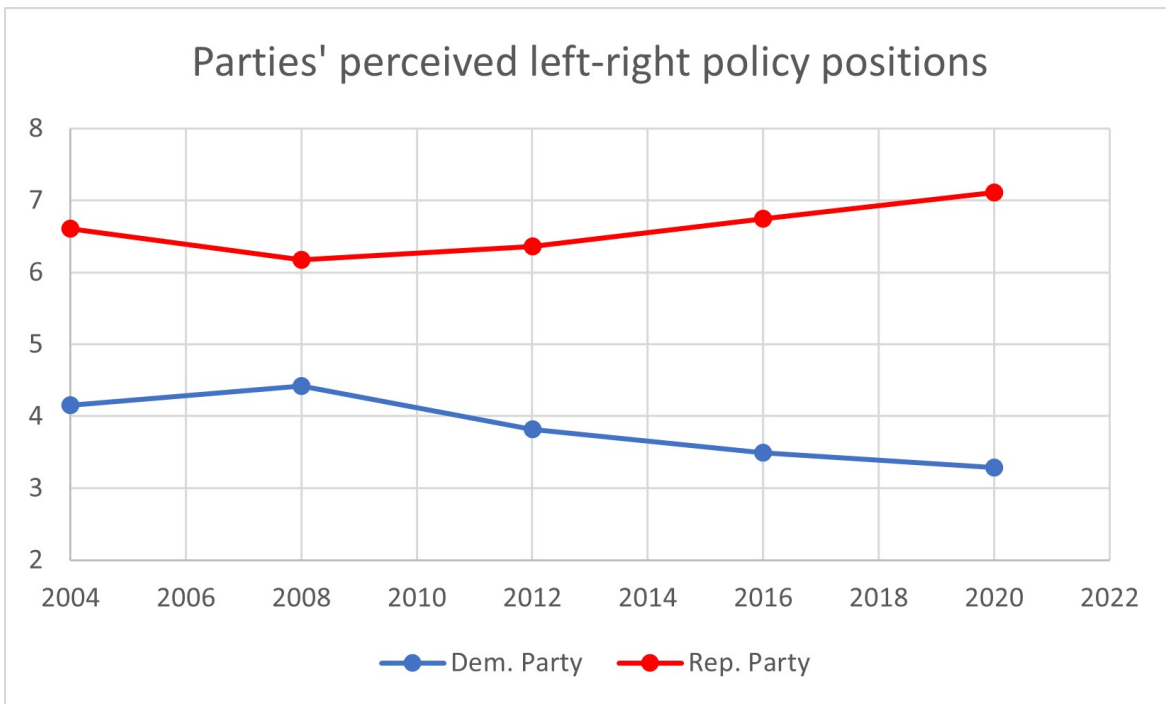Figure 2: Polarization of the US electorate, 2004-2017.

Figure 3: Party polarization in the US, 2004-2020.

What our theory does predict is that, absent a shock outside our model, polarization is irreversible: If society starts from initial conditions with a consensus around a moderate policy, in an environment with frequent (and thus weak) moderate signals and ex-ante rare (and thus strong when they do arise) extreme signals, we show that – otherwise rational– agents with costly recall eventually favor extreme policies, polarize, and never return to moderation. Further, in such a polarized society, party platforms polarize as well, and remain polarized.

Whether polarization is "permanent" depends on the time scale under consideration. In the United States, polarization was permanently high for a few decades during the Gilded Age and the Progressive area (1880s-1910s) (Putnam 2020), and it is once again high a little over a century later... but in between these two crests, during World War II and the post-war period there was a period of a broad national consensus and low polarization that lasted through the Eisenhower Presidency in the 1950s (Trilling 1950, Hofstadter 1964). It thus seems that over long historical time-spans, polarization ebbs and flows in waves. While our theory is dynamic with an infinite horizon, we interpret it as applicable to explain only part of this historical evolution: the rise of polarization over a relatively short period of time. The return to a national consensus from a polarized state responds to factors outside our model. Chief among such potential factors are external security threats (Desch 1995) that trigger a "rally round the flag" unifying effect (Baker and O'Neal 2001, Groeling and Baum 2008).

One could, if desired, embed our theory into a more general account of the rise and fall of polarization over the history of a democratic nation: suppose that at the conclusion of each period, with some small probability, the entire political environment suffers a common-knowledge shock that resets the state of the world, drawing it anew from its common prior distribution. Formally, this is but a small departure from our model: it suffices to interpret the shock as "ending" the game and starting a new one,[18] and to reinterpret the discount factor λ as including both the time discount and the probability that the game ends after each period. Agents then play a new game after each shock. If we observe their behavior over the infinite sequence of such games, we find it cyclical: agents abruptly return to moderation and consensus to start each

---

[18]Just like the Algiers crisis of 1958 triggered the end of the Fourth Republic and the advent of the Fifth Republic in France.

new game, and then in this new game they eventually polarize at the extremes, before returning to moderation at the next game-ending shock.

If we are correct in our prediction that —until the next unifying national crisis, threat or shock— polarization is permanent, social interventions that seek to nudge the electorate back to moderation (such as regulating online content, deplatforming extreme speakers, or seeking to break information bubbles by exposing audiences to both sides of an argument) are unlikely to succeed. Rather, social interventions such as fostering norms of pluralism, support for view-point diversity and tolerance of dissent, and enshrining civil discourse and the democratic process as the means to channel ideological disagreement, may be necessary to manage ongoing political polarization.

6   Appendix: Proofs

In this Appendix we provide all the proofs for the results we stated in sections 3 and 4. We start by formally stating and proving the claim that if extreme signals are fully informative, then any no return 3-state automaton is optimal.

Lemma 6.1. *Under the preferences and the information structure given by Table 1 with $\epsilon = 0$ and Assumption (1), for any α, $\beta \in [0, 1]$, the unconstrained optimal rule can be implemented by $FA_3(\alpha, \beta)$. Moreover, any optimal SFSA with three memory states is in FA_3.*

Proof. The optimal decision rule, denoted by D^*, is such that for any $t \in \mathbb{N}$, and for any sequence $(a^\tau)_{\tau=1}^{t-1}$,

$$D^*(s_1, ..., s_t; a_1, ..., a_{t-1}) = a^M \text{ if } s_\nu = m \text{ for all } \nu = 1, ..., t;$$

$$D^*(s_1, ..., s_t; a_1, ..., a_{t-1}) = a^L \text{ if } s_\nu = \ell \text{ for some } \nu \text{ such that } s_{\nu'} = m \text{ for all } \nu' < \nu;$$

$$D^*(s_1, ..., s_t; a_1, ..., a_{t-1}) = a^R \text{ if } s_\nu = r \text{ for some } \nu \text{ such that } s_{\nu'} = m \text{ for all } \nu' < \nu.$$

$$(9)$$

Under the information structure given by Table 1 with $\epsilon = 0$, it is impossible to have an r-signal followed by an ℓ-signal or vice versa. Thus, the transition probabilities α and β do not matter under $\epsilon = 0$, and $FA_3^p(\alpha, \beta)$ implements D^* for any $\alpha, \beta \in [0, 1]$. $\square$

To establish Proposition 3.1, we first generalize results on multi-self consistency for two states of nature (Wilson 2014), to a general setting with multiple states of nature. We present these results in the general setting with signals from a finite set $\mathcal{S}$; the environment with only only three states and three signals that we use in our theory is a special case. Given a state of nature θ and a memory state $q \in Q$, for an agent i with type b^i, the expected payoff accumulated from q conditional on θ is then

$$\mathbf{1}_{q=q_0} u[d(q_0), \theta, b^i] + \delta \sum_{s_1 \in \mathcal{S}} \tau(q_0, s_1; q) \mu_{s_1}^\theta u[d(q), \theta]$$

$$+ (\delta)^2 \sum_{q_1 \in Q, s_1, s_2 \in \mathcal{S}} \tau(q^o, s_1; q_1) \mu_{s_1}^\theta \tau(q_1, s_2; q) \mu_{s_2}^\theta u[d(q), \theta, b^i] + ... \qquad (10)$$

$$= \frac{1}{1 - \delta} f(q|\theta) u[d(q), \theta, b^i],$$

where

$$f(q|\theta) = \sum_{T=1}^{\infty} (1-\delta)(\delta)^{T-1} \left[\sum_{(q_1,\ldots,q_{T-1}),(s_1,\ldots,s_{T-1}),q_T=q} \mathbf{1}_{q_1=q_0} \prod_{t=1}^{T-1} \mu_{s_t}^{\theta} \tau(q_t, s_t; q_{t+1}) \right]. \quad (11)$$

As noted in Wilson (2014), $f(q|\theta)$ is the stationary distribution under the transition probability from q' to q given by

$$T^{\theta}(q'; q) = \sum_{s \in \mathcal{S}} [(1-\delta)\mathbf{1}_{q=q_0} + \delta\mu_s^{\theta}\tau(q', s; q)]. \quad (12)$$

Extending Piccione and Rubinstein (1997), we can define the "belief" at $q \in Q$ as

$$p(q)(\theta) = \frac{\mathbf{P}_0(\theta)f(q|\theta)}{\sum_{\theta'} \mathbf{P}_0 f(q|\theta')} \text{ and } p(q, s)(\theta) = \frac{\mathbf{P}_0(\theta)f(q|\theta)\mu_s^{\theta}}{\sum_{\theta'} \mathbf{P}_0(\theta)f(q|\theta)\mu_s^{\theta'}}. \quad (13)$$

To characterize an optimal SFSA, we use $V_q(\theta, b^i)$ to denote the continuation value for agent i with type b^i at memory state q conditional on the state of nature being θ.

With this notation, we can now state the first the two lemmas, which extends a modified multi-self consistency result by Wilson (2014) to our environment.

Lemma 6.2. *Let $K \in \mathbb{N}$ and assume $\langle Q, q_0, \tau, d \rangle$ is an optimal SFSA under prior $\mathbf{P}_0$ among those with $|Q| \le K$. Then, for any type $b^i \in (-\bar{b}, \bar{b})$,*

1. *(Multi-self consistency—transition) For each memory state $q \in Q$ with $\sum_{\theta} \mathbf{P}_0(\theta)f(q|\theta) > 0$, each signal s, and any q' such that $\tau(q, s; q') > 0$,*

$$q' \in \arg\max_{q'' \in Q} \sum_{\theta} p(q, s)(\theta)V_{q''}(\theta, b^i); \quad (14)$$

2. *(Multi-self consistency—action) for each memory state $q \in Q$ with $\sum_{\theta} \mathbf{P}_0(\theta)f(q|\theta) > 0$,*

$$d(q) \in \arg\max_{a \in A} \sum_{\theta} p(q)(\theta)u(a, \theta, b^i); and \quad (15)$$

35

3. *(Revelation Principle) for any $q \in Q$,*

$$q \in \arg\max_{q' \in Q} \sum_{\theta} p(q)(\theta) V_{q'}(\theta, b^i). \tag{16}$$

Proof. For any memory states q and q', define the set

$$W_{q,q'} = \bigcup_{n=1}^{\infty} W_{q,q'}^n,$$

where for each $n = 1, 2, \ldots.,$

$$W_{q,q'}^n = \{\mathbf{w} = (q, s_1; q_1, s_2; \ldots.; q_{n-1}, s_n; q') : s_i \in \mathcal{S}, q_i \in Q\},$$

that is, the set of possible state transitions from q to q'. Given a state of nature θ and $\mathbf{w} \in W_{q,q'}^n$, define

$$\mathbb{P}_\theta(\mathbf{w}) = (1 - \delta)\delta^{n-1} \times \prod_{i=1}^{n} \mu_{s_i}^\theta \tau(q_{i-1}, s_i; q_i),$$

where $q_0 = q$ and $q_n = q'$. The expected payoff from the SFSA is then

$$V = \frac{1}{1-\delta} \sum_{q \in Q} \left\{ \sum_{\theta} \mathbf{P}_0(\theta) \sum_{\mathbf{w} \in W_{q^o,q}} \mathbb{P}_\theta(\mathbf{w}) \right\} u[d(q), \theta, b^i]. \tag{17}$$

We now prove (14) and (15).

First, consider (15). Suppose, by contradiction, that for some memory state $\hat{q}$ with $\sum_{\theta} \mathbf{P}_0(\theta) f(\hat{q}|\theta) > 0$ such that (15) does not hold, and hence there is an action $a' \neq d(q) = a$ that solves the problem in (15) with a strict preference. By (11), $f(\hat{q}|\theta) = \sum_{\mathbf{w} \in W_{q^o,\hat{q}}} \mathbb{P}_\theta(\mathbf{w})$, this then implies that

$$\sum_{\theta} \mathbf{P}_0(\theta) \sum_{\mathbf{w} \in W_{q^o,\hat{q}}} \mathbb{P}_\theta(\mathbf{w}) u(a, \theta, b^i) < \sum_{\theta} \mathbf{P}_0(\theta) \sum_{\mathbf{w} \in W_{q^o,\hat{q}}} \mathbb{P}_\theta(\mathbf{w}) u(a', \theta, b^i). \tag{18}$$

Now, consider the alternative SFSA, which differs from the given one only in that $d'(\hat{q}) = a'$. From (17) and (18) it follows that this alternative SFSA gives a strictly higher expected payoff than the given one, a contradiction to the optimality of the given SFSA.

Now consider (14). Suppose, by contradiction, that $\tau(q, s; q') > 0$ and that for some

36

$q'' \neq q'$,

$$\sum_\theta p(q,s)(\theta)V_{q'}(\theta,b^i) < \sum_\theta p(q,s)(\theta)V_{q''}(\theta,b^i). \tag{19}$$

We denote $p' = \tau(q,s;q')$ and $p'' = \tau(q,s;q'')$. Now, fix all other transition probabilities other than p' and p'', each term $\mathbb{P}_\theta(\mathbf{w})$ in V given by (17) is a polynomial of (p',p'') and, since $\eta \in (0,1)$, V is differentiable w.r.t. (p',p''). Since the given SFSA is optimal and $p' = \tau(q,s;q') > 0$, the FOCs require that $\frac{\partial}{\partial p'}V \geq \frac{\partial}{\partial p''}V$. However, we show below that (19) implies that

$$\frac{\partial}{\partial p''}V > \frac{\partial}{\partial p'}V, \tag{20}$$

a contradiction to the optimality of M.

To prove (20), it is straightforward to verify that

$$\frac{\partial}{\partial p'}V = \frac{1}{1-\delta}\sum_{\hat{q}\in Q}\left\{\sum_\theta \mathbf{P}_0(\theta)\sum_{\mathbf{w}\in W_{q^o,\hat{q}}(q,s;q')}\varphi_{(q,s;q')}(\mathbf{w})\frac{\mathbb{P}_\theta(\mathbf{w})}{p'}\right\}u[d(\hat{q}),\theta,b^i], \tag{21}$$

where

$$W_{q^o,\hat{q}}(q,s;q') = \{\mathbf{w}\in W_{q^o,\hat{q}} : (q,s,q') \text{ occurs in } \mathbf{w}\}$$

and $\varphi_{(q,s;q')}(\mathbf{w})$ is the number of repetitions of $(q,s;q')$ within $\mathbf{w}$.

Now, we show that $\frac{\partial}{\partial p'}V$ is proportional to $\sum_\theta p(q,s)(\theta)V_{q'}(\theta,b^i)$:

$$\left[\sum_\theta \mathbf{P}_0(\theta)f(q|\theta)\mu_s^\theta\right]\left[\sum_\theta p(q,s)(\theta)V_{q'}(\theta,b^i)\right] = \sum_\theta \mathbf{P}_0(\theta)f(q|\theta)\mu_s^\theta V_{q'}(\theta,b^i)$$

$$= \frac{1}{1-\delta}\sum_\theta \mathbf{P}_0(\theta)\sum_{\hat{q}\in Q}\left\{\left[\sum_{\mathbf{w}_q\in W_{q^o,q}}\mathbb{P}_\theta(\mathbf{w}_q)\right]\mu_s^\theta\left[\sum_{\mathbf{w}_{q'}\in W_{q',\hat{q}}}\mathbb{P}_\theta(\mathbf{w}_{q'})\right]\right\}u[d(\hat{q}),\theta,b^i]$$

$$= \frac{1}{1-\delta}\sum_\theta \mathbf{P}_0(\theta)\sum_{\hat{q}\in Q}\left\{\sum_{\mathbf{w}_q\in W_{q^o,q},\mathbf{w}_{q'}\in W_{q',\hat{q}}}\frac{\mathbb{P}_\theta[(\mathbf{w}_q,s;\mathbf{w}_q')]}{\tau(q,s;q')}\right\}u[d(\hat{q}),\theta,b^i]$$

$$= \frac{1}{1-\delta}\sum_\theta \mathbf{P}_0(\theta)\sum_{\hat{q}\in Q,a\in A}\left\{\sum_{\mathbf{w}\in W_{q^o,\hat{q}}}\varphi_{(q,s;q')}(\mathbf{w})\frac{\mathbb{P}_\theta(\mathbf{w})}{p'}\right\}d(a,\hat{q})u(a,\theta,b^i) = \frac{\partial}{\partial p'}V,$$

where the last equality follows from (21) and the second last equality follows from $p' = \tau(q,s;q')$ and the fact that for any $\mathbf{w}_q \in W_{q^o,q}$ and any $\mathbf{w}_{q'} \in W_{q',\hat{q}}$, $(\mathbf{w}_q,s;\mathbf{w}_{q'}) \in$

$W_{q^o,\hat{q}}(q,s;q')$ and that each $\mathbf{w} \in W_{q^o,\hat{q}}(q,s;q')$ is counted $\varphi_{(q,s;q')}(\mathbf{w})$ times in that list. We have analogous expression for $\frac{\partial}{\partial p''}V$, and hence (19) implies that (20).

Now we prove (16). By modified multi-self consistency, for any $s \in \mathcal{S}$ and any q_1, q_2 with $\tau(q,s;q_1) > 0$ and $\tau(q,s;q_2) > 0$ and any $q_3 \in Q$,

$$\sum_\theta p(q,s)(\theta)V_{q_1}(\theta, b^i) = \sum_\theta p(q,s)(\theta)V_{q_2}(\theta, b^i) \geq \sum_\theta p(q,s)(\theta)V_{q_3}(\theta, b^i),$$

By (13), this implies that

$$\sum_\theta \mathbf{P}_0(\theta)f(q|\theta)\mu_s^\theta V_{q_1}(\theta, b^i) = \sum_\theta \mathbf{P}_0(\theta)f(q|\theta)\mu_s^\theta V_{q_2}(\theta, b^i) \geq \sum_\theta \mathbf{P}_0(\theta)f(q|\theta)\mu_s^\theta V_{q_3}(\theta, b^i).$$
$$(22)$$

Thus,

$$\sum_\theta p(q)(\theta)V_q(\theta, b^i)$$

$$= \sum_\theta p(q)(\theta)\left\{ u[d(q), \theta, b^i] + \delta \left[\sum_{s \in \mathcal{S}, q'' \in Q} \mu_s^\theta \tau(q,s;q'')V_{q''}(\theta, b^i) \right] \right\}$$

$$= \sum_\theta \{p(q)(\theta)u[d(q), \theta, b^i]\} + \delta \sum_{s \in \mathcal{S}} \left\{ \sum_{q'' \in Q} \frac{\sum_\theta \mathbf{P}_0(\theta)f(q|\theta)\mu_s^\theta V_{q''}(\theta, b^i)}{\sum_{\theta'} \mathbf{P}_0(\theta')f(q|\theta')} \tau(q,s;q'') \right\}$$

$$\geq \sum_\theta \{p(q)(\theta)u[d(q'), \theta, b^i]\} + \delta \sum_{s \in \mathcal{S}} \left\{ \sum_{q'' \in Q} \frac{\sum_\theta \mathbf{P}_0(\theta)f(q|\theta)\mu_s^\theta V_{q''}(\theta, b^i)}{\sum_{\theta'} \mathbf{P}_0(\theta')f(q|\theta')} \tau(q',s;q'') \right\}$$

$$= \sum_\theta p(q)(\theta)V_{q'}(\theta, b^i),$$

where the first equality follows from the recursive equation for $V_q(\theta, b^i)$ for each θ, the second follows from (13), the inequality follows term by term, first the terms without δ follow from (15), the terms starting with δ follows from (22), again term by term for each s: any term with q'' with $\tau(q,s;q'') > 0$ has the same value in the inequality above, and that value is no less than that for the corresponding term with $\tau(q', s; q'') > 0$, and the last equality follows from the recursive equation for $V_{q'}(\theta, b^i)$. $\qquad\square$

We say that two memory states are called *equivalent* if they share the same transition rules to any other states or their equivalents, and have the same decision rule.

With this definition, and reformulating the necessary conditions for an optimal SFSA in Lemma 6.2, we obtain the following, more convenient, necessary conditions for optimality.

Lemma 6.3. *Let $K \in \mathcal{N}$ and assume $\langle Q, q_0, \tau, d \rangle$ is a SFSA without equivalent states that is optimal among those of size $|Q| = K$. For each $q \in Q$, and for each type $b^i \in (-\bar{b}, \bar{b})$, define*

$$\Pi_q(b^i) = \left\{ p \in \Delta(\Theta) : \sum_{\theta \in \Theta} p(\theta) V_q(\theta, b^i) \geq \sum_{\theta \in \Theta} p(\theta) V_{q'}(\theta, b^i) \text{ for all } q' \in Q \right\}. \quad (23)$$

Then, for each $q \in Q$ with $p(q)$ and $p(q, s)$ defined by the pair of expressions (13),

$$\tau(q, s; q') > 0 \quad \Rightarrow \quad p(q, s) \in \Pi_{q'}(b^i), \quad (24)$$

$$d(q) = a^\theta \quad \Rightarrow \quad p(q) \in \Delta^\theta(b^i). \quad (25)$$

Proof. Consider first (24). Suppose that $\tau(q, s; q') > 0$ in the optimal SFSA. Then, (14) implies that $p(q, s) \in \Pi_{q'}(b^i)$. Similarly, (25) follows immediately from (15). □

Lemma 6.3 gives necessary conditions for optimality based on conditions for the optimal transition probabilities, and we use these conditions below to show that for agent i with lean b_i, and for a particular pair of values $\alpha(b^i)$ and $\beta(b^i)$ that depend on b^i, $FA_3(\alpha(b^i), \beta(b^i))$ is the optimal SFSA among those with at most three memory states, for a range of ϵ's above zero.

First we show that for a range of parameters, the optimal stochastic finite state automaton has three states.

Lemma 6.4. *There exist $\bar{\kappa} \in \mathbb{R}_{++}$ and a function $\bar{\epsilon} : (0, \bar{\kappa}) \to \mathbb{R}_{++}$ such that for any $b^i \in (-\bar{b}, \bar{b})$, for any $\kappa \in (0, \bar{\kappa})$ and for any $\epsilon \in (0, \bar{\epsilon}(\kappa))$, the optimal automaton for agent i has three memory states.*

Proof. For any $K \in \mathbb{N}$, let $\bar{V}_K(\epsilon, b^i)$ be the optimal payoff from K-memory-state finite automata under $\epsilon \geq 0$ for agent i with type $b^i \in (-\bar{b}, \bar{b})$. Note that for any $b^i \in (-\bar{b}, \bar{b})$, $\bar{V}_2(0, b^i) > \bar{V}_1(0, b^i)$, $\bar{V}_K(0, b^i) > \bar{V}_2(0, b^i)$ for all $K \geq 3$, and $\bar{V}_K(0, b^i) = \bar{V}_{K'}(0, b^i)$ for all $K \geq 3$ and all $K' \geq K$.

Define $\kappa_1(b^i) \equiv \frac{\bar{V}_3(0,b^i) - \bar{V}_1(0,b^i))}{2}$, $\kappa_2(b^i) \equiv \bar{V}_3(0,b^i) - \bar{V}_2(0,b^i)$, $\kappa(b^i) \equiv \min\{\kappa_1(b^i), \kappa_2(b^i)\}$, and $\bar{\kappa} \equiv \inf_{b^i \in (-\bar{b}, \bar{b})} \kappa(b^i)$, and notice that $k(b^i) > 0$ for any $b^i \in (-\bar{b}, \bar{b})$, with values bounded away from zero over $(-\bar{b}, \bar{b})$, so $\bar{\kappa} > 0$. Assume $\kappa \in (0, \bar{\kappa})$; then every agent prefers the optimal three-state automaton over any automata with fewer states.

Given κ, by continuity of $\bar{V}_k(\epsilon, b^i)$ with respect to ϵ, there exists $\bar{\epsilon}(\kappa) \in \mathbb{R}_{++}$ sufficiently small that

$$\bar{V}_k(\bar{\epsilon}(\kappa), b^i) - \bar{V}_3(\bar{\epsilon}(\kappa), b^i) < k\kappa$$

for all $k \leq \bar{V}_3(0, b^i)/\kappa$ and for any $b^i \in (-\bar{b}, \bar{b})$. Then for any $\epsilon \in (0, \bar{\epsilon}(\kappa))$, the optimal automaton for any agent i has three memory states. $\qquad\square$

We next proceed within the range of parameters for which the optimal automaton has three memory states. We seek to establish that the optimal automaton is a no return three state automaton. Our proof strategy is the following.

We divide the set of SFSA into two groups. The automata in the first group have transition probabilities close to those in $FA_3(\alpha, \beta)$, while the second group consist of all others. We then show that $FA_3(\alpha, \beta)$ with optimal α and β is the unique optimal SFSA within the first group, and outperforms those in the second group. The first claim is proved using Lemma 6.3, while the second follows from the uniqueness in Lemma 6.1 and continuity of the optimal value for ϵ close to zero.

To proceed with this argument, we need to define a distance between SFSA. For any two SFSA $\langle Q, q_0, \tau, d \rangle$ and $\langle Q, q_0, \tau', d \rangle$, define the distance between them as $\max_{q \in Q} \|\tau(q, s) - \tau'(q, s)\|$, where $\|\cdot\|$ is the Euclidean distance over $\Delta(Q)$.

Now, let $\tau_{(\alpha, \beta)}(q, s)$ and d_3 denote the transition probabilities given $(q, s) \in Q \times \mathcal{S}$ and the decision rule of automaton $FA_3(\alpha, \beta)$, and define

$$\mathcal{FA}(\rho) \equiv \left\{ \langle (q^L, q^M, q^R), q^M, \tau, d_3 \rangle : \|\tau(q, s) - \tau_{(\alpha, \beta)}(q, s)\| < \rho \text{ for all } (q, s) \neq (q_R, \ell), (q_L, r) \right\}.$$
(26)

That is, $\mathcal{FA}(\rho)$ consists of SFSA within distance of ρ to some SFSA in class FA_3. Let $\mathcal{FA}^c(\rho)$ denote the set of all SFSA with $|Q| = 3$ not in $\mathcal{FA}(\rho)$.

We are now ready to prove Proposition 3.1.

Proof of Proposition 3.1.

Proof. We proceed in two steps. We first show that — for a range of parameter— some no return three state automata outperform any three state automata in $\mathcal{FA}^c(\rho)$, so if there exists an optimal three state automaton among those in $\mathcal{FA}(\rho)$, then this automation is optimal among all those with three states, and thus, by Lemma 6.4, it is the optimal automaton among all automata. The second step is to establish that there exists a no-return three-state automaton that is optimal among those in $\mathcal{FA}(\rho)$.

Step 1. For any $\rho, \epsilon \in \mathbb{R}_{++}$, and any $b^i \in (-\bar{b}, \bar{b})$, define

$$W(\rho, \epsilon, b^i) \equiv \max_{FA \in \mathcal{FA}^c(\rho)} V(FA, \epsilon, b^i),$$

where $V(FA, \epsilon, b^i)$ is the expected ex ante payoff for agent i with type $b^i \in (-\bar{b}, \bar{b})$ from an arbitrary SFSA FA under ϵ. Notice that since $\mathcal{FA}^c(\rho)$ is compact, and $V(FA, \epsilon, b^i)$ is continuous in FA, the maximum exists and $W(\rho, \epsilon, b^i)$ is well defined.

For any $b^i \in (-\bar{b}, \bar{b})$ and for any $\alpha, \beta \in [0, 1]$, $W(\rho, 0, b^i) < V[FA_3(\alpha, \beta), 0, b^i]$. By continuity and the Theorem of the Maximum, for any $b^i \in (-\bar{b}, \bar{b})$, there exists $\tilde{\epsilon}(b^i) \in (0, \bar{\epsilon}(\kappa)]$ (where $\bar{\epsilon}(\kappa)$ is as defined in Lemma 6.4), such that $W(\rho, \epsilon, b^i)) < V[FA_3(\alpha, \beta), \epsilon, b^i]$ for all $\epsilon \leq \tilde{\epsilon}(b^i)$. Further, $\inf_{b^i \in (-\bar{b}, \bar{b})} \tilde{\epsilon}(b^i) > 0$, so there also exists a common $\tilde{\epsilon}$ such that $W(\epsilon, b^i)) < V[FA_3(\alpha, \beta), \epsilon, b^i]$ for all $\epsilon \leq \tilde{\epsilon}$. Therefore, for sufficiently small ϵ, the optimal automaton in $\mathcal{FA}(\rho)$ (if there is one) is also strictly better than any automaton in $\mathcal{FA}^c(\rho)$, and thus if it exists, it is the optimal 3 state automaton.

Step 2. To show that there exists a no-return three-state automation that is optimal among those in $\mathcal{FA}(\rho)$, first we first compute the continuation values, $V_q(\theta, b^i)$, and the corresponding beliefs, $f(q|\theta)$, in an arbitrary no-return, three-state automaton $FA_3(\alpha, \beta)$. Here we only list the results and the detailed derivation can be found in Online Appendix. We use u^R to denote $1 + b$ and u^L to denote $1 - b$, solving for an

arbitrary $b \in (-\overline{(b)}, \overline{(b)})$.

$$V_{q^R}(R,b) = \frac{[(1-\delta) + \delta(1-\beta)(\mu-\epsilon)]u^R - \delta\epsilon(1-\alpha)c}{(1-\delta)\{1 - \delta[1 - \mu + \beta(\mu-\epsilon) + \epsilon\alpha]\}},$$

$$V_{q^L}(R,b) = \frac{\delta(\mu-\epsilon)(1-\beta)u^R - [(1-\delta) + \delta(1-\alpha)\epsilon]c}{(1-\delta)\{1 - \delta[1 - \mu + \beta(\mu-\epsilon) + \epsilon\alpha]\}},$$

$$V_{q^R}(M,b) = 0 = V_{q^L}(M,b),$$

$$V_{q^L}(L,b) = \frac{[(1-\delta) + \delta(1-\alpha)(\mu-\epsilon)]u^L - \delta\epsilon(1-\beta)c}{(1-\delta)\{1 - \delta[1 - \mu + \alpha(\mu-\epsilon) + \epsilon\beta]\}},$$

$$V_{q^R}(L,b) = \frac{\delta(\mu-\epsilon)(1-\alpha)u^L - [(1-\delta) + \delta(1-\beta)\epsilon]c}{(1-\delta)\{1 - \delta[1 - \mu + \alpha(\mu-\epsilon) + \epsilon\beta]\}},$$

$$V_{q^M}(R,b) = \frac{\delta}{1-\delta} \times \frac{(\mu-\epsilon)[(1-\delta) + \delta\mu(1-\beta)]u^R - \epsilon[(1-\delta) + \delta(1-\alpha)\mu]c}{[1 - \delta(1-\mu)]\{1 - \delta[1 - \mu + \beta(\mu-\epsilon) + \epsilon\alpha]\}},$$

$$V_{q^M}(L,b) = \frac{\delta}{1-\delta} \times \frac{(\mu-\epsilon)[(1-\delta) + \delta\mu(1-\alpha)]u^L - \epsilon[(1-\delta) + \delta(1-\beta)\mu]c}{[1 - \delta(1-\mu)]\{1 - \delta[1 - \mu + \alpha(\mu-\epsilon) + \epsilon\beta]\}},$$

$$V_{q^M}(M,b) = \frac{1}{1 - \delta(1-2\epsilon)},$$

and

$$f(q^M|R) = \frac{(1-\delta)\mathbf{P}_0(R)}{1 - \delta(1-\mu)}, \quad f(q^M|L) = \frac{(1-\delta)\mathbf{P}_0(L)}{1 - \delta(1-\mu)},$$

$$f(q^M|M) = \frac{(1-\delta)\mathbf{P}_0(M)}{1 - \delta(1-2\epsilon)},$$

$$f(q^R|R) = \frac{\delta(\mu-\epsilon)[(1-\delta) + \delta(1-\beta)\mu]\mathbf{P}_0(R)}{\{1 - \delta[1 - \epsilon(1-\alpha) - (\mu-\epsilon)(1-\beta)]\}[1 - \delta(1-\mu)]},$$

$$f(q^L|R) = \frac{\delta\epsilon[(1-\delta) + \delta(1-\alpha)\mu]\mathbf{P}_0(R)}{\{1 - \delta[1 - \epsilon(1-\alpha) - (\mu-\epsilon)(1-\beta)]\}[1 - \delta(1-\mu)]},$$

$$f(q^L|L) = \frac{\delta(\mu-\epsilon)[(1-\delta) + \delta(1-\alpha)\mu]\mathbf{P}_0(L)}{\{1 - \delta[1 - \epsilon(1-\beta) - (\mu-\epsilon)(1-\alpha)]\}[1 - \delta(1-\mu)]},$$

$$f(q^R|L) = \frac{\delta\epsilon[(1-\delta) + \delta(1-\beta)\mu]\mathbf{P}_0(L)}{\{1 - \delta[1 - \epsilon(1-\beta) - (\mu-\epsilon)(1-\alpha)]\}[1 - \delta(1-\mu)]},$$

$$f(q^R|M) = \frac{\delta\epsilon[(1-\delta) + \delta(1-\beta)2\epsilon]\mathbf{P}_0(M)}{\{1 - \delta[1 - 2\epsilon + \epsilon(\alpha+\beta)]\}[1 - \delta(1-2\epsilon)]},$$

$$f(q^L|M) = \frac{\delta\epsilon[(1-\delta) + \delta(1-\alpha)2\epsilon]\mathbf{P}_0(M)}{\{1 - \delta[1 - 2\epsilon + \epsilon(\alpha+\beta)]\}[1 - \delta(1-2\epsilon)]}.$$

Again, the detailed derivations of these expressions can be found in Online Appendix.

Observe that if ϵ is sufficiently small, then $p(q^M, r)(R)$, $p(q^R, r)(R)$ and $p(q^R, m)(R)$

are arbitrarily close to one, and similarly $p(q^M, l)(L)$, $p(q^L, l)(L)$ and $p(q^L, m)(L)$ are arbitrarily close to one. Note as well that $V_{q^R}(R) > V_{q^M}(R) > V_{q^L}(R)$ and $V_{q^L}(L) > V_{q^M}(L) > V_{q^R}(L)$. These two observations, combined together and with Assumption (1), imply that if ϵ is sufficiently small, then under any automaton in FA_3, for any (q, s) other than (q_R, ℓ) and (q_L, r), if $\tau_{(\alpha,\beta)}(q, s) = q'$, then $p(q, s) \in \text{INT}(\Pi_{q'})$.

Since the continuation values and beliefs given by the pair of expressions (13) are continuous in both ϵ and in the transition probabilities, there exist $\rho_0 > 0$ and $\epsilon_0 > 0$ such that for any (q, s) other than (q_R, ℓ) and (q_L, r), if $\tau_{(\alpha,\beta)}(q, s) = q'$, then $p(q, s) \in \text{INT}(\Pi_{q'})$ as well for all $\epsilon \leq \epsilon_0$ and for all SFSA in $\mathcal{FA}(\rho_0)$. Lemma 6.3 then implies that for all $\epsilon \leq \epsilon_0$, among SFSA in $\mathcal{FA}(\rho_0)$,

Lemma 6.3 then implies that for all $\epsilon \leq \epsilon_0$, among SFSA in $\mathcal{FA}(\rho_0)$, any optimal automaton must be $FA_3(\alpha, \beta)$ with optimal α and β.

Further, since the class of automata $FA_3(\alpha, \beta)$ is compact, and utilities are continuous in transition probabilities, a solution to the voter's optimization problem (5) exists, and thus $FA_3(\alpha, \beta)$ with optimal α and β is the optimal SFSA among those in $\mathcal{FA}(\rho_0)$.

As a result, by Step 1, $FA_3(\alpha, \beta)$ with optimal α and β is the optimal SFSA among all 3-state SFSAs, and by Lemma 6.4, if κ is low enough, and ϵ is low enough given κ, then it is also the optimal SFSA among all SFSAs.

We have established that each voter i optimally chooses an automaton in class FA_3. Under such an automaton, for any period t such that $s_\tau \in \{\ell, r\}$ for some $\tau \in \{1, 2, ..., t-1\}$, $q_t^i \in \{q^L, q^R\}$ and $a_t^i \in \{a^L, a^R\}$, thus every voter becomes extreme. $\qquad \square$

Proof of Proposition 3.2.

Proof. Given $\epsilon > 0$, to determine the optimal $\alpha(b^i)$ and $\beta(b^i)$, recall that $FA_3(\alpha, \beta)$ has $q_0 = q^M$, i.e., it starts with memory state q^M, and hence the ex ante payoff from it is given by

$$V(\alpha, \beta) \equiv \mathbf{P}_0(L)V_{q^M}(L, b^i) + \mathbf{P}_0(M)V_{q^M}(M, b^i) + \mathbf{P}_0(R)V_{q^M}(R, b^i).$$

That is, for an agent i with type $b^i = b$,

$$
\begin{aligned}
V(\alpha, \beta; b) \;=\;\; & \mathbf{P}_0(L)\frac{\delta}{1-\delta} \times \frac{(\mu-\epsilon)[(1-\delta)+\delta\mu(1-\beta)]u^R - \epsilon[(1-\delta)+\delta\mu(1-\alpha)]c}{1-\delta[1-\mu+\beta(\mu-\epsilon)+\epsilon\alpha]} \\
+\;\; & \mathbf{P}_0(R)\frac{\delta}{1-\delta} \times \frac{(\mu-\epsilon)[(1-\delta)+\delta\mu(1-\alpha)]u^L - \epsilon[(1-\delta)+\delta\mu(1-\beta)]c}{1-\delta[1-\mu+\alpha(\mu-\epsilon)+\epsilon\beta]} \\
+\;\; & \mathbf{P}_0(M)\frac{1}{1-\delta(1-2\epsilon)}.
\end{aligned}
$$

Recall that we assumed that $\mathbf{P}_0(R) = \mathbf{P}_0(L)$, and note that the last term above does not depend on α or β. So, to maximize $V(\alpha, \beta; b)$, it is equivalent to maximize $G(\alpha, \beta; b) \equiv G^1(\alpha, \beta; b) + G^2(\alpha, \beta; b)$, where

$$
\begin{aligned}
G^1(\alpha, \beta; b) \;\equiv\;\; & \frac{(\mu-\epsilon)[(1-\delta)+\delta\mu(1-\beta)]u^R}{1-\delta[1-\mu+\beta(\mu-\epsilon)+\epsilon\alpha]} + \frac{(\mu-\epsilon)[(1-\delta)+\delta\mu(1-\alpha)]u^L}{1-\delta[1-\mu+\alpha(\mu-\epsilon)+\epsilon\beta]}, \\
G^2(\alpha, \beta; b) \;\equiv\;\; & \frac{-\epsilon[(1-\delta)+\delta\mu(1-\alpha)]c}{1-\delta[1-\mu+\beta(\mu-\epsilon)+\epsilon\alpha]} + \frac{-\epsilon[(1-\delta)+\delta\mu(1-\beta)]c}{1-\delta[1-\mu+\alpha(\mu-\epsilon)+\epsilon\beta]}.
\end{aligned}
$$

Then,

$$
\begin{aligned}
& \frac{\partial}{\partial \alpha} G^1(\alpha, \beta; b) \\
=\;\; & \delta(\mu-\epsilon)\epsilon[(1-\delta)+\delta\mu(1-\beta)] \\
& \times \left\{ \frac{u^R}{\{1-\delta[1-\mu+\beta(\mu-\epsilon)+\epsilon\alpha]\}^2} - \frac{u^L}{\{1-\delta[1-\mu+\alpha(\mu-\epsilon)+\epsilon\beta]\}^2} \right\}, \\
& \frac{\partial}{\partial \beta} G^1(\alpha, \beta; b) \\
=\;\; & \delta(\mu-\epsilon)\epsilon[(1-\delta)+\delta\mu(1-\alpha)] \\
& \times \left\{ \frac{-u^R}{\{1-\delta[1-\mu+\beta(\mu-\epsilon)+\epsilon\alpha]\}^2} + \frac{u^L}{\{1-\delta[1-\mu+\alpha(\mu-\epsilon)+\epsilon\beta]\}^2} \right\}.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\frac{\partial}{\partial \alpha} G^2(\alpha, \beta; b) \\
&= \delta(\mu - \epsilon)\epsilon[(1 - \delta) + \delta\mu(1 - \beta)] \\
&\quad \times \left\{ \frac{c}{\{1 - \delta[1 - \mu + \beta(\mu - \epsilon) + \epsilon\alpha]\}^2} + \frac{-c}{\{1 - \delta[1 - \mu + \alpha(\mu - \epsilon) + \epsilon\beta]\}^2} \right\}, \\
&\frac{\partial}{\partial \beta} G^2(\alpha, \beta; b) \\
&= \delta(\mu - \epsilon)\epsilon[(1 - \delta) + \delta\mu(1 - \alpha)] \\
&\quad \times \left\{ \frac{-c}{\{1 - \delta[1 - \mu + \beta(\mu - \epsilon) + \epsilon\alpha]\}^2} + \frac{c}{\{1 - \delta[1 - \mu + \alpha(\mu - \epsilon) + \epsilon\beta]\}^2} \right\}.
\end{aligned}
$$

In an interior solution the following First Order Condition must hold:

$$
\frac{1 + b + c}{1 - b + c} = \frac{\{1 - \delta[1 - \mu + \beta(\mu - \epsilon) + \epsilon\alpha]\}^2}{\{1 - \delta[1 - \mu + \alpha(\mu - \epsilon) + \epsilon\beta]\}^2} \tag{27}
$$

This then implies that it is optimal to have $\alpha = 1 = 1 - \beta$ if

$$
\frac{1 + b + c}{1 - b + c} = \frac{u^R + c}{u^L + c} \geq \left[\frac{1 - \delta(1 - \mu + \epsilon)}{1 - \delta(1 - \epsilon)} \right]^2, \tag{28}
$$

or, equivalently, if

$$
b \geq (1 + c)\frac{\left[\frac{1 - \delta(1 - \mu + \epsilon)}{1 - \delta(1 - \epsilon)} \right]^2 - 1}{1 + \left[\frac{1 - \delta(1 - \mu + \epsilon)}{1 - \delta(1 - \epsilon)} \right]^2} \equiv b^*; \tag{29}
$$

and, by symmetry, if $b < -b^*$, then $\alpha(b) = 0$ and $\beta(b) = 1$ are optimal. Whereas, if $b \in (-b^*, b^*)$, the optimal solution is interior.

Let $\bar{b} > b^*$. Then according to threshold (29), for all $b \in (b^*, \bar{b}]$, the optimal SFSA has the form $FA_3(0, 1)$ while for all $b \in [-\bar{b}, -b^*)$, the optimal SFSA has the form $FA_3(1, 0)$.

We next find the optimal $\alpha(b)$ and $\beta(b)$ for any $b \in (-b^*, b^*)$ such that the optimal automaton for voter i with $b^i = b$ is $FA_3(\alpha(b), \beta(b))$.

To solve for the optimal randomization in the interior solution explicitly, let

$$
\nu(b) \equiv \sqrt{\frac{1 + b + c}{1 - b + c}}
$$

45

and

$$\bar{\nu} \equiv \frac{1 - \delta(1 - \mu + \epsilon)}{1 - \delta(1 - \epsilon)}, \text{ so } \nu(b^*) = \bar{\nu}.$$

By symmetry, we may consider only $b \geq 0$ and thus equivalently, $\nu(b) \geq 1$. Further, since $b \leq \bar{b}$, we need only consider

$$\nu(b) \in \left[1, \sqrt{(1 + \bar{b} + c)/(1 - \bar{b} + c)}\right],$$

and from inequality (29), for an interior solution we only need to consider $\nu(b) < \bar{\nu}$.

From the FOC of the unconstrained maximization of $G(\alpha, \beta; b)$ with respect to α, and rearranging terms, we obtain:

$$\frac{(1 - \delta) + \delta[\mu - \beta(\mu - \epsilon) - \alpha\epsilon]}{(1 - \delta) + \delta[\mu - \alpha(\mu - \epsilon) - \beta\epsilon]} = \nu(b),$$

that is,

$$\beta = \frac{\delta[\nu(b)(\mu - \epsilon) - \epsilon]\alpha - (\nu(b) - 1)[(1 - \delta) + \delta\mu]}{\delta[\mu - \epsilon - \nu(b)\epsilon]}. \tag{30}$$

Now, first note that since $\nu(b) \geq 1$ for any $b \geq 0$, and that since (by assumption) $\epsilon < \mu/2$, it follows that for any $b \geq 0$,

$$\nu(b)(\mu - \epsilon) - \epsilon > 0, \text{ that is, } \nu(b) \geq 1 > \frac{\epsilon}{\mu - \epsilon}.$$

Similarly, notice that $\bar{\nu} < \frac{\mu - \epsilon}{\epsilon}$, from which it follows that

$$\mu - \epsilon - \nu(b) > 0 \text{ for any } \nu(b) < \bar{\nu}.$$

From Equation (30), we can verify that $\beta = \alpha$ for $\nu(b) = 1$ and that $\beta < \alpha$ for any $\nu(b) > 1$ (or, equivalently, for any $b > 0$). To ensure that $\beta \geq 0$ we also need that the numerator of the right hand side of Equation (30) to be non-negative. That is, we need

$$\alpha \geq \frac{(\nu(b) - 1)[(1 - \delta) + \delta\mu]}{\delta[\nu(b)(\mu - \epsilon) - \epsilon]} \equiv \bar{\alpha}(\nu(b)). \tag{31}$$

46

Notice that $\bar{\alpha}(\nu(b))$ strictly increases with b, with

$$\bar{\alpha}(\nu(0)) = 0 \text{ and } \bar{\alpha}(\nu(b^*)) = 1. \tag{32}$$

Thus, the optimal $\alpha(b)$ and $\beta(b)$ for any $b \in (0, b^*)$ are determined by

$$\alpha(b) \in [\bar{\alpha}(\nu(b)), 1], \ \ \beta(b) = \frac{\delta[\nu(b)(\mu - \epsilon) - \epsilon]\alpha - (\nu(b) - 1)[(1 - \delta) + \delta\mu]}{\delta[\mu - \epsilon - \nu(b)\epsilon]}. \tag{33}$$

To verify an interior solution, we exclude possible corner solutions. First we compute the values at the corners:

$$
\begin{aligned}
G(1, 0; b) &= \frac{(\mu - \epsilon)(1 - \delta + \delta\mu)u^R - \epsilon(1 - \delta)c}{1 - \delta + \delta(\mu - \epsilon)} + \frac{(\mu - \epsilon)(1 - \delta)u^L - \epsilon[(1 - \delta) + \delta\mu]c}{1 - \delta + \delta\epsilon}, \\
G(1, 1; b) &= u^R + u^L - 2\epsilon c = G(0, 0; b), \\
G(0, 1; b) &= \frac{(\mu - \epsilon)(1 - \delta + \delta\mu)u^L - \epsilon(1 - \delta)c}{1 - \delta + \delta(\mu - \epsilon)} + \frac{(\mu - \epsilon)(1 - \delta)u^R - \epsilon[(1 - \delta) + \delta\mu]c}{1 - \delta + \delta\epsilon}.
\end{aligned}
$$

First, $G(1, 0; b) \geq G(0, 1; b)$ if and only if $b \geq 0$. Now, $G(1, 0; b) > G(1, 1; b) = G(0, 0; b)$ if and only if

$$\frac{u^R + c}{1 - \delta + \delta(\mu - \epsilon)} > \frac{u^L + c}{1 - \delta + \delta\epsilon},$$

This is equivalent to $\nu(b) > \sqrt{\nu(b)}$. For lower $\nu(b)$'s, we have to verify that $G(1, \beta; b) > G(1, 1; b)$, where β is given by (33) with $\alpha = 1$. Now, $G(\beta, 1; b) > G(1, 1; b)$ is equivalent to

$$1 - \beta < \frac{(1 - \delta)((\nu(b))^2 - 1)}{\delta(\mu - \epsilon - (\nu(b))^2\epsilon)},$$

which, by plugging in by β given by the solution in Expression (33) with $\alpha = 1$, is equivalent to

$$\frac{(1 - \delta)(\nu(b) - 1)}{\delta(\mu - \epsilon - \nu(b)\epsilon)} < \frac{(1 - \delta)((\nu(b))^2 - 1)}{\delta(\mu - \epsilon - (\nu(b))^2\epsilon)},$$

which holds for all $b > 0$.

Symmetric results apply to $b < 0$.

Finally, we consider the long-run distributions of the memory states and the corresponding actions under the optimal SFSA. The transition matrix of the memory states are given by Table 4 under $FA_3(\alpha, \beta)$. Note that q^M is transitory under all possible

47

states of the world and hence we ignore it. Let $\rho_\theta(q)$ denote the stationary distribution of the memory state q under the state of the world θ. Then, from Table 4 we can compute that

$$\rho_M(q^R) = \frac{1-\beta}{(1-\beta)+(1-\alpha)}, \tag{34}$$

$$\rho_R(q^R) = \frac{(\mu-\epsilon)(1-\beta)}{\epsilon(1-\alpha)+(\mu-\epsilon)(1-\beta)}, \tag{35}$$

$$\rho_L(q^R) = \frac{\epsilon(1-\beta)}{(\mu-\epsilon)(1-\alpha)+\epsilon(1-\beta)}. \tag{36}$$

Note that $\rho_M(q_R)$, $\rho_R(q_R)$, and $\rho_L(q_R)$ are all increasing in $(1-\beta)/(1-\alpha)$. Note also that for any given $\alpha \in [\bar{\alpha}(\nu(b)), 1]$, optimal β is given by Equality (30). Since $\mu > 2\epsilon$ (by assumption), it follows that $1/\rho_L(q^R) \geq 1/\rho_M(q^R) \geq 1/\rho_R(q^R)$ so $\rho_L(q_R) \leq \rho_M(q_R) \leq \rho_R(q_R)$, strictly so whenever $\alpha < 1$. Hence, from Equality (30),

$$\begin{aligned}
\frac{1-\beta}{1-\alpha} &= \frac{\delta(\mu-\epsilon-\nu(b)\epsilon)-\delta[\nu(b)(\mu-\epsilon)-\epsilon]\alpha+(\nu(b)-1)[(1-\delta)+\delta\mu]}{\delta(\mu-\epsilon-\nu(b)\epsilon)(1-\alpha)} \\
&= \frac{\delta(\mu-\epsilon)(1-\nu(b)\alpha)-\delta\epsilon(\nu(b)-\alpha)+(\nu(b)-1)(1-\delta+\delta\mu)}{\delta(\mu-\epsilon-\nu(b)\epsilon)(1-\alpha)} \\
&= \frac{\delta[\nu(b)(\mu-\epsilon)-\epsilon]}{\delta(\mu-\epsilon-\nu(b)\epsilon)}+\frac{(\nu(b)-1)(1-\delta)}{\delta(\mu-\epsilon-\nu(b)\epsilon)(1-\alpha)},
\end{aligned}$$

which increases with α.

Thus, the range of $\rho_\theta(q^R)$ as a function of b is given by $[\bar{\rho}_\theta(q^R; b), 1]$, where $\bar{\rho}_\theta(q^R; b)$ is obtained from (34)-(36) with $\beta = 0$ and $\alpha = \bar{\alpha}(\nu(b))$ (note that β given by Expression (30) is equal to zero if $\alpha = \bar{\alpha}(\nu(b))$). Recall (from Expression 32) that $\bar{\alpha}(\nu(b^*)) = 1$; thus, $\bar{\rho}_\theta(q^R; b^*) = 1$ and $\rho_\theta(q^R; b^*) = 1$. Recall as well that if $b > b^*$, the optimal automaton is $FA_3(0, 1)$. Therefore, if $b \in (b^*, \bar{b})$, then $\rho_\theta(q^R; b) = 1$; that is, voters with a sufficiently high right lean perpetually prefer action a^R after finitely many periods with probability one.

Likewise, symmetric results apply for voters with a left lean: for any $b \in (-\bar{b}, -b^*)$, $\rho_\theta(q^R; b) = 0$; that is, voters with a sufficiently strong left lean perpetually prefer action a^L after finitely many periods with probability one. In other words, some voters polarize at opposite extremes.

48

M	q^L	q^R
q^L	$1-\epsilon(1-\beta)$	$\epsilon(1-\beta)$
q^R	$\epsilon(1-\alpha)$	$1-\epsilon(1-\alpha)$

L	q^L	q^R
q^L	$1-\epsilon(1-\beta)$	$\epsilon(1-\beta)$
q^R	$(\mu-\epsilon)(1-\alpha)$	$1-(\mu-\epsilon)(1-\alpha)$

R	q^L	q^R
q^L	$1-(\mu-\epsilon)(1-\beta)$	$(\mu-\epsilon)(1-\beta)$
q^R	$\epsilon(1-\alpha)$	$1-\epsilon(1-\alpha)$

Table 4: Transition matrix of memory states under state of the world M (top), L (bottom left), and R (bottom right)

Hence,

$$\bar{\rho}_M(q_R; b) = \frac{1}{2-\bar{\alpha}(\nu(b))},$$

$$\bar{\rho}_R(q_R; b) = \frac{(\mu-\epsilon)}{\epsilon[1-\bar{\alpha}(\nu(b))]+(\mu-\epsilon)},$$

$$\bar{\rho}_L(q_R; b) = \frac{\epsilon}{(\mu-\epsilon)[1-\bar{\alpha}(\nu(b))]+\epsilon}.$$

Since $\nu(0)=1$ and $\bar{\alpha}(\nu(0))=0$, $\bar{\rho}_M(q^R; 0)=1/2$ and increases to one as b increases to b^*. Similarly, $\bar{\rho}_R(q^R; 0)=(\mu-\epsilon)/\mu$ and $\bar{\rho}_L(q^R; 0)=\epsilon/\mu$ and both increase to one as b increases to b^*. This then implies that voters with $b \in (-b^*, b^*)$ swing between a^R and a^L, with greater likelihood to support a^R in the long run if $b > 0$, and greater likelihood to support a^L if $b < 0$.

$\square$

We next prove existence of an equilibrium of the electoral competition game, as defined in Definition 4.1.

Lemma 6.5. *An equilibrium exists.*

Proof. From the proof of Proposition 3.2, we obtain that for each $b \in (-\bar{b}, \bar{b})$, a solution to optimization problem (5) for agent i with type $b^i = b$ exists, and the optimal automaton is $FA_3(\alpha(b), \beta(b))$ with $\alpha(b)$ and $\beta(b)$ given by solution (33) if $b \in [0, b^*]$; with $\alpha(b)=1$ and $\beta(b)=0$ if $b \in (b^*, \bar{b}]$; and with optimal values mirroring these for negative types.

Since the solution correspondence $(\alpha^*(b), \beta^*(b)) : (-\bar{b}, \bar{b}) \rightrightarrows [0,1]^2$ is continuous, we can take a continuous function $(\alpha^{**}(b), \beta^{**}(b)) : (-\bar{b}, \bar{b}) \to [0,1]^2$ such that

49

$(\alpha^{**}(b), \beta^{**}(b)) \in (\alpha^{*}(b), \beta^{*}(b))$ and thus $\phi : (-\bar{b}, \bar{b}) \to \mathcal{Q}$ defined by $\phi(b) \equiv FA_3(\alpha(b), \beta(b))$ is a continuous function from the set of types to the set of optimal automata.

Given function ϕ, for any $b \in (-\bar{b}, \bar{b})$ and any agent i with type $b^i = b$, given the automaton $\langle Q^i, q_0^i, \tau^i, d^i, \rangle \equiv \phi(b)$, a support function for agent i such that $a_t^i = d^i(q_t)$ for each t and for any realization of all observables by i satisfies the "sincere support" equilibrium condition.

A voting function for agent i such that the voter votes for the leader on the alternative that the voter supports, and abstains if neither party does so, in every period and for any realization of observables, satisfies the "sincere voting" equilibrium condition.

With respect to the condition on parties' optimization, parties are fully impatient, so in each period, and for any history of previous play they play as if playing a one-period game. Consider the two-player one-period game played between the two parties while taking the optimal support and voting functions of all voters as given and common knowledge among both parties. This one-period, two-player game is a finite game, and thus it has a Nash equilibrium. Construct a strategy profile for parties such that in any period, and for any previous history of the game, parties play a Nash equilibrium of the period game induced by the history up to this period, and by Bayesian updating of beliefs. This strategy profile is sequentially rational and satisfies the party optimization equilibrium condition.

Thus, an equilibrium exists.

$\square$

We next prove that under a range of parameters, party polarization is irreversible: if parties choose platforms at opposite extremes once, they do so forever thereafter.

Lemma 6.6. *For any preference parameters $\bar{b}$ sufficiently close to one and c sufficiently small, there exist a range of memory costs (κ) and probabilities of an incorrect extreme signal (ϵ), such that under all states of the world, in all equilibria, if parties polarize at a period t after an extreme signal has been observed at least once in previous periods, then parties remain polarized in every subsequent period: $a_{t'}^j = a_t^j$ for each $j \in \{1, 2\}$.*

Proof. As established in Proposition 3.1, if κ is small enough, there exists a range of strictly positive values for parameter ϵ such that in any equilibrium, each voter's choice of an alternative to support in each period is as recommended by an optimal 3-state no

return automaton $FA_3(\alpha(b), \beta(b))$. Each voter i with type $b^i \in (-\bar{b}, \bar{b})$, for any period t, and for any realization of all the observables observed by voter i, the alternative that voter i chooses to support is $a_t^i = d^i(q_t)$, where d^i and q_t are respectively the decision rule and the memory state at period t of automaton $FA_3(\alpha(b^i), \beta(b^i))$, and $\alpha(b^i)$ and $\beta(b^i)$ are an optimal randomization for type b^i as computed in Expression (33) in the proof of Proposition 3.2.

With such an optimal no-return 3-state automaton dictating all voters' choices about which alternative to support, all voters initially support alternative a^M. Following the first extreme signal, all voters support the alternative congruent with the signal, and thereafter, voters oscilate between supporting a^L and supporting a^R.

Let $T \in \mathbb{N}$ denote the first period in which $s_t \in \{\ell, r\}$. Given the voters' support decisions, in equilibrium each voter in each period votes for the party that is a leader on the alternative the voter supports; if neither party chooses the alternative supported by the voter, the voter abstains. It follows that in period T and thereafter, no voter would ever vote for a party that chooses platform a^M.

For any period $t \in \mathbb{N}$, let $\zeta_t(b)$ denote the fraction of voters of type b who support a^L. Given the realization of signals $(s_\tau)_{\tau=1}^t$, $\zeta_t(b)$ is continuous in $\alpha(b)$ and $\beta(b)$. Since in equilibrium (by definition), $\alpha(b)$ and $\beta(b)$ are continuous in b, it follows that $\zeta_t(b)$ is also continuous in b. Therefore, $\zeta_t(b) f_b(b)$ is integrable over the range of b, Thus,

$$\zeta_t \equiv \int_{-\bar{b}}^{\bar{b}} f_b(b) \zeta_t(b) \, db$$

is well defined as the fraction of the population of voters who supports a^L. For any $t \geq T$, the fraction of the population of voters who supports a^R is then $1 - \zeta_t$.

If follows that a party that is a leader on a^L in period $t \geq T$ obtains a mass of votes ζ_t, a party that is a leader on a^R obtains $1 - \zeta_t$, and followers or leaders on a^M obtain zero votes.

Assume parties polarize in period $t \geq T$. Without loss of generality (up to relabeling of party labels), assume $a_t^1 = a^L$ and $a_t^2 = a^R$. Each party is a leader on the policy it proposes, and would remain so in the next period, if it sticks to the same policy.

Assume $\bar{b}$ is sufficiently close to one, c sufficiently small, and κ and ϵ are such that the conditions in Proposition 3.1 and Proposition 3.2 hold, and thus the results in

Proposition 3.2 apply, so a positive mass of voters polarize at opposite extremes after the first extreme signal at each extreme has been observed. The probability of winning in period $t+1$ for each of the two parties, as a function of their chosen platforms, is then

$a^1_{t+1} \backslash a^2_{t+1}$	a^L	a^M	a^R
a^L	$(1,0)$	$(1,0)$	$(\pi, 1-\pi)$
a^M	$(0,1)$	$(\frac{1}{2},\frac{1}{2})$	$(0,1)$
a^R	$(1-\pi,\pi)$	$(1,0)$	$(0,1)$

where $\pi \in [0,1]$ is a parameter that depends on ζ_t, on the randomization parameters in the optimal automata chosen by agents, on parties' posterior belief about the state of nature, and on model parameters. The probabilities of winning given platform pairs (a^L, a^R) and (a^R, a^L) are symmetric because party labels are irrelevant to voters; to all voters these two pair of platforms are identical up to relabeling, and under either platform pair, the party that proposes a^L (whichever one it might be) wins with probability π.

If $a^2_{t+1} = a^R$, in period $t+1$ the expected vote share for P^1 is zero if $a^1_{t+1} \in \{a^M, a^R\}$ (zero if $a^1_{t+1} = a^M$ because no voter supports this policy; and zero if $a^1_{t+1} = a^R$ because P^1 would be a follower with no votes). Whereas, if $a^1_{t+1} = a^L$, the expected vote share is strictly positive. Namely, the mass of voters ζ_{t+1} votes for P^1 if $a^1_{t+1} = a^L$, and, by Proposition 3.2 and its proof, if there exists $\tau \in \{1, ..., t+1\}$ such that $s_\tau = \ell$, then $\zeta_{t+1} > 0$. Since the probability that $s_{t+1} = \ell$ is greater than 0, it follows that the expected value of ζ_{t+1} is also strictly positive. Thus, the unique best response for party P^1 to $a^2_{t+1} = a^R$ is $a^1_{t+1} = a^L$ (even if $\pi = 0$).

Analogously, If $a^1_{t+1} = a^L$, in period $t+1$ the expected vote share for P^2 is strictly positive if $a^2_{t+1} = a^R$ (because the mass of voters who support a^R would vote for P^2), and zero otherwise (because no voter supports a^M, and party P^2 would be a follower with no votes on a^L).

Thus, regardless of the exact value of π, the only equilibrium of the two player period $t+1$ game is $(a^1_{t+1}, a^2_{t+1}) = (a^L, a^R)$.

Therefore, by induction, this is as well the only equilibrium in any subsequent period. $\qquad \square$

With Lemma 6.6 in mind, to prove Proposition 4.1, it suffices to show that parties

polarize once after one extreme signal has been observed.

We are now ready for this proof.

Proof of Proposition 4.1.

Proof. Assume $\bar{b}$ is sufficiently close to one, c is sufficiently small, and κ and ϵ are such that the conditions in Proposition 3.1 and Proposition 3.2 apply, and thus a positive mass of voters polarize at opposite extremes with probability converging to one in t.

As established in the first few paragraphs of the proof of Lemma 6.6, letting T denote the first period in which an extreme signal is observed, for any period $t \geq T$, and letting ζ_t denote the share of the population of voters who supports a^L in period t, a party that is a leader on a^L in period t obtains a mass of votes ζ_t, a party that is a leader on a^R obtains $1 - \zeta_t$, and followers on any policy or leaders on a^M obtain zero votes.

Given this voter behavior, equilibrium party platforms in each period must satisfy the following along the equilibrium path.

-In period $t = 1$. Signal s_1 is equal to m with probability

$$\sum_{\theta \in \Theta} \mathbf{P}_0(\theta)\mu_s^\theta \equiv \pi,$$

and for $\theta \in \{\ell, r\}$, it is equal to θ with probability $\frac{1-\pi}{2}$. The 2-player one-period election game between the two parties yields the following probabilities of winning as a function of the platform profile:

$a_1^1 \backslash a_1^2$	a^L	a^M	a^R
a^L	$(\frac{1}{2}, \frac{1}{2})$	$(3\frac{(1-\pi)}{4}, \pi + \frac{1-\pi}{4})$	$(\frac{1}{2}, \frac{1}{2})$
a^M	$(\pi + \frac{1-\pi}{4}, 3\frac{(1-\pi)}{4})$	$(\frac{1}{2}, \frac{1}{2})$	$(\pi + \frac{1-\pi}{4}, 3\frac{(1-\pi)}{4})$
a^R	$(\frac{1}{2}, \frac{1}{2})$	$(3\frac{(1-\pi)}{4}, \pi + \frac{1-\pi}{4})$	$(\frac{1}{2}, \frac{1}{2})$

From our assumptions on the prior $\mathbf{P}_0$ and on parameters μ and ϵ, it follows that $\pi > 1/2$. From $\pi > 1/2$ follows that choosing platform a^M is a strictly dominant strategy for each party in the one-period two-player game. Therefore, in any equilibrium (of the full game), $a_1^1 = a_1^2 = a^M$; Nature randomly identifies one of the two parties as the leader on policy a^M, and the leader wins the election. Without loss of generality (up

to relabeling of parties), let Party 1 be the one chosen as leader by Nature.

-In any period $t \in (2, ..., T)$. Assume Party P^1 was the leader on policy a^M in period $t - 1$. Then it is strictly dominant in the one-period two-player game for Party P^1 to choose $a_t^1 = a^M$. Given that $a_t^1 = a^M$, the set of best responses in the period game for Party P^2 is $\{a^L, a^R\}$; choosing either best response, Party P^2 wins if s_t is extreme in the same direction as the a_t^2, which occurs with probability $\epsilon > 0$, while choosing $a_t^2 = a^M$, Party P^2 would be a follower and would lose. Thus, by induction, in any period $t \in (2, ..., T)$, in any equilibrium, $a_t^1 = a^M$ and $a_t^2 \in \{a^L, a^R\}$.

Case 1: Assume that in period T, $s_T = a_T^2$. Then it is a dominant strategy for party P^2 in the $T + 1$ period game to choose $a_{T+1}^2 = a_T^2$, and given that $a_{T+1}^2 = a_T^2$, the unique best response for party P^1 in the $T + 1$ period game is to commit to the other extreme alternative (i.e. $a_{T+1}^1 = a^L$ and $a_{T+1}^2 = a^R$ if $a_T^2 = a^R$, and $a_{T+1}^1 = a^R$ and $a_{T+1}^2 = a^L$ if $a_T^2 = a^L$).

Thereafter, by Lemma 6.6, parties remain polarized: $a_t^1 = a^L$ and $a_t^2 = a^R$ if $a_T^2 = a^R$, and $a_t^1 = a^R$ and $a_t^2 = a^L$ if $a_T^2 = a^L$, for any period $t \geq T + 1$.

Ex-ante, T is a random variable. For any $t \in \mathbb{N}$, the probability that $t > T$ converges to one in t. Thus, the probability that parties polarize at opposite extremes converges to one over time.

Case 2: In period T, $s_T \neq a_T^2$. Then for any $t > T$, it is strictly dominated in the one-period two player game for either party to choose platform a^M, so they do not choose a^M. If each of the two parties choose a different extreme platform in a given period, then Lemma 6.6 applies, and parties remain polarized thereafter. We are thus left with the case in which either $a_t^1 = a_t^2 = \ell$ or $a_t^1 = a_t^2 = r$.

Consider first the subcase $a_t^1 = a_t^2 = \ell$. If $\zeta_t > 1/2$, then the only equilibrium of the $t + 1$ period game is one in which the leader on a^L in period t again chooses platform a^L in period $t + 1$, while the follower switches to a^R, and thus Lemma 6.6 applies. If $\zeta_t = 1/2$, then we have two possible equilibria of the $t + 1$ period two player game: a polarized platform pair in which the leader repeats the same policy and the follower switches extremes (so Lemma 6.6 applies), and one in which $(a_{t+1}^1, a_{t+1}^2) = (a^R, a^R)$. If only the follower switches extremes, parties have polarized, and Lemma 6.6 applies. If $\zeta_t < 1/2$, then only $(a_{t+1}^1, a_{t+1}^2) = (a^R, a^R)$ holds as an equilibrium of the period $t + 1$ game.

54

Similarly, in the subcase $a_t^1 = a_t^2 = r$, if $\zeta_t < 1/2$, then the only equilibrium of the $t+1$ period game is such that the leader on a^R in period t again chooses platform a^R in period $t+1$, while the follower switches to a^L, and thus Lemma 6.6 applies. If $\zeta_t = 1/2$, then then we have two possible equilibria of the $t+1$ period two player game: a polarized platform pair in which the leader repeats the same policy and the follower switches extremes (so Lemma 6.6 applies), and one in which $(a_{t+1}^1, a_{t+1}^2) = (a^L, a^L)$. If $\zeta_t > 1/2$, then only $(a_{t+1}^1, a_{t+1}^2) = (a^L, a^L)$ holds as an equilibrium of the period $t+1$ game.

Putting these results together, we have that the only sequence of period equilibria that escapes polarization is one in which the two parties converge on the same platform, but they alternate between converging on the left and converging on the right extreme in odd and even periods. This sequence of alternating convergent platforms, as noted above, requires an alternating majority of voters to support it: $(a_{t+1}^1, a_{t+1}^2) = (a^L, a^L)$ requires $\zeta_t \geq 1/2$, and $(a_{t+2}^1, a_{t+2}^2) = (a^R, a^R)$ requires $\zeta_{t+1} \leq 1/2$.

However, alternating majorities of support require signals congruent with this alternation. Crucially, from $\mu > 2\epsilon$ it follows $b^* > 0$, and thus $\zeta_t \geq 1/2$ together with $s_t = \ell$ imply $\zeta_t > 1/2$, and similarly $\zeta_t \leq 1/2$ together with $s_t = r$ imply $\zeta_t < 1/2$.

Thus, whether (a_{t+1}^1, a_{t+1}^2) is equal to (a^L, a^L) or to (a^R, a^R), with probability at least ϵ signal s_t would be such that in period $t+1$, a switch to a convergent platform pair at the opposite extreme cannot be supported in equilibrium, and parties polarize, which then triggers Lemma 6.6.

The probability that one such signal that destroys the sequence of convergence at alternating extremes is observed at least once converges to one in t, so the probability that parties polarize converges to one in t. And once they polarize once, Lemma 6.6 that they will always polarize. $\qquad\square$

We next report, in Table 5, an expanded version of Table 3, with longitudinal data from the five most recent Pew Center surveys that allow us to compute the ideological position of each respondent on an invariant ideological scale. These are: The 2004, 2011, 2014 and 2017 Political Typology surveys (December 1-16, 2004, n=2,000; February 22 to March 14, 2011, n=3,030; January 23 to March 16, 2014, n=10,013; and June 8-18, 2017, n=2,504); and the 2019 Political survey (September

	L	M	R
2004	28%	57%	15%
2011	27%	50%	23%
2014	32%	44%	24%
2017	42%	37%	21%
2019	42%	38%	20%

Table 5: Polarization of the US electorate, 2004-2019.

5-16, 2019, n=2,004).

References

[1] Abramowitz, Alan I., and Kyle L. Saunders. "Is polarization a myth?." *Journal of Politics* 70.2 (2008): 542-555.

[2] Robert Axelrod, Joshua J. Daymude, and Stephanie Forrest. "Preventing extreme polarization of political attitudes." *Proceedings of the National Academy of Sciences* 118.50 (December 2021).

[3] Baker, William D., and John R. Oneal. "Patriotism or opinion leadership? The nature and origins of the "rally'round the flag" effect." *Journal of Conflict Resolution* 45.5 (2001): 661-687.

[4] Baliga, Sandeep, Eran Hanany, and Peter Klibanoff. "Polarization and ambiguity." *American Economic Review* 103.7 (2013): 3071-83.

[5] Banovetz, James, and Ryan Oprea. "Complexity and procedural choice." Working paper, 2020.

[6] Benoît, Jean-Pierre, and Juan Dubra. "Apparent bias: What does attitude polarization show?" *International Economic Review* 60.4 (2019): 1675-1703.

[7] Bol, Damien, Konstantinos Matakos, Orestis Troumpounis, and Dimitrios Xefteris. "Electoral rules, strategic entry and polarization." *Journal of Public Economics* 178: 104065 (2019).

[8] Bowen, Renee, Danil Dmitriev, and Simone Galperti. "Learning from shared news: when abundant information leads to belief polarization." NBER working paper 28465 (2021).

[9] Brennan, Geoffrey, and Alan Hamlin. "Expressive voting and electoral equilibrium." *Public Choice* 95 (1998): 149–175.

[10] Brennan, Geoffrey, and Loren Lomasky. *Democracy and Decision* (1993). Cambridge: Cambridge University Press.

[11] Buisseret, Peter, and Richard Van Weelden. "Polarization, valence, and policy competition." *American Economic Review: Insights* (2021).

[12] Calvert, Randall L. "Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence." *American Journal of Political Science* (1985): 69-95.

[13] Chambers, Christopher P., and Takashi Hayashi. "Bayesian consistent belief selection." *Journal of Economic Theory* 145.1 (2010): 432-439.

[14] Che, Yeon-Koo, and Konrad Mierendorff. "Optimal dynamic allocation of attention." *American Economic Review* 109.8 (2019): 2993-3029.

[15] Coate, Stephen, and Michael Conlin. "A group rule-utilitarian approach to voter turnout: Theory and evidence." *American Economic Review* 94.5 (2004): 1476-1504.

[16] Cover, Thomas M., and Martin E. Hellman, "The two-armed-bandit problem with time-invariant finite memory." *IEEE Transactions on Information Theory* 16.2 (1970): 185-195.

[17] Davis, Otto A., Melvin J. Hinich, and Peter Ordeshook. "An expository development of a mathematical model of the electoral process. " *American Political Science Review* 64.2 (1970): 426-448.

[18] Desch, Michael C. "War and strong states, peace and weak states?" *International Organization* 50.2 (1996): 237-268.

[19] Downs, Anthony. "An economic theory of political action in a democracy " *Journal of Political Economy* 65.2 (1957): 135-150.

[20] Eguia, Jon X., and Francesco Giovannoni. 2019. "Tactical extremism." *American Political Science Review* 113.1 (2019): 282-86.

[21] Feddersen, Timothy, and Alvaro Sandroni. "A theory of participation in elections." *American Economic Review* 96.4 (2006): 1271-1282.

[22] Fryer Jr, Roland G., Philipp Harms, and Matthew O. Jackson. "Updating beliefs when evidence is open to interpretation: Implications for bias and polarization." *Journal of the European Economic Association* 17.5 (2019): 1470-1501.

[23] Gentzkow, Matthew. "Polarization in 2016." Toulouse Network for Information Technology white paper (2016).

[24] Gordon, Sanford C., and Dimitri Landa. "Common problems (or, what's missing from the conventional wisdom on polarization and gridlock)." *Journal of Politics* 79.4 (2017): 1433-1437.

[25] Glazer, Amihai. "A new theory of voting: Why vote when millions of others do." *Theory and Decision* 22.3 (1987): 257–270.

[26] Groeling, Tim, and Matthew A. Baum. "Crossing the water's edge: Elite rhetoric, media coverage, and the rally-round-the-flag phenomenon." *Journal of Politics* 70.4 (2008): 1065-1085.

[27] Hamlin, Alan, and Colin Jennings. "Expressive voting." *The Oxford Handbook of Public Choice* 1 (2018): 333.

[28] Hirsch, Alexander V., and Kenneth W. Shotts. "Competitive policy development." *American Economic Review* 105.4 (2015): 1646–64.

[29] Hofstadter, Richard. "A long view: Goldwater in history." *New York Review of Books*, October 6th, 1964 Issue.

[30] Fryer Jr, Roland G., Philipp Harms, and Matthew O. Jackson. "Updating beliefs when evidence is open to interpretation: Implications for bias and polarization." *Journal of the European Economic Association* 17.5 (2019): 1470-1501.

[31] Kaustia, Markku, and Samuli Knüpfer. "Do investors overweight personal experience? Evidence from IPO subscriptions." *The Journal of Finance* 63 (2008): 2679–2702.

[32] Kogan, Shimon. "Distinguishing the effect of overconfidence from rational best-response on information aggregation." *The Review of Financial Studies* 22.5 (2009): 1889–1914.

[33] Kominers, Scott Duke, Xiaosheng Mu, and Alexander Peysakhovich. "Paying (for) attention: The impact of information processing costs on bayesian inference." SSRN working paper 2857978 (2018).

[34] Krasa, Stephan, and Mattias K. Polborn. "Competition between specialized candidates." *American Political Science Review* 104.4 (2010): 745–65.

[35] Krasa, Stephan, and Mattias K. Polborn. "Political competition between differentiated candidates." *Games and Economic Behavior* 76.1 (2012): 249–71.

[36] Krasa, Stephan, and Mattias K. Polborn. "Social ideology and taxes in a differentiated candidates framework." *American Economic Review* 104.1 (2014): 308–22.

[37] Ledyard, John O. "The pure theory of large two-candidate elections." *Public Choice* 44.1 (1984): 7-41.

[38] Levy, Gilat, and Ronny Razin. "Short-term political memory and the inevitability of polarisation." (2021).

[39] Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt. "Rational inattention: A review." *Journal of Economic Literature*, forthcoming (2021).

[40] Mas-Colell, Andreu, Whinston, Michael D., and Jerry Green. *Microeconomic Theory* (1995). New York: Oxford University Press.

[41] Matějka, Filip, and Guido Tabellini. "Electoral competition with rationally inattentive voters." *Journal of the European Economic Association* 19.3 (2021): 1899-1935.

[42] McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. *Polarized America: The Dance of Ideology and Unequal Riches.* Mit Press, 2016.

[43] McMurray, Joseph. "Polarization and pandering in common-interest elections." Working paper, August 2021.

[44] Nimark, Kristoffer P., and Savitar Sundaresan. "Inattention and belief polarization." *Journal of Economic Theory* 180 (2019): 203-228.

[45] Nunnari, Salvattore, and Jan Zapal. "A model of focusing in political choice." CEPR Discussion paper 12407, (2020).

[46] Oprea, Ryan. "What makes a rule complex?." *American Economic Review* 110.12 (2020): 3913-51.

[47] Perego, Jacopo, and Sevgi Yuksel. "Media competition and social disagreement." *Econometrica* 90.1 (2022): 223-265.

[48] Petrocik, John R. "Issue ownership in presidential elections, with a 1980 case study." *American Journal of Political Science* 40.3 (1996): 825–850.

[49] Pew Research Center. "The partisan divide on political values grows even wider." October 2017.

[50] Piccione, Michele, and Ariel Rubinstein. "On the interpretation of decision problems with imperfect recall." *Games and Economic Behavior* 20.1 (1997): 3-24.

[51] Prato, Carlo, and Stephane Wolton. "The voters' curses: why we need Goldilocks voters." *American Journal of Political Science* 60.3 (2016): 726-737.

[52] Putnam, Robert D. *The upswing: how America came together a century ago and how we can do it again.* Simon and Schuster, (2020).

[53] Riker, William H., and Peter C. Ordeshook. "A theory of the calculus of voting." *American Political Science Review* 62.1 (1968): 25-42.

[54] Schuessler, Alexander A. "Expressive voting." *Rationality and Society* 12.1 (2000): 87-119.

[55] Serra, Gilles. "Polarization of what? A model of elections with endogenous valence." *Journal of Politics* 72.2 (2010): 426-437.

[56] Tolvanen, Juha, James Tremewan, and Alexander K. Wagner. "Ambiguous Platforms and Correlated Preferences: Experimental Evidence." *American Political Science Review* (2021): 1-17.

[57] Tacitus, Cornelius. *Tacitus: Agricola.* Cambridge University Press, 2014.

[58] Trilling, Lionel. *The Liberal Imagination: Essays on Literature and Society.* New York Review of Books (2008 [1950]).

[59] Wilson, Andrea. "Bounded memory and biases in information processing." *Econometrica* 82.6 (2014): 2257-2294.

[60] Wittman, Donald. "Candidate motivation: A synthesis of alternative theories." *American Political Science Review* 77.1 (1983): 142-157.

[61] Yuksel, Sevgi. "Specialized learning and political polarization." *International Economic Review* forthcoming.

A Online Appendix

A.1 Detailed Derivation of Values and Beliefs in $FA_3^p(\alpha, \beta)$

As noted, we use u^R to denote $1 + b$ and u^L to denote $1 - b$, for an arbitrary $b \in (-\bar{b}, \bar{b})$. We first compute the continuation values for $FA_3^p(\alpha, \beta)$ according to the following recursive equations:

$$
\begin{aligned}
V_{q_R}(R, b) &= u^R + \delta\{(1 - \epsilon + \epsilon\alpha)V_{q_R}(R) + \epsilon(1 - \alpha)V_{q_L}(R, b)\}, \\
V_{q_M}(R, b) &= \delta\{(\mu - \epsilon)V_{q_R}(R, b) + \epsilon V_{q_L}(R, b) + (1 - \mu)V_{q_M}(R, b)\}, \\
V_{q_L}(R, b) &= -c + \delta\{[(\mu - \epsilon)\beta + 1 - \mu + \epsilon]V_{q_L}(R, b) + (\mu - \epsilon)(1 - \beta)V_{q_R}(R, b)\}.
\end{aligned}
$$

The above equations give a system of linear equations with three variables, $V_{q_L}(R, b)$, $V_{q_R}(R, b)$, and $V_{q_M}(R, b)$. Note that $V_{q_L}(R, b)$ and $V_{q_R}(R, b)$ do not depend on $V_{q_M}(R, b)$ and can be solved first as follows.

$$
\begin{aligned}
V_{q_L}(R, b) &= \frac{-c + \delta(\mu - \epsilon)(1 - \beta)V_{q_R}(R, b)}{1 - \delta[1 - (\mu - \epsilon)(1 - \beta)]}, \\
V_{q_R}(R, b) &= \frac{u^R}{1 - \delta(1 - \epsilon + \epsilon\alpha)} + \frac{\delta\epsilon(1 - \alpha)}{[1 - \delta(1 - \epsilon + \epsilon\alpha)]} \frac{-c + \delta(\mu - \epsilon)(1 - \beta)V_{q_R}(R, b)}{\{1 - \delta[1 - (\mu - \epsilon)(1 - \beta)]\}}, \\
V_{q_R}(R, b) &= \frac{\{1 - \delta[1 - \mu + \beta(\mu - \epsilon) + \epsilon]\}u^R - \delta\epsilon(1 - \alpha)c}{(1 - \delta)\{1 - \delta[1 - \mu + \beta(\mu - \epsilon) + \epsilon\alpha]\}} \\
&= \frac{[1 - \delta + \delta(1 - \beta)(\mu - \epsilon)]u^R - \delta\epsilon(1 - \alpha)c}{(1 - \delta)\{1 - \delta[1 - \mu + \beta(\mu - \epsilon) + \epsilon\alpha]\}}, \\
V_{q_L}(R, b) &= \frac{\delta(\mu - \epsilon)(1 - \beta)u^R - [(1 - \delta) + \delta(1 - \alpha)\epsilon]c}{(1 - \delta)\{1 - \delta[1 - \mu + \beta(\mu - \epsilon) + \epsilon\alpha]\}}.
\end{aligned}
$$

Symmetrically, we can solve for $V_{q_L}(L, b)$ and $V_{q_R}(L, b)$.

$$
\begin{aligned}
V_{q_L}(L, b) &= \frac{[(1 - \delta) + \delta(1 - \alpha)(\mu - \epsilon)]u^L - \delta\epsilon(1 - \beta)c}{(1 - \delta)\{1 - \delta[1 - \mu + \alpha(\mu - \epsilon) + \epsilon\beta]\}}, \\
V_{q_R}(L, b) &= \frac{\delta(\mu - \epsilon)(1 - \alpha)u^L - [(1 - \delta) + \delta(1 - \beta)\epsilon]c}{(1 - \delta)\{1 - \delta[1 - \mu + \alpha(\mu - \epsilon) + \epsilon\beta]\}}.
\end{aligned}
$$

Finally, we can substitute the solutions from $V_{q_L}(R,b)$ and $V_{q_R}(R,b)$ and $V_{q_L}(L,b)$ and $V_{q_R}(L,b)$ and obtain

$$
\begin{aligned}
V_{q_M}(R,b) &= \frac{\delta[(\mu-\epsilon)V_{q_R}(R,b)+\epsilon V_{q_L}(R,b)]}{1-\delta(1-\mu)} \\
&= \frac{\delta}{1-\delta}\times\frac{(\mu-\epsilon)[(1-\delta)+\delta\mu(1-\beta)]u^R-\epsilon[(1-\delta)+\delta(1-\alpha)\mu]c}{[1-\delta(1-\mu)]\{1-\delta[1-\mu+\beta(\mu-\epsilon)+\epsilon\alpha]\}}, \\
V_{q_M}(L,b) &= \frac{\delta}{1-\delta}\times\frac{(\mu-\epsilon)[(1-\delta)+\delta\mu(1-\alpha)]u^L-\epsilon[(1-\delta)+\delta(1-\beta)\mu]c}{[1-\delta(1-\mu)]\{1-\delta[1-\mu+\alpha(\mu-\epsilon)+\epsilon\beta]\}}.
\end{aligned}
$$

Now we compute the beliefs. The following expressions follow from the recursive equations for beliefs according to (11) and (12) in $FA_3^p(\alpha,\beta)$:

$$
\begin{aligned}
f(q_R|R) &= \delta\{f(q_M|R)(\mu-\epsilon)+f(q_R|R)[1-\epsilon(1-\alpha)]+f(q_L|R)(\mu-\epsilon)(1-\beta)\}, \\
f(q_L|R) &= \delta\{f(q_M|R)\epsilon+f(q_L|R)[(1-\mu+\epsilon)+(\mu-\epsilon)\beta]+f(q_R|R)\epsilon(1-\alpha)\}, \\
f(q_M|R) &= (1-\delta)\mathbf{P}_0(R)+\delta f(q_M|R)(1-\mu), \\
f(q_L|L) &= \delta\{f(q_M|L)(\mu-\epsilon)+f(q_L|L)[1-\epsilon(1-\beta)]+f(q_R|L)(\mu-\epsilon)(1-\beta)\}, \\
f(q_R|L) &= \delta\{f(q_M|L)\epsilon+f(q_R|L)[(1-\mu+\epsilon)+(\mu-\epsilon)\alpha]+f(q_L|L)\epsilon(1-\beta)\}, \\
f(q_M|L) &= (1-\delta)\mathbf{P}_0(L)+\delta f(q_M|L)(1-\mu), \\
f(q_R|M) &= \delta\{f(q_M|M)\epsilon+f(q_R|M)[1-\epsilon(1-\alpha)]+f(q_L|M)\epsilon(1-\beta)\}, \\
f(q_L|M) &= \delta\{f(q_M|M)\epsilon+f(q_L|M)[1-\epsilon(1-\beta)]+f(q_R|M)\epsilon(1-\alpha)\}, \\
f(q_M|M) &= (1-\delta)\mathbf{P}_0(M)+\delta f(q_M|M)(1-2\epsilon).
\end{aligned}
$$

First, we can solve for $f(q_M|R)$, $f(q_M|L)$, and $f(q_M|M)$ directly:

$$
f(q_M|R) = \frac{(1-\delta)\mathbf{P}_0(R)}{1-\delta(1-\mu)},\ \ f(q_M|L)=\frac{(1-\delta)\mathbf{P}_0(L)}{1-\delta(1-\mu)},\ \ f(q_M|M)=\frac{(1-\delta)\mathbf{P}_0(M)}{1-\delta(1-2\epsilon)}.
$$

Then, we can solve for $f(q_R|R)$ and $f(q_L|R)$ simultaneously:

$$
\begin{aligned}
f(q_R|R) &= \frac{\delta(\mu-\epsilon)f(q_M|R)+\delta(\mu-\epsilon)(1-\beta)f(q_L|R)}{1-\delta[1-\epsilon(1-\alpha)]}, \\
f(q_L|R) &= \frac{\delta\epsilon f(q_M|R)+\delta\epsilon(1-\alpha)f(q_R|R)}{1-\delta[1-(\mu-\epsilon)(1-\beta)]},
\end{aligned}
$$

and, by plugging in the solution for $f(q_M|R)$, we obtain

$$
\begin{aligned}
f(q_R|R) &= \frac{\delta(\mu - \epsilon)[(1 - \delta) + \delta(1 - \beta)\mu]}{(1 - \delta)\{1 - \delta[1 - \epsilon(1 - \alpha) - (\mu - \epsilon)(1 - \beta)]\}} f(q_M|R) \\
&= \frac{\delta(\mu - \epsilon)[(1 - \delta) + \delta(1 - \beta)\mu]\mathbf{P}_0(R)}{\{1 - \delta[1 - \epsilon(1 - \alpha) - (\mu - \epsilon)(1 - \beta)]\}[1 - \delta(1 - \mu)]},
\end{aligned}
$$

and

$$
\begin{aligned}
f(q_L|R) &= \frac{\delta\epsilon(1 - \delta)\{1 - \delta[1 - \epsilon(1 - \alpha) - (\mu - \epsilon)(1 - \beta)]\}f(q_M|R)}{\{1 - \delta[1 - (\mu - \epsilon)(1 - \beta)]\}(1 - \delta)\{1 - \delta[1 - \epsilon(1 - \alpha) - (\mu - \epsilon)(1 - \beta)]\}} \\
&+ \frac{\delta^2\epsilon(1 - \alpha)(\mu - \epsilon)[(1 - \delta) + \delta(1 - \beta)\mu]f(q_M|R)}{\{1 - \delta[1 - (\mu - \epsilon)(1 - \beta)]\}(1 - \delta)\{1 - \delta[1 - \epsilon(1 - \alpha) - (\mu - \epsilon)(1 - \beta)]\}} \\
&= \frac{\delta\epsilon[(1 - \delta) + \delta(1 - \alpha)\mu]}{(1 - \delta)\{1 - \delta[1 - \epsilon(1 - \alpha) - (\mu - \epsilon)(1 - \beta)]\}} f(q_M|R) \\
&= \frac{\delta\epsilon[(1 - \delta) + \delta(1 - \alpha)\mu]\mathbf{P}_0(R)}{\{1 - \delta[1 - \epsilon(1 - \alpha) - (\mu - \epsilon)(1 - \beta)]\}[1 - \delta(1 - \mu)]}.
\end{aligned}
$$

Symmetrically, we have

$$
\begin{aligned}
f(q_L|L) &= \frac{\delta(\mu - \epsilon)[(1 - \delta) + \delta(1 - \alpha)\mu]\mathbf{P}_0(L)}{\{1 - \delta[1 - \epsilon(1 - \beta) - (\mu - \epsilon)(1 - \alpha)]\}[1 - \delta(1 - \mu)]}, \\
f(q_R|L) &= \frac{\delta\epsilon[(1 - \delta) + \delta(1 - \beta)\mu]\mathbf{P}_0(L)}{\{1 - \delta[1 - \epsilon(1 - \beta) - (\mu - \epsilon)(1 - \alpha)]\}[1 - \delta(1 - \mu)]}.
\end{aligned}
$$

Finally, by adding up the recursive equations for $f(q_R|M)$ and $f(q_L|M)$, we have

$$
f(q_R|M) + f(q_L|M) = \frac{\delta}{1 - \delta} 2\epsilon f(q_M|M),
$$

and hence we can substitute $f(q_L|M) = \frac{\delta}{1-\delta} 2\epsilon f(q_M|M) - f(q_R|M)$ in the recursive equation for $f(q_R|M)$, putting all terms involving $f(q_R|M)$ on the one side, and obtain

$$
\begin{aligned}
f(q_R|M) &= \frac{\delta\epsilon[(1 - \delta) + \delta(1 - \beta)2\epsilon]}{(1 - \delta)[1 - \delta[1 - 2\epsilon + \epsilon(\alpha + \beta)]]} f(q_M|M) \\
&= \frac{\delta\epsilon[(1 - \delta) + \delta(1 - \beta)2\epsilon]\mathbf{P}_0(M)}{\{1 - \delta[1 - 2\epsilon + \epsilon(\alpha + \beta)]\}[1 - \delta(1 - 2\epsilon)]}, \\
f(q_L|M) &= \frac{\delta\epsilon[(1 - \delta) + \delta(1 - \alpha)2\epsilon]\mathbf{P}_0(M)}{\{1 - \delta[1 - 2\epsilon + \epsilon(\alpha + \beta)]\}[1 - \delta(1 - 2\epsilon)]}.
\end{aligned}
$$

It is then straightforward to confirm that, for ϵ sufficiently small, $p(q_R, m)(R)$ is arbitrarily close to one, and $p(q_L, m)(L)$ is arbitrarily close to one.