Political social-learning: Short-term memory and cycles of polarisation

Gilat Levy and Ronny Razin, LSE[1]

Abstract: This paper investigates the effect of voters' short-term memory on political outcomes by considering politics as a collective learning process. We find that short-term memory may lead to cycles of polarisation and consensus across parties' platforms. Following periods of party consensus, short-term memory implies that there is little variation in voters' data and therefore limited information about the true state of the world. This in turn allows parties to further their own interests and hence polarise by offering different policies. In contrast, periods of polarisation and turnover involve sufficient variation in the data that allows voters to be confident about what the correct policy is, forcing both parties to offer this policy.

[1]Department of Economics, Houghton St. WC2A 2AE, UK. Email addresses: g.levy1@lse.ac.uk, r.razin@lse.ac.uk. We thank the editor and four anonymous referees for their very helpful advice. For useful comments we also thank Steve Callander, Andrew Little, and many seminar participants and discussants. Roberto Maura-Rivero, Raffaele Blasone and Enrico Turri provided helpful research assistance.

1 Introduction

The rise of polarisation in democratic societies in the last few decades has garnered significant attention in recent academic literature.[2] However, polarisation is not a new phenomenon. By looking at longer time periods, a cyclical pattern of consensus and polarisation can be observed. For instance, in the US, policy positions of Senators and Congress members show that polarisation was high at the beginning of the 20th century, decreased in the 1930s, remained low until the late 1970s, and has been increasing since (see Figure 1).

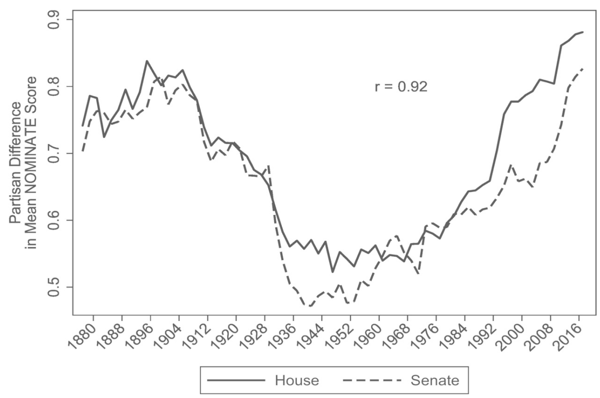

Figure 1: Historical polarisation in the US Senate and Congress (McCarty 2019).

Similar cyclical patterns can be seen in political parties' stated ideologies by examining their manifestos over time. The Manifesto Project decodes policy dimensions into a unidimensional score and tracks changes in these scores over time.[3] The manifestos of the two major parties in the US clearly oscillate between polarisation and relative consensus on economic issues. For example, Figure 2 plots the positions of the Democratic and Republican parties on market regulation, demonstrating periods of little difference in positions from the 1980s to the early 2000s, with relative polarisation before and after these periods.

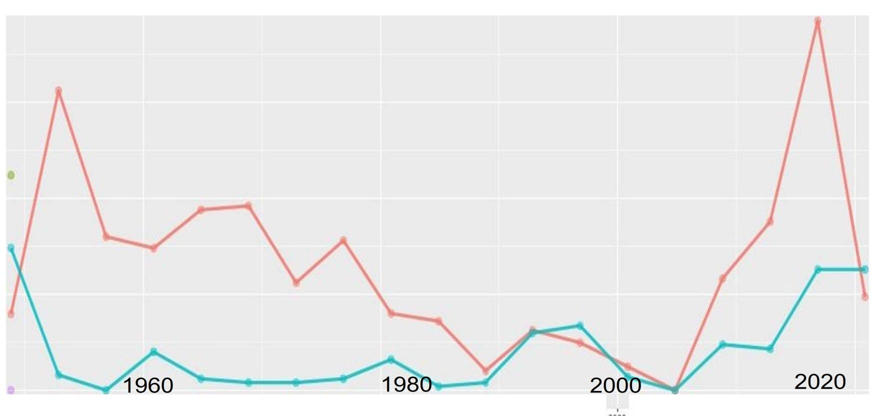

Figure 2: US Democratic (the higher curve) and Republican parties' regulation policies.

[2] See Barber and McCarty (2013) and McCarty (2019).

[3] See https://manifesto-project.wzb.eu/.

We investigate the question of whether cycles of polarisation and consensus in party platforms are an inherent feature of political systems. By studying a dynamic model in which voters learn from history about the optimal policy, we demonstrate that such cycles can emerge when voters have limited or short-term memory. Our findings suggest that short-term memory can imply a systemic cyclical force in politics.

The examination of how short-term memory affects the evolution of policies in society is important for both positive and normative reasons. Previous research, starting from Kramer (1971), has lead to a broad consensus indicating that voters in presidential and congressional elections in the US tend to "myopically ignore any information beyond the recent past" (Peltzman, 1990), a conclusion that is shared by other studies (Achen and Bartels, 2008; Bartels, 2008; Gerber and Green, 1998; Lenz, 2010; Healy and Lenz, 2014, Angelucci and Prat 2023). Short-termism can then hinder learning from past mistakes. One example of repeated mistakes is the cyclical pattern of over-regulation and under-regulation in financial markets. Rajan (2009) notes that "once memories of the current crisis fade and the ideological cycle turns, the political pressure to soften capital requirements or their enforcement will be enormous."

Motivated by the above empirical research, we incorporate the assumption of short-term memory into a dynamic political framework in which voters learn about the true data generating process that influences observable outcomes. In our model, there are two ideologically motivated parties, each with its own political interests, such as a preference for high levels of regulation or deregulation. Voters observe historical experiences and compare the expected utility they will gain from the policies that parties propose. To analyze the electoral competition between these parties, we use a probabilistic voting model and so parties' choice of policies takes into account their partisan interests as well as voters' uncertain preferences. In this model beliefs and policies evolve together over time.

Our model embeds a classic mechanism from the political economy literature, where party positions can diverge when parties are policy motivated and uncertain about voter preferences (as in Calvert 1985). The extent of this polarisation depends in our model on how much voters are able to differentiate between the policies and understand which is better. When their knowledge is relatively weak their vote is determined by other stochastic factors such as party or candidate attachment and this allows parties to polarise. When voters' knowledge allows them to understand well which policy is optimal, their voting behaviour will be less noisy and parties will be disciplined to offer this policy.

What is novel in our paper is that we study this electoral competition between parties in a dynamic model in which both parties' policies as well as voters' knowledge are endogenous. We show how this can lead to cycles of party polarisation and consensus. To see these dynamics, assume first that voters' memory includes only periods in which parties were polarised and so offered different policies. Voters' historical data will then include sufficient variation in policies (resulting from political turnover). This variation provides voters with sufficient information about the effectiveness of different policy options, enabling them to identify the optimal policy. Parties will then be incentivized to offer this policy and will therefore veer towards more consensus. However, after extended periods in which parties are in consensus, voters' memories will consist only of experiences with the same policy. This lack of variation in the observed data will imply that while voters will understand well the benefits of this policy, they might not know how good or bad they have it as compared with an alternative. This weak ability of voters to distinguish among the alternatives will enable parties to credibly offer policies more in line with their partisan inclination, resulting in party polarisation. Ultimately, it is the phase of consensus that can lead to polarisation, and the phase of polarisation that can eventually give way to consensus.

We identify a simple sufficient condition that gives rise to such dynamics. Specifically, cycles emerge when voters are able to differentiate the expected utility of different policies more effectively when they have access to an informative history with policy variation, as compared to a less informative history. For any given prior distribution over the state of the world, there always exist states of the world for which this is satisfied. Additionally, we show how for some prior distributions (featuring "scale-free" learning) this condition is satisfied for almost all states of the world.

In our model, transitions between the phases of the cycle result from the effects of both endogenous policies and exogenous shocks on voter learning. Absent exogenous shocks, voters can learn relatively quickly which policy is optimal, and the cyclical dynamics are such that phases of policy consensus are relatively long compared to phases of party polarisation. This highlights the role of endogenous policies as a systemic force driving the emergence of cycles in political systems. Voters' learning can also be affected by exogenous shocks, such as natural disasters, wars, or pandemics and our analysis illustrates how such shocks together with endogenous policies combine to affect the cycle dynamics (see Section 3.2).

The implication of the forgetfulness of voters is that society will make repeated mistakes; short-term memory is "costly" as voters' knowledge becomes limited ever so often, and so a polity that implements the correct policy in a consensus phase will necessarily desert it after a while. However, in some environments this can actually be beneficial. To see this, note that another way to view polarisation phases is that once in a while, the polity inadvertently "experiments" with different policies. If the true state of the world changes sometimes, such experimentation allows voters to detect it. We show in Section 4 that this feature allows polities with short-term memory to sometimes outperform polities with unbounded memory.

In Section 5 we extend the model to consider an environment in which different voters are exposed to different histories. Such heterogeneity among voters can arise from echo chambers or potentially from cohort effects where voters are affected by different experiences from their youth. We show how such different memories affect the nature of cycles. In Section 6 we consider other extensions, Section 7 discusses the related literature, and all proofs are in an Appendix.

2 The model and preliminary results

A polity is considering a choice between two policies l and r. We first describe the economic environment, which is a simple mapping between policies and outcomes. Specifically, let the observable economic outcome y_t at period t be:

$$(II.1)\ y_t = \begin{cases} \beta_l^* + \varepsilon_t \text{ if policy } l \text{ is chosen} \\ \beta_r^* + \varepsilon_t \text{ if policy } r \text{ is chosen} \end{cases}$$

where ε_t is iid across time and Normally distributed with zero mean and variance σ^2. Our model and analysis are simplified by the assumption that the set of policies is discrete. Our results about cycles can be generalised to the case of continuous policies (see the discussion in Section 6.3).

Voters understand how the data generating process depends on parameters $\boldsymbol{\beta} = (\beta_l, \beta_r)$, but do not know the true value of these parameters, $\boldsymbol{\beta}^* = (\beta_l^*, \beta_r^*)$. They are endowed with a continuous and symmetric prior $G(\beta_l, \beta_r)$ on some compact set $B \in \mathbb{R}^2$, which determines how $\boldsymbol{\beta}^*$ is generated. In the main part of the analysis we consider a fixed $\boldsymbol{\beta}^*$. We discuss the possibility of changing states later on.

The outcome y_t is a common element in the voters' preferences, and so everything else equal, at any period t, all voters prefer policy l if given their information at period t, Ω_t, $E[\beta_l|\Omega_t] > E[\beta_r|\Omega_t]$. This will feed into their voting behaviour, which we will describe below.

2.1 Political parties and electoral competition

There are two parties, each identified with a special interest on a different policy. Party L prefers policy l and party R prefers policy r. The utilities of party L and R from policy $p \in \{l, r\}$, $U^R(p)$ and $U^L(p)$, satisfy:

$$(II.2)\ U^R(r) = U^L(l) = 1, U^R(l) = U^L(r) = 0.$$

In addition, parties also enjoy office rents, denoted by $\alpha > 0$. Thus, given an election at period t, and an implemented policy p_t, party $J's$ utility, for $J \in \{L, R\}$, is

$$U^J(p_t) + I_t^J \alpha$$

where $I_t^J = 1$ if party J won the election and 0 otherwise. Our results also hold for other types of policy motives, e.g., if parties have a legacy motive and so enjoy a higher utility from their ideal policy if they are also the party that implements this policy.

In an election, at any period t, each party offers a policy $p_t^J \in \{l, r\}$. We say that parties polarise when $p_t^R \neq p_t^L$; naturally, this will imply that $p_t^R = r$ and $p_t^L = l$. When $p_t^R = p_t^L$, we say that parties are in consensus.

2.2 Histories and Memory

At each period t voters observe data from only the last $K < \infty$ periods. In particular, we start the model at period 0 with an initial history $H_0 = (p_\tau, y_\tau)_{\tau=-K}^{\tau=-1}$ and denote the history observed by voters at period $t \geqslant 0$ by $H_t = (p_\tau, y_\tau)_{\tau=t-K}^{\tau=t-1}$ where for any τ, $p_\tau \in \{l, r\}$ is the implemented policy in period τ and y_τ is the policy outcome in that period. Thus, at every period t, the information Ω_t held by voters is composed of the prior G, and the K-period history H_t. We model short-term memory by assuming that voters are not able to decode past learning from the behaviour of parties and past electorates. To do so, we use two assumptions: Each period voters use the same original prior G, and treat the history as exogenous. Thus, voters do not ask themselves "why did the parties choose those platforms in the last few periods".

Voters use Bayes rule to compute their posterior distribution G_t on the vector $\boldsymbol{\beta}$, given $\{G, H_t\}$. Specifically, the posterior density distribution $g_t(.)$ on vectors $\boldsymbol{\beta} = (\beta_l, \beta_r)$ is:

$$g_t(\boldsymbol{\beta}) = \frac{g(\boldsymbol{\beta}) \prod_{\tau=t-K}^{\tau=t-1} f(y_\tau - E[y_\tau | p_\tau, \boldsymbol{\beta}])}{\int_{\boldsymbol{\beta}'} g(\boldsymbol{\beta}') \prod_{\tau=t-K}^{\tau=t-1} f(y_\tau - E[y_\tau | p_\tau, \boldsymbol{\beta}']) d\boldsymbol{\beta}'}$$

where $E[y_\tau | l, \boldsymbol{\beta}] = \beta_l$ and $E[y_\tau | r, \boldsymbol{\beta}] = \beta_r$ and $f(.)$ is the (Normal) density of the shock ε_t.

One can consider alternative assumptions on memory. For example, that individuals put different weights on different past periods. All our results hold if instead of short term memory, the weight on events in the past declines fast enough with time.

2.3 Electoral Competition

As already shown in Wittman (1983) and Calvert (1985), party polarisation can arise in an electoral competition when parties are ideological (as we assume) and when they face some uncertainty with regards to voting behaviour. Above we described the element of preferences which is common to voters (the outcome y_t, and observing the same history H_t). Naturally voters may differ on additional aspects which parties may have uncertainty about. For example, voters may react differently to the personal attributes of candidates. We summarise all these other factors, as in the probabilistic voting model (Lindbeck and

Weibull 1987), by a random variable ϕ_t which represents the bias of the median voter towards party L. In period t, the median voter votes for party L if

$$(II.3)\ E[y(p_t^L) - y(p_t^R)|H_t] + \phi_t > 0,$$

where ϕ_t is iid and uniformly distributed on $[-\frac{1}{2\zeta}, \frac{1}{2\zeta}]$, where $\zeta \in (0, \infty)$, and

$$E[y(p_t^L) - y(p_t^R)|H_t] = \left\{ \begin{array}{c} E[(\beta_l - \beta_r)|H_t] = \int_{\boldsymbol{\beta}} (\beta_l - \beta_r) g_t(\boldsymbol{\beta}) d\boldsymbol{\beta} \text{ if parties polarise} \\ 0 \text{ otherwise} \end{array} \right\}.$$

In the case of equality in $(II.3)$, without loss of generality, we assume that the voter votes for party L with probability 0.5.

We assume that voters and parties are myopic. Myopia is a standard assumption in models of electoral competition; we discuss the effects of forward-looking parties and voters in Section 6.2.

Given the above, in period t, the probability that party L wins the election is

$$\begin{aligned} \Pr(L \text{ wins}|p_t^L, p_t^R) &= \Pr(\phi_t > -E[y(p_t^L) - y(p_t^R)|H_t]) \\ &= \begin{cases} 1 \text{ if } \frac{1}{2} + \zeta E[y(p_t^L) - y(p_t^R)|H_t] > 1 \\ 0 \text{ if } \frac{1}{2} + \zeta E[y(p_t^L) - y(p_t^R)|H_t] < 0 \\ \frac{1}{2} + \zeta E[y(p_t^L) - y(p_t^R)|H_t] \text{ otherwise} \end{cases} \end{aligned}$$

Given the platform of party R, p_t^R, party L maximises her expected utility,

$$\Pr(L \text{ wins}|p_t^L, p_t^R)(U^L(p_t^L) + \alpha) + (1 - \Pr(L \text{ wins}|p_t^L, p_t^R))U^L(p_t^R)$$

by choosing p_t^L in equilibrium. Similarly, given p_t^L, party R chooses p_t^R to maximise an analogous expression, where $U^J(p)$ is defined in $(II.2)$ and parties know G, H_t. For completeness we assume that if a party is in equilibrium indifferent between offering l or r then it offers its ideal policy (this assumption plays no role in the analysis as this case will generically not arise).

At any period t, the equilibrium in the electoral competition model is a Nash equilibrium between the parties, where parties choose the platforms p_t^L and p_t^R simultaneously to maximise their utilities detailed above. Following the choice of platforms, ϕ_t is drawn and the election result is in accordance with $(II.3)$.

2.4 Dynamics

The only dynamic link between the periods is the evolving memory of voters. The dynamic model is defined therefore as:

1. There is some initial history H_0 of size K.
2. In period t the party that won the election, party $J \in \{L, R\}$, implements $p_t^J \in \{l, r\}$.
3. Given y_t, history evolves from $H_t = \{p_\tau, y_\tau\}_{\tau=t-K}^{\tau=t-1}$ to $H_{t+1} = \{p_\tau, y_\tau\}_{\tau=t-K+1}^{\tau=t}$.
4. A new electoral competition equilibrium arises in period $t+1$: The two parties offer their equilibrium policy platforms p_{t+1}^J, ϕ_{t+1} is drawn and party L wins the election if

$$E[y(p_{t+1}^L) - y(p_{t+1}^R)|H_{t+1}] + \phi_{t+1} > 0,$$

or with probability 0.5 if the above is satisfied with equality.

2.5 The incentives to polarise: A useful Lemma

Our first result highlights the mechanism by which the level of information in historical data affects electoral competition. Fix a history H_t and consider the one-period political competition game that ensues:

Lemma 1 (Consensus vs Polarisation): *In period t:*
(i) If

$$|E[(\beta_l - \beta_r)|\ H_t]| > \frac{1}{2\zeta(1+\alpha)}$$

then the unique equilibrium involves both parties choosing the same policy (consensus);
(ii) If

$$|E[(\beta_l - \beta_r)|\ H_t]| < \frac{1}{2\zeta(1+\alpha)}$$

then the unique equilibrium involves each party choosing its ideal policy (polarisation).[4]

Generally speaking, parties prefer to pursue their own interests. But they also care about being elected. This implies that sometimes they may be disciplined by voters to choose the policy that voters think is more likely to generate a higher outcome given the historical data. If, given the historical data, voters sufficiently prefer one of the policies, say l, then parties have to offer l and are hence in consensus. In this case, if a party offers r, it will serve neither its own policy interest nor its office motivation, as it will face only a slim probability of being elected. Alternatively, if the historical data does not allow voters to sufficiently discriminate between the different policies, then parties can afford to take the risk and offer platforms that better serve their own policy interests and hence polarise.

Note that a higher office rent α pushes parties to offer the same policy, as they are more eager to get elected and therefore to satisfy the voters' will. This also implies that when parties do polarise, each party must win with a probability that is bounded away from zero;

[4] Both types of equilibria can hold when $|E[(\beta_l - \beta_r)|\ H_t]| = \frac{1}{2\zeta(1+\alpha)}$.

if a party polarises and has only a negligible chance of being elected, it can deviate to the same platform as the other party, gain $\frac{1}{2}\alpha$, and be better off. A higher ζ implies a smaller variance for the shock ϕ_t : This means that parties are more certain about how voters vote, and again this pushes parties to reach consensus more often. As both ζ and α work in the same direction, we can denote the right-hand-side of the expression in the Lemma as

$$\rho \equiv \frac{1}{2\zeta(1+\alpha)},$$

where a lower value for ρ is more conducive to party consensus.

To make learning meaningful in the model, it is reasonable to consider the case where parties are in consensus when voters know the true state. With this in mind, given the Lemma above, we will assume that $|\beta_l^* - \beta_r^*| > \rho$. This is a reasonable requirement in a social learning model to allow voters to reach the optimal policy. For concreteness and without loss of generality, let us assume that the state of the world, a primitive of the model, is such that l is the optimal policy so that $\beta_l^* > \beta_r^*$.

3 Cycles of consensus and polarisation

In what follows we are interested in whether and how a polity transitions between periods of party polarisation and party consensus. To this end the following notation will be useful. Let $\eta_t(polarisation)$ denote the fraction of time in the full history up to time t that the two parties offered different platforms and by $\eta_t(consensus) = 1 - \eta_t(polarisation)$ the fraction of time in the full history up to time t that the two parties offered the same platform. Denote by $\hat{\eta}_t(p)$ the fraction of time in the full history up to time t that policy p was implemented.

Note that the dynamic evolution of policies involves some randomness, given the voting shock ϕ_t and the policy shock ε_t (through the latter's effect on beliefs). This then induces a probability distribution P over the set of infinite paths of history $\mathbb{H}$. Thus, when we write "almost surely" below, here and in the Appendix, we mean P-almost surely on $\mathbb{H}$.

3.1 A benchmark: Unbounded memory

As a benchmark we first consider the case in which the history that voters remember is unlimited, i.e., when $K = \infty$. The model and specifically our definitions of the evolution of histories can be easily extended to this infinite memory case.[5] Our result below shows that with full memory, the two parties will reach a consensus on the same platform.

[5] Specifically, we start with some finite initial history H_0, and at any period t, given y_t, history evolves from H_t to $\{H_t, (p_t, y_t)\}$.

Proposition 1: *Assume that $K = \infty$. Then almost surely the polity experiences long-term consensus, i.e., $\eta_t(consensus) \to 1$ and $\hat{\eta}_t(p) \to 1$ for some $p \in \{l, r\}$.*

To see the intuition for Proposition 1, note first that long-term polarisation cannot arise in equilibrium. When $K = \infty$, the beliefs of voters will converge in the long term as they form a martingale. If beliefs converge so that voters do not sufficiently distinguish between the utilities of the two policies, parties polarise in line with Lemma 1. As parties also have office motivations, it implies that when parties polarise, each has a probability of being elected that is bounded from zero. Given the stochastic political turnover, polarisation implies that voters (unintentionally) "experiment" in the long-term with two different policies: This then allows voters to learn the true state, $\boldsymbol{\beta}^*$. As we assumed that $\beta_l^* - \beta_r^* > \rho$, this will lead to a contradiction as by Lemma 1 the two parties must reach a consensus on l. When $K = \infty$, voters' beliefs in the long-term must converge then to beliefs that induce parties to be in consensus.[6]

Consensus however is *not* guaranteed to be on the optimal policy; as in any learning problem with myopic agents, learning can sometimes lead voters to believe that the wrong policy is optimal due to insufficient "experimentation". For example, a rare and finite series of very bad shocks when l is implemented, may convince voters that r is more likely to be better and therefore induce parties to reach a consensus on r. This consensus can be sustained in the long run; voters will learn β_r^* but might maintain beliefs that the difference between the expected utilities of the two policies is large enough, as they never forget the initial series of bad shocks on l. Thus, long-term memory also implies that long-term outcomes can be history-dependent.

3.2 Short-term memory and cycles

We now consider short-term memory, i.e., finite K. In this case the nature of voters' data can change over time. If, for example, power did not change hands or parties' platforms are the same, history contains little variation in policies and voters' data is rather uninformative. Alternatively when the history involves party polarisation and a high political turnover, voters' data will be relatively informative.

We already know by Lemma 1 that if voters' history is not informative at all, then parties will polarise, as $|E(\beta_l - \beta_r)| = 0$. At the other extreme, if voters' history is fully informative and voters learn the state, parties will reach a consensus on l as $\beta_l^* - \beta_r^* > \rho$. What happens on the equilibrium path with short-term memory will depend then on whether

[6] The condition $|\beta_l^* - \beta_r^*| > \rho$ is only a sufficient condition for long-term consensus. If $|\beta_l^* - \beta_r^*| < \rho$, then beliefs can converge to either allow long-term party polarisation or long-term party consensus.

voters' knowledge allows them to sufficiently distinguish the expected utilities of the two policies. The evolution of the short-term knowledge of voters is further complicated by the fact that policies are themselves endogenous, and by the stochastic noise in the learning process.

To illustrate the mechanism of cycles it is instructive to shut down this stochastic noise and so we first consider the case in which the variance of the outcome shock is infinitesimal. When policy noise vanishes, learning is very fast. Thus, after any period, in the limit as $\sigma \to 0$, voters learn the true effectiveness of the policy that was implemented in that period. As a result, observing two periods with different implemented policies allows voters to learn the state in full.

Consider any $K \geq 2$ and the case in which voters' observed history contains only one implemented policy $p \in \{l, r\}$. As $\sigma \to 0$, voters will learn the true value of the effectiveness of p, β_p^*. The condition below relates to what voters will learn about the other policy (that was not implemented), denoted by $-p$. Let $E[\beta_{-p}|\beta_p^*]$ denote the expected effectiveness of policy $-p$ when voters know β_p^* (and nothing else). The following assumption will allow us to characterise when cycles arise. For completeness, Assumption 1 includes the assumption that we made at the end of Section 2, that $\beta_l^* - \beta_r^* > \rho$.

Assumption 1 The state $\boldsymbol{\beta}^*$ satisfies:

$$\max\{|\beta_l^* - E[\beta_r|\beta_l^*]|, |\beta_r^* - E[\beta_l|\beta_r^*]|\} \ < \rho < \beta_l^* - \beta_r^*$$

Assumption 1 implies that when voters learn the benefit of one policy only, they are less sure about the relative benefit of the two policies, compared to the case in which they learn the benefit of both. We postpone for later the discussion of when Assumption 1 is likely to hold. We henceforth focus on the parameters that satisfy Assumption 1.

In Proposition 2 below we derive systemic cycles of polarisation due to endogenous learning:

Proposition 2: *Let $\sigma \to 0$ and $K \geq 2$. Then the polity experiences perpetual cycles of polarisation and consensus. The unique limit equilibrium has two phases:*

(i) A consensus phase lasting exactly K periods in which both parties espouse the optimal policy l.

(ii) A polarisation phase in which each party espouses its ideal policy. This stage lasts until the party espousing r wins the election. The expected length of the polarisation phase is

$$\frac{1}{1 - (\frac{1}{2} + \zeta(\beta_l^* - E[\beta_r|\beta_l^*])},$$

and so, in the limit equilibrium, the fraction of time the polity implements the correct policy is larger than $1 - \frac{1}{K}$.

The proposition uncovers systemic polarisation cycles. To see the intuition, assume that the polity starts with both parties offering the same policy p. As σ is vanishing, after K periods, the knowledge of voters is perfect about β_p^*, whereas their knowledge on β_{-p} arises from the prior G and their knowledge of β_p^*, with an expectation equal to $E[\beta_{-p}|\beta_p^*]$. By Assumption 1, parties will then polarise. Then, at some point in this polarisation phase, $-p$ will be implemented, allowing the polity to learn from the most recent period and the one before that l is optimal. A consensus phase then arises in which both parties espouse l, and this lasts as long as voters remember their benefit from each policy. Following the Kth period of the consensus phase, voters only remember periods in which l was implemented, which by Assumption 1 means that parties polarise. Thus, society is in an absorbing state of cycles between phases of consensus on l, and polarisation until r is implemented.[7]

Assumption 1 is a sufficient condition that guarantees that cycles can arise: $\beta_l^* - \beta_r^* > \rho$ implies that periods of polarisation will result in consensus on l and $\rho > \max\{|\beta_l^* - E[\beta_r|\beta_l^*]|, |\beta_r^* - E[\beta_l|\beta_r^*]|\}$ implies that periods of consensus on l (and more generally also on r) will result in polarisation. In the Appendix we characterise equilibria when Assumption 1 does not hold. We show that the equilibria have either cycles as above, or perpetual consensus or perpetual polarisation.

A discussion of Assumption 1: When would we expect Assumption 1 to be satisfied? To answer this, note that we can decompose the assumption into two. The first states that $\beta_l^* - \beta_r^* > \max\{|\beta_l^* - E[\beta_r|\beta_l^*]|, |\beta_r^* - E[\beta_l|\beta_r^*]|\}$. This is a condition on the state $\boldsymbol{\beta}^*$ and the prior G. The second part requires that ρ, which captures parameters relating to electoral competition, falls within the range of the above two values.

The condition $\beta_l^* - \beta_r^* > \max\{|\beta_l^* - E[\beta_r|\beta_l^*]|, |\beta_r^* - E[\beta_l|\beta_r^*]|\}$ highlights an intuitive feature of policy evaluation between observed and unobserved policies. In practice, relative policy evaluation between two policies that have both been experienced in the past is easier compared with the case in which only one policy has been experienced. The condition boils down to requiring that when only one policy is experienced it is hard to tell the difference between the two policies. While voters understand well the benefits of this policy, they might not know how good or bad they have it compared with an alternative. This condition is intuitive yet in contrast to standard models of voters with ideal policies, which assume that voters know what is their satiation level.

One way to generate the condition above is in the case in which voters' learning is "scale-free".[8] In particular, suppose that the scale of utility voters experience from one policy is

[7] While here we present the results with a stark model of a finite memory, our results hold more generally with infinite, weighted-memory, as long as weights on the past histories decay fast enough.

[8] We thank a referee for suggesting the idea of scale-free learning.

not too informative about the other; voters might not be sure what is the highest utility scale they can gain, and whether they had reached this level. In the Appendix we show that the condition can then be easily satisfied; we provide an example of a "scale-free" learning environment, for which almost all states satisfy it. In the simple example that we present, $\max\{|\beta_l^* - E[\beta_r|\beta_l^*]|, |\beta_r^* - E[\beta_l|\beta_r^*]|\} \approx 0$, and so learning the utility scale of one policy provides no knowledge about which one is better.

The second part of Assumption 1 involves the level of ρ. In our analysis we assume that ρ is fixed throughout but more generally one can entertain that the parameters that affect the level of ρ, namely ζ and α, might change over time. The level of uncertainty about voters' intentions, ζ, might depend on particular aspects of the election such as political engagement and media consumption. The utility of parties from winning office, α, might also change from year to year depending on politicians' ability to capture rents once in office. As long as the distribution of ρ is such that with strictly positive probability it is in the desired range, the qualitative nature of our results will hold.

The length of consensus and polarisation phases: Note that the parameters discussed above affect the length of the polarisation phase in a natural way. Proposition 2 readily implies that the consensus phase is higher when K is larger as voters retain for longer the knowledge of the effectiveness of both policies.

The average length of the polarisation phase is in turn affected by the probability that policy l is elected in this phase, which increases in $\zeta(\beta_l^* - E[\beta_r|\beta_l^*])$. When $\beta_l^* - E[\beta_r|\beta_l^*]$ is higher (but still lower than ρ), voters are more keen on l, and when ζ is higher, their vote is more certain in this direction, implying that l is more likely to be elected. The more likely l is to be elected in a polarisation phase, the longer is the polarisation phase on average, as it only ends once r is elected for the first time. Note that a longer phase of polarisation (when $\sigma \to 0$) is beneficial for voters as it implies that on average, the correct policy l is implemented more often throughout the cycle.

We have so far abstracted away from exogenous shocks as we have assumed above that $\sigma \to 0$. Naturally these have an important role to play in the dynamics of politics. Our next result establishes that the cycles uncovered above also feature when $\sigma > 0$ and K is large enough:

Proposition 3: *For a large enough K, the polity almost surely experiences cycles of polarisation and consensus. In particular: (i) there exists an $\eta_K > 0$ such that for any $\sigma > 0$, $\liminf_{t\to\infty} \eta_t(\text{polarisation}) > \eta_K$ and $\liminf_{t\to\infty} \eta_t(\text{consensus}) > \eta_K$ almost surely. (ii) There is a strictly positive probability that when the polity is in consensus it is on the wrong policy.*

By comparing the result above to the one in Proposition 2, note that the stochastic element of exogenous outcome shocks affects the nature of cycles in the following ways. First, a large level of noise implies that during a polarisation phase voters may not necessarily learn the true optimal outcome. This implies that parties might reach a consensus on the wrong policy. For example, a finite series of positive shocks when implementing policy r and/or negative shocks when implementing l might convince voters that policy r is substantially better than policy l. While this series of shocks might not happen too frequently, it can arise with a strictly positive probability.

Second, during a polarisation phase, it may take longer to converge to some knowledge (correct or not) that some policy is better than the other. This implies that polarisation phases might be longer and involve more political turnover. Moreover, large policy shocks can trigger abrupt switches from party polarisation to party consensus and vice versa. A large one-off shock during a phase of polarisation can convince voters that policy p is very beneficial and can induce a consensus on p, whereas in such a consensus phase, another shock can dissuade voters from implementing p altogether.

To illustrate further how the systemic element of the cycles works in conjunction with the exogenous shocks to affect the length of the different phases of the cycles, we now present results from simulations of the model. These also allow us to shed some light on the normative implications of the model (e.g., with regard to how often the optimal policy is implemented). For the simulation, we use $B = [0, 6]^2$ with a uniform prior. We set $\beta_l^* = 3.5$ and $\beta_r^* = 2.5$ (so the optimal policy is l), and $\zeta = \alpha = 0.5$ so that $\rho = \frac{2}{3}$. The parameters are chosen to satisfy Assumption 1. We randomise the initial history H_0 and then run the simulation for a hundred periods.

The table below reports positive and normative implications of the model using averages of 20,000 such simulations. For each vector of values for σ and K we measure the average length of a consensus phase, the fraction among the periods of consensus on which consensus is on the optimal policy, and the proportion of time the polity implements the optimal policy (either in a consensus or a polarisation phase).

The variables σ and K affect the results in similar ways as they are both related to how much the voters can learn; when σ decreases, learning is faster and so for any K, the polity is able to implement the optimal policy in more periods and also a greater proportion of consensus periods are on the true optimal policy. As K increases, the polity has a longer memory to learn from, implying similar results. As we can see, the length of the consensus phase is decreasing in σ and gets closer to its limit of K as σ gets smaller, in line with Proposition 2.

Table 1: Empirical Moments from Simulation

	$K=5$			$K=10$		
	$\sigma=0.2$	$\sigma=1.2$	$\sigma=2.5$	$\sigma=0.2$	$\sigma=1.2$	$\sigma=2.5$
Optimal Policy	80.22%	57.51%	48.25%	87.76%	65.34%	53.14%
	(3.31%)	(10.12%)	(9.53%)	(3.20%)	(14.68%)	(14.74%)
Consensus	73.78%	59.12%	61.71%	83.96%	64.95%	63.08%
	(4.05%)	(8.87%)	(6.99%)	(3.85%)	(13.28%)	(11.05%)
Consensus on Optimal Policy	100.00%	88.88%	72.08%	100.00%	93.11%	77.51%
	(0.05%)	(6.92%)	(10.72%)	(0.04%)	(8.31%)	(14.06%)
Length of Consensus Phases	4.69	4.31	3.98	9.22	7.43	6.08
	(0.26)	(0.92)	(0.89)	(0.56)	(2.17)	(2.05)

While, as shown above, both a high K and a small σ increase the information voters have, the comparative statics of these parameters in the limit are more nuanced. Specifically we now explore the cycle outcomes when $\sigma > 0$ but $K \to \infty$:

Proposition 4: *Let $\sigma > 0$. As $K \to \infty$, almost surely:*
(i) Consensus arises, i.e. $\lim_{K\to\infty} \limsup_{t\to\infty} \frac{\eta_t(\textit{polarisation})}{\eta_t(\textit{consensus})} = 0$.
(ii) The wrong policy is implemented in a strictly positive fraction of time, i.e., $\lim_{K\to\infty} \liminf_{t\to\infty} \hat{\eta}_t(r)$ *is bounded from below.*

The result indicates that while cycles always exist, the share of time that the polity spends in a polarisation phase shrinks as $K \to \infty$. But, as in Proposition 1, due to the myopia of voters and parties, consensus may arise on the wrong outcome when $\sigma > 0$. This arises no matter how high K is and in contrast to Proposition 2. Thus, while both a higher K and a lower σ imply better learning on average, they affect learning and hence outcomes in different ways.

Proposition 4 together with the simulations illustrated some of the normative properties of the outcomes of the cycle in terms of the share of time that the optimal policy is implemented. We now elaborate more on the welfare properties of short-term memory.

4 Short-term memory, changing states, and welfare

In this Section we evaluate the welfare properties of a political system with voters who have short-term memory. Our welfare criteria is long-run information aggregation, i.e., the long-run share of time in which the implemented policy accords with the optimal policy. We consider the welfare cost of short-term memory (compared to a polity with $K = \infty$), and then show that, in some environments, short-term memory can actually be beneficial. We focus throughout on the case of $\sigma \to 0$, which implies in our model that a polity with a long-term memory converges to implement the optimal policy almost surely.

To see the cost of short-term memory as compared to a long-term memory, recall that as we had seen in Proposition 2, the optimal policy is implemented for at least $1 - \frac{1}{K}$ of the time. The cost of short-term memory is that every now and then polarisation arises and the polity necessarily implements the suboptimal policy r. However, if short-term memory is not too short, the share of time in which this happens is small.

Short-term memory may however have benefits in some environments. Specifically, forgetfulness can be useful if the state of the world might change over time. Indeed sometimes polities might experience a change in technology which can be triggered by both external factors (e.g., a war, a pandemic) as well as by endogenous factors (e.g., climate change and technological breakthroughs). Remember that voters are myopic in our model and so do not intentionally experiment. But with short-term memory the polity drifts into polarisation phases now and then which, when the state of the world changes, might serve as periods of "unintended experimentation". This can allow the polity to detect changes in the state. Below we identify environments in which voters with short-term memory and who are not aware that the state of the world may change, may gain a higher welfare compared with voters with unbounded memory that are aware that the state of the world may change.

To illustrate this, consider an environment in which the state of the world $\boldsymbol{\beta}^*$ might change over time. Some changes to the fundamentals are easily detected, but some changes may remain undetected at least for a while. For example, if the polity is currently in consensus on l, a change to β_r^* -which may imply that now r is the optimal policy- will not be detected. Voters with short-term memory will detect these changes once they reach a polarisation phase and will switch to the new optimal policy. How will voters with long-term memory fare in this environment? As they are aware of such possible changes to the state they will have a belief that a change has occurred before time t; denote it by $\phi(t)$. Using this, voters can compute the expected utility difference between policies l and r, which generally can be written as:

$$\Delta(t) \equiv \phi(t) E_{t,post-change}[\beta_l - \beta_r] + (1 - \phi(t)) E_{t,pre-change}[\beta_l - \beta_r].$$

By Lemma 1, for any t such that $\Delta(t) > \rho$, parties will remain in consensus on l, and after the first t' for which $\Delta(t') < \rho$, parties will polarise. Once parties polarise, the polity will eventually detect the change and will return to implementing an optimal policy.

Once a change does occur and goes undetected as above, the question is how quickly consensus unravels to allow voters to learn the new optimal policy. As can be seen from the expression $\Delta(t)$, the timing of the unravelling of consensus does not depend only on the level of $\phi(t)$ but also on the expectations of $(\beta_l - \beta_r)$ before and after a change; this may imply that voters may be stuck on the wrong consensus for a very long time even though they believe that there is a high probability that a change in state has already occurred. Example 1 formalises the intuition above:

Example 1 (A simple model with changing states): Assume that there are only two possible states of the world, a and b, with $\boldsymbol{\beta}^a$ and $\boldsymbol{\beta}^b$, where $\beta_l^{a*} - \beta_r^{a*} > \rho$ and $\beta_r^{b*} - \beta_l^{b*} > \rho$. The prior, $\pi \in (0,1)$, is the probability that the state is a. In any period, with Poisson arrival rate λ, a new state is drawn from the prior distribution. Assume further that $\beta_l^{a*} = \beta_l^{b*} = \beta_l^*$; this will imply that a change in the state cannot be observed when policy l is implemented (but can be detected if r is implemented). In the spirit of Assumption 1 assume that $0 < \beta_l^* - (\pi\beta_r^{a*} + (1-\pi)\beta_r^{b*}) < \rho$. This insures that when voters are certain that there was a change, parties will polarise. We then have:

Proposition 5: *Consider Example 1. Let* $\sigma \to 0$ and $K > \frac{1}{1-\pi}$. *Then, when* $\lambda \to 0$, *there are values of* $\{\beta_l^*, \beta_r^{a*}, \beta_r^{b*}\}$ *for which voters with short-term memory (who are not aware of the possibility of changing states) will implement the optimal policy for a higher fraction of time compared to voters with long-term memory (who are aware that the state may change).*

Intuitively, when voters have long-term memory their myopia implies that they do not engage in intentional experimentation; considering their current payoff, even a strong possibility that a change of state had occurred may still not lead to polarisation. This is because, in expectation, they may view the current consensus policy as still optimal. In contrast, voters with short-term memory will always unintentionally experiment as they forget why they have settled on the current consensus policy.

A more general analysis of changing states may yield other implications. For example, if naive voters in our model do realise that the state of the world can change, they can attach less weight to periods in which they believe the state was less "similar" in some sense to the current one, or alternatively, completely discard historical information which is not relevant to the current period. To a degree, this implies that even when voters have full length memory, then the possibility of a changing state may imply that short-term memory can arise endogenously. We leave this for future research.

5 Extension: Different memories

The recollection of history might be different for different voters or groups. One example of this is highlighted in a recent literature which suggests that voters' beliefs are shaped by experiences mostly accumulated during their formative years. Malmendier and Nagel

(2016) show that life-time experiences of inflation significantly affect beliefs about future inflation, and that this channel explains the substantial disagreement between young and old individuals in periods of highly volatile inflation, such as the 1970s. Alternatively, a large literature studies and documents the effect of echo chambers on polarisation, and so different groups that are exposed to different sources of information can end up with different memories and recollection of histories.

As we show below our model is well equipped to analyse the dynamics of political systems when different voter groups have different memories. To see this, we adjust the model above and assume that the voting population is divided into m groups. All the voters in group j observe the same data which constitutes their memory. In period t, the history of each group j is the pairs of policies and outcomes in some selected history with K^j periods, denoted by H_t^j. This allows the voters in group j to compute their posterior distribution G_t^j on the vector $\boldsymbol{\beta}$ in a similar fashion to our model above.

Consider the electoral competition between the two parties at some period in which the groups are exogenously endowed with their respective histories. Below we will omit the subscript t when no confusion arises. We assume that within group j and across the groups, voters may differ on additional dimensions. Thus, voter i in group j votes for party L if:

$$E[y(p^L) - y(p^R)|H^j] + \upsilon^{ij} + \phi > 0,$$

where ϕ is the aggregate shock uniformly distributed on $[-\frac{1}{2\zeta}, \frac{1}{2\zeta}]$ and υ^{ij} is an idiosyncratic group-specific shock distributed on $[-\frac{1}{2\xi^j}, \frac{1}{2\xi^j}]$. The indifferent voter in each group satisfies:

$$\hat{\upsilon}^{ij} = E[y(p^L) - y(p^R)|H^j] + \phi$$

and so party $L's$ overall vote, given each group's share in the population γ^j, where $\sum_{j\in\{1,\ldots,m\}} \gamma^j = 1$, is:

$$\frac{1}{2} + \sum_{j\in\{1,\ldots,m\}} \gamma^j \xi^j [E[y(p^L) - y(p^R)|H^j]] + \phi]$$

Party L wins therefore if:

$$\sum_j \gamma^j \xi^j E[y(p^L) - y(p^R)|H^j] > -\phi \sum_j \gamma^j \xi^j$$

which arises with probability

$$\Pr(L|p^L, p^R) = \begin{cases} 1 \text{ if } \frac{1}{2} + \zeta \sum_j w^j E[y(p^L) - y(p^R)|H^j] > 1 \\ 0 \text{ if } \frac{1}{2} + \zeta \sum_j w^j E[y(p^L) - y(p^R)|H^j] < 0 \\ \frac{1}{2} + \zeta \sum_j w^j E[y(p^L) - y(p^R)|H^j] \text{ otherwise} \end{cases}$$

where $w^j = \frac{\gamma^j \xi^j}{\sum_j \gamma^j \xi^j}$ denotes the political weight of group j in the electoral competition. This weight is increasing in group size as is intuitive and decreases in the variance of the distribution of idiosyncratic shocks of group j. The latter effect is due to the fact that as the variance decreases, this group of voters is more sensitive to policy utility differences.

Anticipating the above, and with the knowledge of H_t^j, parties choose policies in equilibrium to maximise their expected utility as before; we then have an analogous result to Lemma 1:

Lemma 2: *In period t, a consensus on policy $p \in \{l, r\}$ arises if*

$$\sum_{j \in \{1,\dots,m\}} |w^j E[(\beta_p - \beta_{-p}) | H_t^j]| > \rho$$

and polarisation arises if for any policy $p \in \{l, r\}$,

$$\sum_{j \in \{1,\dots,m\}} |w^j E[(\beta_p - \beta_{-p}) | H_t^j]| < \rho$$

When a group has strong preferences towards one policy given their memory, consensus is more likely to arise when this group's political weight is larger; similarly, when a group has weak preferences for policy given their memory, it would be easier for parties to polarise if this group's political weight is larger.

To see the implications of Lemma 2, we now concentrate on a specific case of different memories. Assume that each group j remembers the recent K^j periods. Without loss of generality order groups in terms of the length of their history, i.e., $K^m > K^{m-1} > ... > K^1 \geqslant 2$. An example of such a polity could be age groups, so that older voters have a longer observable history.[9] We maintain Assumption 1 and consider generic distributions so that $w^j > 0$ and $w^j \neq w^{j'}$ for any j, j'. We also consider for simplicity $\sigma \to 0$:

Proposition 6: *The equilibrium in the model with heterogenous (and nested) length of memories is characterised by cycles of polarisation and consensus. Specifically, for all distribution of political weights,*

(i) There is a pivotal group $j^* \in \{1, ..., m\}$ such that the length of the consensus stage is K^{j^*}, the length of memory of j^*.

(ii) The expected length of the polarisation phase depends on the information of the groups.

[9] In Levy and Razin (2023) we study an overlapping generations model and focus on the generational divide in terms of the information of different cohorts of voters. We establish a similar cyclical pattern to that we uncovered in Propositions 2 and 3 in which the polity cycles between periods of polarisation and consensus. For recent literature on polarisation and extremism in different demographic groups see Ortoleva and Snowberg (2015) and Boxell, Gentzkow and Shapiro (2017).

When policy noise is vanishing, any group's knowledge centres on either $|\beta_p^* - E(\beta_p|\beta_{-p}^*)|$ or on $\beta_l^* - \beta_r^*$ as in Proposition 2. Consider just two groups, 1 and 2. Assume a consensus phase, let's say on l, and consider the polity after K^1 such periods. Group 1's knowledge has just switched from $\beta_l^* - \beta_r^* > \rho$ to $|\beta_l^* - E(\beta_r|\beta_l^*)| < \rho$, whereas group 2's knowledge constitutes of more periods and is therefore still at $\beta_l^* - \beta_r^* > \rho$. From Lemma 2 we know that if group 1's political weight is sufficiently large, parties will now polarise and this group will be the pivotal one. If its political weight is too small, parties will be in consensus for K^2 periods, that is, until the knowledge of group 2 also switches from $\beta_l^* - \beta_r^*$ to $|\beta_l^* - E(\beta_r|\beta_l^*)|$. The Proposition shows more generally that there is always one unique such pivotal group.

While the length of the consensus phase only depends on the memory capacity of the pivotal group, the length of the polarisation phase depends on the length of the memory of all groups. To see why, note that once parties polarise, the polarisation phase ends once the two policies had been implemented one after the other. The probability of each policy to be elected depends on $\sum_{j\in\{1,\ldots,m\}} |w^j E[(\beta_p - \beta_{-p})|H_t^j]|$, and so on the knowledge of all groups.

To illustrate what this implies, consider again the case of two groups. Suppose that group 1 is the pivotal one, and now increase K^2 to infinity. As we do this we do not change the length of the consensus phase in equilibrium which is exactly K^1. But by increasing K^2 we guarantee that group 2 will almost surely know the state at any period. This will make group 2 quite "frustrated" about politics as they know what should be done, but because of their de facto weak political power they are not able to "save" the polity from choosing r every now and then in the polarisation phase. Still, all is not lost as group $2's$ superior information affects the outcome even though it does not affect the strategies of parties; specifically, group $2's$ knowledge increases the probability that the correct policy l wins the election in a polarisation phase. While this means that the polarisation phase is longer, this is welfare enhancing: It will take longer to sample policy r during a polarisation phase and so the correct policy l will be implemented for longer in expectation.

Note that the above illustrates that it can be welfare-improving to have heterogeneity of groups' memories. If we start with a society where memory length is homogenous at some level K, all in society can better-off if some parts of society will have less memory and some more. Assume that we create such heterogeneity while maintaining the pivotal group as one with memory length K. As policy noise is vanishing, this implies that the length of the consensus phase is as before, while the length of the polarisation phase increases. This is welfare-improving as then policy l is implemented in equilibrium more often in expectations.

6 Discussion

In this section we discuss our results and assumptions and suggest some extensions of the main model.

6.1 Rationally-Inattentive Voters

In our model, voters' learning is characterised by one key behavioural feature, which is short-term memory. Beyond that, voters are fully attentive to policies and political outcomes, and update their beliefs rationally given the information they have. However, it is reasonable to consider the case that voters are not always fully attentive to politics. And in the context of our model, it is reasonable that voters' attention may differ in times of consensus and in times of polarisation: In times of political polarisation voters might pay more attention in order to understand whom to vote for; the benefits of doing so are more apparent as parties offer different policies. In periods in which parties are in consensus, voters might pay little attention to politics as they see little difference between the parties and will therefore be ill informed. One way to think about this as to consider news content providers that in times of consensus may focus on entertainment news to attract readers, while in times of polarisation, due to voters' heightened demand for information, may put more emphasis on political content.

For simplicity, suppose that we extend our model and assume that in times of consensus voters do not observe the political outcome y, while in time of polarisation all remains as before. Following a consensus on some policy, after K periods, voters' history is null as voters did not pay attention to any outcome in these K periods. Thus the only information voters have is the prior G. As G is symmetric, $E(\beta_l - \beta_r) = 0$, and parties will polarise. Once a polarisation phase starts, at some point voters will acquire sufficient knowledge that will induce a consensus, as before. In such a model, an analogous assumption to Assumption 1 facilitates cycles. Thus, adding rational inattention to the model does not change the qualitative nature of our results.[10]

6.2 Non-Myopic Voters and Parties

In our model voters and parties are fully myopic. The role of myopia was highlighted in Section 4. We now discuss the possibility of having preferences with longer horizons.

First note that the effect of preferences with long-term horizons are potentially restricted due to the short-term memory of future actors. This implies that current voters and parties

[10] We thank a referee for suggesting this alternative version of the model.

can only directly influence a finite number of periods in the future. The only effect that they can have on longer horizons is through the indirect effects of their behaviour today, such as changing the pattern and timing of future cycles.

Consider voters first. If voters are not myopic, they may have some incentive to actively experiment and especially so when they are least informed. This can happen during a phase of consensus in which, as time goes by, voters' information about the difference between the two policies becomes weaker. To see this, consider $\sigma \to 0$. Assume voters' history contains only instances of policy l being implemented. At this point voters compare the current period's benefit of voting for l which is $\beta_l^* - E[\beta_r|\beta_l^*]$ to the future benefit of learning what is the true state tomorrow. As long as $\beta_l^* - E[\beta_r|\beta_l^*] > 0$ our results are robust to voters having a small discount factor. For higher discount factors voters might be tempted to experiment by preferring to vote for policy r; if this motivation is very strong, parties may then switch to both offering r and so sometimes, rather than polarisation, we may have a switch to a different consensus.

For parties, myopia is a standard assumption in models of electoral competition: From their perspective, the stakes are so high that it is reasonable that they concentrate on winning the current election. If parties are not myopic, there are a few things to consider. First, they might want to affect the learning of voters in the future. One example could be that during periods when voters are quite sure that l is the optimal policy, party R might be tempted to offer policy r, rather than offer l. Party R may be elected with some small probability and potentially change the beliefs of voters. Thus, polarisation may be hastened. When parties have longer horizons we might also consider collusion between parties in the long run. This analysis is beyond the scope of this paper and left for future research.

6.3 Continuous Policies

We have used a discrete model with two policies for simplicity. When policies are continuous, it is possible that parties do not fully converge to offer the same policy (unless their office or ego rents, measured by α, are sufficiently large). However, even if they do not fully converge, the same mechanisms that drive our results are still in play with continuous policies and the polity can cycle between periods of substantial party polarisation and periods of low party polarisation.

Let us consider a simple example of continuous policies to illustrate one key mechanism that is behind our results. Suppose that parties choose policies $p \in [0, 1]$, and then the outcome at period t from a policy p is given by:

$$y_t = (1 - p)\beta_l^* + p\beta_r^* + \varepsilon_t$$

The utilities of party L and R from policy $p \in \{l, r\}$, $U^R(p)$ and $U^L(p)$, are now:

$$U^R(p) = p, U^L(p) = 1 - p$$

At period t, we measure polarisation by the distance between parties platforms, i.e., $|p_t^R - p_t^L|$. Everything else remains as in our main model. It is easy to derive from the above that the probability voters vote for L increase the higher is:

$$(p_t^L - p_t^R) E[\beta_r - \beta_l | H_t].$$

From this it can already be seen, as in Lemma 1, that the harder voters find it to substantially differentiate between β_r and β_l, i.e., the lower is $|E[\beta_r - \beta_l | H_t]|$, the more their vote will be based on the random shocks, ε, allowing parties to polarise more. Thus, less informed voters allow parties to substantially polarise and more informed voters induce less polarisation.

To complete the intuition for cycles in our model we also need to show that the political equilibrium affects voters' knowledge. In particular, that a larger degree of parties polarisation $|p_t^R - p_t^L|$ will imply that voters become more informed whereas a small distance between p_t^R and p_t^L will imply that voters are less informed. In our discrete model this relation is simplified as consensus means a complete lack of variation in the data. In a model in which there is always some degree of polarisation (but the degree of polarisation changes), we can show a similar result.

For illustration, suppose that voters observe a political outcome in one period t^R in which some p^R is implemented, and the outcome in one period t^L in which some p^L is implemented. Then they can use

$$\begin{aligned} \Delta y &= y_{t^L}(p^L) - y_{t^L}(p^R) = (p^L - p^R)(\beta_r^* - \beta_l^*) + \varepsilon_{t^L} - \varepsilon_{t^R} \\ &= (p^L - p^R)\Delta\boldsymbol{\beta}^* + \Delta\varepsilon \end{aligned}$$

to form beliefs on $\Delta\boldsymbol{\beta}^* \equiv (\beta_r^* - \beta_l^*)$. Their signal extraction problem is about trying to filter out the stochastic effects of the shocks ε; but note that these shocks obfuscate the true parameters more when $(p^L - p^R)$ is smaller. As a result, periods of low polarisation will involve less learning than periods of high polarisation.

7 Related literature

Our paper contributes to the current literature that focuses on the polarisation of politics in recent decades (for a recent example see Callander and Carbajal 2022). In particular, the analysis shines a light on an inherent feature of democratic political systems that implies the recurrence of polarisation phases. In this way we complement other theories that have focused on more current trends as explanations for the recent polarisation in politics.

Previous literature in political economy that analyzed political cycles focused on cycles between two types of policies rather than cycles of polarisation and consensus as we have here. For example, Battaglini and Coate (2008) show how policy making in legislatures can cycle between a regime in which debt is accumulated by over-redistributing at the expense of future budgets, and a regime in which policies maximize the collective good. Transitions between these regimes arise due to dynamic equilibrium considerations; an incumbent finds it optimal running deficits when it expects future incumbents to be more prudent and vice versa. The polity cycles between "good" and "bad" policies, and the phases of the cycle coincide with electoral cycles. In our analysis we have cycles of polarisation and consensus, and the cycle phases relate to voters' knowledge and last longer than election cycles.

We provide a theoretical model of politics as a process of collective learning. Hall (1993) surveys the literature in political science that views the political process as a learning endeavor. Piketty (2020) provides a historical overview of inequality regimes and ideologies in different countries through the prism of a collective learning process. Piketty (1995) analyses a model in which individuals learn about the true data generating process, but only from their own actions and thus there is no social-learning element. Strulovici (2010) and Messner and Polborn (2004) analyse group strategic experimentation and show that under-experimentation arises as individuals worry about losing their position as the median voter in society. Callander (2011) analyses a political social-learning model with a focus on the dynamics of learnings when the mapping between policies and outcomes is complex.

Several recent papers analyse collective learning in political processes, when voters are behavioural to some degree. In Callander, Izzo and Martin (2021) voters vote for parties according to their beliefs about the effect of a policy; these beliefs are derived in a non-Bayesian manner, as voters adopt one party's interpretation of historical data that has a higher likelihood. Little (2019) studies voter learning problems in which motivated reasoning distorts beliefs.[11] Levy, Razin and Young (2022) analyse a dynamic political social-learning model in which groups in society differ in their subjective model of the true data generating process and so one group has a misspecified model. In that model, power in society changes hands between a group that holds a complex view of the world to one that holds a simple world view. The reason is that perpetual rule by one group implies that the party in opposition becomes more intense in its preferences to win the election due to their subjective interpretation of the outcomes implemented by the ruling party. Eliaz and Spiegler (2020) and

[11] See also Little (2021) and Little, Schankenberg and Turner (2020) who show how motivated reasoning weakens politicians' accountability.

Eliaz, Spiegler and Galperti (2022) analyze political competition when voters are not able to understand the true correlation structure between a political action and a political outcome. Hence voters may consider "false correlations" in their models of the world between some non-relevant variable to the policy outcome. Different groups of politicians offer narratives (models) to voters where in equilibrium these models have to be consistent with the data generated when the winning groups implement policies according to them. A key result in both these papers is that in some cases, the (static) equilibrium outcomes are such that groups share power, as any distribution of outcomes generated by a model advocated by a particular group, can also be explained by another model advocated by another group.[12]

Closer to our focus on short-term memory, Eguia and Hu (2022) assume that voters have bounded memory, and strategically design what information (signals about the state of the world) to retain in their memory (a finite automaton, as in the work of Wilson 2014). They characterise environments in which such memory choices will imply that voters will always polarise even though they observe the same information. In contrast we assume that voters naively treat past history as exogenous. Jehiel and Newman (2010) and Bhaskar and Thomas (2019) analyze social learning with bounded memory. Acemoglu and Wolitzky (2014) analyse dynamic conflicts between groups with limited memory of previous history. A sufficiently long history of a conflict allows the groups to realize that a conflict has started by mistake, and revert to a coordination phase.

8 Conclusion

This paper investigates politics as a collective learning process. We demonstrate how short-term memory of the political players leads to cycles of polarisation and consensus in party platforms. In the model it is both the endogenous policies that are chosen and exogenous shocks that impact the observed outcomes and therefore provide information about optimal policies.

In practice there might be other political institutions both internal as well as external, that might provide other sources of information to the polity. Internally, political systems often have institutions such as the bureaucracy that have both longer horizons and perhaps a longer memory. In many countries tensions between these different political players, with different incentives and memories, have been rising in the last decade. An interesting extension of the analysis would be to understand how these different institutions affect the ability of the

[12] Azzimonti and Fernandez (2018) and Bohren and Hauser (2021) are two additional examples of social learning models in which convergence need not arise; in the former because of bots that provide misinformation that prevents learning, and in the latter due to individuals having misspecified models and hence not able to fully learn under some conditions.

polity to aggregate information and the dynamics of policy making.

Other information sources are more external. Countries can learn from the experiences of other countries. A good example is the Covid-19 pandemic in which both the public and politicians were able to observe and learn from the experiences of countries that were affected by the virus earlier. Similarly, in large countries or unions such as the USA, the Federal government can learn from policy experiments of member states, as exemplified by the concept of "Laboratories of democracies" proposed by Supreme Court justice Louis Brandeis in 1932 (In *New State Ice Co. v. Liebmann*, 285 U.S. 262, 1932). The regulation of industry in the US serves as another example, with state-level experiences influencing both legislation in other states and federal legislation (see Goldin and Libecap 2008). This paper serves as a starting point for further exploration of these issues and a deeper understanding of the study of politics as a collective learning process.

9 Appendix

9.1 Proofs for Section 2

Proof of Lemma 1: Assume that party L offers l. If party R offers l too it attains $\frac{1}{2}\alpha$, whereas if it switches to r it attains $(1 - \Pr(L \text{ wins}|l, r))(1 + \alpha)$, where

$$(1 - \Pr(L \text{ wins}|l, r)) = \begin{cases} 1 \text{ if } \frac{1}{2} + \zeta E[\beta_r - \beta_l|H_t] > 1 \\ 0 \text{ if } \frac{1}{2} + \zeta E[\beta_r - \beta_l|H_t] < 0 \\ \frac{1}{2} + \zeta E[\beta_r - \beta_l|H_t] \text{ otherwise} \end{cases}.$$

Note that if $\frac{1}{2} + \zeta E[\beta_r - \beta_l|H_t] < 0$, then party R indeed offers l, and when $\frac{1}{2} + \zeta E[\beta_r - \beta_l|H_t] > 1$, party R will best respond by offering r. When $\frac{1}{2} + \zeta E[\beta_r - \beta_l|H_t] \in (0, 1)$, then party R will offer l when $(\frac{1}{2} + \zeta E[\beta_r - \beta_l|H_t])(1 + \alpha) < \frac{1}{2}\alpha$, which amounts to $E[\beta_l - \beta_r|H_t] > \frac{1}{2\zeta(1+\alpha)}$. Given the above, whenever $E[\beta_l - \beta_r|H_t] > \frac{1}{2\zeta(1+\alpha)}$, party R offers l when party L offers l. Note that if this is the case, party L for sure offers l. An analogous condition, $E[\beta_r - \beta_l|H_t] > \frac{1}{2\zeta(1+\alpha)}$, guarantees that a consensus on r is the unique equilibrium. In all other cases, polarisation must arise as the unique equilibrium, that is, when

$$|E[\beta_l - \beta_r|H_t]| < \frac{1}{2\zeta(1 + \alpha)}.$$

In the non-generic cases in which $|E[\beta_l - \beta_r|H_t]| = \frac{1}{2\zeta(1+\alpha)}$ both polarisation and consensus on one of the policies will be an equilibrium. ■

We repeat here for convenience some of the notation defined in the text. Denote the expected outcome when policy p is implemented and given degenerate beliefs on some parameters $\boldsymbol{\beta}$, as $E[y|p, \boldsymbol{\beta}]$.

The random history (that arises given the randomness in the election and the randomness of the shock ε, through its effect on beliefs), induces a probability distribution P over the set of infinite paths of histories $\mathbb{H}$. Thus, when we write "almost surely" we mean P-almost surely on $\mathbb{H}$.

Remember that for full history up to time t we define the associated distribution over implemented actions at time t, $\hat{\eta}_t(p)$, as the share of time policy p was implemented, and we let $\eta_t(polarisation)$ and $\eta_t(consensus)$ be the fraction of time in the full history up to time t that the two parties offered different platforms and the same policy respectively.

9.2 Proofs for Section 3

Proof of Proposition 1: Voters' posterior after observing the history H_t satisfies the conditions of the martingale convergence theorem. Therefore, for almost any infinite path, voters' beliefs at any period t, μ_t, converge almost surely to some limit probability distribution μ_∞.

We now consider the measure one of all paths for which the posteriors converge. Consider first all paths for which, in the limit, $\left|E_{\mu_\infty}[\beta_l - \beta_r|H_t]\right| > \rho$. By Lemma 1, for these paths, parties will offer the same policy p, in line with the statement of Proposition 1.

Consider next the paths for which $\left|E_{\mu_\infty}[\beta_l - \beta_r|H_t]\right| < \rho$. By Lemma 1 party polarisation is the unique equilibrium in the limit on these paths. Assume by contradiction that the probability mass of this set of paths is strictly positive. Each of the parties is elected in equilibrium with a strictly positive probability due to $\alpha > 0$ and Lemma 1. We then have a strictly positive measure of paths for which $\lim_{t\to\infty} \inf \hat{\eta}_t(l)$ and $\lim_{t\to\infty} \inf \hat{\eta}_t(r)$ are bounded away from zero.

As the choice of policies in the model is endogenous and as they affect learning, this implies that the process of observed outcomes is not iid. We therefore cannot use standard laws of large numbers to pin down what are the limit beliefs. For this reason, we use below a result from Esponda, Pouzo and Yamamoto (2021), henceforth EPY.

Specifically, note that the Kullback-Leibler divergence at time t between a distribution of posteriors induced by some vector of parameters $\boldsymbol{\beta}$ and the posterior induced by the true parameters $\boldsymbol{\beta}^*$, given the fractions $\hat{\eta}_t(l)$ and $\hat{\eta}_t(r)$, is defined as:

$$KL(\boldsymbol{\beta}|\hat{\eta}_t, \boldsymbol{\beta}^*) = \sum_{p\in\{l,r\}} \hat{\eta}_t(p) \int_{\mathbb{R}} f(\varepsilon) \ln \frac{f(\varepsilon)}{f(E[y|p,\boldsymbol{\beta}^*] + \varepsilon - E[y|p,\boldsymbol{\beta}])} d\varepsilon$$

where $f(\varepsilon)$ is the density over ε, Normal with mean zero. The KL divergence value is always non-negative and the true state, $\boldsymbol{\beta}^*$, is a minimizer of the KL value attaining $KL(\boldsymbol{\beta}^*|\hat{\eta}_t, \boldsymbol{\beta}^*) = 0$ regardless of $\hat{\eta}_t$. But as $\hat{\eta}_t(l)$ and $\hat{\eta}_t(r)$ are bounded away from zero, this means that the true state, $\boldsymbol{\beta}^*$, satisfies $\boldsymbol{\beta}^* = \arg\min_{\boldsymbol{\beta}'} KL(\boldsymbol{\beta}'|\hat{\eta}_t, \boldsymbol{\beta}^*)$, i.e., is the unique such minimiser of the KL divergence. The result in EPY, in the context of our model, implies then that

$$\lim_{t\to\infty}\int_{B_t} KL(\boldsymbol{\beta}|\hat{\eta}_t,\boldsymbol{\beta}^*)d\mu_{t+1}(\boldsymbol{\beta}) = 0 \text{ almost surely.}$$

That is, for any observed frequency of actions, the posterior beliefs will concentrate on values of $\boldsymbol{\beta}$ for which $KL(\boldsymbol{\beta}|\hat{\eta}_t,\boldsymbol{\beta}^*)$ is closest to its minimized value, which is zero.[13] This result implies, by continuity, that if for $\boldsymbol{\beta}\in B$ the KL value is strictly positive, then a ball around $\boldsymbol{\beta}$ must have zero measure in the limit beliefs μ_∞. Thus, beliefs can only concentrate on a ball around $\boldsymbol{\beta}^*$. As $\beta_l^*-\beta_r^*>\rho$, this is in contradiction to the supposition that the beliefs converge to satisfy $\left|E_{\mu_\infty}[\beta_l-\beta_r|H_t]\right|<\rho$.

Finally, consider the infinite paths along which beliefs converge to satisfy $\left|E_{\mu_\infty}[\beta_l-\beta_r|H_t]\right|=\rho$. If either $\hat{\eta}_t(l)$ or $\hat{\eta}_t(r)$ converge to zero in any subsequence, implying consensus almost surely. If $\hat{\eta}_t(l)$ and $\hat{\eta}_t(r)$ are both bounded away from zero, by similar arguments to the argument above, beliefs must converge to satisfy $\left|E_{\mu_\infty}[\beta_l-\beta_r|H_t]\right|>\rho$, a contradiction. This concludes the proof of the proposition.■

The following Lemma will be helpful in the proof of Propositions 2 below and all results that pertain to the case of $\sigma\to 0$. In particular, it will imply that beliefs that arise in the equilibrium sequence as $\sigma\to 0$ converge to equilibrium beliefs when $\sigma=0$. Thus the sequence of equilibria as $\sigma\to 0$ will also converge to the limit equilibrium when $\sigma=0$.

Lemma A.2: *Assume that $K\geq 2$ and $\sigma\to 0$. (i) Suppose that there is a strictly positive measure of histories H_t such that only one policy p was implemented throughout the history. Then almost surely beliefs will concentrate on $(\beta_l^*,E(\beta_r|\beta_l^*))$ in period t. (ii) If there is a strictly positive measure of histories H_t such that both policy l and policy r were implemented, then almost surely beliefs will concentrate on (β_l^*,β_r^*) in period t.*

Proof of Lemma A.2: (i) Assume one policy p is implemented for K periods in a strictly positive measure of histories. Note that with the normal distribution over the shocks, for any $\gamma',\gamma''>0$ there is a $\bar{\sigma}>0$ such that for all $\sigma<\bar{\sigma}$ with probability $1-\gamma'$ all the shocks in the K periods are in $[-\gamma'',\gamma'']$. As $\sigma\to 0$ the distribution of shocks concentrates on its expectation. As a result, when $\sigma\to 0$, with probability arbitrarily close to one, the posterior belief after any path will be concentrated on $(\beta_p^*,E[\beta_{-p}|\beta_p^*])$. (ii) Assume that both l and r have been implemented in a strictly positive measure of histories of K periods. Again, for any $\gamma',\gamma''>0$ there is a $\bar{\sigma}>0$ such that for all $\sigma<\bar{\sigma}$ with probability $1-\gamma'$ all the shocks

[13] Our model satisfies assumptions 1-3 in EPY, which are all technical and relate to the compactness of B and continuity of the outcome function y. In EPY the policy function determining the mapping from beliefs to action is deterministic. In our model the action that is implemented at every period is random when parties polarise, given the shock ϕ_t. But this has no bearing on the proof of the result in EPY.

in the K periods are in $[-\gamma'', \gamma'']$. As a result, when $\sigma \to 0$, with probability arbitrarily close to one, the posterior belief after almost any path will be concentrated on $\boldsymbol{\beta}^*$.■

Proof of Proposition 2: Consider the limit when $\sigma = 0$. After histories H_t that contain two different implemented policies, parties will both offer the optimal policy l. Once there is a K-period history in which only this optimal policy is implemented, parties will polarise, and will continue to do so until two different policies are implemented at which point parties will revert to a consensus on policy l. As the consensus phase is on the correct policy l, the polarisation phase ends once r is selected, which happens with an interior probability $\frac{1}{2} - \zeta(\beta_l^* - E(\beta_r|\beta_l^*))$ at any period (this probability is interior as parties only polarise when each has a strictly positive probability of winning). This probability allows us to calculate the expected length of the polarisation phase and hence the share of time that the correct policy is implemented. As the limit equilibrium is unique, by Lemma A.2, the result also holds for $\sigma \to 0$.■

Analysis of equilibria when $\sigma \to 0$ and when β^* does not satisfy Assumption 1: The result below characterises the equilibria when $\sigma = 0$. We continue to assume that $\beta_l^* - \beta_r^* > \rho$, but now consider violations of Assumption 1.

Proposition A.2 *Let $\sigma = 0$ and $K \geq 2$. If Assumption 1 is violated, then the equilibrium is perpetual consensus or perpetual polarisation, unless $\{|\beta_l^* - E[\beta_r|\beta_l^*]| < \rho$ and $E[\beta_l|\beta_r^*] - \beta_r^* > \rho\}$ in which case the unique equilibrium is the same cycle as in Proposition 2.*

Proof of Proposition A.2:

If Assumption 1 is violated then there are three cases to consider:

1. $|\beta_l^* - E[\beta_r|\beta_l^*]| > \rho$ and $|\beta_r^* - E[\beta_l|\beta_r^*]| < \rho$.
2. $|\beta_l^* - E[\beta_r|\beta_l^*]| > \rho$ and $|\beta_r^* - E[\beta_l|\beta_r^*]| > \rho$.
3. $|\beta_l^* - E[\beta_r|\beta_l^*]| < \rho$ and $|\beta_r^* - E[\beta_l|\beta_r^*]| > \rho$.

Case 1: Remember that the optimal policy is l. Assume that only policy l was implemented in the observed history, then in the next stage the parties will be in consensus on policy l if $\beta_l^* > E[\beta_r|\beta_l^*]$ and on policy r if $\beta_l^* < E[\beta_r|\beta_l^*]$. In the former case, this will be an absorbing state of consensus. In the latter case, after one period of implementing r the polity immediately learns that the optimal policy is l. Parties will then be in consensus on l for K periods, and then switching to a consensus on r for one period, and so on. Thus in this case, we have perpetual consensus.

If the history is composed of observations of the two policies, again we move to a consensus phase in which l is implemented, which takes us to the analysis above and hence we reach perpetual consensus.

Assume now that only r was implemented in the observed history. Then, given that $|\beta_r^* - E[\beta_l|\beta_r^*]| < \rho$, parties polarise, and once l is implemented for the first time, voters learn that the optimal policy is l, implying that l will be implemented for K periods and as we have shown above, this will lead to perpetual consensus.

Case 2: In this case, whatever the history in terms of implemented policies, parties will be in consensus on some policy p as either learning the whole state or learning just one parameter implies a consensus. So in equilibrium we will have consensus forever after some period.

Case 3: If $\beta_r^* - E[\beta_l|\beta_r^*] > \rho$ consensus on r is an absorbing state, as a history in which voters only observed r leads to the beliefs above and parties will converge on offering r by Lemma 1. Depending on the initial history, and specifically if it does not contain only observations of r, an additional equilibrium that is identical to the one in Proposition 2 may arise. This equilibrium is sustained as along its path we never reach a consensus on the policy r. When $E[\beta_l|\beta_r^*] - \beta_r^* > \rho$ the unique equilibrium is the cycle we have in Proposition 2. If we start with an initial history in which only r was implemented and we hence attain these beliefs, in the next period parties will both espouse l and so we revert to the equilibrium cycle of Proposition 2.■

Example for "scale-free" learning discussed in Section 3.2:

To capture the limits of learning from β_p, let

$$\bar{\Delta}(\beta_p) = \sup_{p \in \{l,r\},\boldsymbol{\beta}} |\beta_p - E[\beta_{-p}|\beta_p]|$$

We construct a sequence of priors $\{G_n(\beta_l, \beta_r)\}_{n=1}^{\infty}$ for which in the limit $\bar{\Delta}_n(\beta_p)$ converges to zero for any β_p. This implies that for almost any state of the world $\beta_l^* - \beta_r^* > \bar{\Delta}(\beta_p) \approx 0$, and so Assumption 1 is satisfied whenever $\beta_l^* - \beta_r^* > \rho$. To see this, re-parameterise the distribution $G_n(\beta_l, \beta_r)$ into the parameter space (v, δ) where $v = \frac{\beta_l + \beta_r}{2}$ represents the scale, the mid-point between the utilities, and $\delta = \beta_l - \beta_r$ represents the utility difference. Let $\hat{G}_n(v, \delta) \equiv G_n(v + \frac{\delta}{2}, v - \frac{\delta}{2})$ be the transformed distribution function that satisfies independence between v and δ. Let $\hat{G}_2(\delta|v) = \hat{G}_2(\delta)$ be the marginal over δ, which we assume is independent of v and n and symmetric around zero so that $\int_{-\infty}^{\infty} \delta \hat{g}_2(\delta) d\delta = 0$. Let $\hat{G}_{1,n}(v)$ denote the marginal of this distribution over scale, and we assume that it is uniform on $[-D_n, D_n]$. In addition, we assume that for any v and x, $E[\delta|\delta > x]$ is finite. We now show that as $D_n \to \infty$, $\lim_{n\to\infty} \bar{\Delta}_n(\beta_p) = 0$ for any β_p.

For any β_l^* we have that,

$$\begin{aligned}
& \beta_l^* - E_{G_n(\beta_l,\beta_r)}[\beta_r|\beta_l^*] \\
= & \beta_l^* - \int_{v,\delta \text{ so that } v+\frac{\delta}{2}=\beta_l^*} (v-\frac{\delta}{2})\frac{\hat{g}_{1,n}(v)\hat{g}_2(\delta)}{\int_{v',\delta' \text{ so that } v'+\frac{\delta'}{2}=\beta_l^*}\hat{g}_{1,n}(v')\hat{g}_2(\delta')dv'd\delta'}dvd\delta \\
= & \beta_l^* - \int_{2(\beta_l^*-D_n)}^{2(\beta_l^*+D_n)} (\beta_l^*-\delta)\hat{g}_2(\delta)d\delta \\
\to & {}_{D_n\to\infty} \int_{-\infty}^{\infty} \delta\hat{g}_2(\delta)d\delta = 0\blacksquare
\end{aligned}$$

Proof of Proposition 3:

Step 1: *For a large enough K, there is no positive measure of paths along which there is a subsequence $\{t_n\}_{n=1}^{\infty}$ such that $\hat{\eta}_{t_n}(p) \to 1$ for some $p \in \{l,r\}$.*

Proof of Step 1: To see this, let us assume to the contrary that there exists such a subsequence t_n which on a strictly positive measure of paths satisfies that $\hat{\eta}_{t_n}(p) \to 1$ for some $p \in \{l,r\}$. For any t, denote the preceding K periods of history as the $K-window$ at t.

Claim A.2: *For a large enough K, along the subsequence $\{t_n\}_{n=1}^{\infty}$, after a $K-window$ in which only one policy $p \in \{l,r\}$ was implemented, almost surely the next period will involve parties polarising with a strictly positive probability.*

Proof of Claim A.2: Consider $t_n \to \infty$ and then a large enough K. Then for each $K-window$ with a fixed p, beliefs will concentrate on β_p^* and $E[\beta_{-p}|\beta_p^*]$ with a strictly positive probability. However, as $|\beta_p^* - E[\beta_{-p}|\beta_p^*]| < \rho$ and by Lemma 1, parties will polarise in the next period with strictly positive probability.$\square_{claim\ A.2}$

We can now use Claim A.2 to prove Step 1. As $\hat{\eta}_{t_n}(p) \to 1$, the fraction of these $K-windows$ with only p implemented within the window must be going to one. By Claim A.2, each of these will lead to polarisation with a strictly positive probability almost surely, and so $\eta_{t_n}(polarisation)$ is in the order of $\frac{1}{K}$. But as each party wins with strictly positive probability when there is polarisation, this contradicts the supposition that $\hat{\eta}_{t_n}(p) \to 1$.$\square_{step1}$

Step 2: *For a large enough K, $\liminf_{t\to\infty} \eta_t(polarisation) > 0$ almost surely.*

Proof of Step 2: Suppose not, and so there is a positive measure of paths along which there is a subsequence t_n such that $\eta_{t_n}(polarisation) \to 0$. This implies that if we look at the $K-windows$ along these paths almost all of them include no polarisation. Following from step 1, it cannot be that there is strictly positive measure of $K-windows$ with only one policy implemented as then we would have $\eta_{t_n}(polarisation)$ bounded from zero as we showed above.

Thus the only possibility that remains is that in almost all $K-windows$, at least two policies p and p' are implemented, and that parties will shift from a consensus on one policy p to a consensus on another policy p' (a "consensus-switch").

So assume that in almost all $K-windows$, at least two policies p and p' are implemented. Assume first that in all these K-*windows* the ratio of the share of time that p was implemented compared to the share of time that p' was implemented, converges to some finite $c > 0$. But then, as beliefs in almost all such K-*window* must converge to $\boldsymbol{\beta}^*$ when K grows large, after almost all such K-*window* both parties will choose the optimal policy l and so $\hat{\eta}_{t_n}(l) \to 1$, a contradiction to a finite c.

Thus we must have a strictly positive measure of $K - windows$ for which this ratio of implemented policies converges to zero or infinity. Let p' denote the policy implemented most times in the $K - windows$. Note that this ratio has to converge to infinity slow enough so that overall beliefs do not converge to $(\beta^*_{p'}, E[\beta_p|\beta^*_{p'}])$, as then we would have polarisation after such histories implying a contradiction to $\eta_{t_n}(polarisation) \to 0$.

Let us examine then what happens to $\beta^*_{p'} - E[\beta_p|\beta^*_{p'}, H_{t_n}]$, the beliefs attained for a large K, at a path where mostly p' is implemented. Note that for large K, $E[\beta_{p'}|H_{t_n}] - E[\beta_p|H_{t_n}]$ is arbitrarily close to $\beta^*_{p'} - E[\beta_p|\beta^*_{p'}, H_{t_n}]$. But as $\boldsymbol{\beta}^*$ satisfies Assumption 1, we have that $\left|\beta^*_{p'} - E[\beta_p|\beta^*_{p'}]\right| < \rho$. As we look at a strictly positive measure of paths, we can use iterated expectation to conclude that $E[\beta^*_{p'} - E[\beta_p|\beta^*_{p'}, H_{t_n}]] = \beta^*_{p'} - E[\beta_p|\beta^*_{p'}] \leq \left|\beta^*_{p'} - E[\beta_p|\beta^*_{p'}]\right| < \rho$. This implies that $\beta^*_{p'} - E[\beta_p|\beta^*_{p'}, H_{t_n}] < \rho$ with a strictly positive probability. As a result, for a strictly positive measure of paths we should have polarisation and hence a contradiction to $\eta_{t_n}(polarisation) \to 0.\square_{step\ 2}$

Step 3: *For a large enough K, $\lim_{t\to\infty} \sup_t \eta_t(polarisation) < 1$ almost surely.*

Suppose not, and so there is a positive measure of paths along which there is a subsequence t_n such that $\eta_{t_n}(polarisation) \to 1$. This implies that if we look at all the $K-windows$ almost all of them include polarisation at every period, implying that for all windows there exist at least two different policies p and p' implemented with a strictly positive probability. As a result, for a large enough K and $t_n \to \infty$, after almost all the $K - windows$ we have that, as in Proposition 1, beliefs almost surely concentrate on a ball around $\boldsymbol{\beta}^*$. This implies that both parties must choose the optimal policy after almost all these $K - windows$, a contradiction to $\eta_{t_n}(polarisation) \to 1.\square_{step\ 3}$

Step 4: *For a large enough K, for any $\sigma > 0$, there exists $\eta_K > 0$ such that*

$$\min\{\lim_{t\to\infty} \inf \eta_t(polarisation), \lim_{t\to\infty} \inf \eta_t(consensus)\} > \eta_K.$$

Suppose the statement is not true. Steps 2 and 3 imply that there exists a large enough K such that for any $\sigma > 0$, $\min\{\lim_{t\to\infty} \inf \eta_t(polarisation), \lim_{t\to\infty} \inf \eta_t(consensus)\} > 0$ almost surely. So for the statement to be wrong we must have that

$\lim_{\sigma\to 0}\min\{\lim_{t\to\infty}\inf\eta_t(polarisation),\lim_{t\to\infty}\inf\eta_t(consensus)\}=0$ with strictly positive probability. But in Proposition 2 we have shown that at $\sigma=0$,

$\min\{\lim_{t\to\infty}\inf\eta_t(polarisation),\lim_{t\to\infty}\inf\eta_t(consensus)\}>\frac{1}{K}$. Therefore, by continuity there must be an $\eta_K>0$ that satisfies the statement of Step 4. $\square_{step\ 4}$

The above concludes part (i). To consider part (ii) of the Proposition, note that consensus on policy p arises when:

$$\left|E[(\beta_p-\beta_{-p})|H_t]\right|>\rho.$$

As $K<\infty$ there is always a strictly positive probability that the above inequality arises for the wrong policy.■

Proof of Proposition 4: Assume that $\lim_{K\to\infty}\limsup_{t\to\infty}\frac{\eta_t(polarisation)}{\eta_t(consensus)}>\psi$ for some $\psi>0$, and so there is a strictly positive measure of paths for which for any convergent sequence $\{t_n\}$, $\lim_{K\to\infty}\lim_{t_n\to\infty}\frac{\hat{\eta}_{t_n}(p)}{\hat{\eta}_{t_n}(p')}\geq c$, for any p,p', where $c\in(0,\infty)$. As $K\to\infty$, by continuity, and using similar arguments as in Proposition 1, this implies that $\lim_{K\to\infty}\lim_{t_n\to\infty}E[\beta_l-\beta_r|H_{t_n}]\to\beta_l^*-\beta_r^*>\rho$, and so parties must converge in the long run on l almost surely, and so we must have $\lim_{K\to\infty}\liminf_{t_n\to\infty}\frac{\eta_\infty(polarisation)}{\eta_\infty(consensus)}=0$ almost surely, a contradiction.■

9.3 Proofs for Section 4

Proof of Proposition 5:

We prove three results that together suffice for the proof of Proposition 5.

Claim A.4.1 *Let $\sigma\to 0$. For small enough λ:*

(i) The equilibrium in the model with long-term memory follows the following phases:

(1) A phase of consensus on l that continues $\lfloor\bar{t}(\lambda)\rfloor$ periods, where $\bar{t}(\lambda)$ satisfies $\phi(\bar{t}(\lambda))=\varphi\equiv\frac{(\beta_l^{a*}-\beta_r^{a*})-\rho}{(1-\pi)(\beta_r^{b*}-\beta_r^{a*})}\in(0,1)$. At period $\lfloor\bar{t}(\lambda)\rfloor+1$ we move to phase (2).

(2) A phase of polarisation which lasts until r is implemented. After r is implemented, voters learn the current state. If voters learn that the state is a we move to phase (1). If voters learn that the state is b we move to phase (3).

(3) A phase of consensus on r until the state changes to state a at which point we revert to (1).

(ii) The long run proportion of time implementing the wrong policy can be arbitrarily close to $(1-\pi)$ as λ goes to zero by setting $\beta_l^{a*}-(\pi\beta_r^{a*}+(1-\pi)\beta_r^{b*})$ arbitrarily close to ρ.

Proof of Claim A.4.1: We begin by proving (i). Consider first the case in which voters are sure that the state is a at some point in time which we name time 0, and that there is a consensus on l for t periods after. Note that even though the voters cannot see a change of state in these periods, voters are aware that the state may have changed. They will then compute the probability that a change had occurred in this time frame of t periods. Specifically, the probability of a change at period t (and not before) is $(1-\lambda)^{t-1}\lambda$, and the probability that a state had changed before time t is $\phi(t) = 1-(1-\lambda)^t$. Note that the expected time in which a change actually occurs is $\frac{1}{\lambda}$.

Along this path, voters cannot be certain whether a change has occurred. To vote, the voters will compute the expected difference in outcomes next period, between implementing policy l rather than r:

$$\phi(t)(\beta_l^{a*} - (\pi\beta_r^{a*} + (1-\pi)\beta_r^{b*})) + (1-\phi(t))(\beta_l^{a*} - \beta_r^{a*}).$$

As long as this expression is higher than ρ then parties will remain in consensus on l but once this expression is smaller than ρ parties will polarise and subsequently voters will learn the true state once r is chosen.

As $\phi(t)$ is increasing, the expression above is decreasing, and we can calculate $\bar{t}$ (possibly a non-integer), which satisfies:

$$\phi(\bar{t})(\beta_l^{a*} - (\pi\beta_r^{a*} + (1-\pi)\beta_r^{b*})) + (1-\phi(\bar{t}))(\beta_l^{a*} - \beta_r^{a*}) = \rho$$
$$\Leftrightarrow \phi(\bar{t}) = \frac{(\beta_l^{a*} - \beta_r^{a*}) - \rho}{(1-\pi)(\beta_r^{b*} - \beta_r^{a*})} \equiv \varphi \in (0,1).$$

The last statement that $\varphi \in (0,1)$ follows from our assumptions on the parameters. Note that φ could be made arbitrarily close to one if $(\beta_l^{a*} - (\pi\beta_r^{a*} + (1-\pi)\beta_r^{b*}))$ is close to ρ (from below).

Thus, the equilibrium will have consensus up to the time $\lfloor \bar{t} \rfloor$. Computing the period this happens, $\lfloor \bar{t} \rfloor$, we get:

$$\begin{aligned} \phi(\bar{t}) &= 1-(1-\lambda)^{\bar{t}} = \varphi \\ &\rightarrow \bar{t}(\lambda) = \frac{\ln(1-\varphi)}{\ln(1-\lambda)}. \end{aligned}$$

Consider now that we are in a history in which there was consensus on r for some periods. At any period voters can learn the state. If the state is b the consensus on r continues, and if the state is a then parties switch to a consensus on l and the continuation follows as above. When the polity is in the polarisation phase, once policy r is implemented voters will know the state and we move to one of the consensus periods above. This concludes the proof of (i).

We now prove (ii): We focus on the relevant phase in which mistakes can happen, which is phase 1, on which there is a consensus on l. The expected time in which a change actually occurs is $\frac{1}{\lambda}$. To assess how much time passes until polarisation follows the expected change in the state, we look at $\lim_{\lambda\to 0} \frac{\bar{t}(\lambda)}{(1/\lambda)} = \lim_{\lambda\to 0} \frac{\frac{\ln(1-\varphi)}{\ln(1-\lambda)}}{(1/\lambda)} = -\ln(1-\varphi)$ where the last equality follows, as when $\lambda \to 0$ the expected time of change in state, $\frac{1}{\lambda}$, grows large at the same rate as $\frac{1}{-\ln(1-\lambda)}$. As we can take φ to be as close to one as we want (by choosing $(\beta_l^{a*} - (\pi\beta_r^{a*} + (1-\pi)\beta_r^{b*}))$ close to ρ) this implies that voters, even though they might be quite sure that a change has happened, may delay indefinitely their switching to a polarisation phase and continue to choose the wrong policy.

As the average number of periods between changes of the state is $1/\lambda$, it is instructive to count in terms of λ-periods which are blocks of $\lfloor 1/\lambda \rfloor$ periods. Note that whenever policy r is implemented, then, upon a change, with probability π state a is drawn and then detected, implying that the polity moves to implement l. Alternatively, when l is implemented, the polity will remain stuck on it for many λ-periods as $\lim_{\lambda \to 0} \frac{\bar{t}(\lambda)}{(1/\lambda)} = -\ln(1-\varphi) \to_{\varphi \to 1} \infty$. This implies that eventually the polity is stuck on l and so will implement the wrong decision whenever the state is b, which arises on average with probability $1-\pi$.■

We now consider voters with short-term memory. Since we assume that these voters are not aware that the state may change, we need to characterize how they update their beliefs. We can do this for strictly positive σ and characterise the beliefs as $\sigma \to 0$.

As $\lambda \to 0$ we only consider cases in which there was one change in the K period that the voters observe; the probability of two or more changes is negligible. Note further that by Lemma A.2, when $\sigma \to 0$, observed outcomes will concentrate on β_r^{b*} or on β_r^{a*} when implementing policy r. We now consider what happens to beliefs when voters observe both such outcomes in their history. Note that beliefs can change only when r is implemented as $\beta_l^{b*} = \beta_l^{a*}$.

Claim A.4.2: *Assume a K period history of play in which voters observe outcomes generated by policy* r *for* $K_r \leq K$ *periods in this history. For any* $\frac{|\beta_r^{b*} - \beta_r^{a*}|}{2} > \delta > 0$, *as* $\sigma \to 0$, *voters' posteriors are given by:*

(i) If a majority (minority) of outcomes out of the K_r *are in a* $\delta-$*ball around* β_r^{b*} *then the voters believe that the state is* b (a) *with probability converging to* 1.

(ii) If the number of periods out of the K_r *in which the outcome was in a* $\delta-$*ball around* β_r^{b*} *is equal to the number of periods out of the* K_r *in which the outcome was in a* $\delta-$*ball around* β_r^{a*} *then the voters beliefs converge to the prior.*

Proof of Claim A.4.2: If there are different states in these K_r periods, let us rename and order all observations (of $\{y_t\}_{t=1}^{t=K_r}$) under state b to be implemented at periods $t = 1$ to $t = T$, and all observations under state a to be implemented under renamed periods $t = T+1$

to $t = K_r$. This is without loss of generality. Then:

$$\Pr(\beta_r = \beta_r^{b*}) = \frac{(1-\pi)\Pi_{t=1}^{T} f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{b*})\Pi_{t=T+1}^{K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{b*})}{(1-\pi)\Pi_{t=1}^{T} f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{b*})\Pi_{t=T+1}^{K} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{b*})+\pi\Pi_{t=1}^{T} f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{a*})\Pi_{t=T+1}^{K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{a*})}$$
$$= \frac{1-\pi}{(1-\pi)+\pi\frac{\Pi_{t=1}^{T} f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{a*})\Pi_{t=T+1}^{K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{a*})}{\Pi_{t=1}^{T} f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{b*})\Pi_{T+1}^{K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{b*})}}$$

As $\sigma \to 0$, by the same arguments as in Lemma A2, $\lim_{\sigma\to 0} \frac{f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{a*})}{f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{b*})} = 1$ and $\lim_{\sigma\to 0} \frac{f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{b*})}{f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{a*})} = 1$ and so $\frac{\Pi_{t=1}^{T} f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{a*})\Pi_{t=T+1}^{K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{a*})}{\Pi_{t=1}^{T} f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{b*})\Pi_{t=T+1}^{K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{b*})} \to_{\sigma\to 0} 1$ if $T = K/2$, which implies that voters beliefs converge to the prior. If $T < K/2$,

$$\frac{\Pi_{t=1}^{T} f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{a*})\Pi_{t=T+1}^{K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{a*})}{\Pi_{t=1}^{T} f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{b*})\Pi_{t=T+1}^{K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{b*})} \underset{\sigma\to 0}{\to}$$
$$\frac{\Pi_{t=2T+1}^{t=K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{a*})}{\Pi_{t=2T+1}^{K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{b*})} \underset{\sigma\to 0}{\to} \infty$$

and so $\Pr(\beta_r = \beta_r^{b*}) \underset{\sigma\to 0}{\to} 0$ and similarly if $T > K_r/2$,

$$\frac{\Pi_{t=1}^{T} f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{a*})\Pi_{t=T+1}^{K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{a*})}{\Pi_{t=1}^{T} f_\sigma(\beta_r^{b*}+\varepsilon_t-\beta_r^{b*})\Pi_{t=T+1}^{K_r} f_\sigma(\beta_r^{a*}+\varepsilon_t-\beta_r^{b*})} \underset{\sigma\to 0}{\to} 0$$

and so $\Pr(\beta_r = \beta_r^{b*}) \underset{\sigma\to 0}{\to} 1.$ ■

Claim A.4.3:

When voters have short term memory, in the limit when $\sigma \to 0$, the long run fraction of time implementing the non-optimal policy is of the order of $\frac{1}{K}$, and the equilibrium has the following phases:

(i) A consensus phase of K periods with consensus on l, followed by phase (ii).

(ii) A polarisation phase until r is implemented. Once r is implemented, voters revert to phase (i) if the state is a, and move to phase (iii) if the state is b.

(iii) A phase of consensus on r, that ends with a switch to phase (i) $\lfloor \frac{K}{2} \rfloor + 1$ periods after the state changed to a.

Proof of Claim A.4.3:

Note that when voters know that the state is a, there is a consensus on l for K periods as no new information is generated when l is implemented. After K periods the voters have no knowledge about the state from the history and their belief about the state accords with the prior belief. A polarisation phase arises, and lasts until the first time r is implemented. When r is implemented the voters will learn the state immediately. If the state is a we go back to the consensus on l and such a consensus will last for K periods. If the state is b there will be consensus on r until the state changes to a and then, in accordance with Claim

A.4.2, the polity will revert to the consensus on l after $\lfloor K/2 \rfloor + 1$ periods (as $\lambda \to 0$ we consider only one change of policy per $\lambda-$*period*).

Note that the instances in which the polity makes a mistake are

1. There was a change of state to b within phase (i).

2. There was a change of state to b in phase (ii) when policy r hasn't been chosen yet.

3. The first time r is implemented in phase (ii) and the state is still a.

4. In phase (iii) when the state switches to a in the periods before the polity switches to phase (i).

When $\lambda \to 0$ the mistakes in (1), (2) and (4) are negligible in size as they happen only when a change of state occurs, on average once in any $\lambda - period$. However, mistakes in (3) happen in any phase (ii) which is recurring multiple times in any $\lambda - period$. Still, the probability of mistakes in (3) is smaller than $\frac{1}{K}$. This concludes the proof of the claim.■

This concludes the proof of the proposition.■

9.4 Proofs for Section 6

Proof of Lemma 2: The proof follows that of Lemma 1 by substituting the expression derived in the text in Section 2 for $\Pr(L \text{ wins}|l, r)$.■

Proof of Proposition 6:

Let $\sum_{j \leq \hat{\jmath}} w^j = \mu(\hat{\jmath})$. Without loss of generality assume that $\beta_l^* - E[\beta_r|\beta_l^*]$ and $\beta_l^* - \beta_r^*$ are such that for any $\hat{\jmath} \in \{0, 1, ..., m\}$ we have $L(\hat{\jmath}) \equiv |\mu(\hat{\jmath})(\beta_l^* - E[\beta_r|\beta_l^*]) + (1-\mu(\hat{\jmath}))(\beta_l^* - \beta_r^*)| \neq \rho$. Note that $L(0) > \rho$ and $L(m) < \rho$. In addition, $L(\hat{\jmath})$ is either decreasing in $\hat{\jmath}$ or first decreasing and then increasing. This follows from the fact that $\mu(\hat{\jmath})(\beta_l^* - E[\beta_r|\beta_l^*]) + (1 - \mu(\hat{\jmath}))(\beta_l^* - \beta_r^*)$ is always decreasing in $\hat{\jmath}$, and it is either positive at $\hat{\jmath} = m$ or negative. In the latter case its absolute value $L(\hat{\jmath})$ is then decreasing and then increasing. This means that we can find a unique $j^* \in \{1, ..., m\}$ which is the solution to the following inequalities:

$$L(j^* - 1) > \rho > L(j^*)$$

Now consider the following limit equilibrium cycle (when $\sigma \to 0$):

(i) A consensus phase on l lasting exactly K^{j^*} periods after which we move to phase (ii) below.

(ii) A polarisation phase which continues until the party espousing r wins an election, in this case we go to phase (i).

To see that this is a unique equilibrium, note that along this cycle, in a consensus phase that lasted for $K < K^{j^*}$ periods, we will have by assumption 1, the equilibrium conjecture, and the definition of j^*:

$$
\begin{aligned}
& |\sum_{j \text{ such that } K^j < K^{j^*}} w^j(\beta_l^* - E[\beta_r|\beta_l^*]) + \sum_{j \text{ such that } K^{j^*} \geq K} w^j(\beta_l^* - \beta_r^*)| \\
= \;& L(j|j \leq j^* - 1) \geq L(j^* - 1) > \rho
\end{aligned}
$$

and so by Lemma 2 there will be consensus on l in the next election. After period K^{j^*} of the consensus phase we will have polarisation. As long as l is elected polarisation continues. The expected length of polarisation will depend on the probability that party L wins the election, as:

$$
\Pr(L|p^L, p^R) = \begin{cases} 1 \text{ if } \frac{1}{2} + \zeta \sum_j w^j E[(\beta_l - \beta_r)|H_t^j] > 1 \\ 0 \text{ if } \frac{1}{2} + \zeta \sum_j w^j E[(\beta_l - \beta_r)|H_t^j] < 0 \\ \frac{1}{2} + \zeta \sum_j w^j E[(\beta_l - \beta_r)|H_t^j] \text{ otherwise} \end{cases}
$$

Note that in this expression all groups' information affects the probability that L wins. As there is randomness in the election's outcome, this probability might change over time as $\sum_j w^j E[(\beta_l - \beta_r)|H_t^j]$ changes. In particular, as long as the polarisation phase lasts, more groups' histories will move from $E[(\beta_l - \beta_r)|H_t^j] = \beta_l^* - \beta_r^*$ to $E[(\beta_l - \beta_r)|H_t^j] = \beta_l^* - E[\beta_r|\beta_l^*]$ as their history progresses to contain only l being implemented.■

References

[1] Acemoglu, D. and A. Wolitzky (2014). "Cycles of Conflict: An Economic Model", American Economic Review 104: 1350-1367.

[2] Achen, C., and L. Bartels (2008). "Myopic Retrospection and Party Realignment in the Great Depression". mimeo.

[3] Angelucci, C. and A. Prat (2023). "Is Journalistic Truth Dead? Measuring How Informed Voters Are About Political News?", mimeo.

[4] Azzimonti, M. and M. Fenandes (2018). "Social Media Networks, Fake News, and Polarisation", NBER working paper 24462.

[5] Barber, M. and N. McCarty (2013). "The Causes and Consequences of polarisation". In *Negotiating Agreement in Politics,* eds. Cathie Jo Martin and Jane Mansbridge. APSA task force report.

[6] Bartels, L. (2008). *Unequal Democracy: The Political Economy of the New Gilded Age.* Princeton, NJ: Princeton University Press.

[7] Bhaskar, V. and C. Thomas (2019)."Community Enforcement of Trust with Bounded Memory", Review of Economic Studies 86: 1010-1032.

[8] Battaglini, M. and S. Coate (2008). "A Dynamic Theory of Public Spending, Taxation, and Debt". American Economic Review 98: 201-36.

[9] Bohren, A. and D. Hauser (2021). "Learning with Heterogeneous Misspecified Models: Characterization and Robustness". Econometrica 89: 3025-3077.

[10] Boxell, L., M. Gentzkow and J. Shapiro (2017). "Greater internet use is not associated with faster growth in political polarisation among US demographic groups", PNAS 114: 10612-10617.

[11] Callander, S. (2011). "Searching for good policies", American Political Science Review 105: 643–662.

[12] Callander, S., F. Izzo and G. Martin (2021). "Ideological Competition", mimeo.

[13] Callander, S. and J. Carbajal (2022). "Cause and Effect in Political Polarisation: A Dynamic Analysis", Journal of Political Economy 130: 825-880.

[14] Calvert, R. L. (1985). “Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence,” American Journal of Political Science 29: 69–95.

[15] Eguia, J. and T. Hu (2022). “Voter Polarisation and Extremism", working paper.

[16] Eliaz, K. and R. Spiegler (2020). “A Model of Competing Narratives", American Economic Review 110: 3786–3816.

[17] Eliaz, K., R. Spiegler and S. Galperti (2022). “False Narratives and Political Mobilization”. mimeo.

[18] Esponda, I., D. Pouzo and Y. Yamamoto (2021). “Asymptotic Behavior of Bayesian Learners with Misspecified Models", Journal of Economic Theory 195, Article 105260.

[19] Gerber, A. and D. Green (1998). “Rational Learning and Partisan Attitudes”. American Journal of Political Science 42: 794–818.

[20] Goldin, C. and G. Libecap (2008). *The regulated economy: A historical approach to political economy.* University of Chicago Press, Chicago IL.

[21] Hall, P.A. (1993). “Policy Paradigms, Social Learning, and the State: The Case of Economic Policymaking in Britain”, Comparative Politics 25: 275-296

[22] Healy, A. and G. Lenz (2014). “Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy”, American Journal of Political Science 58: 31-47.

[23] Jehiel, P. and A. Newman (2010). “Loopholes: Social Learning and the Evolution of Contract Form”. mimeo.

[24] Kramer, G. (1971). “Short Term Fluctuations in U.S. Voting Behavior, 1896-1914”. American Political Science Review 65: 131-143.

[25] Levy, G. and R. Razin (2023). “The Dynamics of Cohort Effects in Politics”. mimeo.

[26] Levy, G., R. Razin and A. Young (2022). “Misspecified Politics and the Recurrence of Populism”. American Economic Review 112: 928-62.

[27] Lenz, G. (2010). “(Mis)Remembering Incumbent Performance.” Presented at the annual meeting of the Midwest Political Science Association.

[28] Lindbeck, A. and J. Weibull (1987), “Balanced-Budget Redistribution as the Outcome of Political Competition”. Public Choice 52: 273-297.

[29] Little, A. (2019). “The Distortion of Related Beliefs”, American Journal of Political Science 63: 675-689.

[30] Little, A. (2021). “Directional motives and different priors are observationally equivalent”, mimeo.

[31] Little, A., K. Schnakenberg and I. Turner (2020). “Motivated Reasoning and Democratic Accountability”, mimeo.

[32] Malmendier, U. and S. Nagel (2016). “Learning from Inflation Expectations”, Quarterly Journal of Economics 131: 53-87.

[33] McCarty, Nolan. (2019). *Polarization: What Everyone Needs to Know.* Oxford University Press.

[34] Messner, M. and M. Polborn (2004). "Voting on Majority Rules". Review of Economic Studies 71: 115-132.

[35] Ortoleva, P. and E. Snowberg (2015). “Overconfidence in Political Behavior”. American Economic Review 105: 504-535.

[36] Peltzman, S. (1990). “How Efficient Is the Voting Market?”. The Journal of Law & Economics 33: 27-63.

[37] Piketty, T. (2020). *Capital and Ideology.* Harvard University Press, Cambridge.

[38] Piketty, T. (1995). “Social Mobility and Redistributive Politics”. The Quarterly Journal of Economics 110: 551-584.

[39] Rajan, G.R (2009). “The Credit Crisis and Cycle-Proof Regulation”. Federal Reserve Bank of St. Louis Review 91: 397-402.

[40] Strulovici, B. (2010). “Learning While Voting: Determinants of Collective Experimentation”, Econometrica 78: 933-971.

[41] Wilson, A. (2014). “Bounded Memory and Biases in Information Processing”. Econometrica 82: 2257-2294.

[42] Wittman, D. (1983). Candidate motivation: A synthesis of alternatives. American Political Science Review, 77:142-57