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

Gilat Levy and Ronny Razin, LSE[1]

Abstract: This paper investigates the influence 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 among political 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 involve sufficient variation that allows voters to be confident about what the correct policy is, forcing parties to both offer this policy. We analyse the emergence and the nature of such cycles and examine the impact of crises on political competition through the lens of our framework.

[1]Department of Economics, Houghton St. WC2A 2AE, UK. Email addresses: g.levy1@lse.ac.uk, r.razin@lse.ac.uk. For useful comments we thank Steve Callander, Andrew Little, and many seminar participants and discussants. Roberto Maura-Rivero and Enrico Turri provided helpful research assistance.

1 Introduction

Polarisation in democratic societies has garnered significant attention in recent literature in both economics and political science, with many attempting to understand the increase in polarisation observed in the past few decades.[2] However, political scientists have also noted that 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).

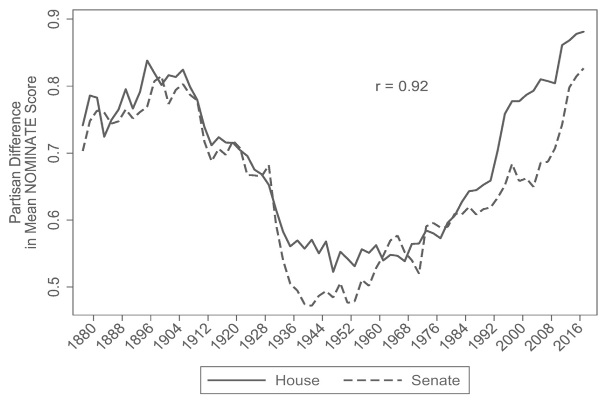

Figure 1: Historical polarisation in the US Senate and Congress (McCarty 2019).

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] By looking at the manifestos of the two major parties in the US, it is apparent that parties oscillate between polarisation and consensus on economic issues. For example, Figure 2 plots the positions of the Democratic and Republican parties on market regulation from 1948 to 2020, demonstrating periods of relative consensus (on deregulation) from the 1980s to the early 2000s, with relative polarisation occurring before and after these periods. It is interesting to note that these periods of consensus among parties seem to coincide with consensus among economists on market deregulation, which began to unravel in the early 2000s.[4] Similar cyclical patterns are also present in public opinion (see Stimson 1999).

[2] See for example the survey of Barber and McCarty (2013).

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

[4] For example, John Williamson's famous "Washington Consensus" paper about the standard (de)regulation reform package for developing countries was published in 1989, and Dani Rodrik's response, entitled "Good bye Washington Consensus, Hello Washington confusion?...", was published in 2006.

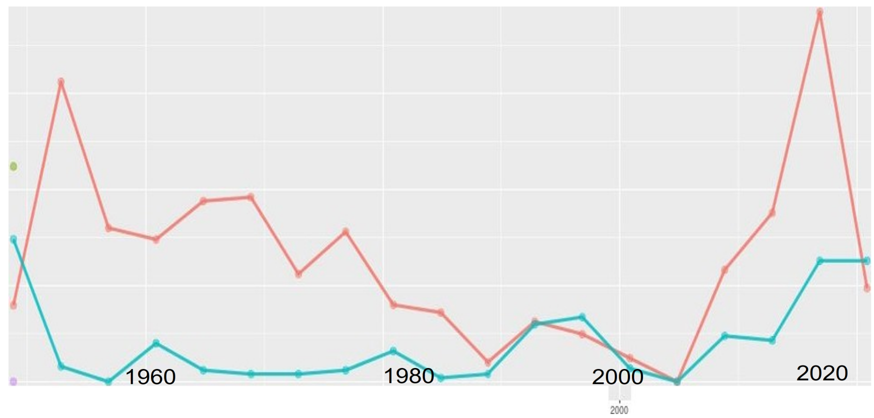

Figure 2: US Democratic (the higher curve) and Republican parties' regulation policies (where a higher level on the y axis indicates a higher support for regulation).

We investigate the question of whether cycles of polarisation and consensus, in party platforms and public opinion, 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 finite memory may be a systemic factor contributing to such cycles; specifically, when voters have short-term memory, periods of consensus often give way to periods of polarisation, and vice versa.

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). This conclusion is shared by other recent studies in political science (Achen and Bartels, 2008; Bartels, 2008; Gerber and Green, 1998; Lenz, 2010; Healy and Lenz, 2014). Additionally, short-termism can hinder learning from past mistakes. One example of repeated mistakes is the cyclical pattern of over-regulation and under-regulation in financial markets. For example, 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... We need to acknowledge these differences and enact cycle-proof regulation."

We incorporate the assumption of short-term memory into a dynamic 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 self-interests as well as voters' preferences. In this model beliefs and policies evolve together over time.

Our results build on two key mechanisms. The first is a political mechanism which implies how the level of voters' knowledge affects the incentives of ideological parties to polarise. When voters are uncertain about which policy is best, parties are more likely to advocate for their own, extreme positions. However, when voters have a stronger conviction about the superiority of a particular policy, both parties are "disciplined" to converge on this policy, as a deviation would elicit a negative response from the voting population.

The second mechanism we identify relates to how short-term memory affects the knowledge of voters when policies are endogenous. Consider first the case where voters' memory includes periods in which parties were polarised and so offered different policies. Such variation in policies (resulting from political turnover) provides voters with information about the effectiveness of different policy options, enabling them to identify the optimal policy. According to the first mechanism described above, parties will then be incentivized to reach a consensus on this policy. However, if consensus among parties persists for an extended period, voters' memories will eventually consist only of periods in which parties offer 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. Parties will then have an easier time convincing voters to try an alternative. 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 can give rise to cycles of polarisation and consensus in political systems. Specifically, we find that these cycles can 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 with little policy variation. 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 phases of the cycle result from the influence of both endogenous policies and exogenous shocks on voter learning. To understand the dynamics of these cycles, consider the case in which the variance of exogenous shocks to economic outcomes is very small. In this scenario, learning from endogenous policies occurs very quickly, as experiencing a particular policy for just a few periods is sufficient for voters to understand its effectiveness. As we show, such a system can oscillate between short periods of party polarisation and long periods of consensus. This result highlights the role of endogenous policies as a systemic force driving the emergence of cycles in political systems.

To give a vivid example of how society learns from implemented policies, it is instructive to look at public opinion over time. Stimson (1999) remarks about the cyclical nature of public opinion that: "...cycle(s) back and forth, left and right, as leaders and followers change their views of government policy over time." Shapiro and Gilroy (1984), write about public opinion of regulation in the US: "Attitudes and opinions are not immovable, but that they change slowly, without wild fluctuations (beyond empirical sampling error), and when they do change, they respond in reasonable ways to events and changes in objective conditions."

In fact, the history of public opinion and legislation on regulation in the US is a good example of the kind of dynamics that happen in the model we analyse. Indeed, economists have used variation in implemented policies to assess the effects of regulatory regimes and have in turn played a crucial role in shaping public opinion as well as future policies. Rose (2014) discusses the regulation of cable TV and observes that: "Cable provides a rich laboratory for economists in search of policy variation... deregulation, re-regulation, and deregulation once again in this sector." Juskow (2005), in a paper summarising 25 years of IO literature, remarks: "The extensive experience with deregulation in the last twenty-five years has created enormous opportunities both to re-examine what we thought we knew about the effects of regulation as well as to provide opportunities to examine the attributes of imperfectly competitive industries after they have been "shocked" by the relaxation or removal of price and entry constraints."

Voters' learning can also be affected by large exogenous shocks, such as natural disasters, wars, or pandemics. These shocks can have varying impacts on public opinion and government policies. For example, some shocks, such as WWII, have led to consensus on the welfare state and increased government investment, while others, like the financial crisis of 2008, have resulted in greater polarisation in public opinion and policy. In our model, transitions between phases of polarisation and consensus can also be triggered by exogenous shocks, as we demonstrate in Sections 3.2 and 3.3. For instance, high variance in shocks can increase the probability that consensus emerges on the wrong policy. Additionally, different types of shocks may have different effects on public opinion and future policies depending on the policy that is in place at the time they occur.

Other crises may be related to changes in the data generating process. The financial crisis of 2008, for example, was partially driven by new technologies in financial derivatives. In Section 3.4 we analyze the case of a changing data generating process and demonstrate how this too can give rise to cyclical behavior. Essentially, when society realises that the state of the world has changed, voters' historical knowledge becomes less valuable, leading parties to polarise their positions. However, as time passes, this polarisation can facilitate the emergence of a new consensus.

In Section 4 we extend the model to consider an environment in which 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. Section 5 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, we discuss this further below (after Proposition 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 $G(\beta_1, \beta_2)$ 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 $\boldsymbol{\beta}^*$, while in Section 3.4 we discuss the case of shifting environments (that is, when $\boldsymbol{\beta}^*$ changes over time).

The outcome y_t is a common element in the voters' preferences, and so 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 enjoy small 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. In the 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 $K \leqslant \infty$ 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: Voters are Bayesian updaters but treat the history H_t as exogenous. Together with the prior G they compute their posterior distribution G_t on the vector $\boldsymbol{\beta}$. 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 far away in the past decline fast enough with time. In sections 3.4 and 4 we discuss other alternative assumptions to short-term memory.

2.3 Electoral Competition

Polarisation can only arise in electoral competition when parties are ideological (as we assume) and when that they face some uncertainty in the election with regards to voting behaviour. To introduce uncertain voting behaviour we adopt the probabilistic voting model, following Lindbeck and Weibull (1987). 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 of the voting choice which parties may have uncertainty about. For example, voters may be affected by personal attributes of candidates, or by some other factors that influence the attachment of voters to particular parties. We summarise all these other factors by the 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. For the voters this means that when they chose who to vote for they do not take into account the experimentation value of voting for a policy. For the parties this assumption implies that in their choice of policy they do not think about manipulating voters' beliefs to affect their winning chances in the future. Myopia is a standard assumption in models of electoral competition. From the parties' perspective, the stakes are so high that it is reasonable that they concentrate on winning the current election. From the perspective of voters, the low incentives to vote and lack of sophistication are often the motivation behind assuming myopia (see also the discussion in section 3.4.3).

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

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

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,

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

where $U^J(p)$ is defined in $(II.2)$ and parties know G, H_t.

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 dynamic model is defined as follows:

1. There is some initial history H_0.

2. In period t, party $J \in \{L, R\}$, the party that won the election, 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.

Note that the only dynamic link between the periods is the evolving memory of voters, that changes in accordance with the new policy that is implemented and the respective outcome, y_t, that depends also on the shock, ε_t.

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 that voters observe at time t and consider the one-period political competition game that ensues. The result below is reminiscent of the result in Calvert (1985):

Lemma 1 (Consensus vs Polarisation): *In period t, if*

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

then both parties choose the same platform (consensus), and otherwise each party chooses its ideal policy (polarisation).

Generally speaking, parties prefer to pursue their own interests, but to successfully do so, must be elected with some probability. This implies that they may be disciplined by voters to choose the policy that is more likely to generate a higher outcome given the historical data. If the historical data makes voters sufficiently confident that one of the policies, say l, is more likely to generate a higher outcome, then parties have to offer this policy and hence reach a consensus on l. In this case, if a party offers r, it will neither serve their own policy interest nor their office motivation, as they will face only a slim probability of being elected. Alternatively, if the historical data does not sufficiently discriminate between the different policies, then parties can afford to offer platforms that better serve their own policy interests and hence polarise.

To see the role that other parameters play, note that a higher office rent α pushes parties to be in consensus, 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; if a party polarises and has only a negligible chance of being elected, it can deviate to be in consensus with the other party, gain $\frac{1}{2}\alpha$, and be better off.

A higher ζ implies a smaller interval for the shock ϕ_t : This means that parties are more certain about how voters vote, and again this pushes parties to be more in consensus. As both ζ and α work in the same direction to affect the possibility of consensus, let

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

be the parameter enabling consensus, so that a lower ρ is more conducive to consensus.

To make learning meaningful in the model, it is reasonable to consider the case where parties reach a consensus when voters know the true state. Given the Lemma above we will henceforth assume that $|\beta_l^* - \beta_r^*| > \rho$, that is, the true state is such that if the voter knew the state, parties will be bound to offer the optimal policy and consensus will arise.

3 Cycles of consensus and polarisation

In what follows we are interested in whether and how a polity transitions between periods of polarisation and consensus. To this end the following notation will be useful. Denote by $\hat{\eta}_t(p)$ the fraction of time in the history up to time t that policy p was implemented. In addition, denote by $\eta_t(polarisation)$ the fraction of time in the 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 history up to time t that the two parties offered the same platform. 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", 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 $K = \infty$. Our result shows that with full memory, the two parties will converge to offer the same platform.

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

To see the intuition for Proposition 1, note first that polarisation cannot be part of a long-term equilibrium. When $K = \infty$, the beliefs of the voters will converge in the long term, as they form a martingale. If they converge to beliefs that 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, as they observe long-term outcomes from each policy. 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 converge on the same platform. Thus, when $K = \infty$ voters' beliefs in the long-term must converge to beliefs that induce parties to reach a consensus and on one specific policy p in the long-term.

Note that consensus is *not* guaranteed to be on the optimal policy; as in any learning problem with agents that discount the future to some degree, 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 r is implemented, may convince voters that l is more likely to be better and therefore induce parties to converge on l. This convergence can be sustained in the long run; voters will learn β_l^* but might maintain beliefs that the difference between expected utilities of the two policies is large enough. Their beliefs on the difference are conditional on knowing β_l^*, and the series of very bad shocks when r was implemented.[5]

3.2 Short-term memory and cycles

We now turn to consider short-term memory, i.e., finite K. In this case voters will never fully learn the state of the world. Moreover, 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 a high political turnover and changes in policies, voters' data will be relatively informative.

We already know that if voters' history is not informative at all, then parties will polarise, as $E(|\beta_l - \beta_r|) = 0$. On the other hand, if voters' history has lots of variation and is very informative, parties will reach a consensus as $|\beta_l^* - \beta_r^*| > \rho$. What happens on equilibrium path with short-term memory will depend then on voters' knowledge and whether this knowledge sufficiently distinguishes between the expected utilities of the two policies. The evolution

[5]Note that $|\beta_l^* - \beta_r^*| > \rho$ is only a sufficient condition for long-term consensus, as, in line with the intuition above, consensus may arise even before β^* is learnt. In addition, as beliefs converge, cycles can never arise even when $|\beta_l^* - \beta_r^*| < \rho$. In this case the polity converges to either long term consensus or long term polarisation.

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 consider first the model in which the variance of the outcome shock, σ^2, is very small. In this case learning is very fast, and after any period, voters learns the true effectiveness of the policy that was implemented in that period. It is therefore enough to observe two periods with different implemented policies to learn everything there is to learn in the model.

Assume therefore that σ^2 is zero. Consider any $K \geq 2$ and the case in which voters' observed history contains only one implemented policy $p \in \{l, r\}$. In this case, voters learn the true value of the effectiveness of policy 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 knows β_p^* (and nothing else). Let $B(\rho) = \{\boldsymbol{\beta}| \max\{|\beta_l - E[\beta_r|\beta_l]|, |\beta_r - E[\beta_l|\beta_r]|\} < \rho\}$. The following proposition characterizes the equilibria in the model for $\sigma^2 \to 0$.

Proposition 2: *Let $\sigma^2 \to 0$ and $K \geq 2$. Then whenever $\boldsymbol{\beta}^* \in B(\rho)$, the polity experiences cycles of polarisation and consensus, with a consensus phase lasting exactly K periods in which the parties converge on the optimal policy. The polarisation phase lasts until two distinct policies are implemented, consecutively.*

When the variance of the shock is small, the state being in the set $B(\rho)$ is a sufficient condition for cycles to arise. Assume that the polity starts with both parties offering the same policy p. As σ^2 is vanishing, after any 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 $E[\beta_{-p}|\beta_p^*]$. As $\boldsymbol{\beta}^* \in B(\rho)$, parties will polarise. In this case, at every period, $-p$ will be implemented with some strictly positive probability. Once this arises for the first time, voters' memory will include two periods in which two different polices were implemented. As σ^2 is vanishing, their beliefs will concentrate on the true $\boldsymbol{\beta}^*$, implying that parties will now both offer the optimal policy. In the proposition we require that $K \geq 2$, as when $K = 1$ voters' history is too short to learn the true state, and when $\boldsymbol{\beta}^* \in B(\rho)$ polarisation is perpetual. Also, in the Appendix we show for completeness that if alternatively $\boldsymbol{\beta}^* \notin B(\rho)$, then the polity is either in perpetual consensus or experiences cycles exactly as stated in the Proposition.

We now turn to the case in which the variance of the shock is substantial. In this case the shocks to the outcome will influence learning, adding an additional stochastic element to the model. Below we show that $\beta^* \in B(\rho)$ is still a sufficient condition for cycles when K is large enough:

Proposition 3: *Suppose that $\beta^* \in B(\rho)$. For large enough K, the polity almost surely experiences cycles of polarisation and consensus. In particular: (i) there exists an $\eta > 0$ such that for any $\sigma^2 > 0$, $\liminf_{t\to\infty} \eta_t(polarisation) > \eta$ and $\liminf_{t\to\infty} \eta_t(consensus) > \eta$. (ii) In any period that the polity is in consensus, there is a strictly positive probability that the consensus is on the wrong policy.*

Consider first the possibility that parties polarise. Following some history of political polarisation (and turnover), if voters' memory is long enough, their data will contain a lot of information and will therefore allow them to learn the truth. As $|\beta_l^* - \beta_r^*| > \rho$, this will induce both parties to espouse the policy that yields the highest expected utility and hence reach a consensus. Following periods in which there is consensus, short-term memory implies that voters at some point will only remember consensus periods. When $\beta^* \in B(\rho)$ observing outcomes generated by only one policy is not sufficient to discriminate between the expected utility of the two policies. Voters may learn the level of utility this policy delivers, but do not know how good or bad they have it, compared with the other policy. In this case, the party whose ideal policy is different from the consensus can "encourage" the voter to try it out, leading to a polarisation phase. One implication of this is that a status quo or a policy "lock-in" cannot last too long: When the history contains little variation and is not sufficiently informative about the true data generating process, there are lower electoral costs for parties to choose their ideal policies, and polarisation must arise at some point.

With Propositions 2 and 3 above we can state a sufficient condition for cycles to arise. Recall that we assumed that $|\beta_l^* - \beta_r^*| > \rho$. Propositions 2 and 3 assume that $\beta^* \in B(\rho)$. To satisfy these two requirements we need that

$$|\beta_l^* - \beta_r^*| > \max\{|\beta_l^* - E[\beta_r|\beta_l^*]|, |\beta_r^* - E[\beta_l|\beta_r^*]|\}.$$

Note that fixing the prior, the state in B which has the maximal $|\beta_l^* - \beta_r^*|$ always satisfies the above. In Section 3.4 below we provide an example of a prior distribution (that features "scale free" learning) for which this condition is satisfied for almost all states of the world.

The stochastic element of outcome shocks affects the nature of cycles in two fundamental ways. First, during consensus phases, the parties might converge on the wrong policy. For example, a finite series of positive shocks when implementing policy l and/or negative shocks when implementing r might convince voters that policy l is substantially better than policy r. In turn this might induce the parties to converge on offering l even though the true optimal policy is r. While this series of shocks might not happen too frequently, it might arise with a strictly positive probability. This inefficient learning did not occur in Proposition 2 above, as learning is perfect when $\sigma^2 \to 0$.

Second, when $\sigma^2 > 0$, the shocks to the outcome can trigger a move from party polarisation to party consensus and vice versa. Intuitively, during a phase of party polarisation, when implementing the policy l, a large, one-off shock can move beliefs towards a strong consensus that policy l is much more effective than policy r. This will induce the parties to choose the same platform in the next election. Alternatively, during a phase of party consensus on policy r, a negative shock might convince voters that policy l might not be so bad relative to policy r, and parties may now polarise. Our result above that the lower bound η does not depend on σ^2, implies that the deterministic cycle characterised in Proposition 2 is always present even when $\sigma^2 > 0$, but that unexpected positive or negative shocks might also affect the transitions between cycles.

Different shocks have different effects on public opinion and future policies. While some shocks like WWII have led to some consensus on the welfare state, others, including the financial crisis of 2008, led to more polarisation. One way to think of the contribution of our paper is that it is the conjunction of both events and the policies that were implemented at the time which affects opinions. In our model there are both exogenous shocks (events) as well as endogenous policies that together figure in to how voters update their beliefs. A large negative shock affects a polity differently, depending on whether they were implementing a left policy or a right policy. Also, the size of the effect on opinions will depend on the history (which itself has both endogenous policies and past exogenous shocks), so if the history contains variation in policies, voters would already be quite knowledgeable about the state as opposed to the case in which the history has little policy variation. In the next subsection we report some results from a simulation that allows us to illustrate further these features of the cycles.

Finally note that while we had simplified our model by focusing on a discrete set of policies, our result generalises to the case of continuous policies. With continuous policies, parties will never choose the same policy; they always have an incentive to polarise slightly. Still, when the history is very informative, parties will have to choose policies that are sufficiently close to each other. Close policies and finite memory will imply little variation and little learning as we have in our simple two-policies model.[6]

3.3 A simulation

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

[6]Formally, the statement of the existence of cycles will be that there exists two values $\bar{a} > \underline{a} > 0$ such that as $t \to \infty$ the $\liminf$ of the fraction of time of the distance between the platforms is above $\bar{a}$ and the $\liminf$ of the fraction of time in which the distance between the platforms is smaller than $\underline{a}$ are both strictly positive in the limit.

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 $\rho = \frac{2}{3}$. Note that $\beta_l^* - \beta_r^* = 1 > \rho = \frac{2}{3}$, implying that parties will reach a consensus on l when voters have full knowledge, and that $\beta^* \in B(\rho)$ as $\beta_l^* - E[\beta_r|\beta_l^*] = E[\beta_l|\beta_r^*] - \beta_r^* = 0.5 < \rho$. 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 ten 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).

$K = 10$	$\sigma = 2.5$	$\sigma = 1.2$	$\sigma = 0.1$
Average length of consensus phase	5.59	6.51	9.89
Fraction of consensus on the true optimal policy	83.24%	95.96%	100%
Proportion of periods implement the true optimal policy	74.8%	84.1%	93.2%

$K = 5$	$\sigma = 2.5$	$\sigma = 1.2$	$\sigma = 0.1$
Average length of consensus phase	3.78	4.15	4.9
Fraction of consensus on the true optimal policy	73.57%	87.63%	100%
Proportion of periods implement the true optimal policy	67.8%	75.8%	89.5%

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.

To gain some insight as to how the polity transitions between phases of consensus and polarisation we now follow one path of shocks (one simulation), when $K = 10$ and $\sigma = 1.2$. For the purpose of the graphs below, we fix the values of the policies at $l = -1$ and $r = 1$ (this plays no role in the model). Figure 3 shows the evolution of $E[\beta_l - \beta_r|H_t]$ over the hundred periods in the simulation, which captures how the knowledge of the voters aligns with the optimal policy l, and hence also determines the probability that party L wins when parties polarise. We also illustrate in the other panels the implemented policy and the policies offered by parties.

If we look for example at period 59, we see a decline in $E[\beta_l - \beta_r|H_t]$ from a level of approximately one, which is consistent with being informed about the state, to a level of approximately 0.5 which is consistent with learning from only one policy. This induces party R to switch to offer r, implying a move from consensus to polarisation. Indeed up to that period, the two parties were in consensus on policy l for several periods, and the transition is instigated by voters forgetting the experiences of policy r being implemented in periods 48 and 49 and having a memory of just policy l being implemented. This transition is consistent with the systemic force for cycles described above.

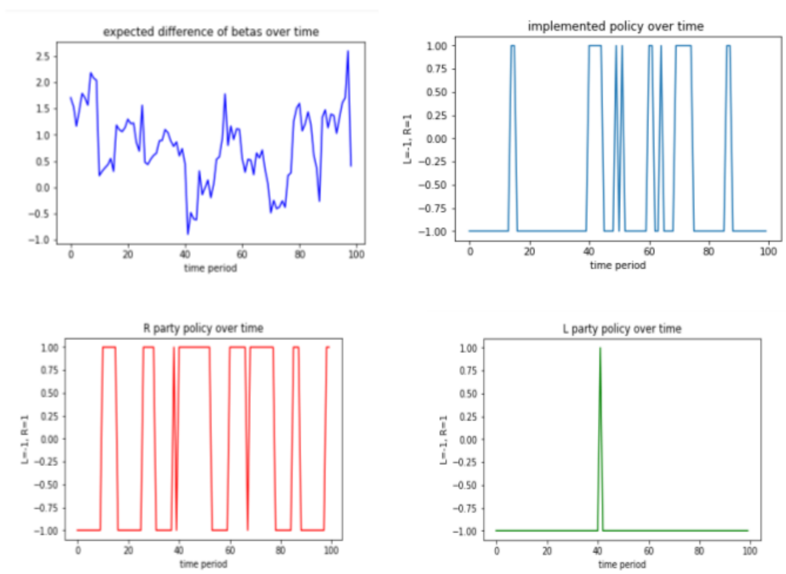

Figure 3: Beliefs (upper left panel), party policies (lower panels) and implemented policies (upper right panel).

Other examples on his path show how large shocks affect transitions: At period 40, there is a large decline in $E[\beta_l - \beta_r|H_t]$ from 0.5 to around -0.8, after a long period in which there was a consensus on l. The fall here is due to a large negative shock which implies that both parties switch from offering l to offering r. Around period 66 there is a similar negative shock but smaller and so we see a transition from consensus to polarisation.

3.4 Discussion

We now discuss the different parameters of the model that may affect the instances of consensus or polarisation.

3.4.1 Changing states

Our analysis so far focused on a fixed state of the world. However, 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 such as technological changes. Changes to the true state of the world can also arise from implementation of wrong policies (e.g., climate change).

Take for example the 2008 financial crisis; following the onset of the crisis, investors and governments realised that the effects of financial innovations have not been fully understood and hence policies were not properly tailored to the evolving technologies. As a result, and after some time, investment banks, governments and economists had to change their old models in favour of new ones. Old models and empirical analysis that relied on many years of data were deemed less relevant. Naturally, if voters in our model realise that the state of the world had changed, they can discard historical information which is potentially less relevant.

Our model can be easily adapted to better understand the effects of such environments on politics. A simple way to accommodate a changing data generating process is to assume that at each period t, with probability λ, nature draws a new vector of parameters β_t^* according to the prior distribution G. For simplicity, let us consider first the situation in which voters are aware once the state of the world has changed (or alternatively one can consider some delay in understanding such change). After a change has happened, the history up to that period becomes uninformative about the future. In the language of our model, this means that the "effective" K, i.e. the number of periods of history that voters learn from, is equal to zero at that period. To a degree, the possibility of a change in the state implies that memory can become short *endogenously*. Once memory is reset, parties will then polarise as $E(|\beta_l - \beta_r|) = 0$ given the prior. Therefore when voters understand that the world has changed, political polarisation will follow.[7]

Alternatively, voters may not be aware when the state had actually changed, but understand the process detailed above. In this case, when looking back at history, voters try to learn when the state has changed as this is informative about what the current state might be. The next proposition analyses this case and shows how for small σ^2, cycles will arise when λ is small enough, under the same sufficient condition as identified in our main analysis.

[7] The implications of the discussion above can shed light on recent empirical literature on crises and economic and political uncertainty. Mian, Sufi and Trebbi (2014) show that financial crises lead to political polarisation. Funke, Schularick and Trebesch (2016) show that financial crises create more political polarisation (mainly right-wing extremism) compared to normal recessions, and that these effects diminish after around ten years.

Proposition 4: *Assume that in each period t, with probability λ, $\boldsymbol{\beta}_t^*$ is drawn from $B(\rho)$ using the conditional distribution G on $B(\rho)$, and with the remaining probability $\boldsymbol{\beta}_t^* = \boldsymbol{\beta}_{t-1}^*$. Let $\sigma^2 \to 0$ and $2 \le K \le \infty$. Then there exists $\bar{\lambda} > 0$ such that for all $\lambda \le \bar{\lambda}$, the polity experiences cycles of polarisation and consensus.*

Intuitively, when the variance of the policy shock is vanishing, learning is very fast. During a consensus phase it is easy to detect when the state had changed, and so parties can then immediately polarise. In times of polarisation, as long as λ is small and so the chance that the state has changed is small, voters learn about the current state of the world and parties then reach a consensus.

The simple model above illustrates how short (and recent) memory can be derived endogenously. Alternatively, in a more elaborate example of the dynamics of the state of the world, one can think that there is more regularity in how states change, and that some states may recur throughout time. In this case, to learn, voters may attach more weight to periods in which they believe that the associated past states were more similar in some sense to the current state. This is an interesting extension of the model which we leave for future research.

3.4.2 Scale-free relative judgement

We now look a little closer at the conditions for cyclical patterns to emerge in equilibrium. Specifically, a sufficient condition for cycles to emerge is that $|\beta_l^* - \beta_r^*| > \max\{|\beta_l - E[\beta_r|\beta_l]|, |\beta_r - E[\beta_l|\beta_r]|\}$. In general, whether this condition holds or not depends on the prior G and the true state $\boldsymbol{\beta}^*$. We now identify one particular learning environment, with "scale-free" learning, for which almost all states satisfy the condition for cycles to emerge.[8]

Specifically, for $\left|\beta_p^* - E[\beta_{-p}|\beta_p^*]\right| < \left|\beta_p^* - \beta_{-p}^*\right|$ to hold the knowledge of β_p^* does not reveal too much information about how good or how bad she has it compared with the other policy $-p$. To capture the limits of learning from the scale of β_p^*, let

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

If $\bar{\Delta}$ is bounded then whenever $|\beta_l^* - \beta_r^*| > \bar{\Delta}$, the condition is satisfied, in which case cycles will occur in line with Propositions 2-4. Below we push this point to show that if learning about the relative benefit is close to being completely scale-free, cycles will always arise.

Example 1: We construct a sequence of priors $\{G_n(\beta_l, \beta_r)\}_{n=1}^{\infty}$ for which in the limit $\bar{\Delta}_n$ converges to zero. This implies that for almost any state of the world $|\beta_l^* - \beta_r^*| > \bar{\Delta}$, and

[8] We thank a referee for suggesting to explore "scale-free" learning.

hence cycles arise. 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 $\hat{G}_n(v, \delta) = G_n(v+\frac{\delta}{2}, v-\frac{\delta}{2})$ be the transformed distribution function. We denote by $\hat{G}_n(v)$ the marginal of this distribution over scale, and we assume that it is uniform on $[-D_n, D_n]$. Let $\hat{G}_n(\delta|v) = \hat{G}_n(\delta)$ be the marginal over δ, which we assume is independent of v and symmetric around zero. In addition, we assume that for any v and x, $E_n[\delta|v, \delta > x]$ is finite. In the appendix we show that as $D_n \to \infty$, $\lim_{n\to\infty} \bar{\Delta}_n = 0$.

In this simple example, when a voter learns β_l^*, she knows that β_r^* is either $\beta_l^* + \delta$, or $\beta_l^* - \delta$ for some δ which is symmetrically distributed. In the limit, when the distribution over the mid-point has no boundary points, it becomes point-wise uninformative about δ implying that $E[\beta_r|\beta_l^*] = \beta_l^*$ for any β_l^*. As a result, in the limit her learning about the relative difference between the two policies is completely scale-free.

3.4.3 Alternative assumptions on voters' behaviour

We have shown that short-term memory can lead to cycles. Within the context of our model, alternative assumptions about voters' behaviour can increase the emergence of such cycles:

Rational Inattention: To learn, voters need to pay attention to politics. In times of political polarisation voters might be inclined to pay more attention; the benefit of doing so is more apparent and news content providers in different media outlets will put more emphasis on political content. 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. This will lead to similar cyclical patterns as we established in our main results; specifically, the combination of short-term memory and rational inattention implies that after K periods of consensus, voters may end up with very little information, and potentially only the prior to guide them, leading parties to polarise, and so on.[9]

Political narratives: In a companion working paper, Levy and Razin (2022a), we explore another behavioral environment in which voters are more passive in their learning. In that model, beyond policies, parties also offer "narratives" -model specifications- to explain voters' history. These narratives justify their policy choices, and voters vote with a higher probability for the party who proposes a narrative that has a better fit to the historical data. When there is little variation in the data, e.g., in consensus phases, many narratives or models can fit very well the historical observations. Thus, this alternative model also has a feature of scale-free relative learning and generates cycles as in our current model.

[9] We thank a referee for suggesting this extension of the model.

Non-myopic behaviour of voters and risk-loving: In our model voters are fully myopic; if this is not the case, they may have some incentive to actively experiment and especially so when they are least informed. As a result, after periods of consensus, when their information is weaker (compared to periods of polarisation), these incentives to try other policies will potentially lead parties to polarise. More generally if voters have a taste for risk when they have less information we can sustain the same types of dynamics.

4 Extension: Different memories

The recollection of history might be different across voters; a recent literature suggests that voters' beliefs are disproportionately shaped by their own experiences mostly accumulated during their formative years. Guilliano and Spilimbergo (2016) show how individuals' beliefs and preferences are affected by their experiences of recessions in their youth, and 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.[10] Alternatively, a large literature studies and documents the effect of echo chambers on polarisation,[11] and so different groups that are exposed to different sources of information can end up with different memories and recollection of histories.

We now extend the model to consider heterogeneity in memories, as suggested above. Assume that the voting population is divided into m groups, where within each group, voters share the same memory. Before voting, voters in group j observe some data which constitutes their history, potentially different for each group. The history of each group j is a time-ordered finite set of pairs of policies and outcome, denoted by H^j. Thus, the information held by group j is composed of the prior G, and the $K^j - period$ history H^j. This allows the voters in group j to compute their posterior distribution G^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 for the period. As above, we assume that within group j and across the groups, voters may differ on additional dimensions. Thus, a voter i in group j votes for party L if:

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

[10] This can then translate into policy making; Malmendier, Nagel and Yang (2021) show how personal experiences of inflation strongly influence the hawkish or dovish leanings of central bankers.

[11] See Levy and Razin (2019) for a survey of this literature and Levy, Moreno de Barreda and Razin (2021) for a theoretical link between echo chambers, targeted political campaigns and polarisation.

where ϕ is an aggregate shock uniformly distributed on $[-\frac{1}{2\zeta}, \frac{1}{2\zeta}]$ and υ^{ij} is an idiosyncratic group-specific shock distributed on $[-\frac{1}{2\xi^j}, \frac{1}{2\xi^j}]$.[12] The indifferent voter in each group satisfies

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

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

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

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. Note that 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^j, parties choose policies in equilibrium to maximise their expected utility as before; we then have an analogous result to Lemma 1:

Lemma 2: *A consensus on policy $p \in \{l, r\}$ arises iff*

$$\sum_j^j w^j [E[(\beta_p - \beta_{-p})|H^j]] > \frac{1}{2\zeta(1+\alpha)}$$

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.

Lemma 2 is a useful result to study both generational and group conflicts as they evolve together with the policies and economic and other shocks. In Levy and Razin (2022b) we use this result to study an overlapping generations model and focus on the generational divide

[12] In case of an equality, wlog we assume that the voter votes for party L with probability 0.5.

in terms of the polarisation 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. One additional insight is that cohorts' opinions are influenced not just by the shocks that occurred in their lifetime, as described in the empirical literature cited above, but also by the actual policies that were implemented during this time. For example, an exogenous shock can affect the young of today; but the policies they will implement in response to this shock will in turn affect the learning of the future young generation. Living through the generous welfare state for baby boomers -a response of the previous generation to the WWII shock- might have had a different effect on their beliefs compared with those developed by the following generations living through more polarised times.[13]

5 Related literature

Our paper contributes to the current literature that focuses on the polarisation of politics in recent decades.[14] 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. Rogoff (1990) considers a model where an incumbent politician that tries to signal her ability chooses policies that pander to the public's will in election years, assuming that retrospective voters focus on a very short period of time, while the politician switches to "normal" policy making in other times. Rogoff (1990) therefore also assumes short-term memory for voters which implies a cycle between good policies far from election and potentially worse policies prior to election, and the phases of the cycles coincide with the timing of elections. 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. As in Rogoff (1990), here too the cycles are of "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

[13] For recent literature on polarisation and extremism in different demographic groups see Ortoleva and Snowberg (2015) and Boxell, Gentzkow and Shapiro (2017).

[14] For a recent example see Callander and Carbajal (2022).

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 (where in our case the behavioural assumption is the short-term memory). In Callander, Izzo and Martin (2022) voters vote for parties according to their beliefs about the effect of a policy; these beliefs are learnt in a passive manner, as in an initial stage parties compete by advocating these different effects of a policy and voters adopt the one that has a higher likelihood given their (finite and exogenous) data set. Little (2019) studies voter learning problems in which motivated reasoning distorts beliefs.[15] 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 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.[16]

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

[16] Azzimonti and Fernandez (2018) and Bohren and Hauser (2021) are two additional examples of social

There is some theoretical work that considers learning with bounded memory, such as Kocer (2010) and Wilson (2014). Jehiel and Newman (2010) and Bhaskar and Thomas (2019) analyze social learning with bounded memory. Related to our work is the analysis in Acemoglu and Wolitzky (2014) of 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.

6 Conclusion

In this paper, we investigate the concept of politics as a collective learning process. We demonstrate how the presence of short-term memory leads to cycles of polarization and consensus in party platforms and public attitudes. Our model incorporates both endogenous policies and exogenous shocks that impact the observed outcome and provide information about the state of the world. In certain situations, external signals from other polities can also influence learning. For instance, during the Covid-19 pandemic, the public and politicians were able to learn from the experiences of countries that were affected by the virus earlier. Similarly, US states have a history of learning from one another, as exemplified by the concept of "Laboratories of democracies" proposed by Supreme Court justice Louis Brandeis in 1932.[17] The regulation of industry in the US serves as another example, with state-level experiences influencing both legislation in other states and federal legislation.[18] 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.

7 Appendix

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

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

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.

[17] In *New State Ice Co. v. Liebmann*, 285 U.S. 262 (1932).

[18] See Goldin, C. and G. D. Libecap (2008).

Therefore, 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:

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

In this case, party L prefers to do so as well, yielding a consensus on l. An analogous condition guarantees a consensus on r. As a result, polarisation must arise when neither of these conditions hold, that is when

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

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

Proof of Proposition 1: Voters' posterior after observing the history H_t satisfies the conditions of the martingale convergence theorem. Therefore, voters' beliefs converge almost surely to some limit probability distribution μ_∞. Let the support of these beliefs be denoted by B_∞. This means that $\frac{\mu_\infty(\beta)}{\mu_\infty(\beta')} \to \infty$ for all $\beta \in B_\infty$ and $\beta' \notin B_\infty$. Note that as beliefs converge, cycles can never emerge, as beliefs will either converge to satisfy the condition in Lemma 1 for polarisation or for consensus.

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 a result from Esponda, Pouzo and Yamamoto (2021), henceforth EPY. Specifically, remember that $\hat{\eta}_t(p) \in [0,1]$ is the fraction of times that policy p has been implemented in H_t. The result in EPY implies, in the context of our model, that the posterior distribution of the beliefs, denoted by μ_t, will concentrate in the limit on some set $B_\infty^{\hat{\eta}}$ whose elements have a Kullback-Leibler (KL) value that is close to the minimum such value. The KL value of some vector of parameters $\boldsymbol{\beta}$ given a fraction $\hat{\eta}_t$ 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 $\boldsymbol{\beta}^*$ is a minimizer of KL for which $KL(\boldsymbol{\beta}^*|\hat{\eta}_t, \boldsymbol{\beta}^*) = 0$ regardless of $\hat{\eta}_t$.

Step 1: In this step we show that if for some $\boldsymbol{\beta} \in B$ the KL value is strictly positive, then a ball around $\boldsymbol{\beta}$ must have zero measure in the limit beliefs μ_∞.

EPY show in Theorem 1 that no matter whether actions (implemented policies in our model) converge or not, for any observed frequency of actions, the posterior beliefs will concentrate on values of $KL(\boldsymbol{\beta}|\hat{\eta}_t, \boldsymbol{\beta}^*)$ that are closest to its minimized value.[19] The result in EPY, in the context of our model, implies that, as $\boldsymbol{\beta}^*$ is a minimizer and $KL(\boldsymbol{\beta}^*|\hat{\eta}_t, \boldsymbol{\beta}^*) = 0$, there exists a set $E \subset \mathbb{H}$ such that $P(E) = 0$ and that for all $H \in \mathbb{H} \setminus E$,

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

We now show that this implies that if for some $\boldsymbol{\beta}$ the KL value is strictly positive then a ball around $\boldsymbol{\beta}$ must have zero measure in the limit beliefs. Suppose $\boldsymbol{\beta} \in B$ is such that $KL(\boldsymbol{\beta}|\hat{\eta}_t, \boldsymbol{\beta}^*)$ does not converge to zero. Then there exists $\psi > 0$ and a subsequence t_n such that for all t_n, $KL(\boldsymbol{\beta}|\hat{\eta}_{t_n}, \boldsymbol{\beta}^*) \geq \psi$. By continuity there is a $\delta-$ball $B^\delta(\boldsymbol{\beta})$, around $\boldsymbol{\beta}$, such that, when integrating over this ball that satisfies $KL(\boldsymbol{\beta}|\hat{\eta}_{t_n}, \boldsymbol{\beta}^*) \geq \psi$, we have

$$\int_{B^\delta(\boldsymbol{\beta})} KL(\boldsymbol{\beta}'|\hat{\eta}_{t_n}, \boldsymbol{\beta}^*) \mu_{t_n+1}(\boldsymbol{\beta}') d\boldsymbol{\beta}' \geq \int_{B^\delta(\boldsymbol{\beta})} \psi \mu_{t_n+1}(\boldsymbol{\beta}') d\boldsymbol{\beta}' = \psi \int_{B^\delta(\boldsymbol{\beta})} \mu_{t_n+1}(\boldsymbol{\beta}') d\boldsymbol{\beta}'$$

But $\int_B KL(\boldsymbol{\beta}'|\hat{\eta}_{t_n}, \boldsymbol{\beta}^*) \mu_{t_n+1}(\boldsymbol{\beta}') d\boldsymbol{\beta}' \geq \int_{B^\delta(\boldsymbol{\beta})} KL(\boldsymbol{\beta}'|\hat{\eta}_{t_n}, \boldsymbol{\beta}^*) \mu_{t_n+1}(\boldsymbol{\beta}') d\boldsymbol{\beta}'$, and

$\lim_{t\to\infty} \int_B KL(\boldsymbol{\beta}'|\hat{\eta}_t, \boldsymbol{\beta}^*) \mu_{t_n+1}(\boldsymbol{\beta}') d\boldsymbol{\beta}' = 0$. Thus we must have $\int_{B^\delta(\boldsymbol{\beta})} \mu_{t_n+1}(\boldsymbol{\beta}') d\boldsymbol{\beta}' \to 0$ and so $\mu_{t_n}(\boldsymbol{\beta}') \to 0$ almost surely in $B^\delta(\boldsymbol{\beta})$, as $t_n \to \infty$.

Step 2: 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 Proposition 1. So we focus on 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

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

a strictly positive probability due to $\alpha > 0$ (otherwise parties do not polarise). We then have a strictly positive measure of paths for which $\hat{\eta}_t(l)$ and $\hat{\eta}_t(r)$ are bounded away from zero. This means that $\boldsymbol{\beta}^* = \arg\min_{\boldsymbol{\beta}'} KL(\boldsymbol{\beta}'|\hat{\eta}_t, \boldsymbol{\beta}^*)$, i.e. $\boldsymbol{\beta}^*$ is the unique minimizer of the KL divergence as $t \to \infty$. By continuity and Step 1, beliefs can only concentrate on a ball around $\boldsymbol{\beta}^*$. By the convergence of beliefs we must have that beliefs along these paths have $B_\infty = \{\boldsymbol{\beta}^*\}$. But, as $|\beta_l^* - \beta_r^*| > \rho$, this is in contradiction to the supposition that $\left|E_{\mu_\infty}(\beta_l - \beta_r|H_t)\right| < \rho$.

Steps 1 and 2 imply the result in the Proposition.■

The following Lemma will be helpful in the proofs of Propositions 2 and 4 below.

Lemma A2: *Assume that* $K \geq 2$ and $\sigma^2 \to 0$, *and that* $\beta^* \in B(\rho)$. *(i) Suppose that there is a strictly positive measure of histories* H_t *such that only one policy was implemented throughout the history. Then with probability one parties polarise at 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 with probability one parties are in consensus on the optimal policy in period* t.

Proof of Lemma A2: (i) Assume one policy p is implemented for K periods in a strictly positive measure of histories. 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 a result, when $\sigma^2 \to 0$, with probability arbitrarily close to one, the posterior belief will be concentrated on $(\beta_p^*, E[\beta_{-p}|\beta_p^*])$. As $\boldsymbol{\beta}^* \in B(\rho)$, parties will polarise. (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 in the K periods are in $[-\gamma'', \gamma'']$. As a result, when $\sigma^2 \to 0$, with probability arbitrarily close to one, the posterior belief will be concentrated on $\boldsymbol{\beta}^*$. In this case, parties will converge on the optimal policy.■

Proof of Proposition 2: Following Lemma A2, after histories H_t that contain two different implemented policies, parties will both offer the optimal policy. Parties will continue to both this policy as long as the history they observe contains two different implemented policies. 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. This concludes the proof.■

Analysis of equilibria when $\sigma \to 0$ and $\boldsymbol{\beta}^* \notin B(\rho)$:

For completeness, we explore our result in Proposition 2 for the case in which $\boldsymbol{\beta}^* \notin B(\rho)$. As we show, in this case, the polity will either be in perpetual consensus or will experience cycles.

If $\boldsymbol{\beta}^* \notin B(\rho)$ 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 $|\beta_r^* - E[\beta_l|\beta_r^*]| > \rho$.

Let us start with case 3. In this case, if only one policy p was implemented in the observed history, then in the next stage the parties will be in consensus on policy p if $\beta_p^* > E[\beta_{-p}|\beta_p^*]$ and on policy $-p$ if $\beta_p^* < E[\beta_{-p}|\beta_p^*]$. So in equilibrium we will have consensus forever after some period.

Cases 1 and 2 are analogous so we focus on case 1. Suppose first that the optimal policy is r. Assume that only policy r was implemented in the observed history, then in the next stage the parties will be in consensus on policy r if $\beta_r^* > E[\beta_l|\beta_r^*]$ and on policy l if $\beta_r^* < E[\beta_l|\beta_r^*]$. In the former case, this will be an absorbing state of consensus. In the latter case, after one period of implementing l the polity immediately learns that the optimal policy is r. Parties will be in consensus on r for K periods, and so on. Thus in this case, we have perpetual consensus.

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

Assume now that l is the optimal policy. Assume that only policy r was implemented in the observed history, then in the next stage the parties will be in consensus on policy r if $\beta_r^* > E[\beta_l|\beta_r^*]$ and on policy l if $\beta_r^* < E[\beta_l|\beta_r^*]$. In the former case, this will be an absorbing state of consensus. In the latter case, the polity immediately learns that the optimal policy is l. Parties will then be in consensus on l for K periods, following which parties polarise. Once r is sampled, parties will be in consensus on l. And so on, implying cycles of consensus on l and polarisation. Thus, in cases 1 and 2, we will either have cycles or a perpetual consensus (in which parties may switch between consensus on one policy to another).

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 such that $\hat{\eta}_{t_n}(p) \to 1$* for some $p \in \{l, r\}$.

To see this, let us assume to the contrary that there exists such a subsequence t_n which on a 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 1: *For a large enough K, almost surely after a $K-window$ in which only one policy $p \in \{l, r\}$ was implemented, the next period will involve parties polarising with a strictly positive probability.*

Proof of Claim 1: Note that we have assumed that there is a positive measure of paths such that $\hat{\eta}_{t_n}(p) \to 1$ and therefore, for any such path, we will have infinite $K-windows$ defined as above. When K is large and t_n goes to infinity, then on a strictly positive measure of paths, each $K-window$ with only a fixed p will almost surely imply that beliefs concentrate on β^*_p, and on $E[\beta_{-p}|\beta^*_p]$ for β_{-p}. However, as $\beta^* \in B(p)$, $|\beta^*_p - E[\beta_{-p}|\beta^*_p]| < \rho$, and by Lemma 1, for high enough K parties polarise in the next period with strictly positive probability.$\square_{claim\ 1}$

We can now use Claim 1 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 1, each of these will lead to polarisation almost surely. But as each party wins with strictly positive probability when there is polarisation, this contradicts the supposition that $\hat{\eta}_{t_n}(p) \to 1$ as $\eta_{t_n}(polarisation)$ is in the order of $\frac{1}{K}$.$\square_{step1}$

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

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 $K-windows$ 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* we have that the ratio of the share of time that p was implemented compared to the share of time that p' was implemented, converges, with a large K, to some finite $c_n > 0$. But then, as beliefs in almost all such K-*window* must converge to $\boldsymbol{\beta}^*$, after almost all such K-*window* both parties will choose the optimal policy, denote it by p^* and so $\hat{\eta}_{t_n}(p^*) \to 1$, a contradiction to $c_n > 0$.

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 $K-windows$. But this ratio has to converge to zero or infinity slowly 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 that $\eta_{t_n}(polarisation)$ is bounded from zero.

We now examine 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, $\left|E[\beta_{p'}|H_{t_n}] - E[\beta_p|H_{t_n}]]\right|$ is arbitrarily close to $\left|\beta^*_{p'} - E[\beta_p|\beta^*_{p'}, H_{t_n}]]\right|$. But as $\boldsymbol{\beta}^* \in B(\rho)$, and so $\left|\beta^*_{p'} - E[\beta_p|\beta^*_{p'}]\right| < \rho$. As we look at a strictly positive measure of paths, we can use iterated expectation to conclude that $E[\beta^*_{p'} - E[\beta_p|\beta^*_{p'}, H_{t_n}]] = \beta^*_{p'} - E[\beta_p|\beta^*_{p'}] < \left|\beta^*_{p'} - E[\beta_p|\beta^*_{p'}]\right| < \rho$. This implies that $\left|\beta^*_{p'} - E[\beta_p|\beta^*_{p'}, H_{t_n}]]\right| < \rho$ with a strictly positive probability. As a result, for a strictly positive measure of paths we should have polarisation and hence a contradiction to $\eta_{t_n}(polarisation) \to 0.\square_{step2}$

Step 3: *For a large enough* K, $\limsup_t \eta_t(polarisation) < 1$.

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 $\eta_{t_n}(polarisation) \to 1.\square_{step3}$

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

$$\min\{\liminf \eta_t(polarisation), \liminf \eta_t(consensus)\} > \eta.$$

Suppose the statement is not true. Steps 2 and 3 imply that there exists a large enough K such that for any σ, $\min\{\liminf \eta_t(polarisation), \liminf \eta_t(consensus)\} > 0$. So for the statement to be wrong we must have that $\lim_{\sigma\to 0} \min\{\liminf \eta_t(polarisation), \liminf \eta_t(consensus)\} = 0$. But in Proposition 2 we have shown that at $\sigma = 0$, $\min\{\liminf \eta_t(polarisation), \liminf \eta_t(consensus\} > \frac{1}{K}$. Therefore, by continuity there must be an $\eta > 0$ that satisfies the statement of Step 4. $\square_{step4}$

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. More generally, as consensus on the wrong policy can arise when $K = \infty$, it can *a fortiori* arise when $K < \infty$.$\blacksquare$

Proof of Proposition 4: From lemma A2, as $\sigma^2 \to 0$ voters can learn precisely the values of $\beta^*_{p,t}$ in a period in which p was implemented. When voters look back on historical outcomes, they can see when the change of the state has occurred if it did between two periods in which the same policy p was implemented. When this happens, in period t, voters only know $\beta^*_{p,t}$. In this case, parties will polarise as from the voters perspective there are two options. Either the state remains (in which case, as $\boldsymbol{\beta}^*_t \in B(\rho)$, and voters know only $\beta^*_{p,t}$, parties can polarise), or the state changes and then according to the symmetric prior (which stays symmetric when restricted to $B(\rho)$), parties can polarise. When parties polarise, at some point at time t the policy that is implemented is different from that in time $t-1$. Then voters' posteriors concentrate on $\boldsymbol{\beta}^*_t$ if λ is small enough, as then it is more likely that the state did not change at period t. When λ is small, it is also the case that the voters expect the state to remain the same at period $t+1$, and so there is a $\bar{\lambda} > 0$ such that if $\lambda < \bar{\lambda}$ the parties will form a consensus on the optimal policy given that the state is $\boldsymbol{\beta}^*_t$. The above implies that the polity will cycle between polarisation and consensus; when the state change is detected, parties polarise, and after two different policies are implemented, parties will be in consensus. Note that the above proof holds for all $K \geq 2$, and specifically also for $K = \infty$. In the latter case, with a strictly positive probability the state changes and parties polarise, but after two consecutive periods in which different policies were implemented, parties will switch to consensus on the optimal policy, and will remain in consensus on this policy until the state changes.■

Proof for result in Example 1: For any β^*_l we have that,

$$
\begin{aligned}
& \beta^*_l - E_{G_n(\beta_l,\beta_r)}[\beta_r|\beta^*_l] \\
= \; & \beta^*_l - \int_{v,\delta \text{ so that } v+\frac{\delta}{2}=\beta^*_l} (v - \frac{\delta}{2}) \frac{\hat{g}(v)\hat{g}(\delta)}{\int_{v,\delta \text{ so that } v+\frac{\delta}{2}=\beta^*_l} \hat{g}(v)\hat{g}(\delta) dv d\delta} dv d\delta \\
= \; & \beta^*_l - \int_\delta (\beta^*_l - \delta) \frac{\hat{g}(\beta^*_l - \frac{\delta}{2})\hat{g}(\delta)}{\int_\delta \hat{g}(\beta^*_l - \frac{\delta}{2})\hat{g}(\delta)} d\delta \\
= \; & \int_{2(x-D_n)}^{2(x+D_n)} \delta \frac{\hat{g}(\delta)}{\int_{2(x-D_n)}^{2(x+D_n)} \hat{g}(\delta)} d\delta \\
\rightarrow_{D_n \rightarrow \infty} \; & \int_{-\infty}^{\infty} \delta \frac{\hat{g}(\delta)}{\int_{-\infty}^{\infty} \hat{g}(\delta)} d\delta = 0.■
\end{aligned}
$$

Proof of Lemma 2: The proof follows that of Lemma 1 by substituting the expression derived in the text in Section 4 for $\Pr(L \text{ wins}|l, r)$.■

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

[4] 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.

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

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

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

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

[9] 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.

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

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

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

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

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

[15] Eliaz, K., R. Spiegler and S. Galperti (2022). "False Narratives and Political Mobilization". mimeo.

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

[17] Funke, M., M. Schularick, and C.Trebesch (2016). "Going to extremes: Politics after financial crises, 1870–2014", European Economic Review 88: 227-260.

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

[19] Ghitza, Y., Gelman, A. and J. Auerbach (2022). "The Great Society, Reagan's Revolution, and Generations of Presidential Voting", American Journal of Political Science https://doi.org/10.1111/ajps.12713.

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

[21] Giuliano, P. and A. Spilimbergo (2014), "Growing up in a Recession". Review of Economic Studies 81: 787–817.

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

[23] 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.

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

[25] Juskow, P. (2005), "Regulation and Deregulation after 25 Years: Lessons Learned for Research in Industrial Organization". Review of Industrial Organization 26: 169-193.

[26] Kocer, Y. (2010). "Endogenous Learning with Bounded Memory". mimeo.

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

[28] Levy, G. and R. Razin (2022a). “Political social-learning: short-term memory and cycles of polarisation.” mimeo.

[29] Levy, G. and R. Razin (2022b). “The Political Economy of cohort effects, the intergenerational conflict and polarised generations”. mimeo.

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

[31] Levy, G., Moreno de Barreda, I., and R. Razin (2021). "Polarized Extremes and the Confused Centre: Campaign Targeting of Voters with Correlation Neglect". Quarterly Journal of Political Science 16: 139-155.

[32] Levy, G. and R. Razin (2019). "Echo Chambers and Their Effects on Economic and Political Outcomes". Annual Review of Economics 11: 303-328 .

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

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

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

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

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

[38] Malmendier, U., S. Nagel, and Z. Yang (2021). “The Making of Hawks and Doves”, Journal of Monetary Economics 117: 19-42.

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

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

[41] Mian, A., A. Sufi and F. Trebbi (2014). “Resolving Debt Overhang: Political Constraints in the Aftermath of Financial Crises”. American Economic Journal: Macroeconomics 6: 1-28.

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

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

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

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

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

[47] Reinhart, C. and K. Rogoff (2011). *This Time is Different.* Princeton University Press, Princeton, NJ.

[48] Rogoff, K. (1990). "Equilibrium Political Budget Cycles". American Economic Review 80: 21-36.

[49] Rose, N. (2014). "Learning from the Past: Insights for the Regulation of Economic Activity". In *Economic Regulation and its Reform: What Have We Learned*, ed. by Nancy Rose, University of Chicago Press, Chicago, IL.

[50] Shapiro, R.Y. and J.M. Gilroy (1984). "The Polls: Regulation-Part I". The Public Opinion Quarterly 48: 531-542.

[51] Stimson, J. (1991). *Public Opinion in America. Moods, Cycles, and Swings*. Boulder, CO: Westview.

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

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