

Political Social-learning: Short Term Political Memory and the Cycles of Polarisation

Gilat Levy and Ronny Razin LSE

July 2023

- We often say: "History repeats itself"
 - Polarisation high today, but also beginning of 20th century.
 - Regulation/Deregulation
 - "The Washington consensus" Williamson 1989, "Goodbye Washington Consensus, Hello Washington Confusion?" Rodrik 2006.
 - Financial regulation cycles: Rajan (2009): "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."

- Politics as a social learning endeavor by society:
 - Voters update beliefs given past observations of political actions,
 - Political actions depend endogenously on these beliefs.
- We, economists, do it:
 - Rose (2014) discussing the regulation of cable TV: “Cable provides a rich laboratory for economists in search of policy variation... deregulation, re-regulation, and deregulation once again in this sector.”
 - Juskow (2005), summarising 25 years of IO literature: “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.”

- The basic questions we hope to shed light on:
 - How does society fair when voters have short term memory?
 - How are consensus and polarisation related to the level of knowledge in the polity?
 - Does (political) history repeat itself? Are there political mechanisms that imply the recurrence of polarisation?
 - Is short-term memory all that bad?

Introduction: Key assumptions

- Parties represent polarised interests.
- Voters are inclined to vote for the party that offers the correct platform.
- Electoral competition:
 - Parties choose policy platforms.
 - Voters vote (probabilistic voting).
 - Consensus or polarisation.
- Voters (try to) learn over time what is the correct platform but might have short term memory.

Introduction: Examples

- The common goal of voters is to maximise welfare. But what are the effects of taxation and redistribution on Economic welfare?
- The common goal of voters is to have a functioning economy. But what is the correct level of regulation? What are the effects of regulation on consumer and producer surpluses?

- Two policy choices, $p \in \{l, r\}$, with outcomes:

$$y_t = \begin{cases} \beta_l^* + \varepsilon_t & \text{if } p = l \\ \beta_r^* + \varepsilon_t & \text{if } p = r \end{cases}$$

- ε_t is iid across time and normally distributed with zero mean and variance σ^2 .
- Voters understand the data generating process, but do not know the true value of these parameters, $\beta^* = (\beta_l^*, \beta_r^*)$.
- They have a continuous prior $G(\beta_1, \beta_2)$ on β with support on some compact set B in R^2 .

The Model: Preferences

- Fixing some stochastic shocks, all voters commonly prefer a higher outcome y_t , where policy l is considered better at period t iff

$$E[\beta_l|\Omega_t] > E[\beta_r|\Omega_t].$$

- Parties: Party L prefers l , party R prefers r .

$$U_R(l) = 0, U_R(r) = 1, U_L(p) = 1 - U_R(p).$$

- Parties enjoy small office-rents when they win the election, and so parties' utility is:

$$U_J(p) + \alpha I_J$$

The Model: Histories

- At each period t , the voters observe data from only the last K periods and treat them **as exogenous**.
- H_t , is the set of implemented policy vectors and policy outcomes, $\{p_\tau, y_\tau\}_{\tau=t-K}^{\tau=t-1}$ where $p_\tau \in \{l, r\}$.
- At every period t , $\Omega_t = \{G, H_t\}$.
- This allows the voters to compute their posterior distribution G_t on the vector β .

The Model: Voting

- Given the information they have, voters are inclined to vote for the party that offers a higher expected outcome y_t ;
- Probabilistic voting:
 - The median voter votes for party L if

$$E[y(p_t^L)|\Omega_t] - E[y(p_t^R)|\Omega_t] + \phi_t > 0,$$

where ϕ_t is uniformly distributed on $[-\frac{1}{2\zeta}, \frac{1}{2\zeta}]$.

Summary of model:

- There is an initial history H_0
- At period t , party J_t that won the election implements its platform $p_t^{J_t} \equiv p_t \in \{l, r\}$.
- Outcome y_t is realized and history is updated so that $H_{t+1} = \{p_\tau, y_\tau\}_{\tau=t-K+1}^t$.
- At period $t+1$, each party J offers a platform p_{t+1}^J , ϕ_t is drawn and party L wins the election if

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

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

A Lemma and an assumption

- Fix a history H_t and consider the one-period political competition game that ensues at period t .
- Let

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

- **Lemma 1 (Consensus vs Polarisation):** *At period t , if $\exists j$ for which*

$$E(\beta_j|\Omega_t) - E(\beta_{-j}|\Omega_t) > \rho$$

then both parties are in consensus over this expected-outcome maximizing policy, and otherwise parties polarise.

- **Assumption 1:**

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

- More information allows you to differentiate more between the policies.

Collective learning with full memory

- When voters observe the full history, the polity converges to parties offering the same platform (consensus).
- Intuitively, as long as parties polarise:
 - This implies experimentation, which improves learning,
 - And so voters will have strong beliefs about the correct policy.
 - As a result, parties must converge.
- The consensus is not necessarily on the correct policy (history matters).

- **Proposition 1:** *Assume that $K = \infty$. Then almost surely the polity converges to party consensus. The policy the polity converges to is not necessarily the optimal policy.*
 - Beliefs must converge by the martingale convergence theorem.
 - If beliefs converge to a set that allows parties to polarise, they will do so.
 - Parties polarising in the long run means that history is fully informative
 - That is, long-run beliefs must converge to β^* by Assumption 1.
 - Policy might not be optimal due to myopia (more generally $\delta < 1$).

- For finite memory, the nature of the voter's data can change over time.
 - May depend on how much variation in policies there was in the last K periods.
- How does endogenous history and short term memory interact?

Short term memory: Political Cycles

- Short-term memory implies that society might oscillate between periods of polarisation and consensus.
- Polarisation \Rightarrow consensus:
 - After a phase of polarisation, there is sufficient variation in implemented policies
 - Voters will have a better idea of what is the correct policy and politicians are forced to adopt it.
- Consensus \Rightarrow polarisation:
 - After a phase of consensus politics, learning is hindered by too little variation of policies.
 - Voters are less sure about the state of the world and so Politicians can push their self-interested platforms.

- To show the systemic force behind cycles in the model we take the variance of the policy noise to be small so that learning is fast.
- **Proposition 2:** *Let $\sigma^2 \rightarrow 0$ and $K \geq 2$. Then:*
 - *The polity experiences perpetual cycles of polarisation and consensus;*
 - *The consensus phase is on the correct policy and lasts exactly K periods,*
 - *The polarisation phase lasts until two distinct policies are implemented.*

Prevalence of Cycles

- Can always find environments for which Assumption 1 holds.
- A feature that is conducive for cycles is "scale-free" learning:
 - The magnitude of $\beta_j - E[\beta_{-j}|\beta_j]$, which depends on the joint distribution prior G , can be interpreted as a scale effect; experiencing different levels of utility of one policy affects a voters' perception about another policy.
 - "Scale-free": the voter's perception of the average distance between her utility from the two policies is fixed and does not depend on the scale (e.g., the value of β_j for some j).
 - We provide an example (below) in which learning becomes completely scale-free in the limit, i.e., $\sup_{\beta_j} |\beta_j - E[\beta_{-j}|\beta_j]| \rightarrow 0$.
 - For this environment cycles will arise for almost any state.
- Alternative "behavioral" models of voting will enlarge the set under which we have cycles:
 - Rational Inattention.
 - Voting according to likelihoods of narratives.

Welfare: Could forgetting be good?

- No: In the current model no benefit of short term memory but...
 - As K grows large does as good as $K = \infty$.
 - As K grows large we get the same distribution over correct policies as you do in $K = \infty$:
 - With $K = \infty$ a probability χ of being in the correct state and staying there forever.
 - With finite large K a probability χ of being in the correct state at any point in the future.

Welfare: Could forgetting be good?

- YES: Short term memory might be better if there is a small probability the state changes.
 - Assume a probability λ that the state changes.
 - Myopia implies that in the model with $K = \infty$, might be stuck for a long time on the wrong policy.
 - With finite K we constantly switch to the correct state with probability higher than a half.
 - When K is large enough, short term memory can outperform a model with infinite memory.

- Cohort effects: Voters' beliefs may not necessarily be shaped by the most recent K periods, but by the periods that consist their formative years. Malmendier and Nagel (2016), Malmendier et al (2021), Aksoy et al (2020).
- Derive similar cycles in an OLG model with cohorts.
- Empirical analysis showing differences in the second order distributions of opinions of different cohorts.
 - Related to different levels of political knowledge and engagement.

- Learning from other polities.
 - History of regulation in the US: Federal learning from states.
 - China.
 - But...covid.
- Long term versus short term institutional memory.
 - Elected decision makers: Short-termism
 - Government Bureaucracies: Long-term learning
 - Key is how to link the information and the decision making.
- Diversity to increase the depth of collective memory

- Short term memory (Kramer 1971, Peltzman, 1990, Achen and Bartels, 2008, Bartels 2008, Gerber and Green 1998, Lenz, 2010, Healy and Lenz, 2014).
- Social learning in politics (Piketty 1995, Strulovici 2010, Callander 2011, Little 2019, Levy Razin and Young 2022, Jehiel and Newman 2010).
- Political Cycles (Rogoff 1990, Battaglini and Coate, Levy et al 2022, Wolitzky and Acemoglu 2014, Morelli and Foarta 2022)
- Learning with short-term memory (Kocer 2010, Wilson 2004, Pavan 2014).
- Long short-term memory (LSTM) in Neural Networks.