Department of Economics, University of Warwick

Monash Business School, Monash University

as part of

Monash Warwick Alliance

The Dynamics of Polarisation and Revolutions

Ruilang Qin

Warwick-Monash Economics Student Papers

September 2024 **No: 2024/77**

ISSN 2754-3129 (Online)

The Warwick Monash Economics Student Papers (WM-ESP) gather the best Undergraduate and Masters dissertations by Economics students from the University of Warwick and Monash University. This bi-annual paper series showcases research undertaken by our students on a varied range of topics. Papers range in length from 5,000 to 8,000 words depending on whether the student is an undergraduate or postgraduate, and the university they attend. The papers included in the series are carefully selected based on their quality and originality. WM-ESP aims to disseminate research in Economics as well as acknowledge the students for their exemplary work, contributing to the research environment in both departments.

Recommended citation: Qin, R. (2024). The Dynamics of polarisation and revolutions. *Warwick Monash Economics Student Papers* 2024/77.

[1] Warwick Economics would like to thank Gianna Boero and Samuel Obeng for their contributions towards the selection process.

The Dynamics of Polarisation and Revolutions

Ruilang Qin*

Abstract

Political polarisation has become a prevalent phenomenon in the past decades. Parallelly, citizens have increasingly resorted to collective actions to demand change, resulting in incidents such as the "Jan 6" US Capitol riot. Evidence suggests that such public remonstrations exacerbated the extent of opinion divergence. This paper therefore presents a model that explains the dynamic connection between political polarisation and collective actions. In the setup, voting, abstention, and participation in collective actions are novelly modelled as individual components of a citizen's political toolkit. With endogenous voter preferences alone, polarisation has an exacerbating but limited effect on the level of collective actions. In turn, collective actions accelerate the process of polarisation for the election-losing partisans, creating asymmetry in the voter distribution. It is only when combined with strategic behaviour of the parties that polarisation may lead to substantially intensified collective actions.

JEL codes: D72, D74

Keywords: political polarisation, collective actions, electoral contest

*My most sincere gratitude goes to Kirill Pogorelskiy for his mentorship and continuous encouragement. I have also benefited from the helpful comments of Alex Dobson, Samuel Obeng, Anna Hedgecock, and Mikhail Chebotar. I thank the audience at the International Conference of Undergraduate Research 2024. Online appendix can be found here.

Department of Economics, University of Warwick, UK. Email: ruilangqin@gmail.com

1 Introduction

Polarisation has become a prevalent phenomenon in democratic countries over the recent decades. In the United States, the divergence is no longer limited to public attitudes towards policy issues, but perceptions of factual reality (Alesina, Miano, and Stantcheva 2020). Similar trends have emerged in other OECD countries, in particular those in Western Europe, making the political environment increasingly antagonistic (Canen, Kendall, and Trebbi 2020). As a result, citizens who hold extreme views are enjoying growing influence over election outcomes over those with moderate views: extreme citizens are more likely to self-select into politics by donating to campaigns, attending rallies, and ultimately turning out to vote (Mitchell et al. 2014). In general, polarisation has been linked to increased uncertainty in government policies including taxation, spending, and investments in infrastructure (Baker et al. 2014). These consequences have motivated the economic literature to grant the topic substantial consideration.

Parallel to the trend of polarisation, citizens have increasingly resorted to collective actions outside the formal representation framework to demand change. Compared to voting, actions such as protesting and striking are perceived to have a higher influence on political issues (Oliver 2001). In the West, the most recent high-profile case has been the "Jan 6" riot at the US Capitol in 2021. Gramlich (2022) of the Pew Research Centre found that in the months following the riot, both Democrats and Republicans became less accepting of party officials who openly criticised their respective party leaders. This finding suggests that public remonstrations may exacerbate the state of opinion divergence.

This paper provides a model to explain the dynamic connection between polarisation and revolts, focusing on a democratic context. I modify a spatial model of electoral competition to investigate whether there exists a two-way causal loop between the intensity of collective actions and the state of divergence in public opinion. The foundation of my model is Hotelling's (1929) spatial competition, which was extended by Smithies (1941), then Callander and Carbajal (2022) to allow for abstention by alienation. In my model, collective actions are incorporated as one of the citizens' tools of opinion expression besides voting. Turning out to vote and participating in a revolt both have different impacts on a citizen's ideal point in subsequent elections.

In the model, two endogenous forces drive voters' ideal points towards the extremes, causing polarisation. The first is each voter's actions, which have substantive impacts on her preferences. This setup is rooted in the findings in behavioural literature about post-decision attitude change. In contrast to the classical view that a fixed set of preferences dictates an individual's actions, recent behavioural findings suggest the act of voting may cause a voter's preferences to shift. Evidence collected from various elections—across the US, Canada, and the UK—shows that voters ex-post voting attempt to rationalise their choice by updating their preferences a little closer to the party for which they voted (Beasley

and Joslyn 2001, McGregor 2013, Bølstad, Dinas, and Riera 2013). This reverse causality serves as the foundation of the dynamic updating in my model: after each election, citizens who turn out to vote update their ideal points towards the party they support. Analogously, participating in a revolt moves a citizen's ideal point towards that of the opposition party.

The second and more powerful driving force is the positioning of strategic parties, who are both office- and policy-motivated. In the model, the risk of collective actions fundamentally changes the parties' strategic calculus. Following an election, the winning party is installed in office and governs for a set term, before voters decide whether to rebel against the election outcome. The parties begin at their ideal policy points and converge strategically to mobilise centralist abstainers. As voters' preferences update between election cycles, gaps emerge in the parties' support. Such gaps lead to diminished voter mass in the middle of the distribution, incentivising the parties to polarise.

Without strategic parties, the polarisation process can lead to once-moderate voters rebelling, but not a meaningful change in the intensity of collective actions. The acts of voting and rebelling both compress a citizen's ideal point towards that of her preferred party, leading to the formation of homogenised voting blocs. But since abstainers do not undergo such compression, the final state of this evolution is a bimodal voter distribution where only the original partisans engage in political actions.

It is only when combined with the strategic behaviours of the parties that polarisation results in a substantial increase in collective actions. When the reward of office is high, the parties' forward-looking attribute incentivises them to moderate at first in an attempt to maximise voter capture and minimise the probability of revolts. The parties converge incrementally in each election cycle, mobilising central abstainers to vote. As these voters update towards the party's ideal points, the parties re-polarise. Collective actions abate during the period of convergence and intensify during divergence, resulting in the citizenry plunging into persistent infighting that in turn exacerbates polarisation.

A surprise finding is that revolts may intensify even when the parties moderate. I show that fundamentally, mass polarisation is what matters. Elite polarisation—the state of divergence between the parties—does not directly influence a citizen's propensity to revolt. Rather, the parties' positioning determines whether the masses moderate or polarise, which in turn drives the intensity of collective actions.

The rest of the paper begins with an overview of the literature. I then outline the model and discuss its assumptions, theoretical and empirical foundation. Next, I characterise the equilibrium in the first election cycle, which is followed by an analysis of the full game with fixed party locations. Lastly, I analyse the equilibria with strategic parties, before offering discussions and a conclusion.

1.1 Related Literature

This paper follows the approach of Callander and Carbajal (2022) in modelling the movements of voter preferences, but is distinguished in several key aspects. Extending a classic one-dimensional voting model, Callander and Carbajal introduced a mechanism by which voters' preferences update following the act of voting. Specifically, a voter's ideal point moves closer towards the party for which she voted. Their model provided a way to rationalise the trend of polarisation in the US since the 1950s, and characterised a dynamic path for strategic candidates' campaign positioning. In this paper, I adopt their dynamic updating approach in a spatial electoral contest, but otherwise depart meaningfully from their model.

The first distinction of this paper is the focus on revolt as a driver of polarisation. In economics, studies on the causes of polarisation have produced a sizeable literature, analysing the effects of misinformation, media bias, and selective sharing in social networks (for a review see Tucker et al. 2018). However, in economic theory, relatively little attention has been paid to the effects of collective actions on citizens' preferences. My model addresses this research gap directly and provides one way to rationalise the extent and the speed at which unconventional political actions influence public ideologies.

Secondly, this paper's theoretical framework helps explain how polarisation influences the dynamics of collective actions. There is a rich political science literature that discusses why polarisation—often in narrower definitions—drives revolts. For example, Gurr (1970) argued that an increase in the variance of grievances between social groups leads to a surge in both violent and non-violent protests. Using Gurr's model of relative deprivation, Kleiner (2018) found a positive correlation between polarisation and protest behaviour in Europe. Yet, existing theories fall short of establishing a dynamic feedback loop between revolts and polarisation, in broader definitions of both concepts. Showing the existence of this feedback loop and examining its robustness can be crucial in advancing our understanding of voter behaviours.

Another contribution is investigating the dynamics of collective actions. Previous research on the topic has mostly modelled citizens as agents with a binary action space, who need to simply decide between revolting or not revolting (e.g. Shadmehr and Bernhardt 2011, Barberà and Jackson 2019). My model approaches the issue differently by treating the acts of voting, abstaining, and rebelling as components of a citizen's political toolkit. This setup reveals insight into the behaviour of individuals in choosing between adhering to and departing from a formal political system. In the game, the dynamic evolution of voter preferences leads to abstainers in earlier elections turning out, or even revolting in later ones.

In conjunction, I ask under what conditions does a one-shot revolt turn into persistent infighting between diverging citizen camps. Take protests as an example: despite a sub-

stantial proportion of recurring movements, the majority of protests around the world are one-shot events (Cantoni et al. 2023). Previous studies on the dynamics of protests have focussed on the citizen learning process and the interplay between protesters and the state (for instance Meirowitz and Tucker 2013, Kricheli, Livne, and Magaloni 2011). Yet, I know of no formal model that explores this topic in the context of polarising citizens. The dynamic updating in my model provides a way to rationalise this effect of polarisation on the persistence of collective action movements.

2 The Model

A two-period election cycle is repeated on a discrete timeline $t = 1, 2, ...$, starting from $t = 1$. Each cycle comprises an election followed by a post-election period.[1] In the n-th election, two political parties C and L simultaneously announce their campaign policies $c_n, l_n \in \mathbb{R}$. The winner takes office immediately and implements its policy for that election period.

A continuum of citizens with mass 1 has ideal points uniformly distributed in the voter space $\mathcal{V} = [-1, 1]$. Citizens are myopic and maximise periodic payoffs, separately for each election and post-election period. At the n-th election, an individual with ideal point $v_n \in \mathcal{V}$ observes the parties' campaign platforms and decides between voting for one of the parties or abstaining. Abstaining gives her a utility of 0, while voting for party $\mathrm{P} \in \{\mathrm{C}, \mathrm{L}\}$ with announced policy p_n, gives a periodic utility of

$$U^{\mathrm{V}}_{\text{election}}(\mathrm{P}; v_n) = \lambda - |p_n - v_n|,$$

where $\lambda > 0$ is the voter's policy tolerance. Hence, a citizen votes for

$$\begin{aligned} &\mathrm{C} \text{ if } |c_n - v_n| \leq |l_n - v_n| \text{ and } |c_n - v_n| \leq \lambda, \\ &\mathrm{L} \text{ if } |c_n - v_n| > |l_n - v_n| \text{ and } |l_n - v_n| \leq \lambda; \end{aligned}$$

and abstains otherwise.[2]

The key novelty of the model is the incorporation of citizen revolts. During each post-election period, an individual chooses whether to accept the result of election n, or challenge its legitimacy by participating in a revolt. A citizen's pure strategy in this period is a function $r : \mathcal{V} \times \{c_n, l_n\} \mapsto \{0, 1\}$, where 1 indicates rebelling and 0 not rebelling.

There are three exhaustive outcomes in each post-election period. Firstly, in the case of no revolution, the election n winner continues to implement its election policy. Secondly, if some citizens revolt but fail to upend the government, they are penalised while the implemented policy remains unaffected. Lastly, if a successful revolution takes place, the elected party is overthrown and replaced by the previous election loser. The latter then

[1] For instance, the 10th election cycle includes the election at $t = 19$ and the post-election period $t = 20$.
[2] The tie-breaking rule for a voter indifferent between the parties is unimportant.

implements a policy of its choice for the period. Regardless of the realised outcome, the parties compete anew in the election cycle $n+1$.

The parties have mixed motivations and are two-period forward-looking. Each derives utility from both holding office and implementing a policy that aligns with its ideology. Let $C, L \in [-1,1]$ respectively denote the ideal policy points of parties C and L. The stage game utility function for party C is therefore:

$$U_n^{\mathrm{C}} = \begin{cases} (1+\delta)(-|c_n - C| + \pi) & \text{if C is elected and remains in office} \\ -|c_n - C| + \pi + \delta(-|l_{\text{post-election}} - C|) & \text{if C is elected but overthrown} \\ -|l_n - C| + \delta(-|c_{\text{post-election}} - C| + \pi) & \text{if L is elected but overthrown} \\ (1+\delta)(-|l_n - C|) & \text{if L is elected and remains in office,} \end{cases}$$

where $\pi \geq 0$ is the benefit of office and $\delta \in (0,1)$ is a common discount factor. The utility function of party L is defined analogously with ideal point L.

Citizens balance two opposite incentives when deciding whether to revolt. On the one hand, participation in revolts incurs a cost of $\psi > 0$, which is homogeneous for all citizens. On the other hand, following a successful revolt, each rebel receives a reward based on her distance from the new government's policy given by

$$\rho(v_n; q_{\text{post-election}}, b, \gamma) = b - \gamma |q_{\text{post-election}} - v_n|,$$

where $q_{\text{post-election}}$ is the election n loser's policy implementation if it comes to power, and $b, \gamma > 0$ are constants.[3] Subsection 2.1 discusses the basis for this setup in detail.

Thus overall, a voter's utility in each post-election period is given by

$$U^{\mathrm{V}}_{\text{post-election}} = \begin{cases} -|p_n^* - v_n| & \text{if no revolt happens} \\ -|p_n^* - v_n| - \psi \mathbb{I}\{r = 1\} & \text{if a revolt takes place but fails} \\ -|q_{\text{post-election}} - v_n| + (\rho(\cdot) - \psi)\mathbb{I}\{r = 1\} & \text{if a revolt is successful,} \end{cases}$$

where $p_n^* \in \{c_n, l_n\}$ is the policy implemented at election n.

The outcome of elections and revolts in reality are both to some degree uncertain. In this paper, I follow Callander and Carbajal (2022) and define the probability of a party winning as the share of votes it receives. The foundation of this setup is the reduced-form approach of Calvert (1985), where candidates in an election of incomplete information are each assigned a subjective probability of winning.[4] Analogously, the probability of a successful revolt in the n-th election cycle is defined as the share of the citizens that participate,

[3] There are two additional technical assumptions on b, γ and ψ that ensure the boundaries of any rebel set are contained in the voter space. They do not impact the robustness of the results. See Appendix A.2.

[4] This formulation is additionally supported by empirical observation that some citizens vote spatially while others simply cast their votes ramdomly (Jessee 2009).

which can be written as

$$w_n(c_n, l_n; b, \gamma, \psi) = \int_{v \in \mathcal{K}_n} f_n(v)\, dv$$

where $f_n(\cdot)$ is the density function of the citizen distribution at election n, and $\mathcal{K}_n \subset \mathbb{R}$ is the set of rebels against the n-th election outcome.

Finally, the dynamics in the model arise from how the acts of voting and rebelling lead to movement in citizens' ideal points. For a citizen who voted in period t but did not participate in a revolt in period $t+1$, her ideal point at the following election cycle becomes

$$v_{n+1} = v_n + \tau_v(p_n - v_n),$$

where $\tau_v \in (0,1)$ is the speed of updating for voters. Analogously, for a citizen who participated in the revolt—regardless of her voting history at election t—her ideal point at the next election cycle becomes

$$v_{n+1} = v_n + \tau_r(q_n - v_n),$$

where $\tau_r \in [\tau_v, 1)$ is the speed of updating for rebels. There is no updating within cycles. Further discussion of this updating rule is contained in subsection 2.3.

2.1 Assumptions

To keep the model tractable, I impose the following four assumptions throughout. Firstly, the parties have symmetric ideal points $C = -L > 0$. Secondly, following Callander and Carbajal (2022), I assume these ideal points are sufficiently large so that the parties' campaign policies c_n, l_n are elements of the compact, non-empty policy space $\mathcal{G} = [L, C] \subset \mathcal{V}$. Thirdly, I assume that $\lambda < \frac{C-L}{2}$ so there is abstention in the middle in election 1. Lastly, to prevent the parties' support from overlapping, I assume c_n, l_n are bounded away from zero by λ.

As with many participation games, there may be a multiplicity of equilibria where different levels of revolt participation can be sustained. For tractability, I introduce the following definition and focus exclusively on *stable, outcome-invariant* revolt equilibria.

Definition 1. In each election cycle $n \geq 2$, a revolt probability $w_n^*(\cdot)$ is said to be:

1. *stable*, if it can be realised notwithstanding every citizen's expected revolt probability being $w_n^*(\cdot) + \epsilon$, where $\epsilon > 0$ is an arbitarily small constant; and

2. *outcome-invariant*, if $w_n^*(\cdot) = k \in \left[0, \frac{1}{2}\right]$ regardless of which party wins election n.[5]

[5]Note that by construction, $0 < w_n^*(\cdot) \leq \frac{1}{2}$ for all n. See appendix A.2 for more detail.

2.2 *Foundation of the Revolt Payoff Structure*

A voter's decision to rebel depends on the location of her ideal point in each election cycle. Her revolt payoff has three elements: a reward function $\rho(\cdot)$, a homogeneous cost $\psi > 0$, and the probability of success $w_n(\cdot)$. This setup is inspired by a rich line of costly voting literature that traces back to Riker and Ordeshook (1968).

The reward function $\rho(\cdot)$ is in principle comparable to the concept of *consumption benefits* in various turnout games (for a review see Feddersen 2004). As will become clear later on, in every post-election period, $\rho(\cdot)$ is single-peaked at the ideal point of the election-losing party. Thus, the marginal benefits of revolt are the highest for the most avid partisans, who have the strongest beliefs in their ideologies.

An alternative interpretation of $\rho(\cdot)$ draws from ethical agent models. In such games, a citizen's political engagements are assumed to be at least in part motivated by some altruistic considerations. Applied to my context, $\rho(\cdot)$ can be seen as each voter's level of conviction in her party's policy. Naturally, a citizen closer to the party's ideal point would have a more favourable belief in the benefit of her party's platform, and more willing to suffer the cost of a failed revolution. This interpretation is consistent with empirical findings: for example, Blais (2000) has shown that a sense of civic duty is an important driver for voter turnouts.

The probability of a successful revolt $w_n(\cdot)$ dictates each voter's strategy and is simultaneously determined by the aggregate actions of the voters. This seemingly paradoxical setup draws from Ledyard (1984), who showed that positive turnout equilibria exist in large elections if the voter distribution is common knowledge. Translated to this model, each voter's beliefs about others' strategies are correct in any Nash equilibrium. In conjunction, the indicator functions contained in $U^{\mathrm{V}}_{\text{post-election}}$ represent a voter's marginal utility of revolting, conditional upon failure and success.[6]

2.3 *Foundation of the Updating Rule*

The linear updating rule of citizen ideal points is inspired by Callander and Carbajal (2022). Many interpretations of this rule are valid, but two theories from behavioural science are the most relevant.

The first is Festinger's (1957) cognitive dissonance theory. According to his formulation, individuals unconsciously reevaluate their beliefs when presented with unchangeable external facts. In a political economy context, Festinger's theory implies a voter's preferences are not exogenously fixed. The second interpretation sprouts from Converse (1969). In this alternative concept, a voter's partisan preferences are simply the cumulation of her electoral experience. As one participates in more elections, her partisan outlook solidifies.

[6] This interpretation is consistent with Barberà and Jackson (2019).

Indeed, behavioural findings support the theories. During American presidential elections, Mullainathan and Washington (2009) documented that voters polarised two to three times more than non-voters. A similar effect was found by McGregor (2013) in Canadian federal elections. In addition, Bølstad, Dinas, and Riera (2013) has shown that in the UK, even purely tactical voters' preferences shifted towards the party for whom they voted.

Strong empirical evidence similarly suggests that collective actions have a substantive effect on preferences. Studying Ukraine's Euromaidan protests, Pop-Eleches, Robertson, and Rosenfeld (2022) found that participants' preferences over the core protest issues became more closely aligned. In post-Mubarak Egypt, Ketchley and El-Rayyes (2021) observed that street protests led to a deterioration in citizens' attitudes towards democracy.

3 The First Election Cycle

The first election cycle begins with a classic electoral contest, and proceeds to a game of participation in which citizens decide whether to challenge the election outcome.

With myopic voters, the parties' platforms in the first election balance the competing incentives of office and risk of collective actions. By moderating towards the centre, they reduce the probability of being overthrown in the post-election period. But moderation comes at the cost of departing from their ideal points, reducing the attractiveness of the policy outcome.

Proposition 1 establishes the equilibrium in the first election cycle. Each party chooses to locate at its ideal point and the election is split with equal probability. Its proof is contained in appendix A.

Proposition 1. *In the first election cycle, a unique symmetric Nash equilibrium exists with the parties locating at their ideal points: $c_1^* = C$ and $l_1^* = L$. The voters play the revolt strategy*

$$r^*(v_1; c_1, l_1) = \begin{cases} 1 & \text{if } v_1 \in \left[P - w_1^*(b,\gamma,\psi),\ P + w_1^*(b,\gamma,\psi)\right] \\ 0 & \text{otherwise,} \end{cases}$$

where $P \in \{C, L\}$ is the election losing party's ideal point and

$$w_1^*(b,\gamma,\psi) = \frac{b + \sqrt{b^2 - 4\gamma\psi}}{2\gamma}. \tag{1}$$

In the equilibrium, elite convergence is precluded by the assumption about voter tolerance. Because $\lambda < \frac{C-L}{2}$, the parties' support intervals do not intersect and abstention happens both among the central voters and the extreme ones. No citizen is indifferent between the two candidates, and the electoral competition does not revolve around the median voter. Rather, parties compete by mobilising their core bases and abandon the effort to appeal to the centrists.

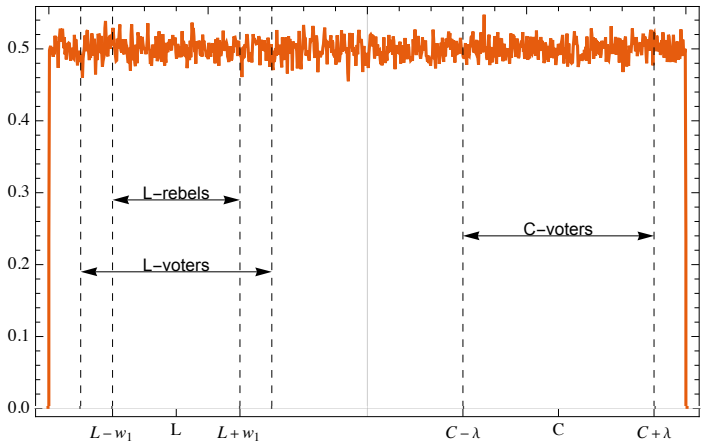


Figure 1: Simulated ideal-point distribution at the first cycle, assuming C wins election 1.

Figure 1 depicts the equilibrium configuration. The parties' positions represent an idealised "mobilising the base" strategy that has become dominant in American presidential elections (Panagopoulos 2016). Since 2000, there has been an explosive growth in the amount of individual-level data that can reveal a citizen's preferences and characteristics. Political campaigns are increasingly utilising big data technologies to precisely target voters, sending tailored messages to precisely defined subsets of their bases (Sides et al. 2023). For the candidates, encouraging these voters to turn out may be seen as a more reliable strategy than persuading swing voters.

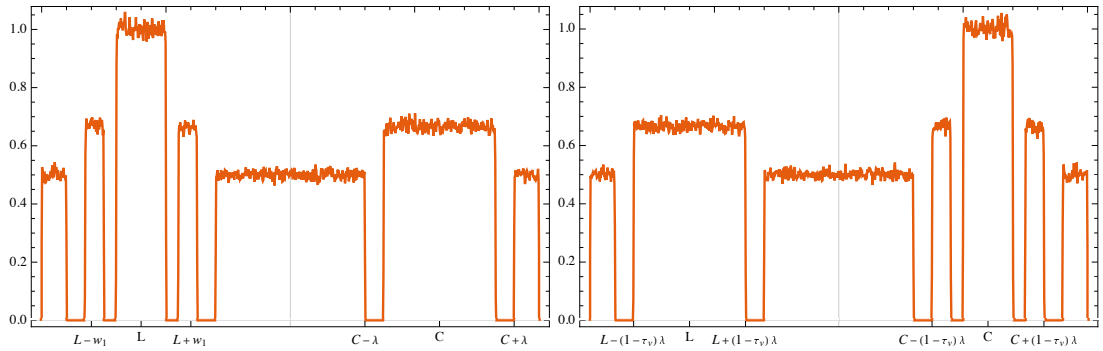
Following a loss, only the most avid partisans revolt in the post-election period. The singlepeakedness of the reward function $\rho(\cdot)$ means that the rebel set is symmetric about the election-losing party's ideal point, flanked by partisan voters on either side. In the subsequent cycles, these rebels and voters update their preferences, which leads to polarisation and greater intensity of revolt.

4 Long-run Evolution with Non-strategic Parties

The movements of voter ideal points are relatively intuitive when party ideal points are fixed. At the beginning of the second election cycle, voters and rebels update their preferences. This changes the distribution of the support, and affects the parties differently. For the election 1 winner, voter updating simply leads to its partisans compressing towards the party's ideal point. The marginal voters who were λ away from the party's ideal point now converge to positions only $(1-\tau_v)\lambda$ away. For the election loser, however, the cycle 1 rebels updated faster than the rest of the voters. They leave behind empty intervals within the party's support. Figure 2 visualises the two possible scenarios.

The long-run revolt probability can be similarly specified. First, one notes that the unique stable, outcome-invariant revolt probability at the second cycle is $w_2^*(\cdot) = \lambda$.[7] It

[7]For a proof see Appendix B.2.



(a) C wins election 1.

(b) L wins election 1.

Figure 2: Updated ideal-point distribution at the second election cycle.

follows that in each cycle $n \geq 2$, the rebel set contains precisely the voting population of the losing party. In other words, revolt becomes a complement to voting for election losers. An election-losing partisan polarises faster than a winner, with $\tau_r - \tau_v$ being the periodic difference in updating speeds. No abstainer moves into the voting range for either party, and the voter distribution becomes bimodal, with peaks at the parties' ideal points.

Combining the above observations, Proposition 2 describes the progression of voter ideal points with fixed party positions. Its proof is immediate.

Proposition 2. *Fix the parties' positions at $\hat{c} = C$ and $\hat{l} = L$. For all elections $n \geq 2$,*

1. *voters converge towards their preferred party's ideal point monotonically;*
2. *the periodic convergent speed is τ_v if a voter's preferred party wins and τ_r if it loses;*
3. *the average width of the empty intervals in $\mathcal{V}$ approaches λ, as $n \to \infty$;*
4. *the stable revolt probability $w_n^*(\cdot)$ is constant at λ; and*
5. *turnout remains at 2λ throughout.*

The key insight here is that without strategic candidates, polarisation alone does not lead to more revolts. The endogenous voter updating process eventually sorts the citizens into two homogenised blocs concentrated around the elites' ideal points. Neither the moderate nor extreme citizens would turn out, or have any sway in the outcomes of rebellions. In the following section, I reintegrate the parties' strategic behaviours. Their positioning choices lead to even more complex dynamics.

5 Reincorporating Strategic Parties

Strategic candidates seek to maximise their utility, which is a product of the election outcome, implemented policy, and risk of collective actions. Beyond the first election cycle, symmetric equilibria do exist and are differentiated by the reward of office π. When π

is low, parties are incentivised to converge in the short term to maximise the probability of winning the election. But this incentive quickly dissipates as voters compress towards the partisan policy points, which allows the parties to re-polarise. The incentive to further converge is dominated by the policy motivation, which pushes the parties back to their ideal points. When π is high, however, the office inventive dominates the policy incentive, pushing the parties to fully converge at first, but diverge again after mobilising all centralist abstainers.

The analysis in this section assumes party C wins the first election. Since both the initial voter distribution and the parties' ideal points are symmetric, the other equilibrium following C's loss is also exactly symmetrical. All proofs are contained in appendix B.

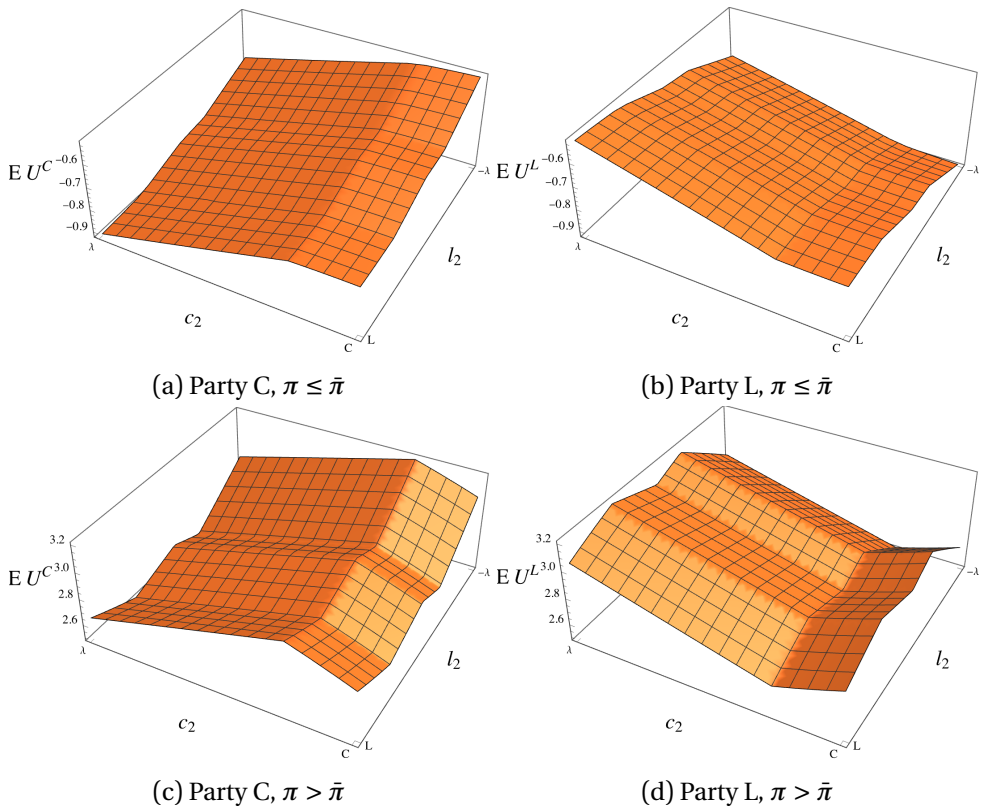
5.1 The Second Election Cycle

To begin, Lemma 1 defines the threshold office reward that separates the two equilibria.

Lemma 1. *There exists a critical value $\bar{\pi} > 0$ such that*

$$1 - \left(C - \lambda\tau_v + \frac{M(\bar{\pi})}{2A}\right)\left(\frac{1}{(2+\tau_v)\lambda}\right) = 0, \tag{2}$$

where shorthands $A \equiv 1 + (1-\lambda)\delta$ and $M(\bar{\pi}) = -2\lambda C\delta + [1 + \delta(1-2\lambda)]\bar{\pi}$.



(a) Party C, $\pi \leq \bar{\pi}$ (b) Party L, $\pi \leq \bar{\pi}$

(c) Party C, $\pi > \bar{\pi}$ (d) Party L, $\pi > \bar{\pi}$

Figure 3: Parties' expected payoffs at $n = 2$ as functions of policies.

Figure 3 shows the parties' expected payoffs as functions of the policies, for both cases of π. Given that voters updated after the first cycle, gaps formed in the voter distribution. These intervals of discontinuity imply the parties' utility functions are no longer smooth. Therefore, staying at the ideal point ceases to be the strategy that maximises winning probability. To further the analysis, Lemma 2 allows one to overcome the smoothness challenge in parties' payoffs.

Lemma 2. *For all $c_2 \in [l_2 + 2\lambda, +\infty) \cap \mathcal{G}$, the unique stable, outcome-invariant revolt probability is $w_2^*(b, \gamma, \psi) = \lambda$.*

This lemma says that the mass of the rebel set in cycle 2 is equal to that of a party's voting base, as long as the parties' support intervals do not intersect.[8] The intuition is twofold. First, $w_2^*(\cdot)$ is at least as high as $w_1^*(\cdot)$, since the previous rebels compressed towards their party and would rebel again should they continue to lose. On the other hand, the revolt probability would not be higher than λ if the election 1 winner loses, since no voters would join the revolt once the rebel set reaches the empty intervals in $\mathcal{V}$.

Given Lemma 2, the second election cycle equilibria are characterised by the following two propositions. In either case, strategic candidates converge towards the median voter, in order to increase their probability of winning. Each equilibrium is unique in its corresponding range of π and they are collectively exhaustive.

Proposition 3. *At the second election, the unique symmetric equilibrium for $0 \leq \pi \leq \bar{\pi}$ has the parties locating at $c_2^* = -l_2^* \in [C - \lambda\tau_v, C]$. Party C's position is implicitly defined by*

$$1 + \left(c_2^* + \frac{M(\pi)}{2A}\right)\left(\frac{1}{c_2^* - 2\lambda - C}\right) = 0. \tag{3}$$

Proposition 3 establishes the interior solution to the parties' maximisation problem. Each candidate converges slightly from their ideal points so that some middle abstainers are mobilised. Given that each party's most extreme voter moved inwards by a distance of $\lambda\tau_v$, the parties can converge by that same distance without losing any votes. In effect, the equilibrium position is where the elites balance two competing incentives: by moving inwards, it can mobilise more central abstainers and reduce the risk of collective actions; however, moving comes at the cost of distancing itself from its ideal policy.

As the reward of office increases, the equilibrium eventually switches to a corner solution. Proposition 4 states that parties simply maximise their total votes when the reward of office is large. The extent of each party's moderation is only constrained by the convergence of its most extreme voter.

[8] Recall that this holds by the fourth assumption in subsection 2.1.

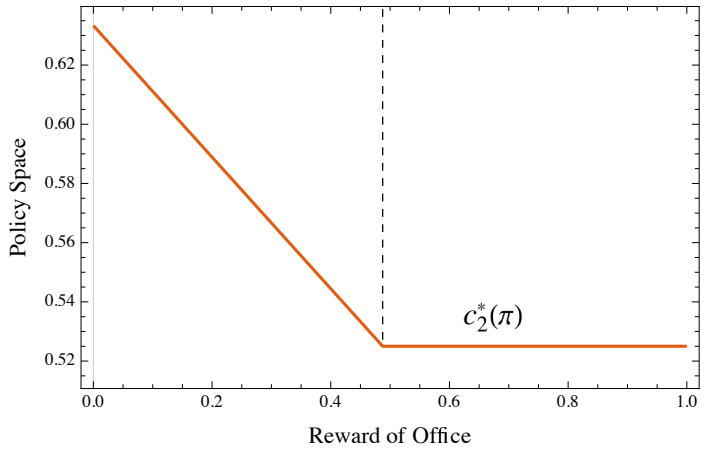


Figure 4: Equilibrium position of party C at the second cycle. The dashed line shows $\bar{\pi}$. See appendix for parameter values.

Proposition 4. *For $\pi > \bar{\pi}$, the unique symmetric equilibrium at the second election cycle has the parties locating at*

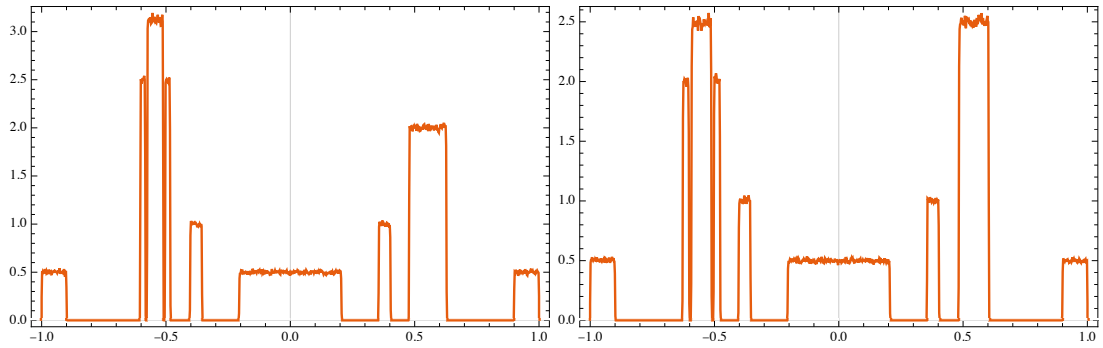
$$c_2^* = C - \lambda\tau_v$$
$$l_2^* = L + \lambda\tau_v.$$

The intuition of the corner equilibrium relates to the elites' trade-off between winning office and implementing ideal policies. When the reward of office outweighs the policy incentive, the parties become more willing to converge away from their ideal points. This policy moderation enables each to attract the voters in the centre, who had abstained in the first election. However, the incentive to converge is not unlimited. Since the boundary voter on the outer flanks only moved inwards by $\lambda\tau_v$, converging any more than this distance loses a party these votes on the outside. Candidates therefore stabilise at the positions in the above proposition for any higher values of π, as shown in Figure 4.

5.2 The Third Election Cycle and Beyond

At the third election cycle, voters' preferences update again. Figure 5 describes the distribution following each outcome of the second election. Notably, for both scenarios, the mobilised centralist voters polarise by $\lambda\tau_v$. Their updating allows the parties to re-polarise. When $\pi \leq \bar{\pi}$, this happens immediately: the parties' incentive to further converge is dominated by the policy motivation, which leads them to return to their ideal points. Revolt probability thus increases, as the newly mobilised voters polarise towards the party ideal points and become core partisans. These observations are summarised by Proposition 5.

Proposition 5. *For $0 \leq \pi \leq \bar{\pi}$, the unique symmetric equilibrium in election cycle $n \geq 3$ exists with the parties returned to their ideal points: $c_n^* = C$ and $l_n^* = L$. Revolt probability stabilises at $w_n^*(\cdot) = \lambda + \frac{1}{2}\left(C - c_2^*\right)$, as $n \to \infty$.*



(a) Party C re-elected at election 2.

(b) Party C not re-elected.

Figure 5: Simulated ideal-point distribution at the third election cycle.

This result illustrates the profound impact of strategic elite behaviour. Surprisingly, even a transient period of convergence leads to a permanent, non-trivial increase in collective actions. Compared to the benchmark with non-strategic parties, the candidates' movements mobilise central citizens into political actions. These citizens then updates closer to the elite ideal points in each cycle, eventually becoming polarised partisans perfectly aligned with the elites' ideologies. Counter-intuitively, mass polarisation increases in the long run despite the temporary decrease in elite polarisation. Figure 6a visualises the path of the parties.

On the other hand, the equilibrium with high office incentive is more complicated, and may not be symmetric. In this configuration, parties' incentive to win office outweighs their policy motivation, resulting in policy moderation in every election until they reach either the lower bound or the interior maximiser. Each election sees the parties attempt to mobilise centralist abstainers without losing any existing voters on the outer flank.

Proposition 6. *If $\pi > \bar{\pi}$, there exists a symmetric equilibrium iff $\tau_v = \tau_r = \tau$. Parties locate at $c_n^* = -l_n^* \in [\lambda, C]$ for all $n \geq 3$, and the evolution of c_n^* and $w_n^*(\cdot)$ satisfy*

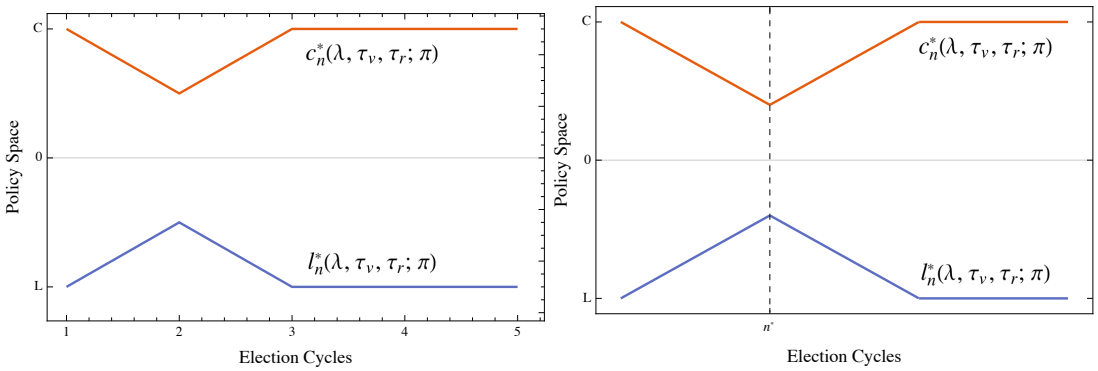
$$c_n^* = \max\left\{c_{n-1}^* - \lambda\tau, c^\circ, \lambda\right\} \quad \textit{and} \quad w_n^*(\cdot) \leq w_{n-1}^*(\cdot)$$

for election cycles $3 \leq n \leq n^*$, while

$$c_n^* = \min\left\{c_{n-1}^* + \lambda\tau, C\right\} \quad \textit{and} \quad w_n^*(\cdot) \geq w_{n-1}^*(\cdot)$$

for $n > n^*$. The position c° is the implicit solution to (B1) and $n^* \in \mathbb{Z} \cap [3, \infty)$. Moreover, $w_n^*(\cdot) = \min\left\{\lambda + \frac{1}{2}(C - c^\circ), \frac{1}{2}\right\}$ as $n \to \infty$.

Proposition 6 describes a dynamic progression of two stages. As a direct consequence of elite convergence, more voters update towards zero than the extremes in the short-run. Revolt probability decreases during this period as voters move away from the party ideal points. After the parties reach the minimum, however, the mechanism is reversed. With



(a) $\pi \leq \bar{\pi}$ (b) $\pi > \bar{\pi}$, dashed line is n^*

Figure 6: Evolution of campaign platforms.

the newly mobilised central voters, each party's dominant strategy is to incrementally re-polarise towards its ideal point. Figure 6b depicts this movement trend.

6 Empirical Discussion

The model can rationalise several empirical observations about polarisation and collective actions. In particular, the key prediction about party positioning—that they converge before polarising—fits the data reasonably well. As an example, the United Kingdom experienced a moderation in public opinion during 1990-2000, followed by a period of divergence from 2000-2016. More broadly, evidence from other OECD countries suggests that the trends of preference movements are often monotonic, if not linear (Boxell, Gentzkow, and Shapiro 2024). However, the empirical literature on polarisation remains thin and more detailed studies may be needed to confirm the model's predicted evolution.

Furthermore, the model provides one way to explain the relatively stable turnouts in US presidential elections despite the worsening of polarisation (McDonald and Popkin 2001). The low π equilibrium path given by Proposition 3 and Proposition 5 can explain this phenomenon well. With low voter tolerance and small reward of office, the mobilisation effect of elite convergence is limited: only a small fraction of abstainers turn out to vote in subsequent elections, and both the turnout and the probability of revolt remain stable. However, polarisation within the voting population still increases monotonically over time.

An ancillary implication relates to the observation that an individual is more likely to vote in future elections if her preferred candidate lost (Kanazawa 2000). For the election-losing partisans in the model, collective action eventually becomes a complement to voting when they become sufficiently compressed around their party. Therefore, even the once-moderate citizens find it strictly more profitable to vote and participate in revolts in an attempt to achieve a "lesser of two evils" policy outcome. Rebelling leads to the individuals polarising faster, which in turn can rationalise the asymmetrical extent of polarisation in the US (Bartels 2016).

The main limitation of the model is that it's only tractable to define the relative evolution of revolt probability and its value at the limits. A more refined model should consider a configuration that will allow us to characterise revolt probability in every period.

7 Conclusion

This paper has constructed a model to explain the two-way causal loop between political polarisation and collective actions. I have shown that collective action accelerates the process of mass polarisation, particularly so for election-losing partisans. With a higher degree of polarisation, more voters resort to collective actions to voice their political opinions, and revolt becomes a complement to voting.

Without strategic parties, endogenous updates in voter preferences alone have a limited impact on the probability of revolt in the long run. In such a configuration, abstainers are not tempted to vote or participate in collective action. Rather, gaps open up in parties' support as voters compress towards their preferred party's ideal point. Polarisation only exacerbates for the voters, while abstainers remain depolarised.

It is only when the parties strategically converge to capture voters, that mass polarisation leads to a meaningful increase in the intensity of collective actions. Surprisingly, this two-way causal loop holds even when elite polarisation decreases. When the reward of office is low, parties moderate their campaign positions in the short run to mobilise centralist abstainers but quickly diverge again as voters compress towards their positions. These voters eventually polarise close enough to the party's ideal points that they start participating in collective actions, which in turn exacerbates their intensity.

When the reward of office is high, parties have an even stronger incentive to converge, ultimately mobilising most of the central abstainers. Both polarisation and collective action abate during the period of convergence and exacerbate during the period of divergence.

The model also offers several possible extensions for future research. To begin, empirical work can examine whether other forms of political participation have a comparable effect on voter preferences. If actions such as attending rallies or volunteering in campaigns can also induce cognitive dissonance, future models may seek to consider even more heterogeneous voter characteristics beyond only distinguishing between voters, rebels, and abstainers. Secondly, it is worth verifying the asymmetry in voter polarisation in elections around the world and whether an election-losing partisan group display a higher propensity to participate in revolts. Thirdly, future studies on collective actions can incorporate the degree of political participation into a citizen's characteristics profile. In this model, voting and revolt participation are each on their own binary actions. A more refined model can consider whether a core partisan votes or participates in collective actions more rigorously than a voter further away from a party's ideal point.

References

Alesina, Alberto, Armando Miano, and Stefanie Stantcheva (2020). "The Polarization of Reality". *AEA Papers and Proceedings*, 110, pp. 324–328. DOI: 10.1257/pandp.20201072.

Baker, Scott R., Nicholas Bloom, Brandice Canes-Wrone, Steven J. Davis, and Jonathan Rodden (2014). "Why Has US Policy Uncertainty Risen Since 1960?" *American Economic Review*, 104 (5), pp. 56–60. DOI: 10.1257/aer.104.5.56.

Barberà, Salvador and Matthew O. Jackson (2019). "A Model of Protests, Revolution, and Information". *SSRN Electronic Journal*, DOI: 10.2139/ssrn.2732864.

Bartels, Larry (2016). "Failure to Converge: Presidential Candidates, Core Partisans, and the Missing Middle in American Electoral Politics". *The Annals of the American Academy of Political and Social Science*, 667 (1), pp. 143–165. DOI: 10.1177/0002716216661145.

Beasley, Ryan K. and Mark R. Joslyn (2001). "Cognitive Dissonance and Post-Decision Attitude Change in Six Presidential Elections". *Political Psychology*, 22 (3), pp. 521–540. DOI: 10.1111/0162-895X.00252.

Blais, André (2000). *To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory*. University of Pittsburgh Press. DOI: 10.2307/j.ctt5hjrrf.

Bølstad, Jørgen, Elias Dinas, and Pedro Riera (2013). "Tactical Voting and Party Preferences: A Test of Cognitive Dissonance Theory". *Political Behavior*, 35 (3), pp. 429–452. DOI: 10.1007/s11109-012-9205-1.

Boxell, Levi, Matthew Gentzkow, and Jesse M. Shapiro (2024). "Cross-Country Trends in Affective Polarization". *Review of Economics and Statistics*, 106 (2), pp. 557–565. DOI: 10.1162/rest_a_01160.

Callander, Steven and Juan Carlos Carbajal (2022). "Cause and Effect in Political Polarization: A Dynamic Analysis". *Journal of Political Economy*, 130 (4), pp. 825–880. DOI: 10.1086/718200.

Calvert, Randall L. (1985). "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence". *American Journal of Political Science*, 29 (1), pp. 69–95. DOI: 10.2307/2111212.

Canen, Nathan, Chad Kendall, and Francesco Trebbi (2020). "Unbundling Polarization". *Econometrica*, 88 (3), pp. 1197–1233. DOI: 10.3982/ECTA16756.

Cantoni, Davide, Andrew Kao, David Y. Yang, and Noam Yuchtman (2023). *Protests*. Working Paper. DOI: 10.3386/w31617.

Converse, Philip E. (1969). "Of Time and Partisan Stability". *Comparative Political Studies*, 2 (2), pp. 139–171. DOI: 10.1177/001041406900200201.

Feddersen, Timothy J (2004). "Rational Choice Theory and the Paradox of Not Voting". *Journal of Economic Perspectives*, 18 (1), pp. 99–112. DOI: 10.1257/089533004773563458.

Festinger, Leon (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Gramlich, John (2022). *A look back at Americans' reactions to the Jan. 6 riot at the U.S. Capitol*. URL: https://www.pewresearch.org/short-reads/2022/01/04/a-look-back-at-americans-reactions-to-the-jan-6-riot-at-the-u-s-capitol/.

Gurr, Ted Robert (1970). *Why Men Rebel*. Princeton University Press.

Hotelling, Harold (1929). "Stability in Competition". *The Economic Journal*, 39 (153), pp. 41–57. DOI: 10.2307/2224214.

Jessee, Stephen A. (2009). "Spatial Voting in the 2004 Presidential Election". *American Political Science Review*, 103 (01), pp. 59–81. DOI: 10.1017/S000305540909008X.

Kanazawa, Satoshi (2000). "A New Solution to the Collective Action Problem: The Paradox of Voter Turnout". *American Sociological Review*, 65 (3), pp. 433–442. DOI: 10.2307/2657465.

Ketchley, Neil and Thoraya El-Rayyes (2021). "Unpopular Protest: Mass Mobilization and Attitudes to Democracy in Post-Mubarak Egypt". *The Journal of Politics*, 83 (1), pp. 291–305. DOI: 10.1086/709298.

Kleiner, Tuuli-Marja (2018). "Public opinion polarisation and protest behaviour". *European Journal of Political Research*, 57 (4), pp. 941–962. DOI: 10.1111/1475-6765.12260.

Kricheli, Ruth, Yair Livne, and Beatriz Magaloni (2011). *Taking to the Streets: Theory and Evidence on Protests under Authoritarianism*. SSRN Scholarly Paper. Rochester, NY. URL: https://papers.ssrn.com/abstract=1642040.

Ledyard, John O. (1984). "The Pure Theory of Large Two-Candidate Elections". *Public Choice*, 44 (1), pp. 7–41. URL: https://www.jstor.org/stable/30023935.

McDonald, Michael P. and Samuel L. Popkin (2001). "The Myth of the Vanishing Voter". *American Political Science Review*, 95 (4), pp. 963–974. DOI: 10.1017/S0003055400400134.

McGregor, R. Michael (2013). "Cognitive dissonance and political attitudes: The case of Canada". *The Social Science Journal*, 50 (2), pp. 168–176. DOI: 10.1016/j.soscij.2013.01.004.

Meirowitz, Adam and Joshua A. Tucker (2013). "People Power or a One-Shot Deal? A Dynamic Model of Protest". *American Journal of Political Science*, 57 (2), pp. 478–490. DOI: 10.1111/ajps.12017.

Mitchell, Amy, Jeffrey Gottfried, Jocelyn Kiley, and Katerina Eva Matsa (2014). *Political Polarization and Media Habits*. URL: https://www.pewresearch.org/journalism/2014/10/21/political-polarization-media-habits/.

Mullainathan, Sendhil and Ebonya Washington (2009). "Sticking with Your Vote: Cognitive Dissonance and Political Attitudes". *American Economic Journal: Applied Economics*, 1 (1), pp. 86–111. DOI: 10.1257/app.1.1.86.

Oliver, J. Eric (2001). *Democracy in Suburbia*. Princeton University Press. DOI: 10.1515/9780691223360.

Panagopoulos, Costas (2016). "All about that base: Changing campaign strategies in U.S. Presidential elections". *Party Politics*, 22 (2), pp. 179–190. DOI: 10.1177/1354068815605676.

Pop-Eleches, Grigore, Graeme Robertson, and Bryn Rosenfeld (2022). "Protest Participation and Attitude Change: Evidence from Ukraine's Euromaidan Revolution". *The Journal of Politics*, 84 (2), pp. 625–638. DOI: 10.1086/716302.

Riker, William H. and Peter C. Ordeshook (1968). "A Theory of the Calculus of Voting". *American Political Science Review*, 62 (1), pp. 25–42. DOI: 10.2307/1953324.

Shadmehr, Mehdi and Dan Bernhardt (2011). "Collective Action with Uncertain Payoffs: Coordination, Public Signals, and Punishment Dilemmas". *American Political Science Review*, 105 (4), pp. 829–851. DOI: 10.1017/S0003055411000359.

Sides, John, Daron R. Shaw, Matthew Grossmann, and Keena Lipsitz (2023). *Campaigns and elections: rules, reality, strategy, choice*. Fourth edition, election update. New York London: W.W. Norton & Company.

Smithies, Arthur (1941). "Optimum Location in Spatial Competition". *Journal of Political Economy*, 49 (3), pp. 423–439. DOI: 10.1086/255724.

Tucker, Joshua A., Andrew Guess, Pablo Barbera, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan (2018). *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature*. SSRN Scholarly Paper. Rochester, NY. DOI: 10.2139/ssrn.3144139.