DISCUSSION PAPER SERIES

DP16037

Using Re-election Thresholds to Curb Political Polarization

Hans Gersbach, Philippe Muller and Oriol Tejada

ORGANIZATIONAL ECONOMICS

PUBLIC ECONOMICS

ISSN 0265-8003

Using Re-election Thresholds to Curb Political Polarization

Hans Gersbach, Philippe Muller and Oriol Tejada

Discussion Paper DP16037
Published 14 April 2021
Submitted 01 April 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Organizational Economics
- Public Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Hans Gersbach, Philippe Muller and Oriol Tejada

Using Re-election Thresholds to Curb Political Polarization

Abstract

We examine how tightening reelection hurdles for incumbents can curb political polarization and increase welfare. We use a two-period model in which a politician is elected for office in the first period and enacts a new policy. In the second period, elections take place between the incumbent and a challenger, and the winning candidate chooses the extent to which the first-period policy is reformed. Reforming a policy is costly, and such costs increase with the policy shift and are borne by parties and voters. We show that raising the vote-share needed for re-election above one half reduces policy polarization and increases welfare. Moreover, the latter measures depend on the re-election threshold in a non-monotonic way and a particular (intermediate) threshold simultaneously minimizes policy polarization and maximizes welfare.

JEL Classification: C72, D72, D78, H4

Keywords: Elections, Political Polarization, costs of change, re-election hurdles

Hans Gersbach - hgersbach@ethz.ch
ETH Zurich and CEPR

Philippe Muller - pmuller@ethz.ch
ETH Zurich

Oriol Tejada - toriol@ethz.ch
ETH Zurich

Acknowledgements
We would like to thank Marco Battaglini, Dan Bernhardt, Volker Britz, Volker Hahn, Helios Herrera, Niall Hughes, Matthew Jackson, Aniol Llorente-Saguer, César Martinelli, Rebecca Morton, Santiago Oliveros, Marco Ottaviani, Thomas Palfrey, Mattias Polborn, Ronny Razin, Maik Schneider, Guido Tabellini, Christoph Vanberg, Bauke Visser, Richard van Weelden, Jan Zapal, participants in various seminars in Zurich and Barcelona, in the meeting of the Society of Social Choice and Welfare, and in the Priorat Workshop in Theoretical Political Science for helpful comments and suggestions.

Using Re-election Thresholds to Curb Political Polarization*

Hans Gersbach
CER-ETH–Center of Economic
Research at ETH Zurich and CEPR
Zuerichbergstrasse 18
8092 Zurich, Switzerland
hgersbach@ethz.ch

Philippe Muller**
CER-ETH–Center of Economic
Research at ETH Zurich
Zuerichbergstrasse 18
8092 Zurich, Switzerland
pmuller@ethz.ch

Oriol Tejada
CER-ETH–Center of Economic
Research at ETH Zurich
Zuerichbergstrasse 18
8092 Zurich, Switzerland
toriol@ethz.ch

This Version: April 2021

Abstract

We examine how tightening reelection hurdles for incumbents can curb political polarization and increase welfare. We use a two-period model in which a politician is elected for office in the first period and enacts a new policy. In the second period, elections take place between the incumbent and a challenger, and the winning candidate chooses the extent to which the first-period policy is reformed. Reforming a policy is costly, and such costs increase with the policy shift and are borne by parties and voters. We show that raising the vote-share needed for re-election above one half reduces policy polarization and increases welfare. Moreover, the latter measures depend on the re-election threshold in a non-monotonic way and a particular (intermediate) threshold simultaneously minimizes policy polarization and maximizes welfare.

Keywords: elections; political polarization; costs of change; re-election hurdles

JEL Classification: C72, D72, D78, H4

*We would like to thank Marco Battaglini, Dan Bernhardt, Volker Britz, Volker Hahn, Helios Herrera, Niall Hughes, Matthew Jackson, Aniol Llorente-Saguer, César Martinelli, Rebecca Morton, Santiago Oliveros, Marco Ottaviani, Thomas Palfrey, Mattias Polborn, Ronny Razin, Maik Schneider, Guido Tabellini, Christoph Vanberg, Bauke Visser, Richard van Weelden, Jan Zapal, participants in various seminars in Zurich and Barcelona, in the meeting of the *Society of Social Choice and Welfare*, and in the *Priorat Workshop in Theoretical Political Science* for helpful comments and suggestions. All errors are ours.

**Philippe Muller has been funded by SNF (Project no. 100010_153482/1).

1 Introduction

Motivation

Polarization in policy as well as in party and voter preferences—generically called *political polarization*—is a phenomenon in democracy, and has been on the rise in many countries for the last decades. This development has drawn a lot of attention from media as well as from the public at large. From an academic perspective, an extensive strand of literature has documented a significant increase in political polarization in the U.S. and examined its causes and potential welfare implications (Prior, 2013; McCarty et al., 2016; Gentzkow, 2016; Boxell et al., 2017; Benkler et al., 2018, see Section 2).

In this paper, we examine a complementary issue. We suggest that high(er) levels of political polarization may justify re-election hurdles (equivalently, re-election thresholds) that are higher than the standard hurdle used for the election for a first term, viz. the majority rule. Higher re-election hurdles can increase welfare in various ways, and we argue that one way in which they can do so is by inducing office-holders to adopt more moderate policy positions. This lowers policy polarization. Re-election huddles can also improve welfare if they are suitably chosen to moderate policy while avoiding the excessive costs associated with turnovers.

Re-election hurdles that are tighter than the majority rule have been *explicitly* used in practice. In the Swiss canton of Zurich, long-term Social Democrat deputies of the Swiss National Council currently need two thirds of the votes to be eligible for the next term (see Art. 7 in SP Kanton Zürich (2010)).[1] In Argentina, in a text of the constitution that failed to be approved in 1819, the Head of State could only be re-elected for a second term if s/he obtained a two-third majority of votes (Article 73) (see e.g. Presa, 2019). In Liberia, in the constitutional text of 1955, *(a)* a second term is prohibited unless the majority votes for it and *(b)* if voted on and elected, the incumbent's second term is shortened (Article 3.1) (see e.g. Steinberg, 1962): *"The Supreme Executive Power shall be vested in a President who shall be elected by the people, and shall hold his/her office for a term of eight years. No President may be elected for two consecutive terms of eight years, but should a majority of the ballots cast at a second or any other succeeding election by all of the electors voting thereat elect him/her, his/her second or any other succeeding term of office shall be for four years"*. Implicit higher re-election hurdles have also emerged *de facto*, e.g. in elections for mayors in the US through the combination of term limits and supermajority needed for constitutional changes. These and other examples are discussed in Section 5.3.

[1] For the 2011 elections to the Swiss National Council, three candidates had to reach this two-thirds majority in order to be nominated for the elections. One of them did not make it, see `http://www.nzz.ch/aktuell/startseite/sp-nationalraetin-anita-thanei-quorum-1.10577885` (retrieved 13 May 2019).

While our paper mainly adopts a normative approach, our (mathematical) analysis and results can also be interpreted from a positive perspective since re-election thresholds differing from 50% can manifest themselves *implicitly* as a result of different features of elections, as well as a result of politician and voter behavior; we also refer to Section 5.3. This extends the scope of our paper.

Model

To see whether higher re-election hurdles for incumbents could reduce policy polarization, we analyze a two-period model in which reforming a freshly-enacted policy is costly. The *costs of change* increase (linearly) with the magnitude of the policy reform. The linearity assumption simplifies the analysis but our results are not knife-edge. The crucial assumption is that the "degree of convexity" of the cost-of-change term is not larger than that of the policy utility term. This suffices to create an electoral advantage for the incumbent under the standard re-election rule.[2] The origins of costs of change are manifold, but basically they either originate because additional resources are needed to change the status quo or because the policy reform itself creates disutilities for citizens; we refer to Gersbach et al. (2020b) for an in-depth description of such costs (see also Section 2).

At the beginning of each period, two candidates with ideal policies coinciding with the ideal policies of their parties compete in an election for an executive office. Once in office, the winning candidate faces issues in two different dimensions. First, s/he has to make a decision with regard to a *policy* on which citizens have dissenting preferences. Second, s/he must carry out a *public project* that benefits every citizen, including himself or herself, in the same way. Candidates cannot commit to a policy before the election. Additionally, candidates diverge in their *ability* to carry out the public project. If they belong to different political parties, they also diverge in their preferred policy position. The ability of a candidate can be low, average, or high, and this only becomes apparent (for the candidate himself/herself, the other candidate, and the electorate) at the end of his/her first term, i.e. after the public project has been carried out and the policy has been implemented. This means that the public project serves the sole purpose of revealing the office-holder's competence or valence. Considering different abilities enriches the model and yields valuable insights regarding tail events. The ability dimension can be interpreted in forms that are not candidate-specific but simply establish a difference between the incumbent party and the challenger party.

We pay particular attention to one key variable of democratic elections: the *re-election hurdle*. This variable denotes the vote-share threshold that an incumbent has to attain to be re-elected. Although in most elections, the re-election hurdle is set at 50%, it could be—and has been—changed to different values, either directly or indirectly, as we have mentioned. If an incumbent does not obtain enough

[2] This is *not* a universal property of costs of change (see Gersbach et al., 2020a).

votes to overcome the re-election hurdle, the challenger is elected. In the main body of the paper, we only consider re-election hurdles equal to or higher than 50% and assume that they are exogenously fixed. Lower re-election hurdles are discussed in Section 5.3, while in Section 6 we allow politicians to individually choose the re-election threshold to which they commit. Throughout the paper, we also use the term *extra-hurdle* to denote a share of votes that, when added to a 50% share of the votes, constitutes the re-election hurdle.[3]

Main results

In the presence of costs of policy change, we find three main results regarding the effects of extra-hurdles on policy polarization and welfare. First, higher re-election hurdles induce office-holders to adopt policies in the first period that are closer to the median voter's ideal policy, so *policy polarization*—i.e. the average distance between the policies chosen by the two parties—declines with non-zero extra-hurdles. Rather than the median voter, a more partisan voter is decisive for re-election when extra-hurdles are positive. Policy moderation then follows from the fact that the critical voter and the incumbent are located on opposite sides of the median voter. The magnitude of the moderating effect depends on the exact value of the positive extra-hurdle. For low values, the incumbent has to moderate his/her first-period policy choice to remain attractive to the critical voter and thus to maintain his/her electoral advantage. If the extra-hurdle is above some critical level, however, s/he is less prone to compromising because his/her chances for re-election are low. Still, to reduce the costs of change incurred if s/he is ousted, s/he chooses a more moderate policy than his/her bliss point.

The second result concerns the impact of extra-hurdles on welfare. The welfare of the median voter depends on three components: the policies chosen in both periods, the public project carried out in each period, and the costs associated with policy changes. Raising the extra-hurdle from 0% to some level below the critical level referred to above neither affects expected the benefits from public projects nor the likelihood that an incumbent is ousted. But it reduces policy polarization and the expected costs of change and thus increases welfare. If an extra-hurdle above the critical level is implemented, policy polarization is still reduced compared to no extra-hurdle being implemented. However, in addition to this, the likelihood of turnover rises, which leads to higher expected costs of change and thus to lower welfare. Nevertheless, we find that, compared to a customary 50% majority, an extra-hurdle above the critical level increases welfare, as the reduction in policy polarization offsets the increase in the expected costs of change.

[3]The implementation of an extra-hurdle strictly larger than 0% but below 50% can be viewed as a weaker and flexible form of term limits. Similarly to term limits, extra-hurdles treat incumbents and challengers differently.

Our third result is that the critical extra-hurdle mentioned above is *optimal* in the sense that it maximizes welfare and minimizes policy polarization. The optimal extra-hurdle moderates policy without increasing the costs associated with turnover. It depends on several parameters of the model. This dependence yields insightful comparative statics. Moreover, since below the optimal extra-hurdle a marginal increase of the extra-hurdle always weakly increases welfare and weakly reduces policy polarization, marginal changes in re-election hurdles can be seen as a cautious approach for experimentation in democracy focused on directional welfare improvements.

Organization of the paper

The paper is organized as follows: In Section 2 we review the papers related to our article. In Section 3 we outline our baseline two-period model, where costs of change are linear in the absolute difference between the policies chosen in the two periods. In Section 4 we compute equilibrium policy choices and re-election probabilities as a function of the extra-hurdle. In Section 5 we study how this parameter affects policy polarization and welfare. In Section 6 we allow candidates to endogenously commit to a re-election threshold. In Section 7 we investigate the role of social polarization in the optimal design of re-election hurdles. Section 8 concludes. The proofs are given in Appendix A. Finally, Appendix B contains the analysis of large costs of change, while Appendix C contains extensions of the baseline model that include the existence of a status-quo policy at the beginning of the first period and non-linear costs of change (see supplementary material).

2 Relation to the Literature

The present paper is related to several strands of the literature.

Re-election thresholds

Some papers have investigated whether re-election thresholds that differ from the standard majority rule can be used to increase social welfare through different channels such as learning-by-doing, public good provision, effort, and selection (see e.g. Gersbach and Liessem, 2008; Gersbach et al., 2010; Gersbach and Müller, 2017; Gersbach, 2017, 2020). We contribute to this topic by examining yet another potential benefit of higher hurdles for incumbents: if there exist significant costs of change, such hurdles can moderate policy without increasing costly turnover.

Electoral competition

Our model of candidate competition for winner-take-all elections shares features with the classic literature on electoral competition. As in Hansson and Stuart (1984), Duggan and Fey (2005), and Krasa and Polborn (2010a,b), for example, each candidate is both office-motivated and policy-motivated and has some exogenous characteristics. In our model, these are his/her most preferred policy and his/her randomly determined ability. As is also standard, voters and candidates have quadratic—and hence single-peaked—preferences about policies chosen in the two periods. Due to the existence of costs of change, however, they also care about the difference between the two policies. This makes agents' preferences not separable in time.

Costs of change in policies

Few models have analyzed electoral competition by explicitly incorporating costs associated with changes in policies (see e.g. Glazer et al., 1998; Gersbach et al., 2019a; Gersbach and Tejada, 2018; Gersbach et al., 2020a,b). The so-called costs of change can moderate policy if politicians lack commitment power and they are alike regarding how efficiently they can carry out reforms, but they can also lead to more extreme policies in other setups. We contribute to the (growing) literature investigating costs of change by showing that in the presence of such costs, raising the bar for re-election can curb policy polarization.

Dynamic links

Many papers have investigated models where choices in one period affect either choices or utilities in the next period, from which we have learned how this possibility can sometimes be exploited for partisan goals (see Persson and Svensson, 1989; Alesina and Tabellini, 1990; Callander et al., 2017; Bowen et al., 2014; Baron and Bowen, 2015; Chen and Eraslan, 2017; Bowen et al., 2017; Dziuda and Loeper, 2016; Bouton et al., 2016; Zápal, 2016; Buisseret and Bernhardt, 2017; Dziuda and Loeper, 2018; Baron, 2018, among many many others). In this context, it is worth mentioning Forand (2014) and Nunnari and Zápal (2017), as they are closest to our paper in that candidates fully commit to a policy before the election and policies are bound not to change as long as the office-holder stays in office. Although our model features policy persistence for the incumbent's choice, this occurs only in equilibrium. Moreover, our approach is different because costs of change affect all subsequent office-holders, not just the incumbent.

Political polarization

There is a large body of literature that deals with political polarization, both from a theoretical and an empirical point of view. Most papers have analyzed the existence of, and trends in, political polarization, focusing on causes and consequences. In the case of the United States, the existence and rapid increase of policy polarization is well-established, as demonstrated by the evolution in the voting patterns of Democratic and Republican legislators of the House of Representatives and the

Senate (se e.g. Poole and Rosenthal, 1984, 2001; McCarty et al., 2006). These and other papers generally inform us that three manifestations of political polarization (party, social, policy) can influence each other.[4] Our paper adds to this literature by proposing a novel institutional feature for winner-take-all elections that curbs policy polarization without creating welfare losses.

This can be important because, following the rapid increase of political polarization, scholarly concerns about the role of polarization in democracies are now far-ranging. Although there is some controversy about whether high levels of political polarization are generally beneficial or detrimental for a society, there evidently may be strong negative consequences. For instance, increased party polarization may undermine trust in the policy-making process and lead to legislative gridlock (Jones, 2001; Binder, 1999). Moreover, high party polarization encourages disinterest in politics, party disidentification, and a decline in turnout (Fiorina et al., 2005b).[5]

3 The Model

3.1 General setup

We examine a two-period model ($t = 1, 2$) in which, before each period begins, a society elects an office-holder for an executive position to whom it delegates policy-making. The society is composed of a continuum of voters of measure one, each voter being indexed by $i \in [0, 1]$. Policy-makers differ in their *policy orientation*: there are *right-wing* and *left-wing* policy-makers. We use R and L to denote the pool of all right-wing and left-wing policy-makers, respectively. We assume that there are at least two candidates in each pool. At both election dates, two candidates—one from R and one from L—compete for office. The defeated candidate from the first election does not run for office in the second election but is replaced by another candidate of the same party (or pool). Hence, while the first election is an open race, i.e. a race between two new candidates, in the second election one candidate is the incumbent and the other candidate is the challenger. We thus rule out the possibility

[4]For the US Congress, the potential causes for the increase in party polarization are diverse: geographical changes in partisan alignment (see e.g. Rohde, 2010; Theriault, 2004); incumbent-friendly redistricting (Layman et al., 2006); the structure of roll-call votes and the issues decided by those roll calls (Theriault, 2006); changes in the legislative agenda and the strategies of party leaders (Roberts and Smith, 2003); the leadership selection system (Heberlig et al., 2006), in which extreme candidates tend to excel.

[5]The vast literature on political polarization has pointed to a number of other effects triggered by high political polarization levels, some of which we mention here. Increased party polarization has driven the political debate in the US to stark confrontation, which discourages open deliberation on policy issues, especially in the media (Sinclair, 2002) but also in Congress (Jamieson and Falk, 2000). Testa (2012) shows that societies where social polarization is large are more likely to be characterized by high party polarization. Increasing party and policy polarization levels have nonetheless clarified party importance (Hetherington, 2001) and have made it easier for citizens to vote ideologically because party platforms are significantly different from each other.

of the incumbent being replaced in the second election by a candidate of the same ruling party. This is a plausible assumption, as incumbents tend to have great influence on their parties' internal decisions. Throughout the paper, we denote policy-makers by k, k', or k''. Independently of his/her policy orientation, each candidate is one of three *types*: candidate $k \in R \cup L$ is characterized by his/her *ability* a_k, which is drawn independently from a discrete distribution with support $\{-A, 0, A\}$, where $A > 0$. We assume that

$$P\{a_k = 0\} = \rho \quad \text{and} \quad P\{a_k = A\} = P\{a_k = -A\} = \frac{1-\rho}{2}, \tag{1}$$

where $\rho \in (0,1)$. Results similar to those presented in this paper could be obtained assuming a continuous distribution on the ability component. Assuming discrete ability levels, however, facilitates the analysis greatly, while ensuring the fundamental feature that the ability support is non-degenerate and that the first-period policy choice is made under uncertainty.[6] A candidate with ability A has outstanding ability. If a candidate has ability $-A$, s/he is considered to have very low skills. This means that $1-\rho$ captures the tail probability that a candidate has extraordinarily high or low ability. We note that $E[a_k] = 0$, so the ability of a candidate with zero ability coincides with expected ability.[7] In each period $t \in \{1, 2\}$, the office-holder, denoted by $k \in R \cup L$, faces issues in two different dimensions:

- *Public Project P:* s/he undertakes a public project denoted by g_{kt}. For simplicity, we assume that the output of the project is directly proportional to the ability of the office-holder k, and for simplicity we assume $g_{kt} = a_k$. The public project includes all business-as-usual activities of the government and must be carried out by the office-holder.

- *Policy I:* s/he chooses a policy from a one-dimensional policy space $[0,1]$, which impacts each voter and each candidate differently. We use $i_{kt} \in [0,1]$ to denote his/her policy choice.

Candidates in their first term observe their ability after the public project P has been realized and once policy I has been implemented, and not before. As already mentioned, this means that all decisions by politicians and citizens involve uncertainty about the ability of the office-holder, with the exception of the policy choice in $t = 2$ by an office-holder who has been re-elected—see Figure 1. For each period $t \in \{1, 2\}$, voters observe both i_{kt} and g_{kt} during the term. From g_{kt} voters can infer

[6] If all politicians have the same average ability, and in particular abilities are common knowledge, office-holders choose their desired policies in the first period, and then are re-elected and choose their bliss point again. We note that in a large body of literature it is assumed that politicians know their type but voters do not. If we assumed this, the main thrust of the results would still remain intact. Yet, the incentives to choose moderate policies would be reduced for policy-makers with above-average ability.

[7] We normalize expected ability to zero, but doing so does not affect our results. We also stress that the ability of the challenger in the second election is again drawn from $\{-A, 0, A\}$ according to the distribution defined by (1).

the ability of the incumbent in the first term. Accordingly, at the end of period $t = 1$ and before the second election takes place, the ability of the incumbent is common knowledge. By contrast, at the time of the second election, all agents only know that the challenger's ability is drawn according to the distribution defined by (1). The entire sequence of events is summarized in Figure 1.

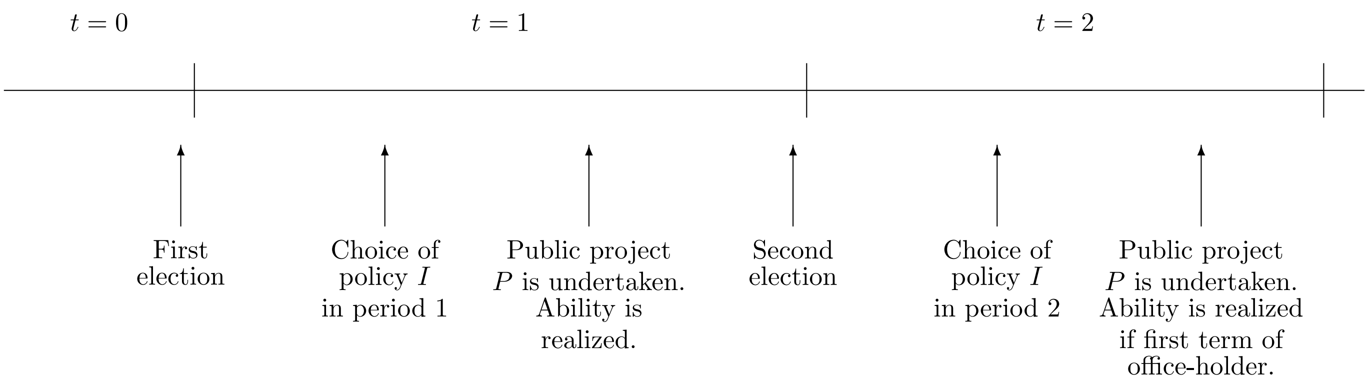


Figure 1: Timeline of the sequence of events.

Finally, it is worth noting that from a mathematical perspective, the public project dimension offers alternative interpretations for which our results carry over and yield further insights. For instance, one can consider that with probability $1 - \rho$ one of two tail events can occur: one event that benefits the incumbent party and another event that benefits the challenger party. These events can be a financial crisis or a natural disaster. Parameter ρ then measures the stability and/or lack of fluctuation of the (political) system. The important assumption is that tail events can be publicly observed and occur *after* the first policy choice is made but *before* the second election takes place.

3.2 Utilities of voters and policy-makers

3.2.1 Instantaneous utilities

Voters and policy-makers derive utility from P and I. Suppose that $k \in R \cup L$ is in office in period t. First, all voters and policy-makers derive the same utility from the public project, given by the instantaneous utility function

$$U^P(g_{kt}) = g_{kt} = a_k.$$

Second, agents have dissenting preferences on policy I. We arrange voters according to their most preferred choice of I, so voter i's ideal policy regarding I is i. Then voter $i \in [0, 1]$ derives utility from a choice $i_{kt} \in [0,1]$ in period t according to[8]

$$U_i^I(i_{kt}) = -(i_{kt} - i)^2.$$

In the main part of the paper, we further assume that voters' ideal policies are uniformly distributed in the electorate. For their part, arbitrary candidates $k' \in R$ and $k'' \in L$ have ideal policies $\mu_{k'} = \mu_R$ and $\mu_{k''} = \mu_L$, respectively, with μ_R and μ_L being fixed. We also assume for simplicity that

$$\frac{1}{2} < \mu_R \leq 1 \quad \text{and} \quad \mu_L = 1 - \mu_R. \tag{2}$$

Hence, the ideal policies of the candidates are distributed symmetrically around the median position. Condition (2) allows us to investigate the effect on policy and turnover of extra-hurdles in isolation, but it can be relaxed without affecting the equilibrium analysis (see Gersbach et al., 2019b). This is because, under majority rule, costs of change create an electoral advantage for the incumbent that only disappears if the median voter's peak is very close to the challenger's peak. If $i_{kt} \in [0,1]$ has been chosen by office-holder k in period t, a candidate $k' \in R \cup L$ derives utility

$$U_{k'}^I(i_{kt}) = -(i_{kt} - \mu_{k'})^2.$$

Additionally, an office-holder obtains a fixed private benefits $b > 0$ from holding office in a particular period. The benefits b account for all sources of utility that politicians derive from office beyond policy choices. For instance, they include ego rents and satisfaction from holding power and heading a branch of government, plus additional career opportunities after office-holding.

3.2.2 Costs of change

A key feature in our model is that policy changes are costly for voters and candidates. More precisely, given a policy $i_1 \in [0,1]$ chosen in the first period, the policy choice in the second period, $i_2 \in [0,1]$, imposes additional costs (or utility losses) on voters and policy-makers alike, equal to

$$U^c(i_1, i_2) = -c \cdot |i_1 - i_2|. \tag{3}$$

Hence, unless $c = 0$, the so-called *costs of change* increase linearly with the absolute difference between the policies adopted in both periods. The parameter $c \geq 0$ is the *marginal cost of a policy change*. Assuming that costs of change are linear is a first-order approximation of the general case where costs associated with policy changes increase arbitrarily in the extent of the policy shift. It facilitates the analysis, but it does not drive our results, as we show in Appendix C (see also Gersbach

[8]The assumption that preferences are quadratic simplifies the exposition of the results but is not crucial.

et al., 2020a). Costs of change that are less convex than the policy utility loss function—linear costs of change and quadratic utility loss from policy, in particular—suffice for the incumbent to have an electoral advantage compared to the challenger under majority rule *in equilibrium*. This is *not* a universal property of costs of change and is a critical feature of our model. On the other hand, although c is common for voters and policy-makers, politicians and citizens diverge in one crucial aspect: the former obtain a reward from being in office, as captured by parameter b.[9] We note that even with the same marginal-cost-of-change parameter, in equilibrium this feature makes new office-holders more prone to policy changes than citizens, since office-holders want to stay in office.

3.2.3 Lifetime utility

Voters and policy-makers discount utility in the second period with a common factor β, where $0 < \beta \leq 1$. Let candidates $k \in R \cup L$ and $k' \in R \cup L$ be in power in $t = 1$ and $t = 2$, respectively. Then the lifetime utility of voter $i \in [0, 1]$ is

$$U^P(g_{k1}) + U_i^I(i_{k1}) + \beta \cdot \left[U^P(g_{k'2}) + U_i^I(i_{k'2}) + U^c(i_{k1}, i_{k'2})\right]. \tag{4}$$

Similarly, the lifetime utility of any policy-maker $k'' \in R \cup L$ is

$$U^P(g_{k1}) + U_{k''}^I(i_{k1}) + \mathbb{1}_{k''}(k) \cdot b + \beta \cdot \left[U^P(g_{k'2}) + U_{k''}^I(i_{k'2}) + U^c(i_{k1}, i_{k'2}) + \mathbb{1}_{k''}(k') \cdot b\right],$$

where $\mathbb{1}_{k''}(\cdot)$ is an indicator variable for holding (or not holding) power. That is,

$$\mathbb{1}_{k''}(x) = \begin{cases} 1 & \text{if } k'' = x, \\ 0 & \text{otherwise.} \end{cases}$$

3.3 Election thresholds and tie-breaking rules

In the first election, a candidate is elected according to the simple majority rule. We assume that, in the case of a tie, each candidate wins the election with a probability equal to $\frac{1}{2}$. In the second election, the incumbent is re-elected if his/her vote-share is equal to or larger than $\frac{1}{2} + \delta$, where $\delta \in \left[0, \frac{1}{2}\right]$. The parameter δ measures the increase of the re-election hurdle for the incumbent, so we call it the *extra-hurdle*. For instance, $\delta = 0$ corresponds to the simple majority rule, while $\delta = \frac{1}{2}$ implies that an incumbent is re-elected only if s/he obtains unanimous support. If the incumbent's

[9] This is not a critical assumption since standard continuity arguments can be applied for small (not necessarily infinitessimal) differences in marginal costs between voters and politicians. Considering large differences in marginal costs between voters and politicians implies assuming away that politicians are regular citizens (they care about policies and about policy shifts), with the only difference stemming from the fact that they derive additional utility from office tenure. If marginal costs were very low for office-holders, in particular, they would be more inclined to engineer excessive policy changes.

vote-share is strictly less than $\frac{1}{2}+\delta$, the challenger wins the election. For ease of presentation, the incumbent is assumed to win the election if s/he receives a vote-share of exactly $\frac{1}{2}+\delta$.[10] We also assume that every voter who is indifferent between the incumbent and the challenger votes for the incumbent.[11]

3.4 Assumptions on the parameters

For the sake of analytical tractability, we assume that politicians' benefits from holding office, b, are sufficiently large to ensure that, under any constellation of parameters analyzed in the paper, all policy-makers prefer being in office to not being in office. That is, the relative weight put on policy changes, as well as on policies themselves, is lower for politicians than for citizens. It suffices to assume $b > 2 + c + 4A$. This assumption guarantees that, in equilibrium, the office-holder's policy choice always maximizes his/her re-election probability. Moreover, A is assumed to be large enough to ensure that, in equilibrium, an incumbent with ability A is always re-elected, while an incumbent with ability $-A$ is never re-elected. For that, it suffices to consider $A > 1 + c$. To simplify notation, we further suppose that $\beta = 1$, as results can easily be extended to any value of $\beta \in (0, 1]$.

Finally, we focus throughout on the case where costs of change are not too large, since this leads to significant policy changes by office-holders with opposing preferences and thus interesting dynamics. Specifically, we assume that

$$0 \leq \frac{c}{2} < \mu_R - \frac{1}{2}. \tag{5}$$

The upper bound in the above inequality,

$$\Pi := \mu_R - \frac{1}{2} \in \left(0, \frac{1}{2}\right],$$

measures the degree of *party polarization*. If Π is close to a half, parties' ideal policies are very far away, so the interests of both parties are opposed. By contrast, if Π is close to zero, parties' ideal policies are very close, so both parties primarily focus on the interests of those voters located close to the median voter. Condition (5) therefore ensures that costs of change are small compared to the level of party polarization.[12]

[10] Without this assumption, we would need to discretize the set of possible policy choices in order to guarantee the existence of a best response of the office-holder in $t = 1$.

[11] This assumption is only necessary for the results that consider large values of c, given in Appendix B.

[12] The case where $c \geq 2\Pi$ is dealt with in Appendix B.

3.5 Notion of equilibrium and informational assumptions

Before the end of each period, voters observe the policy choice and the output of the public project. From the latter they immediately infer the ability of the incumbent. This resolves the asymmetry of information regarding the incumbent's ability and thus makes it impossible for the incumbent to use the asymmetry to his/her advantage. An equilibrium in our model is a perfect Bayesian Nash equilibrium of the game displayed in Figure 1, denoted by $\mathcal{G}$, in which we apply the standard refinement that all citizens vote for the candidate from whom they expect the highest utility. This rules out implausible equilibria.

4 Analysis

In this section we determine the equilibria of game $\mathcal{G}$ and assess how key parameters and variables determine policy choices. Recall that we have assumed for simplicity that left- and right-wing candidates' ideal policy positions are symmetrically distributed around the median—see (2). Since the distribution of a_k is independent of the policy orientation of candidate k, the median voter is indifferent between the left-wing and the right-wing candidate when the first election takes place. Thus, each candidate receives a vote-share of $\frac{1}{2}$ in the first election.[13] According to the election rule, both candidates have a 50% chance of winning the first election. Let $\mathcal{G}^R$ (resp. $\mathcal{G}^L$) be the game that starts after a candidate from pool R (resp. L) has been chosen as first office-holder. We assume, without loss of generality, that a right-wing politician is in office in the first period. Once the equilibria of $\mathcal{G}^R$ have been determined, the equilibria of $\mathcal{G}^L$, and hence of $\mathcal{G}$, immediately follow by symmetry. Next, we solve $\mathcal{G}^R$ backwards.

4.1 The second period

We start with the analysis of the policy-makers' behavior in the second period.

Proposition 1 (Gersbach et al. (2019a))

Let $k \in R$ and $k' \in R \cup L$ be the office-holders in $t = 1$ and $t = 2$, respectively. In $t = 2$, the best response of k' to a policy i_{k1} chosen in the first period is given by

$$i_{k'2}(i_{k1}) = min\left\{max\left\{\mu_{k'} - \frac{c}{2}, i_{k1}\right\}, \mu_{k'} + \frac{c}{2}\right\} =: \begin{cases} i^*_{R2}(i_{k1}) & \text{if } k' \in R, \\ i^*_{L2}(i_{k1}) & \text{if } k' \in L, \end{cases} \quad (6)$$

where $\mu_{k'}$ is the ideal policy of policy-maker k'.

[13] Whether indifferent voters vote for the left- or the right-wing candidate has no influence on the outcome of the election. Since the median voter has measure zero, the election is tied, independently of his/her behavior.

Accordingly, the best response of the office-holder in $t = 2$ is to match the status quo, but only when the latter lies within a certain range around his/her preferred policy position. Otherwise, his/her policy choice is bounded by $\mu_{k'} + \frac{c}{2}$ from above and by $\mu_{k'} - \frac{c}{2}$ from below. In these two policy points, the marginal disutility from changing policy, c, is equal to the marginal disutility from choosing a policy farther away from the office holder's bliss point. The best response of the office-holder in $t = 2$ is illustrated by Figure 2.

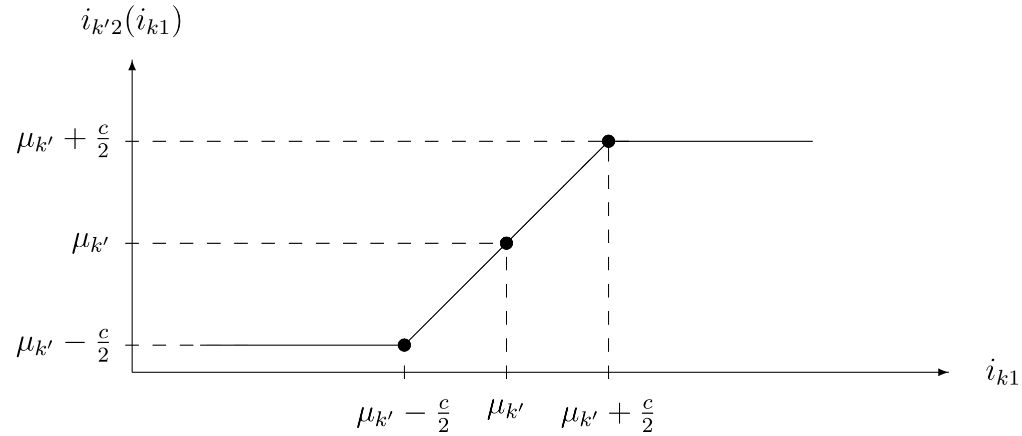


Figure 2: Best response $i_{k'2}(i_{k1})$ of an office-holder with bliss point $\mu_{k'}$ to policy i_{k1}.

4.2 The first period

We next study the decisions taken in the first period. First we analyze the election at the end of this period, in which all citizens select either the incumbent, $k \in R$, or the new left-wing candidate, $k' \in L$. Accordingly, suppose that, in period $t = 1$, k chooses a policy to the right of the median voter's ideal point. Then the critical voter who is ideologically to the left of the median voter faces a trade-off: on the one hand, k' would choose a policy that is closer to his/her ideal point; on the other, this policy would result in larger costs of change. The proposition below analyzes this trade-off and as a result identifies the first-period policy choices for which the incumbent receives a vote-share of at least $\frac{1}{2} + \delta$ in the elections that take place right before period $t = 2$ starts, provided his/her ability is a_k.

Proposition 2

Let $\delta \in \left[0, \frac{1}{2}\right]$ and $k \in R$ be in office in $t = 1$. Then, in any equilibrium of $\mathcal{G}^R$, k is re-elected in the second election if and only if his/her policy choice in $t = 1$, denoted by i_{k1}, and his/her ability level, a_k, satisfy

$$a_k \geq a_\delta(i_{k1}), \tag{7}$$

where

$$a_\delta(i_{k1}) = \left(i^*_{R2}(i_{k1}) - \left(\frac{1}{2} - \delta\right)\right)^2 + c \cdot |i^*_{R2}(i_{k1}) - i_{k1}|$$
$$- \left(i^*_{L2}(i_{k1}) - \left(\frac{1}{2} - \delta\right)\right)^2 - c \cdot |i^*_{L2}(i_{k1}) - i_{k1}| \qquad (8)$$

and $i^*_{R2}(\cdot)$ *and* $i^*_{L2}(\cdot)$ *are given in (6).*

The function $a_\delta(\cdot)$ we have just defined determines the proportion of citizens who prefer the incumbent to the challenger in the second election, and it is crucial for understanding the result of the above proposition. We start by illustrating the electorate's decision graphically as a function of i_{k1}. Figure 3 depicts the graphs of $a_\delta(\cdot)$ for three different values of δ.

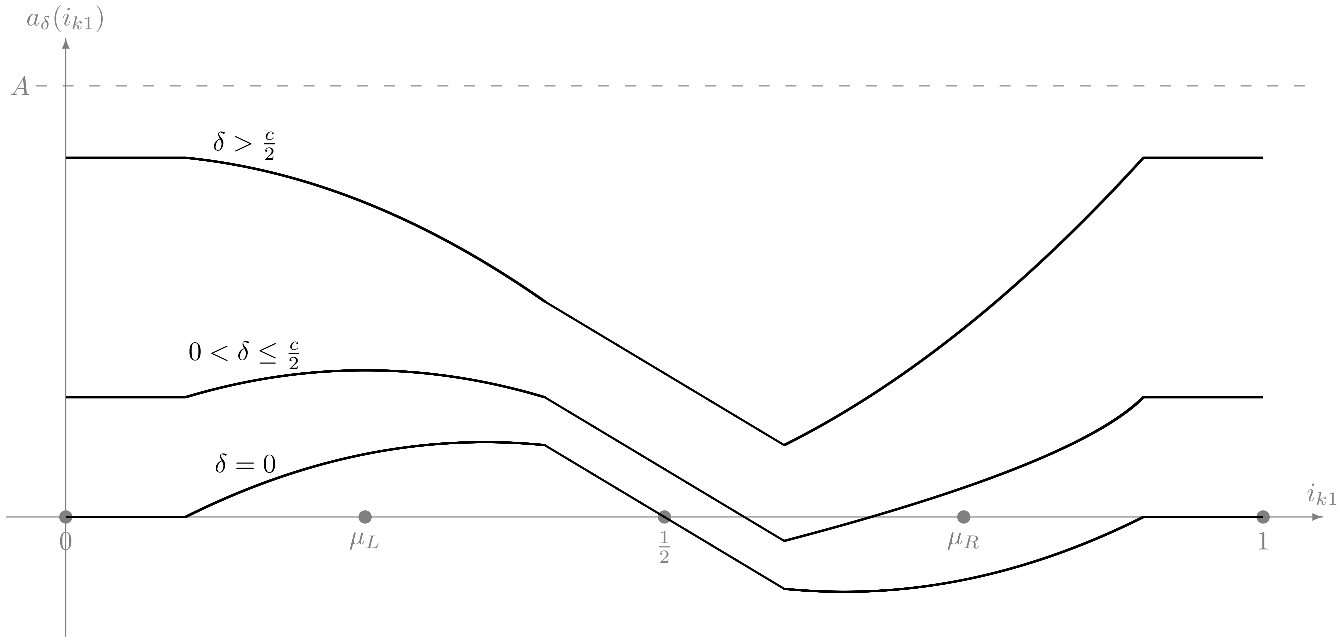


Figure 3: Illustration of the electorate's decision in the second election for different extra-hurdles $\delta \in \left[0, \frac{1}{2}\right]$. For δ fixed, a right-wing incumbent k with average ability is re-elected precisely when his/her policy choice satisfies $a_\delta(i_{k1}) \leq 0$.

Since A has been assumed to be sufficiently large, the incumbent k is re-elected independently of his/her policy choice i_{k1} if his/her ability is A. In Figure 3, this follows from the fact that $a_\delta(i_{k1}) < A$ for any $\delta \in \left[0, \frac{1}{2}\right]$ and any $i_{k1} \in [0, 1]$. Similarly, an incumbent with ability $-A$ is always ousted. For incumbents of average ability, i.e. if $a_k = 0$, Figure 3 provides more compelling insights. Let the extra-hurdle $\delta \in \left[0, \frac{1}{2}\right]$ be fixed, and assume that k is a right-wing candidate. Then, s/he is re-elected whenever s/he chooses i_{k1} such that $a_\delta(i_{k1}) \leq 0$. With the standard majority rule, i.e. for $\delta = 0$, k is re-elected if $i_{k1} \geq \frac{1}{2}$, and in particular for any policy close to his/her bliss point μ_R. This is not the case for $\delta > 0$. If $0 < \delta \leq \frac{c}{2}$, the set of policies that secure office-holder k's re-election is a strict non-empty subset of $\left[\frac{1}{2}, 1\right]$. It then turns out that

$$\frac{da_\delta(i_{k1})}{d\delta} = 2(i^*_{R2}(i_{k1}) - i^*_{L2}(i_{k1})) > 0,$$

where the inequality follows from (5). Hence, as we increase δ, the rightmost policy ensuring that the incumbent from party R is re-elected becomes more moderate. The explanation is as follows: The critical voter in the election is no longer the median voter but a more partisan voter whose ideal policy is closer to the challenger's ideal policy than to the incumbent's, that is, $\frac{1}{2} - \delta$. Accordingly, a positive extra-hurdle raises the office-holder's interest in choosing a policy closer to the median voter's preferred policy. By doing so, the incumbent can become more attractive for voters on the other side of the political spectrum, as policy changes in the second period are milder if the incumbent wins the election. This occurs despite the fact that the challenger's ideal policy is closer to the critical voter's ideal policy. The reason is that the challenger would cause significant costs of change. But if the first-period policy is too close to the median, costs of change become so mild than the decisive voter prefers to elect that challenger instead of the incumbent. That is why the leftmost policy ensuring that the incumbent from party R is re-elected becomes more extreme as we increase δ. This means that the set of policies that make re-election possible shrinks with δ. In fact, if $\delta > \frac{c}{2}$, i.e. for sufficiently large values of the extra-hurdle, an average-ability incumbent can never achieve re-election because such a set of policies becomes empty.

Next, we analyze how the policy choice in $t = 1$ depends on $\delta \in \left[0, \frac{1}{2}\right]$. At this point, it is useful to recall that the office-holder chooses his/her policy under uncertainty with regard to his/her own ability and that, as the latter has discrete support, the expected re-election probability only takes discrete values. Since the reward for gaining office is very large, the office-holder proceeds as follows: First, s/he determines the set of policies that grant him/her the highest re-election probability, i.e., s/he only considers $i_{k1} \in [0, 1]$ such that $a_\delta(i_{k1}) \leq 0$, provided that such policies exist. Otherwise, s/he considers the whole set of alternatives $[0, 1]$. Second, s/he maximizes his/her expected utility on this subset of policies. The next result yields the solution to this two-stage problem.

Proposition 3

Let $\delta \in \left[0, \frac{1}{2}\right]$, and let $k \in R$ be in office in $t = 1$. Then k's policy choice in $t = 1$, in any equilibrium of $\mathcal{G}^R$, is

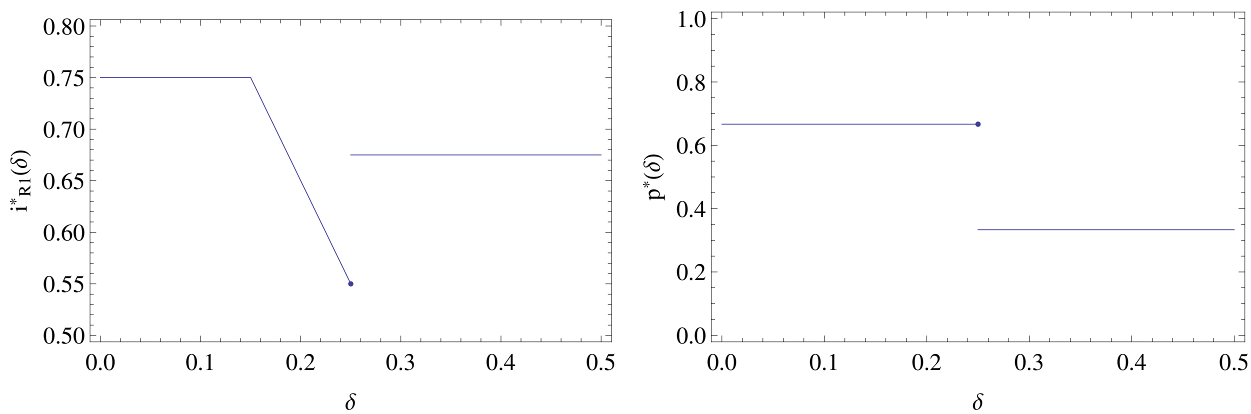
$$i_{k1}(\delta) = i^*_{R1}(\delta) := \begin{cases} \mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho} & \text{if } \delta \in \left[0, \frac{c}{3+\rho}\right], \\ \mu_R + \frac{c}{2} - 2\delta & \text{if } \delta \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right], \\ \mu_R - \frac{c}{2} \cdot \frac{1+\rho}{3-\rho} & \text{if } \delta \in \left(\frac{c}{2}, \frac{1}{2}\right]. \end{cases} \quad (9)$$

Hence, while the policy chosen by $k \in R$ in the first period increases with μ_R, the effect of c and ρ is ambiguous, as it depends on the value of δ. If δ is small, the first-period policy decreases with c and increases with ρ. If δ is intermediate, the first-period policy increases with c and does not change with ρ. If δ is large, the first-period policy decreases with c and with ρ. It is worth noting that the three regions for δ (small, intermediate, and large) depend themselves on c. Unsurprisingly, policy becomes more extreme with party polarization Π.

As for our main (design) variable of interest, viz. δ, in Figure 4(a) we illustrate Proposition 3 by plotting $i^*_{R1}(\delta)$ for $\mu_R = 0.8$, $c = 0.5$, and $\rho = \frac{1}{3}$. It shows that the first-period policy choice does not change monotonically with δ. For the same parameters, Figure 4(b) depicts k's re-election probability in equilibrium, $p^*(\delta)$. It can be shown—see the proof of Proposition 3—that as a function of δ this probability is given by

$$p^*(\delta) = \begin{cases} \frac{1+\rho}{2} & \text{if } \delta \in \left[0, \frac{c}{2}\right], \\ \frac{1-\rho}{2} & \text{if } \delta \in \left(\frac{c}{2}, \frac{1}{2}\right]. \end{cases} \tag{10}$$

That is, the probability that the incumbent is re-elected decreases with δ. Moreover, the drop in the re-election probability as a function of δ increases with ρ and thus decreases with the tail event probability that the incumbent has an ability that is out of the ordinary. Finally, the point at which such a drop takes place is pinned down by c.



(a) Equilibrium policy choice of right-wing policy-maker in $t = 1$.

(b) Re-election probability in equilibrium.

Figure 4: Illustration of the equilibrium analysis in $t = 1$ using the parameter values $\mu_R = 0.8$, $c = 0.5$, $A = 2$, and $\rho = \frac{1}{3}$.

It is instructive to look more closely at Proposition 3. First, any non-zero extra-hurdle δ weakly moderates the first-period policy choice compared to $\delta = 0$. Second, $i^*_{R1}(\delta)$ is closest to the median voter at $\delta = \frac{c}{2}$. Third, the re-election probability of an incumbent is larger than $\frac{1}{2}$ when extra-hurdles are low, in particular zero. This reflects an incumbency advantage associated with the standard majority rule.[14] In our model, this advantage grows with ρ. Fourth, the fact that positive

[14]There is a large body of literature on the existence and causes of incumbency advantages for US congressmen (see e.g. Alford and Brady (1989), Gelman and King (1990), and Levitt and Wolfram (1997)). Erikson et al. (1993) find that governors have similar advantages when seeking re-election.

extra-hurdles yield weakly more moderate first-period policies than a zero extra-hurdle is triggered by two forces.

On the one hand, when $0 < \delta \leq \frac{c}{2}$, the incumbent chooses a more moderate policy than his/her bliss point to maximize his/her re-election probability. This behavior is illustrated by Figure 3 and has already been discussed. On the other, as we can also see from Figure 3, if $\delta > \frac{c}{2}$, it is impossible for the incumbent to be re-elected unless his/her ability is high. This may in principle induce him/her to try steamroller tactics and choose a policy in the first period that is as close as possible to his/her preferred policy, as this would subsequently keep his/her own policies and the challenger's policies closer to his/her ideal policy position. However, as the latter undertakes a significant change of policy if s/he is elected, the office-holder prefers to moderate his/her policy choice in exchange for reducing the costs associated with the likelier event of the challenger being elected in the second election.

4.3 Equilibrium outcomes

Sections 4.1 and 4.2 enable us to describe the equilibrium outcomes of game $\mathcal{G}$. From Propositions 1 and 3 it immediately follows that the equilibrium policy choices in $\mathcal{G}^R$ are given by

$$i^*_{R1}(\delta) = i^*_{R2}(\delta) = \begin{cases} \mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho} & \text{if } \delta \in \left[0, \frac{c}{3+\rho}\right], \\ \mu_R + \frac{c}{2} - 2\delta & \text{if } \delta \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right], \\ \mu_R - \frac{c}{2} \cdot \frac{1+\rho}{3-\rho} & \text{if } \delta \in \left(\frac{c}{2}, \frac{1}{2}\right] \end{cases} \quad \text{and} \quad i^*_{L2}(\delta) = \mu_L + \frac{c}{2}, \tag{11}$$

for each given $\delta \in \left[0, \frac{1}{2}\right]$. Some comments are helpful. First, for all $\delta \in \left[0, \frac{1}{2}\right]$, $i^*_{L2}(\delta)$ is strictly smaller than $i^*_{R1}(\delta)$. This means that although costs of change are the same for all citizens including politicians, the incumbent's party affiliation matters in explaining policy choices in the ideological dimension. Second, if the incumbent is re-elected, s/he does not shift the policy. Hence, persistence in the incumbent's choice arises endogenously in our model. Third, the feature that the challenger's second-period policy choice is independent of δ follows from the assumption that the marginal cost of change is small—as expressed in (5). According to Proposition 1, this implies that, being a member of R, the incumbent chooses a policy to the right of $\frac{1}{2}$, while the challenger, a member of L, chooses a policy to the left of $\frac{1}{2}$. Notably, the challenger chooses the upper bound of his/her best-response function as given in (6), which does not depend on the policy choice in $t = 1$.

Due to the various symmetry assumptions for parties and voters, and because the distribution of a_k is independent of the policy orientation of candidate k, the equilibrium policy choices of $\mathcal{G}^R$ and $\mathcal{G}^L$

are symmetrically distributed around $\frac{1}{2}$. Thus, equilibrium policy choices in $\mathcal{G}^L$ are given by

$$i^{**}_{L1}(\delta) = i^{**}_{L2}(\delta) = \begin{cases} \mu_L + \frac{c}{2} \cdot \frac{1-\rho}{3+\rho} & \text{if } \delta \in \left[0, \frac{c}{3+\rho}\right], \\ \mu_L - \frac{c}{2} + 2\delta & \text{if } \delta \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right], \\ \mu_L + \frac{c}{2} \cdot \frac{1+\rho}{3-\rho} & \text{if } \delta \in \left(\frac{c}{2}, \frac{1}{2}\right] \end{cases} \quad \text{and} \quad i^{**}_{R2}(\delta) = \mu_R - \frac{c}{2}. \tag{12}$$

In $\mathcal{G}$, there are four possible voting outcomes: (R, R), (R, L), (L, L), and (L, R), depending on the types of first- and second-period office-holders. Since, for each $\delta \in \left[0, \frac{1}{2}\right]$ the equilibrium policy choices for these voting outcomes are determined once the policy orientations of both office-holders are known, we also denote the four different equilibrium outcomes of game $\mathcal{G}$ by (R, R), (R, L), (L, L), and (L, R). For instance, in (R, R) and (R, L) the equilibrium policy choices are given by $i^*_{R1}(\delta)$, $i^*_{R2}(\delta)$, and $i^*_{L2}(\delta)$ from (11). We note that the realization of ability in the first period only influences whether the office-holder gets re-elected. It has no impact on the policy choice for a given office-holder in the second period. While the probability of winning the first election is equal to $\frac{1}{2}$ for both initial candidates, the re-election probability is given by (10) for both candidates. Hence, the equilibrium outcomes of $\mathcal{G}$ occur with probabilities

$$p_{(R,R)}(\delta) = p_{(L,L)}(\delta) = \frac{p^*(\delta)}{2} \quad \text{and} \quad p_{(R,L)}(\delta) = p_{(L,R)}(\delta) = \frac{1 - p^*(\delta)}{2}, \tag{13}$$

where $p^*(\delta)$ is given by (10). We conclude this section by noting that when $c = 0$, both candidates choose their ideal policies whenever they are in office, since there are no costs to prevent them from indulging their own preferences. This means that the interesting dynamics happen when $c > 0$.

5 Policy Polarization and Welfare

In this section we explore how the levels of *policy polarization* and *welfare* depend on the (institutional) variable δ.

5.1 Policy polarization and welfare

First, we define *ex-post policy polarization.*

Definition 1

Let $i_1, i_2 \in [0, 1]$ be the policy choices in $t = 1$ and $t = 2$. Then, ex-post policy polarization is[15]

$$EPP(i_1, i_2) = \frac{\left|i_1 - \frac{1}{2}\right| + \left|i_2 - \frac{1}{2}\right|}{2}.$$

[15] Ex-post policy polarization could be defined in different ways. As long as ex-post policy polarization increases in the distance of each period's policy choice from the median voter's ideal position, however, the results of the paper would qualitatively be the same.

Ex-post policy polarization measures how policy choices in both periods deviate from the median citizen's preferred policy. Building on the previous definition, we define ex-ante policy polarization:

Definition 2

Ex-ante policy polarization is the expected value of ex-post policy polarization.

From Section 4.3, it follows by straightforward algebra and the symmetry of the model that ex-ante policy polarization for game $\mathcal{G}$ coincides with that of game $\mathcal{G}^R$. Thus, we can write

$$EAP(\delta) = p^*(\delta) \cdot EPP(i^*_{R1}(\delta), i^*_{R2}(\delta)) + (1 - p^*(\delta)) \cdot EPP(i^*_{R1}(\delta), i^*_{L2}(\delta)). \tag{14}$$

Second, to measure welfare, we take the integral over all voters' expected lifetime utilities. As voters' ideal policy positions are symmetrically distributed around the median voter's ideal policy, $i = \frac{1}{2}$, and since all voters are affected by the public project and by costs of change in the same way, maximizing the utilitarian welfare function is equivalent to maximizing the median voter's expected lifetime utility. According to (4), expected lifetime utility of the median voter consists of three terms: expected lifetime utility from public projects (EU^P), expected lifetime utility from policies ($EU^I_{\frac{1}{2}}$), and expected lifetime utility from costs of change (EU^c). As in the case of policy polarization, it suffices to consider expected lifetime utility over all equilibrium outcomes of $\mathcal{G}^R$. Thus, expected welfare as a function of $\delta \in \left[0, \frac{1}{2}\right]$ is given by

$$W(\delta) = EU^P(\delta) + EU^I_{\frac{1}{2}}(\delta) + EU^c(\delta). \tag{15}$$

If $k \in R$ is the first-period office-holder and $k' \in L$ is the challenger in the second election, the three components of welfare are given by

$$\begin{aligned}
EU^P(\delta) &= p^*(\delta) \cdot 2E\left[a_k | k \text{ re-elected and } p_\delta(i_{k1}) = p^*(\delta)\right] \\
&\quad + (1 - p^*(\delta)) \cdot \{E\left[a_k | k \text{ not re-elected and } p_\delta(i_{k1}) = p^*(\delta)\right] + E[a_{k'}]\}, && (16)\\
EU^I_{\frac{1}{2}}(\delta) &= -\left(i^*_{R1}(\delta) - \frac{1}{2}\right)^2 - p^*(\delta) \cdot \left(i^*_{R2}(\delta) - \frac{1}{2}\right)^2 - (1 - p^*(\delta)) \cdot \left(i^*_{L2}(\delta) - \frac{1}{2}\right)^2, && (17)\\
EU^c(\delta) &= p^*(\delta) \cdot [-c \cdot |i^*_{R1}(\delta) - i^*_{R2}(\delta)|] + (1 - p^*(\delta)) \cdot [-c \cdot |i^*_{R1}(\delta) - i^*_{L2}(\delta)|], && (18)
\end{aligned}$$

with $p^*(\delta)$ from (10), $i^*_{R1}(\delta)$, $i^*_{R2}(\delta)$, and $i^*_{L2}(\delta)$ from (11), and $p_\delta(i_{k1})$ denotes k's re-election probability in the second election as a function of i_{k1}.

5.2 Polarization and welfare as functions of the extra-hurdle

To compare different extra-hurdles, it is useful to introduce the following definitions:

Definition 3

An extra-hurdle $\delta \in \left[0, \frac{1}{2}\right]$ *is called*

- *W-optimal if it maximizes* $W(\delta)$*, and P-optimal if it minimizes* $EAP(\delta)$*,*
- *W-increasing if* $W(\delta) > W(0)$*, and weakly W-increasing if* $W(\delta) \geq W(0)$*,*
- *P-reducing if* $EAP(\delta) < EAP(0)$*, and weakly P-reducing if* $EAP(\delta) \leq EAP(0)$*.*

The reference to optimality in the above definition is justified because δ is a variable that can be modified through institutional arrangements. With the above definitions at hand, we can now state our main result on the role of δ with regard to policy polarization and welfare.[16]

Theorem 1

In equilibrium of $\mathcal{G}$*,*

(i) Any $\delta \in \left(0, \frac{1}{2}\right]$ *is both weakly W-increasing and weakly P-reducing.*

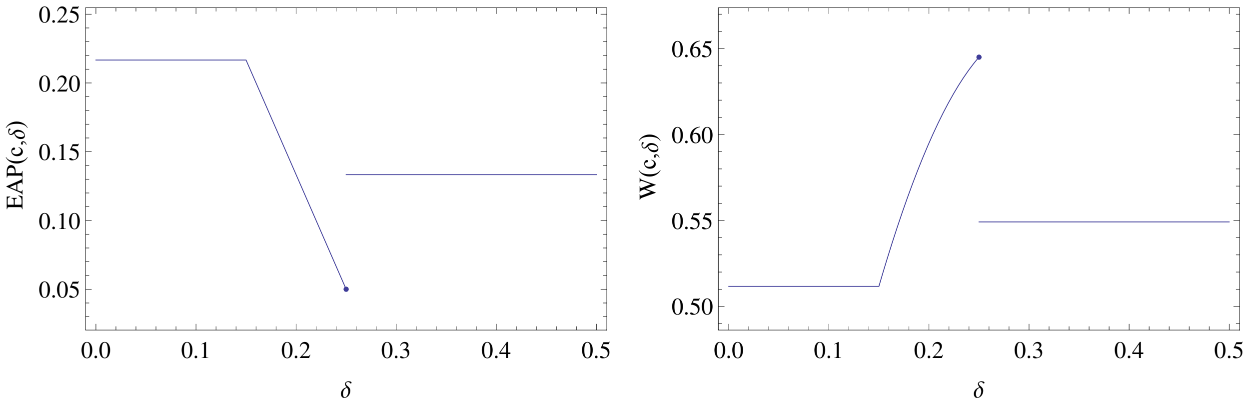
(ii) Extra-hurdle $\delta^* = \frac{c}{2}$ *is W- and P-optimal. Moreover, if* $c \neq 0$ *and* $\delta \in \left[0, \frac{1}{2}\right]$ *is either W- or P-optimal, then* $\delta = \delta^*$*.*

In the following we discuss this result.[17] On the one hand, assume that $c = 0$. Then $EAP(\delta)$ and $W(\delta)$ are constant in δ. The reason is that the reduction in re-election probability $p^*(\delta)$ has no effect on policy polarization and welfare, because changing policies is not costly. Thus, the incumbent implements his/her ideal policy independently of δ in each period. For $c > 0$, on the other hand, the intuition and understanding of Theorem 1 is developed with the illustrations in Figures 5(a) and 5(b), where $EAP(\delta)$ and $W(\delta)$ are plotted for $\mu_R = 0.8$, $c = 0.5$, $A = 2$, and $\rho = \frac{1}{3}$.

The non-monotonic behavior of ex-ante policy polarization as a function of δ follows from the shape of $i^*_{R1}(\delta)$, which is depicted in Figure 4(a) and driven by the two mechanisms described in Section 4.2. The behavior of $W(\delta)$ is more subtle. First, $EU^P(\delta)$ is constant and equal to $\frac{A(1-\rho)}{2}$, independently of δ. Second, $EU^I_{\frac{1}{2}}(\delta)$ is increasing (resp. decreasing) when $EAP(\delta)$ is decreasing (resp. increasing). For $\delta \in [0, \delta^*]$, with $\delta^* = \frac{c}{2}$ as given from Theorem 1, expected costs of change are weakly decreasing because the re-election probability $p^*(\delta)$ is constant, the first-period policy choice approaches the median when δ increases, and $i^*_{R1}(\delta) = i^*_{R2}(\delta)$. Hence, welfare is weakly increasing for $\delta \in [0, \delta^*]$. For $\delta > \delta^*$, expected costs of change are larger than for $\delta = 0$ because the incumbent's probability of being ousted is larger. Nevertheless, since the increase in expected costs of change is outweighed by the reduction in ex-ante policy polarization, any extra-hurdle above δ^* is W-increasing. Moreover, $W(\delta)$

[16] The counterpart of Theorem 1 for $c \geq 2\Pi$ is given by Theorem 4 in Appendix B. Note that for $c \geq 2\Pi$ the set of extra-hurdles that are both W- and P-optimal is not necessarily a singleton.

[17] We note that for $c = 0$, all $\delta \in \left[0, \frac{1}{2}\right]$ are both W- and P-optimal.



(a) Ex-ante policy polarization. (b) Welfare.

Figure 5: Illustration of the equilibrium values of ex-ante policy polarization and welfare, using the parameter values $\mu_R = 0.8$, $c = 0.5$, $A = 2$ and $\rho = \frac{1}{3}$.

is uniquely maximized at δ^*, as $EU^P(\delta)$ is constant and both $EU^I_{\frac{1}{2}}(\cdot, \delta)$ and $EU^c(\delta)$ are maximized at δ^*.

For completeness, we emphasize that the proof of Theorem 1 allows us to find closed-form expressions for policy polarization and welfare. These are respectively

$$EAP(\delta) = \begin{cases} \mu_R - \frac{1}{2} - \frac{c(1-\rho)}{4} & \text{if } \delta \in \left[0, \frac{c}{3+\rho}\right], \\ \mu_R - \frac{1}{2} + \frac{c(1+\rho)}{4} - \frac{(3+\rho)\delta}{2} & \text{if } \delta \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right], \\ \mu_R - \frac{1}{2} - \frac{c(1+\rho)}{4} & \text{if } \delta \in \left(\frac{c}{2}, \frac{1}{2}\right] \end{cases}$$

and

$$W(\delta) = \begin{cases} \frac{A(1-\rho)}{2} - \frac{(2\mu_R-1)^2}{2} + \frac{c^2(1-\rho)}{2(3+\rho)} & \text{if } \delta \in \left[0, \frac{c}{3+\rho}\right], \\ \frac{A(1-\rho)-(1-2\mu_R-c)^2}{2} - 2(3+\rho)\delta^2 + \left[4c + (3+\rho)(2\mu_R - 1)\right]\delta & \text{if } \delta \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right], \\ \frac{A(1-\rho)}{2} - \frac{(2\mu_R-1)^2}{2} + \frac{c^2(1+\rho)}{2(3-\rho)} & \text{if } \delta \in \left(\frac{c}{2}, \frac{1}{2}\right]. \end{cases}$$

The above expressions offer precise comparative statics of policy polarization and welfare with respect to the main model parameters (ρ, c, and Π). These results follow the comparative statics of the first-period policy choice that have been discussed in Section 4.2.

5.3 Implicit and negative extra-hurdles

In the Introduction we have discussed real-world examples of re-election rules featuring $\delta > 0$, and our theory can be used to provide a rationale for such rules. However, most elections in democracy operate under the standard majority rule, i.e., they assume $\delta = 0$. Being strict, our main result only suggests a guideline for experimentation for the latter electoral systems. Yet, it is worth mentioning that in these cases re-election hurdles for incumbents different from zero can also emerge implicitly as a combination of two existing institutional features of the political system: term limits and qualified majorities to amend the rules regulating such limitations. To stay in office beyond the limit imposed by law, the incumbent usually needs a majority of more than 50% to remove such limitations. One instance of legislative modification aimed at enabling an extra term for the incumbent occurred in the elections for the mayor of New York in 2009. In this example, Michael Bloomberg succeeded in staying in office for a third term after the City Council of New York agreed to modify the two-term limit that was previously in place.

On the other hand, our main result—Theorem 1—can be easily extended to negative values of the extra-hurdle.[18] It can be easily seen that the case $\delta < 0$ yields the same policy choices—and hence the same levels of policy polarization and welfare—as $\delta = 0$. This means that the analysis set out in the main body of the paper—i.e., non-negative values of δ—captures all the relevant effects of varying re-election thresholds on polarization and welfare. In particular, we obtain that $\delta \leq 0$ is never W- nor P-optimal.

From a real-world perspective, negative extra-hurdles are meaningful in at least two ways. First, we can imagine the set of voters $[0, 1]$ as comprising only voters who do not inelastically vote for either candidate. Then, besides these voters, there could be a significant number of *partisan voters* who vote for the party they prefer the most, regardless of any policy or ability consideration.[19] In particular, suppose that the share of partisan voters favoring the incumbent is larger than the share of partisan voters supporting the challenger because the former is more capable of mobilizing citizens due to his/her control of the public resources. Theorem 1 then warns us that, in the presence of costs of change, non-policy- and non-ability-related phenomena that allow the incumbent to mobilize a larger fraction of partisan voters have negative effects on welfare and on polarization. Second, thresholds that require less than half of the votes for re-election can arise in a majoritarian multi-district system. In this case, Theorem 1 suggests that in terms of welfare and policy polarization a proportional system would be better than a majoritarian multi-district system if there exist significant costs of change. The reason is that a one-district, proportional system does not allow the possibility that the incumbent wins the election with less than half of the votes, say because s/he has the power to redefines the voting districts.[20] From this perspective, our theory suggests that the power of gerrymandering in a multi-district system should be in the hands of the opposition.

[18] A proof can be provided upon request.

[19] From a mathematical perspective, given that the first candidate has been elected, our model with δ different from zero is equivalent to a model with no re-election hurdle in which the electorate preferences are skewed (in favor of the incumbent if $\delta < 0$, in favor of the challenger if $\delta > 0$).

[20] The relation between polarization and gerrymandering has been studied by several authors—see e.g. Fiorina et al. (2005a) or McCarty et al. (2009).

6 Endogenous Re-election Hurdles

In analyzing the impact of extra-hurdles on policies, we have assumed that the value of δ is exogenously given to the politicians. Introducing such re-election hurdles should be done at the constitutional level. Yet this might be problematic, as the value of c that determines the optimal value for the extra-hurdle may change over time. Requiring re-election rules that differ from the majority rule and apply only to the incumbent (at a given election) might also be objectionable from a democratic perspective, as different candidates would hold on to different requirements. In this section we offer a way to handle these two objections. We do so by extending our model so that the choice of the extra-hurdle δ is made by the candidates themselves. More specifically, before the first election, candidates $k \in R$ and $k' \in L$ choose $\delta_k \in \left[0, \frac{1}{2}\right]$ and $\delta_{k'} \in \left[0, \frac{1}{2}\right]$, respectively. Then they both announce their choices publicly and commit to them. Such commitments can occur in environments where they are credible, say through *re-election hurdle contracts*. These are special political contracts outlined in Gersbach (2012). But commitments of this nature can also arise in the absence of such contracts if the politicians have the credibility to do so.[21] In the first election, the simple majority rule applies; in the case of a tie, each candidate wins the election with probability $\frac{1}{2}$. All voters know the announced values of δ_k and $\delta_{k'}$ when they vote in the first election. In the second election, if k (resp. k') is the office-holder, s/he is re-elected if his/her vote-share is equal to or larger than $\frac{1}{2} + \delta_k$ (resp. $\frac{1}{2} + \delta_{k'}$). We use $\mathcal{G}'$ to denote this new game. The sequence of events in the extended model is shown in Figure 6.

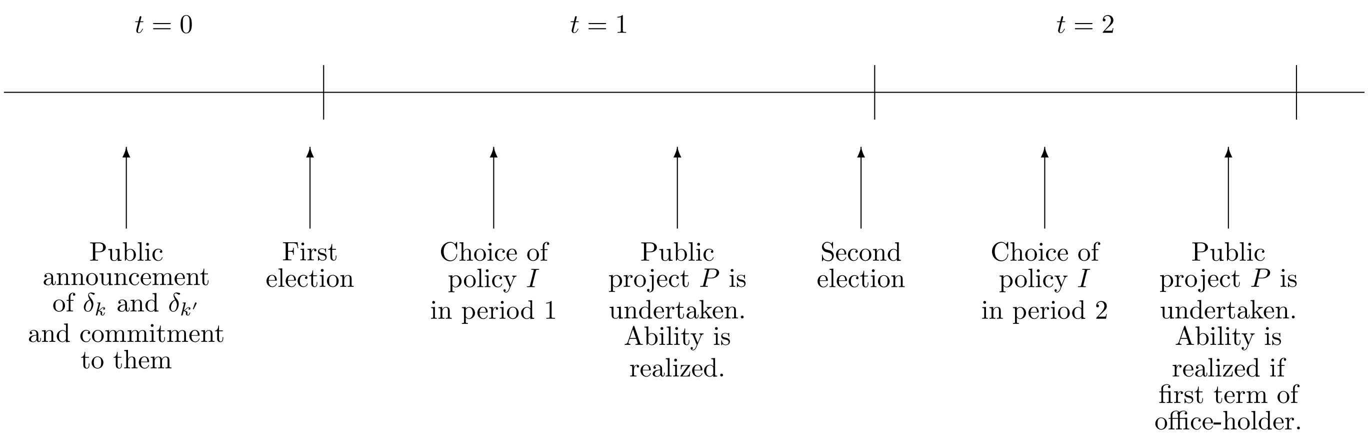


Figure 6: Timeline of the sequence of events in the extended model.

Under the mechanism described above, the following result holds:

[21] There exist cases where a politician has made commitments of this nature. For instance, José María Aznar (PP) committed to a two-term tenure before he was elected President of Spain in 1996 and only run once for re-election, despite the absence of any term limitation in the Spanish legislation.

Theorem 2

In equilibrium of the game $\mathcal{G}'$, $k \in R$ and $k' \in L$ commit to $\delta^* = \frac{c}{2}$, and both win the first election with probability equal to $\frac{1}{2}$.

Theorem 2 shows that in equilibrium the two candidates voluntarily commit to $\delta^* = \frac{c}{2}$, which is the unique extra-hurdle that is both W- and P-optimal. As a consequence, $\delta^* = \frac{c}{2}$ is implemented. This makes re-election hurdle contracts a legitimate method for implementing the concept of extra-hurdles in practice, even if c changes over time. Moreover, it does not entail violating any democratic principle. The intuition for this result is straightforward. If one candidate commits to a different value, say $\delta_k \neq \delta^*$, the other candidate can secure election (with probability one) by announcing δ^*. The reason is that δ^* is the unique maximizer of the expected lifetime utility of the median voter. Hence, the choice $\delta_k \neq \delta^*$ cannot be a best response to the extra-hurdle to which the other candidate commits, because any of the candidates is elected with (at least) probability $\frac{1}{2}$ if s/he chooses δ^*.

7 Social Polarization and Re-election Hurdles

So far we have assumed that the electorate's ideal policies were distributed uniformly across the interval $[0,1]$. In the present section we relax this assumption and allow for different levels of *social polarization*.[22] For tractability, we focus on the following parametric family of density functions, where $\alpha \in [0,1]$:

$$f_\alpha(i) = \begin{cases} -4\alpha i + (1+\alpha) & \text{if } i \in \left[0, \frac{1}{2}\right], \\ 4\alpha i + (1-3\alpha) & \text{if } i \in \left(\frac{1}{2}, 1\right]. \end{cases} \tag{19}$$

Any distribution in (19) with $\alpha > 0$ is bimodal, with kinks in $i = 0$ and $i = 1$. This is the description of an electorate with voters centered around two extreme positions. Parameter α thus reflects the degree of social polarization. In particular, if $\alpha = 0$, the distribution defined by $f_\alpha(\cdot)$ reduces to a uniform distribution on $[0,1]$, whereas $\alpha = 1$ captures the situation where there are few centrist voters. The density functions defined in (19) are illustrated in Figure 7.

Since the degree of social polarization can be directly measured by α, we denote welfare and ex-ante policy polarization by $W_\alpha(\cdot)$ and $EAP_\alpha(\cdot)$, respectively. The following theorem contains the main results on how extra-hurdles impact $EAP_\alpha(\cdot)$ and $W_\alpha(\cdot)$:

[22] We mentioned in Section 2 that the existence and rapid increase of policy polarization is a well-established fact. Whereas the preferences of an ample fraction of citizens may have been less subject to polarization, the preferences of citizens with high partisan identification have become more polarized. This has led to an increase in social polarization and, in turn, to a rise in party polarization, as the views of partisan citizens help shape party ideology (McCarty et al., 2006). Van Weelden (2015) finds that even in environments with low social polarization, strongly polarized party preferences can be welfare-enhancing.

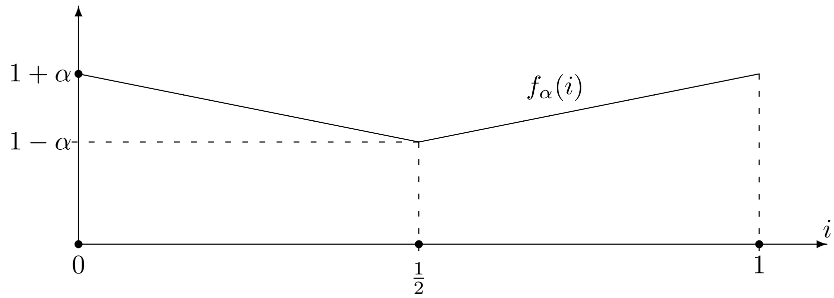


Figure 7: The density functions $f_\alpha(i)$ describe a socially polarized electorate. The parameter $\alpha \in [0,1]$ measures the degree of social polarization.

Theorem 3

Let $\alpha \in [0,1]$ and the distribution of the voters' ideal policies have a density $f_\alpha(\cdot)$. Then,

(i) Theorem 1 holds, with possibly a different W- and P-optimal extra-hurdle.

(ii) The unique W- and P-optimal extra-hurdle δ^*_α is continuous and decreasing in $\alpha \in [0,1]$.

(iii) Moreover,

(iii.a) $EAP_\alpha(0)$ and $W_\alpha(0)$ do not depend on α.

(iii.b) If some fixed $\delta \in \left[0, \frac{c^2}{2}\right]$ is chosen, $EAP_\alpha(\delta)$ is weakly decreasing in α and $W_\alpha(\delta)$ is weakly increasing in α.

Part *(i)* in Theorem 3 establishes the robustness of our main results for varying degrees of social polarization. According to part *(ii)*, the uniquely determined W- and P-optimal δ^*_α is not equal to $\frac{c}{2}$ for each value of $\alpha \in [0,1]$, as the effect of increasing the extra-hurdle is leveraged by the degree of social polarization. The net effect leads to the feature that the more polarized a society is, the lower the optimal extra-hurdle is. This implies that δ^*_α is decreasing in α. Finally, part *(iii)* looks at changes on α, keeping fixed the extra-hurdle δ. On the one hand, according to *(iii.a)*, policy polarization is not driven by the exact level of social polarization when the simple majority rule is applied in all stages, as is customary in current democratic systems. On the other hand, according to *(iii.b)*, policy polarization and welfare are both decreasing in the degree of social polarization for general extra-hurdles.

8 Concluding Remarks

We have analyzed a model of electoral competition in which changes of policies are costly for voters and candidates and these costs increase with the magnitude of the policy change. The model has allowed us to investigate a way to curb ex-ante policy polarization without reducing welfare: by setting the re-election hurdle higher than 50%. This result may have broader implications for democracy since significant levels of policy polarization may have negative mid- and long-term consequences on the functioning of democracy and the quality of its outcomes.

Numerous extensions of our model can be pursued. For instance, exploring the impact of costs of change in democracies with more than two parties, where governments typically consist of a coalition of parties, could reveal how such costs influence the say of parties in government policy-making. This and other extensions can be expected to further enrich our understanding of policy-making in democracy.

References

Alesina, A. and Tabellini, G. (1990). A positive theory of fiscal deficits and government debt. *Review of Economic Studies*, 57(3):403–414.

Alford, J. R. and Brady, D. W. (1989). Personal and partisan advantage in US congressional elections. In Dodd, L. C. and Oppenheimer, B. I., editors, *Congress Reconsidered*. Praeger Press, New York, NY.

Baron, D. P. (2018). Elections and durable governments in parliamentary governments. *Journal of Theoretical Politics*, 30(1):74–118.

Baron, D. P. and Bowen, T. R. (2015). Dynamic coalitions.

Benkler, Y., Faris, R., and Roberts, H. (2018). Network propaganda: Manipulation, disinformation, and radicalization in american politics.

Binder, S. A. (1999). The dynamics of legislative gridlock, 1947–96. *American Political Science Review*, 93(03):519–533.

Bouton, L., Lizzeri, A., and Persico, N. (2016). The political economy of debt and entitlements. Technical report, National Bureau of Economic Research.

Bowen, T. R., Chen, Y., and Eraslan, H. (2014). Mandatory versus discretionary spending: The status quo effect. *American Economic Review*, 104(10):2941–2974.

Bowen, T. R., Chen, Y., Eraslan, H., and Zápal, J. (2017). Efficiency of flexible budgetary institutions. *Journal of Economic Theory*, 167:148–176.

Boxell, L., Gentzkow, M., and Shapiro, J. M. (2017). Greater internet use is not associated with faster growth in political polarization among us demographic groups. *Proceedings of the National Academy of Sciences*, 114(40):10612–10617.

Buisseret, P. and Bernhardt, D. (2017). Dynamics of policymaking: Stepping back to leap forward, stepping forward to keep back. *American Journal of Political Science*, 61(4):820–835.

Callander, S., Raiha, D., et al. (2017). Durable policy, political accountability, and active waste. *Quarterly Journal of Political Science*, 12(1):59–97.

Chen, Y. and Eraslan, H. (2017). Dynamic agenda setting. *American Economic Journal: Microeconomics*, 9(2):1–32.

Duggan, J. and Fey, M. (2005). Electoral competition with policy-motivated candidates. *Games and Economic Behavior*, 51(2):490–522.

Dziuda, W. and Loeper, A. (2016). Dynamic collective choice with endogenous status quo. *Journal of Political Economy*, 124(4):1148–1186.

Dziuda, W. and Loeper, A. (2018). Dynamic pivotal politics. *American Political Science Review*, pages 1–22.

Erikson, R. S., Wright, G. C., and McIver, J. P. (1993). *Statehouse Democracy: Public Opinion and Policy in the American States*. Cambridge University Press, Cambridge.

Fiorina, M. P., Abrams, S. J., and C, P. J. (2005a). Culture war? In *The myth of a polarized America*. New York, Pearson Longman.

Fiorina, M. P., Abrams, S. J., and Pope, J. (2005b). *Culture War?* Pearson Longman, New York, NY.

Forand, J. G. (2014). Two-party competition with persistent policies. *Journal of Economic Theory*, 152:64–91.

Gelman, A. and King, G. (1990). Estimating incumbency advantage without bias. *American Journal of Political Science*, 34(4):1142–1164.

Gentzkow, M. (2016). Polarization in 2016. Toulouse Network for Information Technology Whitepaper.

Gersbach, H. (2012). Contractual democracy. *Review of Law & Economics*, 8(3):823–851.

Gersbach, H. (2017). Vote thresholds with signaling of competence. In *Redesigning Democracy*, pages 55–81. Springer.

Gersbach, H. (2020). History-bound reelections. *American Economic Journal: Microeconomics*, 12(3):33–75.

Gersbach, H. et al. (2010). On higher hurdles for incumbents. *Economics Bulletin*, 30(1):774–785.

Gersbach, H., Jackson, M. O., Muller, P., and Tejada, O. (2020a). Electoral competition with costly policy changes: A dynamic perspective. *Available at SSRN*.

Gersbach, H., Jackson, M. O., and Tejada, O. (2020b). Optimal term-length and the structure of democracy. *Available at SSRN*.

Gersbach, H. and Liessem, V. (2008). Reelection threshold contracts in politics. *Social Choice and Welfare*, 31(2):233–255.

Gersbach, H. and Müller, M. (2017). Higher bars for incumbents and experience. *Journal of Theoretical Politics*, 29(3):492–513.

Gersbach, H., Muller, P., and Tejada, O. (2019a). Costs of change and political polarization. *European Journal of Political Economy*, 60:101801.

Gersbach, H., Muller, P., and Tejada, O. (2019b). Costs of change and political polarization. *European Journal of Political Economy*, 60:101801.

Gersbach, H. and Tejada, O. (2018). A reform dilemma in polarized democracies. *Journal of Public Economics*, 160:148–158.

Glazer, A., Gradstein, M., and Konrad, K. A. (1998). The electoral politics of extreme policies. *Economic Journal*, 108(451):1677–1685.

Hansson, I. and Stuart, C. (1984). Voting competitions with interested politicians: Platforms do not converge to the preferences of the median voter. *Public Choice*, 44(3):431–441.

Heberlig, E., Hetherington, M., and Larson, B. (2006). The price of leadership: Campaign money and the polarization of congressional parties. *Journal of Politics*, 68(4):992–1005.

Hetherington, M. J. (2001). Resurgent mass partisanship: The role of elite polarization. *American Political Science Review*, 95(3):619–632.

Jamieson, K. H. and Falk, E. (2000). Continuity and change in civility in the House. In Bond, J. R. and Fleisher, R., editors, *Polarized Politics: Congress and the President in a Partisan Era*, pages 96–108. CQ Press, Washington, DC.

Jones, D. R. (2001). Party polarization and legislative gridlock. *Political Research Quarterly*, 54(1):125–141.

Krasa, S. and Polborn, M. (2010a). The binary policy model. *Journal of Economic Theory*, 145(2):661–688.

Krasa, S. and Polborn, M. K. (2010b). Competition between specialized candidates. *American Political Science Review*, 104(4):745–765.

Layman, G. C., Carsey, T. M., and Horowitz, J. M. (2006). Party polarization in American politics: Characteristics, causes, and consequences. *Annual Review of Political Science*, 9:83–110.

Levitt, S. D. and Wolfram, C. D. (1997). Decomposing the sources of incumbency advantage in the US House. *Legislative Studies Quarterly*, 22(1):45–60.

McCarty, N., Poole, K. T., and Rosenthal, H. (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. MIT Press, Cambridge, MA.

McCarty, N., Poole, K. T., and Rosenthal, H. (2009). Does gerrymandering cause polarization? *American Journal of Political Science*, 53(3):666–680.

McCarty, N., Poole, K. T., and Rosenthal, H. (2016). *Polarized America: The Dance of Ideology and Unequal Riches*. MIT Press Books. The MIT Press.

Nunnari, S. and Zápal, J. (2017). Dynamic elections and ideological polarization. *Political Analysis*, 25(4):505–534.

Persson, T. and Svensson, L. E. (1989). Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences. *Quarterly Journal of Economics*, 104(2):325–345.

Poole, K. T. and Rosenthal, H. (1984). The polarization of American politics. *Journal of Politics*, 46(04):1061–1079.

Poole, K. T. and Rosenthal, H. (2001). D-nominate after 10 years: A comparative update to Congress: a political-economic history of roll-call voting. *Legislative Studies Quarterly*, 26(1):5–29.

Presa, D. G. (2019). A 200 años de la constitución de 1819. breve comentario de su historia. *Derechos en Acción*, 11(11):280–280.

Prior, M. (2013). Media and political polarization. *Annual Review of Political Science*, 16(1):101–127.

Roberts, J. M. and Smith, S. S. (2003). Procedural contexts, party strategy, and conditional party voting in the U.S. House of Representatives, 1971–2000. *American Journal of Political Science*, 47(2):305–317.

Rohde, D. W. (2010). *Parties and Leaders in the Postreform House.* University of Chicago Press, Chicago, IL.

Sinclair, B. (2002). The dream fulfilled? Party development in Congress, 1950-2000. In Green, J. C. and Herrnson, P. S., editors, *Responsible Partisanship? The evolution of American political parties since 1950*, pages 121–140. University Press of Kansas, Lawrence, KS.

SP Kanton Zürich (2010). Reglement Nominationsverfahren Ständerats- und Nationalratswahlen. November 30, 2010. `http://www.spkantonzh.ch/icc.asp?oid=33448` (retrieved 27 January 2015).

Steinberg, S. H. (1962). *Liberia*, pages 1197–1202. Palgrave Macmillan UK, London.

Testa, C. (2012). Is polarization bad? *European Economic Review*, 56(6):1104–1118.

Theriault, S. M. (2004). The case of the vanishing moderates: Party polarization in the modern Congress. In *Annual Meeting of the Midwest Political Science Association*, 15–18 April 2004, Chicago, IL.

Theriault, S. M. (2006). Party polarization in the U.S. Congress: Member replacement and member adaptation. *Party Politics*, 12(4):483–503.

Van Weelden, R. (2015). The welfare implications of electoral polarization. *Social Choice and Welfare*, 45(4):653–686.

Zápal, J. (2016). Markovian equilibria in dynamic spatial legislative bargaining: Existence with three players. *Games and Economic Behavior*, 98:235–242.

Appendix A: Proofs

This appendix contains all proofs from the main body of the paper.

Proof of Proposition 2

Let $i \in [0,1]$ be an arbitrary voter. The ability of the incumbent, $k \in R$, is common knowledge when citizens vote the second time. However, the ability of the challenging left-wing candidate, say $k' \in L$, is not known, because s/he is a new candidate. According to (1), $E[a_{k'}] = 0$. If k wins the second election, i expects the policy to be $i^*_{R2}(i_{k1})$, while i expects $i^*_{L2}(i_{k1})$ if k' wins. Thus, since i votes for k if s/he prefers office-holder k to win the second election, s/he votes for k if

$$a_k - (i^*_{R2}(i_{k1}) - i)^2 - c|i^*_{R2}(i_{k1}) - i_{k1}| \geq E[a_{k'}] - (i^*_{L2}(i_{k1}) - i)^2 - c|i^*_{L2}(i_{k1}) - i_{k1}|.$$

The above inequality is equivalent to

$$a_k > \underbrace{(i^*_{R2}(i_{k1}) - i)^2 + c|i^*_{R2}(i_{k1}) - i_{k1}| - (i^*_{L2}(i_{k1}) - i)^2 - c|i^*_{L2}(i_{k1}) - i_{k1}|}_{=:(\dagger)}. \tag{20}$$

Under (2) and (5), we have

$$i^*_{R2}(i_{k1}) > i^*_{L2}(i_{k1}) \text{ for all } i_{k1} \in [0,1]. \tag{21}$$

Indeed, by (2) and (5), $\mu_R - \frac{c}{2} > \mu_L + \frac{c}{2}$ holds, and by (6) we know that $i^*_{R2}(i_{k1}) \geq \mu_R - \frac{c}{2}$ and $i^*_{L2}(i_{k1}) \leq \mu_L + \frac{c}{2}$. From (21) it follows that (†) defined in (20) is strictly decreasing in i, since

$$\frac{d(\dagger)}{di} = 2\left(i^*_{L2}(i_{k1}) - i^*_{R2}(i_{k1})\right) < 0.$$

Thus, because an incumbent is re-elected in the second election if s/he receives a vote-share of $\frac{1}{2} + \delta$ or larger, the critical voter in the second election is[23]

$$i = \frac{1}{2} - \delta. \tag{22}$$

Inserting (22) into (20) proves that office-holder k is re-elected if and only if (7) holds.[24]

□

Proof of Proposition 3

We prove Proposition 3 for $c > 0$. The proof for $c = 0$ is similar and thus omitted here. For given δ, office-holder k aims at maximizing his/her expected utility at the beginning of $t = 1$, which depends on i_{k1}, his/her policy choice in the first period. It is useful to introduce the following notation:

[23] This builds on the assumption that voters are uniformly distributed on $[0,1]$. Note that, qualitatively, the main results of our paper would also hold for any other symmetric distribution of the voters, as detailed in Appendix 7.

[24] Whether indifferent voters vote for the incumbent or the challenger does not influence the outcome of the election. Indeed, the critical voter has measure zero, so the election is tied whenever s/he is indifferent between both candidates, independently of his/her voting behavior.

- $EU_\delta(i_{k1})$ denotes k's expected utility at the beginning of $t = 1$ as a function of i_{k1} and parametrized by $\delta \in \left[0, \frac{1}{2}\right]$, and

- $p_\delta(i_{k1})$ denotes k's re-election probability in the second election as a function of i_{k1} and parametrized by $\delta \in \left[0, \frac{1}{2}\right]$.[25]

By Proposition 2, we have

$$p_\delta(i_{k1}) = P[a_k \geq a_\delta(i_{k1})].$$

Moreover, we have assumed in Section 3.4 that A is sufficiently large, so

$$-A < a_\delta(i_{k1}) < A$$

for all $\delta \in \left[0, \frac{1}{2}\right]$ and all $i_{k1} \in [0, 1]$. Therefore, for every fixed value of δ, $p_\delta(\cdot)$ is a piecewise constant function with $p_\delta(i_{k1}) \in \left\{\frac{1-\rho}{2}, \frac{1+\rho}{2}\right\}$ for all $i_{k1} \in [0, 1]$. More precisely,

$$p_\delta(i_{k1}) = \begin{cases} \frac{1+\rho}{2} & \text{if } a_\delta(i_{k1}) \leq 0, \\ \frac{1-\rho}{2} & \text{if } a_\delta(i_{k1}) > 0. \end{cases} \tag{23}$$

We can now formulate the maximization problem that k faces in $t = 1$. For a given $\delta \in \left[0, \frac{1}{2}\right]$, k chooses $i_{k1} \in [0, 1]$ so that

$$\begin{aligned} EU_\delta(i_{k1}) =& p_\delta(i_{k1}) \cdot \Big\{2b + 2E[a_k | a_k \geq a_\delta(i_{k1})] \\ & - (i_{k1} - \mu_R)^2 - (i^*_{R2}(i_{k1}) - \mu_R)^2 - c|i_{k1} - i^*_{R2}(i_{k1})|\Big\} \\ & + (1 - p_\delta(i_{k1})) \cdot \Big\{b + E[a_k | a_k < a_\delta(i_{k1})] + E[a_{k'}] \\ & - (i_{k1} - \mu_R)^2 - (i^*_{L2}(i_{k1}) - \mu_R)^2 - c|i_{k1} - i^*_{L2}(i_{k1})|\Big\} \end{aligned} \tag{24}$$

is maximized, where k' is the challenger in the second election. Since b is large and the ability distribution is discrete, maximization of re-election probability $p_\delta(i_{k1})$ is a necessary condition for maximizing $EU_\delta(i_{k1})$. Hence office-holder k's re-election probability is given by

$$p^*(\delta) = \max_{i_{k1} \in [0,1]} p_\delta(i_{k1}). \tag{25}$$

We now define $I^*(\delta)$ as

$$I^*(\delta) = \operatorname*{argmax}_{i_{k1} \in [0,1]} p_\delta(i_{k1}). \tag{26}$$

We next observe that solving

$$\operatorname*{argmax}_{i_{k1} \in [0,1]} EU_\delta(i_{k1})$$

[25] Note that this is re-election probability as perceived before a_k has been realized. Furthermore, $p_\delta(i_{k1})$ is to be understood as conditional on k being in office in $t = 1$.

is equivalent to solving

$$\underset{i_{k1}\in I^*(\delta)}{\operatorname{argmax}}\, EU_\delta(i_{k1}|p^*(\delta)),$$

where $EU_\delta(i_{k1}|p^*(\delta))$ denotes k's expected utility at the beginning of $t=1$, conditional on $p_\delta(i_{k1})$ being equal to $p^*(\delta)$. Accordingly, we now proceed in three steps to maximize $EU_\delta(i_{k1})$. In Step 1, we compute $p^*(\delta)$ and $I^*(\delta)$ as a function of $\delta\in\left[0,\frac{1}{2}\right]$. Then, in Step 2, we solve

$$\underset{i_{k1}\in [0,1]}{\operatorname{argmax}}\, EU_\delta(i_{k1}|p^*(\delta)).$$

Finally, in Step 3, we restrict the solutions of Step 2 to $I^*(\delta)$, that is we compute

$$\underset{i_{k1}\in I^*(\delta)}{\operatorname{argmax}}\, EU_\delta(i_{k1}|p^*(\delta)).$$

Step 1: Computation of $p^*(\delta)$ and $I^*(\delta)$ as functions of $\delta\in\left[0,\frac{1}{2}\right]$

First, it is useful to define

$$\mathcal{A}^\delta := \{i_{k1}\in[0,1] \mid a_\delta(i_{k1})\le 0\}\,, \tag{27}$$

for each $\delta\in\left[0,\frac{1}{2}\right]$. $\mathcal{A}^\delta$ is the set of policies that ensure re-election when the office-holder turns out to have zero ability.[26] By (23), (25), and (26), we know that

$$p^*(\delta)=\begin{cases}\frac{1+\rho}{2} & \text{if } \mathcal{A}^\delta\neq\emptyset,\\ \frac{1-\rho}{2} & \text{if } \mathcal{A}^\delta=\emptyset\end{cases} \tag{28}$$

and

$$I^*(\delta)=\begin{cases}\mathcal{A}^\delta & \text{if } \mathcal{A}^\delta\neq\emptyset,\\ [0,1] & \text{if } \mathcal{A}^\delta=\emptyset.\end{cases} \tag{29}$$

Second, to determine $\mathcal{A}^\delta$, it is useful to define the sets

$$I_1 := \left[0, max\left\{\mu_L-\frac{c}{2},0\right\}\right), \tag{30}$$
$$I_2 := \left[max\left\{\mu_L-\frac{c}{2},0\right\}, \mu_L+\frac{c}{2}\right], \tag{31}$$
$$I_3 := \left(\mu_L+\frac{c}{2}, \mu_R-\frac{c}{2}\right), \tag{32}$$
$$I_4 := \left[\mu_R-\frac{c}{2}, min\left\{\mu_R+\frac{c}{2},1\right\}\right], \tag{33}$$
$$I_5 := \left(min\left\{\mu_R+\frac{c}{2},1\right\},1\right], \tag{34}$$

since the expressions in (6) for $i^*_{R2}(i_{k1})$ and $i^*_{L2}(i_{k1})$ depend on whether $i_{k1}\in I_1$, $i_{k1}\in I_2$, $i_{k1}\in I_3$, $i_{k1}\in I_4$ or $i_{k1}\in I_5$.[27] Since

$$I_1\cup I_2\cup I_3\cup I_4\cup I_5=[0,1],$$

[26] We stress that the office-holder does not know his/her ability when s/he chooses i_{k1}. Moreover, recall that an office-holder with ability A is always re-elected and an office-holder with ability $-A$ never is.

[27] Note that $[0,0)$ and $(1,1]$ are empty sets.

we can use the decomposition

$$\mathcal{A}^\delta = \underbrace{\left\{\mathcal{A}^\delta \cap I_1\right\}}_{\text{Step 1a}} \cup \underbrace{\left\{\mathcal{A}^\delta \cap I_2\right\}}_{\text{Step 1b}} \cup \underbrace{\left\{\mathcal{A}^\delta \cap I_3\right\}}_{\text{Step 1c}} \cup \underbrace{\left\{\mathcal{A}^\delta \cap I_4\right\}}_{\text{Step 1d}} \cup \underbrace{\left\{\mathcal{A}^\delta \cap I_5\right\}}_{\text{Step 1e}} \tag{35}$$

to determine $\mathcal{A}^\delta$. Subsequently, in Steps 1a to 1e, we compute expressions for the subsets of $\mathcal{A}^\delta$ listed in (35). In Step 1f, we use Steps 1a to 1e to obtain expressions for $p^*(\delta)$ and $I^*(\delta)$, for all $\delta \in \left[0, \frac{1}{2}\right]$.

Step 1a: Computation of $\mathcal{A}^\delta \cap I_1$

If $i_{k1} \in I_1$, then, by (6),

$$i^*_{R2}(i_{k1}) = \mu_R - \frac{c}{2} > i_{k1} \quad \text{and} \quad i^*_{L2}(i_{k1}) = \mu_L - \frac{c}{2} > i_{k1}.$$

Thus, by (8),

$$\begin{aligned} a_\delta(i_{k1}) =& \left(\mu_R - \frac{c}{2} - \left(\frac{1}{2} - \delta\right)\right)^2 + c\left(\mu_R - \frac{c}{2} - i_{k1}\right) \\ & - \left(\mu_L - \frac{c}{2} - \left(\frac{1}{2} - \delta\right)\right)^2 - c\left(\mu_L - \frac{c}{2} - i_{k1}\right) \\ =& \left(\mu_R - \frac{c}{2}\right)^2 - \left(\mu_L - \frac{c}{2}\right)^2 - 2\left(\frac{1}{2} - \delta\right)(\mu_R - \mu_L) + c(\mu_R - \mu_L) \\ =& (2\mu_R - 1)(1 - c) - (2\mu_R - 1)(1 - c) + 2\delta(2\mu_R - 1) \\ =& 2\delta(2\mu_R - 1), \end{aligned}$$

where we have used $\mu_L = 1 - \mu_R$. By (27), this yields

$$\mathcal{A}^\delta \cap I_1 = \begin{cases} I_1 & \text{if } \delta = 0, \\ \emptyset & \text{if } \delta \in \left(0, \frac{1}{2}\right]. \end{cases} \tag{36}$$

Step 1b: Computation of $\mathcal{A}^\delta \cap I_2$

If $i_{k1} \in I_2$, then, by (6),

$$i^*_{R2}(i_{k1}) = \mu_R - \frac{c}{2} > i_{k1} \quad \text{and} \quad i^*_{L2}(i_{k1}) = i_{k1}.$$

Hence, by (8),

$$\begin{aligned} a_\delta(i_{k1}) =& \left(\mu_R - \frac{c}{2} - \left(\frac{1}{2} - \delta\right)\right)^2 + c\left(\mu_R - \frac{c}{2} - i_{k1}\right) - \left(i_{k1} - \left(\frac{1}{2} - \delta\right)\right)^2 \\ =& - i_{k1}^2 + (1 - 2\delta - c)i_{k1} + \left(\mu_R - \frac{c}{2}\right)^2 - \left(\mu_R - \frac{c}{2}\right)(1 - 2\delta - c). \end{aligned}$$

The discriminant of

$$-i_{k1}^2 + (1-2\delta-c)i_{k1} + \left(\mu_R - \frac{c}{2}\right)^2 - \left(\mu_R - \frac{c}{2}\right)(1-2\delta-c)$$

is

$$\Delta := (2\delta + 2\mu_R - 1)^2.$$

Hence,

$$\begin{aligned}
\mathcal{A}^\delta \cap I_2 &= \left\{\left(-\infty, \mu_L - \frac{c}{2} - 2\delta\right] \cup \left[\mu_R - \frac{c}{2}, +\infty\right)\right\} \cap I_2 \\
&= \begin{cases} \left\{\mu_L - \frac{c}{2}\right\} & \text{if } \delta = 0 \text{ and } \mu_L - \frac{c}{2} \geq 0, \\ \emptyset & \text{otherwise.} \end{cases}
\end{aligned} \tag{37}$$

Step 1c: Computation of $\mathcal{A}^\delta \cap I_3$

If $i_{k1} \in I_3$, then, by (6),

$$i_{R2}^*(i_{k1}) = \mu_R - \frac{c}{2} > i_{k1} \quad \text{and} \quad i_{L2}^*(i_{k1}) = \mu_L + \frac{c}{2} < i_{k1},$$

and, by (8),

$$\begin{aligned}
a_\delta(i_{k1}) =& \left(\mu_R - \frac{c}{2} - \left(\frac{1}{2} - \delta\right)\right)^2 + c\left(\mu_R - \frac{c}{2} - i_{k1}\right) \\
&- \left(\mu_L + \frac{c}{2} - \left(\frac{1}{2} - \delta\right)\right)^2 + c\left(\mu_L + \frac{c}{2} - i_{k1}\right) \\
=& \underbrace{\left(\mu_R - \frac{c}{2}\right)^2 - \left(\mu_L + \frac{c}{2}\right)^2}_{=2\mu_R - 1 - c} - 2\left(\frac{1}{2} - \delta\right)(2\mu_R - 1 - c) + c(1 - 2i_{k1}) \\
=& 2\delta(2\mu_R - 1 - c) - 2i_{k1}c + c.
\end{aligned}$$

Hence,

$$a_\delta(i_{k1}) \leq 0 \Leftrightarrow i_{k1} \geq \frac{2\delta(2\mu_R - 1 - c) + c}{2c} \tag{38}$$

and

$$\mathcal{A}^\delta \cap I_3 = \left[\frac{2\delta(2\mu_R - 1 - c) + c}{2c}, +\infty\right) \cap I_3.$$

We observe that

$$\frac{2\delta(2\mu_R - 1 - c) + c}{2c} \geq \frac{1}{2} > \mu_L + \frac{c}{2},$$

for all $\delta \geq 0$, and

$$\frac{2\delta(2\mu_R - 1 - c) + c}{2c} < \mu_R - \frac{c}{2} \Leftrightarrow \delta < \frac{c}{2}.$$

Therefore,

$$\mathcal{A}^\delta \cap I_3 = \begin{cases} \left[\frac{2\delta(2\mu_R-1-c)+c}{2c}, \mu_R - \frac{c}{2}\right) & \text{if } \delta \in \left[0, \frac{c}{2}\right), \\ \emptyset & \text{if } \delta \in \left[\frac{c}{2}, \frac{1}{2}\right]. \end{cases} \tag{39}$$

Step 1d: Computation of $\mathcal{A}^\delta \cap I_4$

If $i_{k1} \in I_4$, then, by (6),

$$i^*_{R2}(i_{k1}) = i_{k1} \quad \text{and} \quad i^*_{L2}(i_{k1}) = \mu_L + \frac{c}{2} < i_{k1},$$

and thus, by (8),

$$\begin{aligned} a_\delta(i_{k1}) &= \left(i_{k1} - \left(\frac{1}{2} - \delta\right)\right)^2 - \left(\mu_L + \frac{c}{2} - \left(\frac{1}{2} - \delta\right)\right)^2 + c\left(\mu_L + \frac{c}{2} - i_{k1}\right) \\ &= i_{k1}^2 + (2\delta - 1 - c)i_{k1} - \left[\left(\mu_L + \frac{c}{2}\right)^2 + \left(\mu_L + \frac{c}{2}\right)(2\delta - 1 - c)\right]. \end{aligned}$$

The discriminant of

$$i_{k1}^2 + (2\delta - 1 - c)i_{k1} - \left[\left(\mu_L + \frac{c}{2}\right)^2 + \left(\mu_L + \frac{c}{2}\right)(2\delta - 1 - c)\right]$$

is

$$\Delta := (2\delta - 2\mu_R + 1)^2.$$

Hence,

$$\begin{aligned} \mathcal{A}^\delta \cap I_4 &= \left[\frac{1 + c - 2\delta - |2\delta - 2\mu_R + 1|}{2}, \frac{1 + c - 2\delta + |2\delta - 2\mu_R + 1|}{2}\right] \cap I_4 \\ &= \begin{cases} \left[\mu_L + \frac{c}{2}, \mu_R + \frac{c}{2} - 2\delta\right] \cap I_4 & \text{if } \delta < \mu_R - \frac{1}{2}, \\ \left[\mu_R + \frac{c}{2} - 2\delta, \mu_L + \frac{c}{2}\right] \cap I_4 & \text{if } \delta \geq \mu_R - \frac{1}{2}. \end{cases} \\ &= \begin{cases} \left[\mu_L + \frac{c}{2}, \mu_R + \frac{c}{2} - 2\delta\right] \cap I_4 & \text{if } \delta < \mu_R - \frac{1}{2}, \\ \emptyset & \text{if } \delta \geq \mu_R - \frac{1}{2}, \end{cases} \end{aligned} \tag{40}$$

where the last equality follows from $\mu_L + \frac{c}{2} < \mu_R - \frac{c}{2}$. Moreover, since

$$\mu_R + \frac{c}{2} - 2\delta \geq \mu_R - \frac{c}{2} \Leftrightarrow \delta \leq \frac{c}{2}$$

and $\mu_L + \frac{c}{2} < \mu_R - \frac{c}{2}$, the case where $\delta < \mu_R - \frac{1}{2}$ can be reformulated as

$$\mathcal{A}^\delta \cap I_4 = \begin{cases} \left[\mu_R - \frac{c}{2}, min\left\{\mu_R + \frac{c}{2} - 2\delta, 1\right\}\right] & \text{if } \delta \in \left[0, \frac{c}{2}\right], \\ \emptyset & \text{if } \delta \in \left(\frac{c}{2}, \mu_R - \frac{1}{2}\right). \end{cases} \tag{41}$$

Finally, combining (40) and (41) yields

$$\mathcal{A}^\delta \cap I_4 = \begin{cases} \left[\mu_R - \frac{c}{2}, min\left\{\mu_R + \frac{c}{2} - 2\delta, 1\right\}\right] & \text{if } \delta \in \left[0, \frac{c}{2}\right], \\ \emptyset & \text{if } \delta \in \left(\frac{c}{2}, \frac{1}{2}\right]. \end{cases} \tag{42}$$

Step 1e: Computation of $\mathcal{A}^\delta \cap I_5$

If $i_{k1} \in I_5$, then, by (6),

$$i^*_{R2}(i_{k1}) = \mu_R + \frac{c}{2} < i_{k1} \quad \text{and} \quad i^*_{L2}(i_{k1}) = \mu_L + \frac{c}{2} < i_{k1}.$$

Thus, by (8),

$$\begin{aligned} a_\delta(i_{k1}) =& \left(\mu_R + \frac{c}{2} - \left(\frac{1}{2} - \delta\right)\right)^2 - c\left(\mu_R + \frac{c}{2} - i_{k1}\right) \\ &- \left(\mu_L + \frac{c}{2} - \left(\frac{1}{2} - \delta\right)\right)^2 + c\left(\mu_L + \frac{c}{2} - i_{k1}\right) \\ =& \left(\mu_R + \frac{c}{2}\right)^2 - \left(\mu_L + \frac{c}{2}\right)^2 - 2\left(\frac{1}{2} - \delta\right)(\mu_R - \mu_L) - c(\mu_R - \mu_L) \\ =&(2\mu_R - 1)(1 + c) - (2\mu_R - 1)(1 + c) + 2\delta(2\mu_R - 1) \\ =&2\delta(2\mu_R - 1). \end{aligned}$$

By (27), this yields

$$\mathcal{A}^\delta \cap I_5 = \begin{cases} I_5 & \text{if } \delta = 0, \\ \emptyset & \text{if } \delta \in \left(0, \frac{1}{2}\right]. \end{cases} \tag{43}$$

Step 1f: Computation of $p^*(\delta)$ and $I^*(\delta)$

Let us now combine (36), (37), (39), (42), and (43) to obtain compact expressions for $\mathcal{A}^\delta$. These expressions enable us to compute $p^*(\delta)$ and $I^*(\delta)$ with the help of (28) and (29). We distinguish three different cases.

Case 1: $\delta = 0$

We consider two subcases depending on the value of the parameter c.

Case 1a: $c \leq 2(1 - \mu_R)$

In this case, $\mu_L - \frac{c}{2} \geq 0$, so

$$\mathcal{A}^0 = \left[0, \mu_L - \frac{c}{2}\right] \cup \left[\frac{1}{2}, 1\right],$$

which, by (28) and (29), yields

$$p^*(0) = \frac{1 + \rho}{2} \quad \text{and} \quad I^*(0) = \left[0, \mu_L - \frac{c}{2}\right] \cup \left[\frac{1}{2}, 1\right]. \tag{44}$$

Case 1b: $c > 2(1 - \mu_R)$

In this case, $\mu_L - \frac{c}{2} < 0$, so

$$\mathcal{A}^0 = \left[\frac{1}{2}, 1\right],$$

which, by (28) and (29), yields

$$p^*(0) = \frac{1+\rho}{2} \quad \text{and} \quad I^*(0) = \left[\frac{1}{2}, 1\right]. \tag{45}$$

<u>*Case 2:*</u> $\delta \in \left(0, \frac{c}{2}\right]$

In this case, (36), (37), (39), (42), and (43) yield

$$\mathcal{A}^\delta = \left[\frac{2\delta(2\mu_R - 1 - c) + c}{2c}, min\left\{\mu_R + \frac{c}{2} - 2\delta, 1\right\}\right],$$

which in turn implies

$$p^*(\delta) = \frac{1+\rho}{2} \quad \text{and} \quad I^*(\delta) = \left[\frac{2\delta(2\mu_R - 1 - c) + c}{2c}, min\left\{\mu_R + \frac{c}{2} - 2\delta, 1\right\}\right], \tag{46}$$

by (28) and (29).

<u>*Case 3:*</u> $\delta \in \left(\frac{c}{2}, \frac{1}{2}\right]$

For such extra-hurdles

$$\mathcal{A}^\delta = \emptyset,$$

so, by (28) and (29),

$$p^*(\delta) = \frac{1-\rho}{2} \quad \text{and} \quad I^*(\delta) = [0, 1]. \tag{47}$$

Step 2: Solution to $\text{argmax}_{i_{k1}\in[0,1]} EU_\delta(i_{k1}|p^*(\delta))$

From (24) we obtain

$$\begin{aligned} EU_\delta(i_{k1}|p^*(\delta)) =& p^*(\delta) \cdot \Big\{2b + 2E[a_k|k \text{ re-elected and } p_\delta(i_{k1}) = p^*(\delta)] \\ & - (i_{k1} - \mu_R)^2 - (i^*_{R2}(i_{k1}) - \mu_R)^2 - c|i_{k1} - i^*_{R2}(i_{k1})|\Big\} \\ & + (1 - p^*(\delta)) \cdot \Big\{b + E[a_k|k \text{ not re-elected and } p_\delta(i_{k1}) = p^*(\delta)] + E[a_{k'}] \\ & - (i_{k1} - \mu_R)^2 - (i^*_{L2}(i_{k1}) - \mu_R)^2 - c|i_{k1} - i^*_{L2}(i_{k1})|\Big\}. \end{aligned} \tag{48}$$

This expression of $EU_\delta(i_{k1}|p^*(\delta))$ can be simplified using

$$E[a_k|k \text{ re-elected and } p_\delta(i_{k1}) = p^*(\delta)] = \frac{A(1-\rho)}{2p^*(\delta)} \tag{49}$$

and

$$E[a_k|k \text{ not re-elected and } p_\delta(i_{k1}) = p^*(\delta)] = -\frac{A(1-\rho)}{2(1 - p^*(\delta))}. \tag{50}$$

Recall that, by (23),

$$p^*(\delta) \in \left\{\frac{1-\rho}{2}, \frac{1+\rho}{2}\right\}.$$

Then, (49) follows from

$$\begin{aligned}
E[a_k|k \text{ re-elected and } p_\delta(i_{k1}) = p^*(\delta)] &= \begin{cases} E[a_k|a_k = A] & \text{if } p^*(\delta) = \frac{1-\rho}{2}, \\ E[a_k|a_k \in \{0, A\}] & \text{if } p^*(\delta) = \frac{1+\rho}{2}, \end{cases} \\
&= \begin{cases} A & \text{if } p^*(\delta) = \frac{1-\rho}{2}, \\ \frac{A(1-\rho)}{1+\rho} & \text{if } p^*(\delta) = \frac{1+\rho}{2}, \end{cases} \\
&= \frac{A(1-\rho)}{2p^*(\delta)}.
\end{aligned}$$

The proof of (50) is similar. Inserting (49) and (50) into (48) yields

$$\begin{aligned}
EU_\delta(i_{k1}|p^*(\delta)) =& \frac{A(1-\rho)}{2} + b \cdot (1 + p^*(\delta)) \\
&+ p^*(\delta) \cdot \left[-(i_{k1} - \mu_R)^2 - (i^*_{R2}(i_{k1}) - \mu_R)^2 - c|i_{k1} - i^*_{R2}(i_{k1})|\right] \\
&+ (1 - p^*(\delta)) \cdot \left[-(i_{k1} - \mu_R)^2 - (i^*_{L2}(i_{k1}) - \mu_R)^2 - c|i_{k1} - i^*_{L2}(i_{k1})|\right].
\end{aligned} \tag{51}$$

We now analyze the behavior of $EU_\delta(i_{k1}|p^*(\delta))$, given by (51), as a function of $i_{k1} \in [0, 1]$. We observe from (6) and (51) that, for each $\delta \in \left[0, \frac{1}{2}\right]$, $EU_\delta(i_{k1}|p^*(\delta))$ is continuous in i_{k1} on $[0, 1]$ and differentiable on

$$(0, 1)\backslash \left\{\mu_L - \frac{c}{2}, \mu_L + \frac{c}{2}, \mu_R - \frac{c}{2}, \mu_R + \frac{c}{2}\right\}.$$

Accordingly, in Steps 2a to 2e we analyze the sign of

$$\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}}$$

for $i_{k1} \in I^\circ_j$ and $j = 1, 2, \ldots, 5$. Recall that $I_1, I_2, \ldots, I_5$ are defined in (30)-(34) and, for each $j \in \{1, 2, \ldots, 5\}$, let I°_j denote the interior of I_j. In Step 2f, we combine Steps 2a to 2e to solve $\text{argmax}_{i_{k1}\in[0,1]} EU_\delta(i_{k1}|p^*(\delta))$.

Step 2a: $\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}}$ **for** $i_{k1} \in I^\circ_1$

If $i_{k1} \in I^\circ_1$, then, by (6),

$$i^*_{R2}(i_{k1}) = \mu_R - \frac{c}{2} > i_{k1} \quad \text{and} \quad i^*_{L2}(i_{k1}) = \mu_L - \frac{c}{2} > i_{k1}.$$

Thus, by (51),

$$\begin{aligned}
\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}} =& p^*(\delta) \cdot [-2(i_{k1} - \mu_R) + c] + (1 - p^*(\delta)) \cdot [-2(i_{k1} - \mu_R) + c] \\
=& -2(i_{k1} - \mu_R) + c \\
>& 0,
\end{aligned} \tag{52}$$

for all $i_{k1} < \mu_R + \frac{c}{2}$. From this property, it immediately follows that

$$\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}} > 0 \quad \text{for all} \quad i_{k1} \in I_1^\circ. \tag{53}$$

Step 2b: $\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}}$ for $i_{k1} \in I_2^\circ$

If $i_{k1} \in I_2^\circ$, then, by (6),

$$i^*_{R2}(i_{k1}) = \mu_R - \frac{c}{2} > i_{k1} \quad \text{and} \quad i^*_{L2}(i_{k1}) = i_{k1}.$$

Hence, by (51),

$$\begin{aligned}\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}} =& p^*(\delta) \cdot [-2(i_{k1} - \mu_R) + c] \\ & + (1 - p^*(\delta)) \cdot [-2(i_{k1} - \mu_R) - 2(i_{k1} - \mu_R)] \\ =& - 2(i_{k1} - \mu_R)(2 - p^*(\delta)) + c \cdot p^*(\delta) \\ >& 0,\end{aligned} \tag{54}$$

for all $i_{k1} < \mu_R + \frac{c}{2} \cdot \frac{p^*(\delta)}{2-p^*(\delta)}$. Therefore,

$$\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}} > 0 \quad \text{for all} \quad i_{k1} \in I_2^\circ. \tag{55}$$

Step 2c: $\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}}$ for $i_{k1} \in I_3^\circ$

If $i_{k1} \in I_3^\circ$, then, by (6),

$$i^*_{R2}(i_{k1}) = \mu_R - \frac{c}{2} > i_{k1} \quad \text{and} \quad i^*_{L2}(i_{k1}) = \mu_L + \frac{c}{2} < i_{k1},$$

and by (51),

$$\begin{aligned}\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}} =& p^*(\delta) \cdot [-2(i_{k1} - \mu_R) + c] + (1 - p^*(\delta)) \cdot [-2(i_{k1} - \mu_R) - c] \\ =& - 2(i_{k1} - \mu_R) + c(2p^*(\delta) - 1) \\ >& 0,\end{aligned} \tag{56}$$

for all $i_{k1} < \mu_R + \frac{c(2p^*(\delta)-1)}{2}$. Since $\mu_R + \frac{c(2p^*(\delta)-1)}{2} \geq \mu_R - \frac{c}{2}$, it follows that

$$\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}} > 0 \quad \text{for all} \quad i_{k1} \in I_3^\circ. \tag{57}$$

Step 2d: $\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}}$ for $i_{k1} \in I_4^\circ$

If $i_{k1} \in I_4^\circ$, then, by (6),

$$i^*_{R2}(i_{k1}) = i_{k1} \quad \text{and} \quad i^*_{L2}(i_{k1}) = \mu_L + \frac{c}{2} < i_{k1},$$

and thus, by (51),

$$\begin{aligned}\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}} =& p^*(\delta)\cdot[-2(i_{k1}-\mu_R)-2(i_{k1}-\mu_R)] \\ &+(1-p^*(\delta))\cdot[-2(i_{k1}-\mu_R)-c] \\ =& -2(i_{k1}-\mu_R)(1+p^*(\delta))-c\cdot(1-p^*(\delta)),\end{aligned} \tag{58}$$

which is strictly positive if and only if $i_{k1} < \mu_R - \frac{c}{2}\cdot\frac{1-p^*(\delta)}{1+p^*(\delta)}$. Since, for $p^*(\delta)\in(0,1)$,

$$\mu_R - \frac{c}{2}\cdot\frac{1-p^*(\delta)}{1+p^*(\delta)} \in \left(\mu_R-\frac{c}{2},\mu_R\right)\subset I_4^\circ,$$

it follows that, for $i_{k1}\in I_4^\circ$,

$$\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}}\begin{cases} >0 & \text{if } i_{k1} < \mu_R-\frac{c}{2}\cdot\frac{1-p^*(\delta)}{1+p^*(\delta)}, \\ =0 & \text{if } i_{k1} = \mu_R-\frac{c}{2}\cdot\frac{1-p^*(\delta)}{1+p^*(\delta)}, \\ <0 & \text{if } i_{k1} > \mu_R-\frac{c}{2}\cdot\frac{1-p^*(\delta)}{1+p^*(\delta)}. \end{cases} \tag{59}$$

Step 2e: $\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}}$ for $i_{k1}\in I_5^\circ$

If $i_{k1}\in I_5^\circ$, then

$$i^*_{R2}(i_{k1}) = \mu_R+\frac{c}{2} < i_{k1} \quad \text{and} \quad i^*_{L2}(i_{k1}) = \mu_L+\frac{c}{2} < i_{k1},$$

by (6). Thus, by (51),

$$\begin{aligned}\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}} =& p^*(\delta)\cdot[-2(i_{k1}-\mu_R)-c]+(1-p^*(\delta))\cdot[-2(i_{k1}-\mu_R)-c] \\ =& -2(i_{k1}-\mu_R)-c \\ <& 0,\end{aligned} \tag{60}$$

for all $i_{k1} > \mu_R - \frac{c}{2}$. Therefore,

$$\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}} < 0 \quad \text{for all} \quad i_{k1}\in I_5^\circ. \tag{61}$$

Step 2f: Solution to $\operatorname{argmax}_{i_{k1}\in[0,1]} EU_\delta(i_{k1}|p^*(\delta))$

Combining (53), (55), (57), (59), and (61) yields

$$\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}} > 0 \quad \text{for all} \quad i_{k1} < \mu_R - \frac{c}{2}\cdot\frac{(1-p^*(\delta))}{(1+p^*(\delta))}$$

and

$$\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}} < 0 \quad \text{for all} \quad i_{k1} > \mu_R - \frac{c}{2}\cdot\frac{(1-p^*(\delta))}{(1+p^*(\delta))},$$

whenever $\frac{dEU_\delta(i_{k1}|p^*(\delta))}{di_{k1}}$ exists. Due to the continuity of $EU_\delta(i_{k1}|p^*(\delta))$ in $i_{k1} \in [0,1]$, this yields

$$\underset{i_{k1}\in[0,1]}{\operatorname{argmax}} EU_\delta(i_{k1}|p^*(\delta)) = \left\{\mu_R - \frac{c}{2} \cdot \frac{(1-p^*(\delta))}{(1+p^*(\delta))}\right\} \tag{62}$$

and

$$EU_\delta(i|p^*(\delta)) > EU_\delta(j|p^*(\delta)), \quad \text{for all } i \in \left[0, \mu_R - \frac{c}{2} \cdot \frac{(1-p^*(\delta))}{(1+p^*(\delta))}\right] \text{ and } j < i. \tag{63}$$

Step 3: Solution of $\operatorname{argmax}_{i_{k1}\in I^*(\delta)} EU_\delta(i_{k1}|p^*(\delta))$

From Step 1f we know that the value of $p^*(\delta)$ depends on whether $\delta \leq \frac{c}{2}$ or $\delta > \frac{c}{2}$. Let us therefore distinguish two cases.

<u>*Case 1:*</u> $\delta \in \left[0, \frac{c}{2}\right]$

In this case, (44), (45), and (46) yield

$$p^*(\delta) = \frac{1+\rho}{2}.$$

Hence, by (62),

$$\underset{i_{k1}\in[0,1]}{\operatorname{argmax}} EU_\delta(i_{k1}|p^*(\delta)) = \left\{\mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho}\right\}. \tag{64}$$

From (44) and (45), we obtain that, for $\rho \in (0,1)$,

$$\mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho} \in \left(\mu_R - \frac{c}{6}, \mu_R\right) \subset \left[\frac{1}{2}, 1\right] \subset I^*(0).$$

Moreover, since

$$\frac{2\delta(2\mu_R - 1 - c) + c}{2c} \leq \mu_R - \frac{c}{2} < \mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho},$$

for all $\delta \leq \frac{c}{2}$, and

$$\mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho} \leq \mu_R + \frac{c}{2} - 2\delta \Leftrightarrow \delta \leq \frac{c}{3+\rho},$$

it follows from (46) that

$$\mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho} \in I^*(\delta) \quad \text{for } \delta \in \left[0, \frac{c}{3+\rho}\right]. \tag{65}$$

Combining (64) and (65) yields

$$\underset{i_{k1}\in I^*(\delta)}{\operatorname{argmax}} EU_\delta(i_{k1}|p^*(\delta)) = \left\{\mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho}\right\} \quad \text{for } \delta \in \left[0, \frac{c}{3+\rho}\right]. \tag{66}$$

For $\delta \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right]$,

$$\underset{i_{k1}\in[0,1]}{\operatorname{argmax}} EU_\delta(i_{k1}|p^*(\delta)) \cap I^*(\delta) = \emptyset,$$

or, more precisely, the unique element in $\operatorname{argmax}_{i_{k1}\in[0,1]} EU_\delta(i_{k1}|p^*(\delta))$ is larger than every element in $I^*(\delta)$. Hence, by (63), the office-holder chooses his/her policy by moving to the right as long as s/he stays in $I^*(\delta)$, the area that guarantees high re-election probability. That is,

$$\underset{i_{k1}\in I^*(\delta)}{\operatorname{argmax}} EU_\delta(i_{k1}|p^*(\delta)) = \left\{\mu_R + \frac{c}{2} - 2\delta\right\} \quad \text{for } \delta \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right]. \tag{67}$$

<u>*Case 2:*</u> $\delta \in \left(\frac{c}{2}, \frac{1}{2}\right]$

In this case,

$$p^*(\delta) = \frac{1-\rho}{2},$$

by (47), and (62) implies that

$$\underset{i_{k1}\in[0,1]}{\operatorname{argmax}} EU_\delta(i_{k1}|p^*(\delta)) = \left\{\mu_R - \frac{c}{2}\cdot\frac{1+\rho}{3-\rho}\right\}.$$

Since $I^*(\delta) = [0,1]$, by (47), this immediately implies that

$$\underset{i_{k1}\in I^*(\delta)}{\operatorname{argmax}} EU_\delta(i_{k1}|p^*(\delta)) = \left\{\mu_R - \frac{c}{2}\cdot\frac{1+\rho}{3-\rho}\right\} \quad \text{for } \delta \in \left(\frac{c}{2}, \frac{1}{2}\right]. \tag{68}$$

Combining (66), (67), and (68) yields (9).

□

Proof of Theorem 1

The case where $c = 0$ is straightforward, because $EAP(\delta)$ and $W(\delta)$ are constant in δ. Here we show that Theorem 1 holds for any $c > 0$. Throughout this proof, we omit the dependence of $EAP(\cdot)$ and $W(\cdot)$ on c for ease of notation. We show that

(a) $EAP(\delta) \leq EAP(0)$, for all $\delta \in \left(0, \frac{1}{2}\right]$,

(b) $W(\delta) \geq W(0)$, for all $\delta \in \left(0, \frac{1}{2}\right]$,

(c) $\operatorname{argmin}_{\delta\in\left[0,\frac{1}{2}\right]} EAP(\delta) = \left\{\frac{c}{2}\right\}$,

(d) $\operatorname{argmax}_{\delta\in\left[0,\frac{1}{2}\right]} W(\delta) = \left\{\frac{c}{2}\right\}$,

from which Theorem 1 follows, with $\delta^* = \frac{c}{2}$. In Part 1, we analyze ex-ante policy polarization and prove *(a)* and *(c)*. Part 2 is devoted to welfare and the proofs of *(b)* and *(d)*.

Part 1: Analysis of $EAP(\delta)$

From (11) we know that

$$i^*_{R1}(\delta) = i^*_{R2}(\delta) > \frac{1}{2}, \quad \text{and}$$
$$i^*_{L2}(\delta) = \mu_L + \frac{c}{2} = 1 - \mu_R + \frac{c}{2} < \frac{1}{2}.$$

Inserting these equalities into (14) yields

$$EAP(\delta) = \frac{1 + p^*(\delta)}{2}\left(i^*_{R1}(\delta) - \frac{1}{2}\right) + \frac{1 - p^*(\delta)}{2}\left(\mu_R - \frac{c}{2} - \frac{1}{2}\right). \tag{69}$$

Thus, by (9) and (10), $EAP(\delta)$ is constant for $\delta \leq \frac{c}{3+\rho}$ and $\delta > \frac{c}{2}$ and decreasing in δ for $\delta \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right]$. Moreover, $EAP(\delta)$ has a single discontinuity at $\delta = \frac{c}{2}$. To show that *(a)* and *(c)* hold, it therefore suffices to compare $EAP\left(\frac{1}{2}\right)$ to both $EAP(0)$ and $EAP\left(\frac{c}{2}\right)$. First, inserting (9) and (10) into (69) implies that

$$\begin{aligned}
EAP\left(\frac{1}{2}\right) - EAP(0) =& \frac{3-\rho}{4} \cdot \left(\mu_R - \frac{c}{2} \cdot \frac{1+\rho}{3-\rho} - \frac{1}{2}\right) + \frac{1+\rho}{4} \cdot \left(\mu_R - \frac{c}{2} - \frac{1}{2}\right) \\
&- \frac{3+\rho}{4} \cdot \left(\mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho} - \frac{1}{2}\right) - \frac{1-\rho}{4} \cdot \left(\mu_R - \frac{c}{2} - \frac{1}{2}\right) \\
=& -\frac{\rho}{2}\left(\mu_R - \frac{1}{2}\right) - \frac{\rho}{2} \cdot \frac{c}{2} + \frac{\rho}{2}\left(\mu_R - \frac{c}{2} - \frac{1}{2}\right) \\
=& -\frac{\rho c}{2} < 0,
\end{aligned} \tag{70}$$

which establishes statement *(a)*. Second, again by inserting (9) and (10) into (69), we obtain

$$\begin{aligned}
EAP\left(\frac{1}{2}\right) - EAP\left(\frac{c}{2}\right) =& \frac{3-\rho}{4} \cdot \left(\mu_R - \frac{c}{2} \cdot \frac{1+\rho}{3-\rho} - \frac{1}{2}\right) + \frac{1+\rho}{4} \cdot \left(\mu_R - \frac{c}{2} - \frac{1}{2}\right) \\
&- \frac{3+\rho}{4} \cdot \left(\mu_R - \frac{c}{2} - \frac{1}{2}\right) - \frac{1-\rho}{4} \cdot \left(\mu_R - \frac{c}{2} - \frac{1}{2}\right) \\
=& -\frac{\rho}{2}\left(\mu_R - \frac{1}{2}\right) + \frac{c}{4} + \frac{\rho}{2}\left(\mu_R - \frac{c}{2} - \frac{1}{2}\right) \\
=& \frac{(1-\rho)c}{4} > 0.
\end{aligned} \tag{71}$$

This proves statement *(c)*. In this proof, we did not need an explicit expression for $EAP(\delta)$. However, for the sake of completeness, we can obtain an explicit expression for ex-ante policy polarization by inserting (9) and (10) into (69) and using simple algebraic manipulations:

$$EAP(\delta) = \begin{cases} \mu_R - \frac{1}{2} - \frac{c(1-\rho)}{4} & \text{if } \delta \in \left[0, \frac{c}{3+\rho}\right], \\ \mu_R - \frac{1}{2} + \frac{c(1+\rho)}{4} - \frac{(3+\rho)\delta}{2} & \text{if } \delta \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right], \\ \mu_R - \frac{1}{2} - \frac{c(1+\rho)}{4} & \text{if } \delta \in \left(\frac{c}{2}, \frac{1}{2}\right]. \end{cases}$$

Part 2: Analysis of $W(\delta)$

Now we investigate how welfare behaves as a function of δ. First, we compute $EU^P(\delta)$. For this purpose, suppose that $k \in R$ is the office-holder in $t = 1$ and that $k' \in L$ is his/her challenger in the second election. Because $E[a_{k'}] = 0$, and from (49) and (50),

$$E\left[a_k | k \text{ re-elected and } p_\delta(i_{k1}) = p^*(\delta)\right] = \frac{A(1-\rho)}{2p^*(\delta)} \quad \text{and}$$
$$E\left[a_k | k \text{ not re-elected and } p_\delta(i_{k1}) = p^*(\delta)\right] = -\frac{A(1-\rho)}{2(1-p^*(\delta))},$$

it follows from (16) that expected lifetime utility from the public projects is constant:

$$EU^P(\delta) = \frac{A(1-\rho)}{2}.$$

Second, inserting (11) into (17) and (18) yields

$$EU^I_{\frac{1}{2}}(\delta) = -(1+p^*(\delta))\left(i^*_{R1}(\delta) - \frac{1}{2}\right)^2 - (1-p^*(\delta))\left(\mu_L + \frac{c}{2} - \frac{1}{2}\right)^2$$

and

$$EU^c(\delta) = -(1-p^*(\delta)) \cdot c\left(i^*_{R1}(\delta) - \left(\mu_L + \frac{c}{2}\right)\right),$$

respectively. Therefore, according to (15),

$$\begin{aligned} W(\delta) =& \frac{A(1-\rho)}{2} - (1+p^*(\delta))\left(i^*_{R1}(\delta) - \frac{1}{2}\right)^2 \\ & - (1-p^*(\delta))\left\{\left(\mu_L + \frac{c}{2} - \frac{1}{2}\right)^2 + c\left[i^*_{R1}(\delta) - \left(\mu_L + \frac{c}{2}\right)\right]\right\}. \end{aligned} \tag{72}$$

From (9) and (10) it follows that $W(\delta)$ is constant for both ranges $\delta \leq \frac{c}{3+\rho}$ and $\delta > \frac{c}{2}$ and that $W(\delta)$ has a single discontinuity at $\delta = \frac{c}{2}$. Moreover, $W(\delta)$ is increasing in δ for $\delta \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right]$, since

$$\frac{dW(\delta)}{d\delta} = \underbrace{\left[-2(1+p^*(\delta)) \cdot \left(i^*_{R1}(\delta) - \frac{1}{2}\right) - c(1-p^*(\delta))\right]}_{<0} \cdot \underbrace{\frac{di^*_{R1}(\delta)}{d\delta}}_{<0} > 0,$$

which follows from (72), and (9) and (10) for $\delta \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right]$. To prove statements *(b)* and *(d)*, it is therefore sufficient to compare $W\left(\frac{1}{2}\right)$ to both $W(0)$ and $W\left(\frac{c}{2}\right)$. First, inserting (9) and (10) into (72), yields

$$\begin{aligned} W\left(\frac{1}{2}\right) - W(0) =& -\frac{3-\rho}{2}\left(\mu_R - \frac{c(1+\rho)}{2(3-\rho)} - \frac{1}{2}\right)^2 \\ & - \frac{1+\rho}{2}\left\{\left(\mu_L + \frac{c}{2} - \frac{1}{2}\right)^2 + c\left[\mu_R - \frac{c(1+\rho)}{2(3-\rho)} - \left(\mu_L + \frac{c}{2}\right)\right]\right\} \\ & + \frac{3+\rho}{2}\left(\mu_R - \frac{c(1-\rho)}{2(3+\rho)} - \frac{1}{2}\right)^2 \\ & + \frac{1-\rho}{2}\left\{\left(\mu_L + \frac{c}{2} - \frac{1}{2}\right)^2 + c\left[\mu_R - \frac{c(1-\rho)}{2(3+\rho)} - \left(\mu_L + \frac{c}{2}\right)\right]\right\}. \end{aligned}$$

Thus,

$$
\begin{aligned}
& W\left(\frac{1}{2}\right)-W(0) \\
=& -\frac{3-\rho}{2}\left(\mu_R-\frac{1}{2}-\frac{c(1+\rho)}{2(3-\rho)}\right)^2 \\
& -\frac{1+\rho}{2}\left\{\left(\mu_R-\frac{1}{2}-\frac{c}{2}\right)^2+c\left[2\left(\mu_R-\frac{1}{2}\right)-\frac{2c}{3-\rho}\right]\right\} \\
& +\frac{3+\rho}{2}\left(\mu_R-\frac{1}{2}-\frac{c(1-\rho)}{2(3+\rho)}\right)^2 \\
& +\frac{1-\rho}{2}\left\{\left(\mu_R-\frac{1}{2}-\frac{c}{2}\right)^2+c\left[2\left(\mu_R-\frac{1}{2}\right)-\frac{2c}{3+\rho}\right]\right\} \\
=& \left(\mu_R-\frac{1}{2}\right)^2\cdot\left(-\frac{3-\rho}{2}-\frac{1+\rho}{2}+\frac{3+\rho}{2}+\frac{1-\rho}{2}\right) \\
& +\left(\frac{c}{2}\right)^2\cdot\left[-\frac{(1+\rho)^2}{2(3-\rho)}-\frac{1+\rho}{2}+\frac{4(1+\rho)}{3-\rho}+\frac{(1-\rho)^2}{2(3+\rho)}+\frac{1-\rho}{2}-\frac{4(1-\rho)}{3+\rho}\right] \\
& +c\left(\mu_R-\frac{1}{2}\right)\cdot\left[\frac{1+\rho}{2}+\frac{1+\rho}{2}-(1+\rho)-\frac{1-\rho}{2}-\frac{1-\rho}{2}+(1-\rho)\right] \\
=& \left(\frac{c}{2}\right)^2\cdot\frac{(1+\rho)(3+\rho)(-1-\rho-3+\rho+8)+(1-\rho)(3-\rho)(1-\rho+3+\rho-8)}{2(3-\rho)(3+\rho)} \\
=& \left(\frac{c}{2}\right)^2\cdot\frac{2(1+\rho)(3+\rho)-2(1-\rho)(3-\rho)}{(3-\rho)(3+\rho)} \\
=& \frac{4\rho c^2}{(3-\rho)(3+\rho)},
\end{aligned}
\tag{73}
$$

which is strictly positive for $\rho \in (0,1)$. This proves statement *(b)*. Second, again by making use of (9) and (10) in (72), we obtain

$$
\begin{aligned}
W\left(\frac{1}{2}\right)-W\left(\frac{c}{2}\right)=& -\frac{3-\rho}{2}\left(\mu_R-\frac{c(1+\rho)}{2(3-\rho)}-\frac{1}{2}\right)^2 \\
& -\frac{1+\rho}{2}\left\{\left(\mu_L+\frac{c}{2}-\frac{1}{2}\right)^2+c\left[\mu_R-\frac{c(1+\rho)}{2(3-\rho)}-\left(\mu_L+\frac{c}{2}\right)\right]\right\} \\
& +\frac{3+\rho}{2}\left(\mu_R-\frac{c}{2}-\frac{1}{2}\right)^2 \\
& +\frac{1-\rho}{2}\left\{\left(\mu_L+\frac{c}{2}-\frac{1}{2}\right)^2+c\left[\mu_R-\frac{c}{2}-\left(\mu_L+\frac{c}{2}\right)\right]\right\}.
\end{aligned}
$$

So

$$
\begin{aligned}
&W\left(\frac{1}{2}\right)-W\left(\frac{c}{2}\right)\\
=&-\frac{3-\rho}{2}\left(\mu_R-\frac{1}{2}-\frac{c(1+\rho)}{2(3-\rho)}\right)^2\\
&-\frac{1+\rho}{2}\left\{\left(\mu_R-\frac{1}{2}-\frac{c}{2}\right)^2+c\left[2\left(\mu_R-\frac{1}{2}\right)-\frac{2c}{(3-\rho)}\right]\right\}\\
&+\frac{3+\rho}{2}\left(\mu_R-\frac{1}{2}-\frac{c}{2}\right)^2\\
&+\frac{1-\rho}{2}\left\{\left(\mu_R-\frac{1}{2}-\frac{c}{2}\right)^2+c\left[2\left(\mu_R-\frac{1}{2}\right)-c\right]\right\}\\
=&\left(\mu_R-\frac{1}{2}\right)^2\cdot\left(-\frac{3-\rho}{2}-\frac{1+\rho}{2}+\frac{3+\rho}{2}+\frac{1-\rho}{2}\right)\\
&+\left(\frac{c}{2}\right)^2\cdot\left[-\frac{(1+\rho)^2}{2(3-\rho)}-\frac{1+\rho}{2}+\frac{4(1+\rho)}{3-\rho}+\frac{3+\rho}{2}+\frac{1-\rho}{2}-2(1-\rho)\right]\\
&+c\left(\mu_R-\frac{1}{2}\right)\cdot\left[\frac{1+\rho}{2}+\frac{1+\rho}{2}-(1+\rho)-\frac{3+\rho}{2}-\frac{1-\rho}{2}+(1-\rho)\right]\\
=&\frac{c^2\left(-\rho^2+4\rho+1\right)}{2(3-\rho)}-c\left(\mu_R-\frac{1}{2}\right)(1+\rho)\\
=&\frac{c^2\left(-\rho^2+4\rho+1\right)-c(2\mu_R-1)\left(-\rho^2+2\rho+3\right)}{2(3-\rho)}<-\frac{c^2\left(1-\rho\right)}{3-\rho}<0,
\end{aligned}
\tag{74}
$$

where the first inequality in the last line holds since $-\rho^2+2\rho+3>0$ for all $\rho\in(0,1)$ and $c<(2\mu_R-1)$. This establishes statement *(d)*. For the sake of completeness, we derive the following explicit expression for welfare by inserting (9) and (10) into (72):

$$
W(\delta)=\begin{cases}
\frac{A(1-\rho)}{2}-\frac{(2\mu_R-1)^2}{2}+\frac{c^2(1-\rho)}{2(3+\rho)} & \text{if } \delta\in\left[0,\frac{c}{3+\rho}\right],\\
\frac{A(1-\rho)-(1-2\mu_R-c)^2}{2}-2(3+\rho)\delta^2+\left[4c+(3+\rho)(2\mu_R-1)\right]\delta & \text{if } \delta\in\left(\frac{c}{3+\rho},\frac{c}{2}\right],\\
\frac{A(1-\rho)}{2}-\frac{(2\mu_R-1)^2}{2}+\frac{c^2(1+\rho)}{2(3-\rho)} & \text{if } \delta\in\left(\frac{c}{2},\frac{1}{2}\right].
\end{cases}
$$

□

Proof of Theorem 2

For $i\in[0,1]$ let $EU_i^k(\delta_k)$ (resp. $EU_i^{k'}(\delta_{k'})$) denote the expected utility of voter i, immediately after candidate k (resp. k') has been elected for the first term in period $t=1$, and k (resp. k') has offered the extra-hurdle δ_k (resp. $\delta_{k'}$). Voter i strictly prefers k to win the first election if and only if $EU_i^k(\delta_k)-EU_i^{k'}(\delta_{k'})>0$. The proof of the theorem is now divided into two steps. In Step 1, we show that $i=\frac{1}{2}$ is the critical voter in the first election. That is, the candidate whom voter $i=\frac{1}{2}$ supports in the first election is in office in $t=1$.[28] In Step 2, we use this result to show which extra-hurdles

[28]If the critical voter is indifferent between both candidates, the first election is tied.

are chosen in equilibrium and what the outcome of the first election is.

Step 1: Critical voter in the first election

Analogously to (72), we can represent $EU_i^k(\delta_k)$ by

$$\begin{aligned} EU_i^k(\delta_k) =& \frac{A(1-\rho)}{2} - (1+p^*(\delta_k))\left(i^*_{R1}(\delta_k) - i\right)^2 \\ & - (1-p^*(\delta_k))\left\{\left(\mu_L + \frac{c}{2} - i\right)^2 + c\left[i^*_{R1}(\delta_k) - \left(\mu_L + \frac{c}{2}\right)\right]\right\}. \end{aligned} \tag{75}$$

By (2), and because a_k and $a_{k'}$ are drawn from the same distribution, the latter expression also holds for candidate k':

$$\begin{aligned} EU_i^{k'}(\delta_{k'}) =& \frac{A(1-\rho)}{2} - (1+p^*(\delta_{k'}))\left(1 - i^*_{R1}(\delta_{k'}) - i\right)^2 \\ & - (1-p^*(\delta_{k'}))\left\{\left(1 - \left(\mu_L + \frac{c}{2}\right) - i\right)^2 + c\left[i^*_{R1}(\delta_{k'}) - \left(\mu_L + \frac{c}{2}\right)\right]\right\}. \end{aligned} \tag{76}$$

Voter i strictly prefers k to win the first election if and only if $EU_i^k(\delta_k) - EU_i^{k'}(\delta_{k'}) > 0$. Deducting (76) from (75) yields

$$\begin{aligned} & EU_i^k(\delta_k) - EU_i^{k'}(\delta_{k'}) = \\ & -(1+p^*(\delta_k))\left(i^*_{R1}(\delta_k) - i\right)^2 + (1+p^*(\delta_{k'}))\left(1 - i^*_{R1}(\delta_{k'}) - i\right)^2 \\ & - (1-p^*(\delta_k))\left\{\left(\mu_L + \frac{c}{2} - i\right)^2 + c\left[i^*_{R1}(\delta_k) - \left(\mu_L + \frac{c}{2}\right)\right]\right\} \\ & + (1-p^*(\delta_{k'}))\left\{\left(1 - \left(\mu_L + \frac{c}{2}\right) - i\right)^2 + c\left[i^*_{R1}(\delta_{k'}) - \left(\mu_L + \frac{c}{2}\right)\right]\right\}. \end{aligned} \tag{77}$$

From (9) and (10) we know that (77) yields different expressions, depending on whether $\delta_k \in \left[0, \frac{c}{3+\rho}\right]$, $\delta_k \in \left(\frac{c}{3+\rho}, \frac{c}{2}\right]$, or $\delta_k \in \left(\frac{c}{2}, \frac{1}{2}\right]$, and similarly for $\delta_{k'}$. Analyzing the nine resulting expressions of (77) separately shows that $EU_i^k(\delta_k) - EU_i^{k'}(\delta_{k'})$ is strictly increasing in i for all $(\delta_k, \delta_{k'}) \in \left[0, \frac{1}{2}\right] \times \left[0, \frac{1}{2}\right]$. For instance, if $\delta_k \in \left[0, \frac{c}{3+\rho}\right]$ and $\delta_{k'} \in \left[0, \frac{c}{3+\rho}\right]$, then, by (9),

$$i^*_{R1}(\delta_k) = i^*_{R1}(\delta_{k'}) = \mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho}, \tag{78}$$

and, by (10),

$$p^*(\delta_k) = p^*(\delta_{k'}) = \frac{1+\rho}{2}. \tag{79}$$

Inserting (78) and (79) into (77), and after some simple algebraic manipulations, we find that

$$\frac{d\left[EU_i^k(\delta_k) - EU_i^{k'}(\delta_{k'})\right]}{di} = 2(2\mu_R - 1)(1+\rho),$$

which is strictly positive, since $2\mu_R - 1 > 0$ and $\rho > 0$. The proofs of the other eight cases are analogous. Hence, we conclude that the critical voter in the first election is $i = \frac{1}{2}$.

Step 2: Choice of extra-hurdle in equilibrium and outcome of first election

By the definition of welfare given in (15), we know that

$$EU^k_{\frac{1}{2}}(\delta) = EU^{k'}_{\frac{1}{2}}(\delta) = W(\delta).$$

Therefore, by Step 1:

- k wins the first election with probability 1, if $W(\delta_k) - W(\delta_{k'}) > 0$.
- k' wins the first election with probability 1, if $W(\delta_k) - W(\delta_{k'}) < 0$.
- k and k' each wins the first election with probability equal to $\frac{1}{2}$, if $W(\delta_k) - W(\delta_{k'}) = 0$.

Because b is assumed to be large (see Section 3.4), candidate k (resp. k') offers δ_k (resp. $\delta_{k'}$) such that, given the extra-hurdle of the other candidate, his/her probability of winning the first election is maximized. Therefore, in equilibrium, candidate k suggests some $\delta_k \in \text{argmax}_{\delta \in \left[0, \frac{1}{2}\right]} W(\delta)$, and candidate k' suggests some $\delta_{k'} \in \text{argmax}_{\delta \in \left[0, \frac{1}{2}\right]} W(\delta)$, because otherwise one candidate could increase his/her re-election chances by offering a different extra-hurdle. Hence, by Section 5.2, both candidates commit to $\delta^* = \frac{c}{2}$ and win the election with a probability equal to $\frac{1}{2}$.

□

Proof of Theorem 3

Let

$$F_\alpha(x) := \int_0^x f_\alpha(t)dt$$

denote the cumulative distribution function of $f_\alpha(\cdot)$, and let $k \in R$ be in office in $t = 1$. Then, by the same reasoning as in the proof of Proposition 2, it can be shown that the critical voter in the second election is $i^* = F_\alpha^{-1}\left(\frac{1}{2} - \delta\right)$. By simple algebraic manipulations, it follows that, for any $\delta \in \left[0, \frac{1}{2}\right]$,

$$i^* = \frac{1}{2} - h_\alpha(\delta),$$

where we define

$$h_\alpha(\delta) := \begin{cases} \delta & \text{if } \alpha = 0, \\ \frac{\sqrt{(1-\alpha)^2 + 8\alpha\delta} + \alpha - 1}{4\alpha} & \text{if } \alpha \in (0, 1]. \end{cases}$$

Again, by the same arguments as in the proof of Proposition 2, it can be verified that $k \in R$ is re-elected in the second election if and only if $a_k \geq a_{\delta,\alpha}(i_{k1})$, where

$$\begin{aligned} a_{\delta,\alpha}(i_{k1}) = &\left(i^*_{R2}(i_{k1}) - \left(\frac{1}{2} - h_\alpha(\delta)\right)\right)^2 + c\left|i^*_{R2}(i_{k1}) - i_{k1}\right| \\ &- \left(i^*_{L2}(i_{k1}) - \left(\frac{1}{2} - h_\alpha(\delta)\right)\right)^2 - c\left|i^*_{L2}(i_{k1}) - i_{k1}\right|. \end{aligned} \tag{80}$$

Moreover, for any $\alpha \in [0,1]$, $h_\alpha(\cdot)$ is a bijection from $\left[0, \frac{1}{2}\right]$ to $\left[0, \frac{1}{2}\right]$ and hence invertible. Indeed, $h_\alpha(0) = 0$, $h_\alpha\left(\frac{1}{2}\right) = \frac{1}{2}$, and $h_\alpha(\delta)$ is strictly increasing in δ. Using (80) and the invertibility of $h_\alpha(\cdot)$ in the proof of Proposition 3, yields

$$i^*_{R1,\alpha}(\delta) := \begin{cases} \mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho} & \text{if } \delta \in \left[0, h_\alpha^{-1}\left(\frac{c}{3+\rho}\right)\right], \\ \mu_R + \frac{c}{2} - 2h_\alpha(\delta) & \text{if } \delta \in \left(h_\alpha^{-1}\left(\frac{c}{3+\rho}\right), h_\alpha^{-1}\left(\frac{c}{2}\right)\right], \\ \mu_R - \frac{c}{2} \cdot \frac{1+\rho}{3-\rho} & \text{if } \delta \in \left(h_\alpha^{-1}\left(\frac{c}{2}\right), \frac{1}{2}\right] \end{cases} \tag{81}$$

and

$$p^*_\alpha(\delta) := \begin{cases} \frac{1+\rho}{2} & \text{if } \delta \in \left[0, h_\alpha^{-1}\left(\frac{c}{2}\right)\right], \\ \frac{1-\rho}{2} & \text{if } \delta \in \left(h_\alpha^{-1}\left(\frac{c}{2}\right), \frac{1}{2}\right], \end{cases} \tag{82}$$

where $i^*_{R1,\alpha}(\delta)$ and $p^*_\alpha(\delta)$ denote the equilibrium values of k's first-period policy choice and his/her re-election probability in the second election, respectively. Analogously to (69) and (72), we obtain

$$EAP_\alpha(\delta) = \frac{1 + p^*_\alpha(\delta)}{2}\left(i^*_{R1,\alpha}(\delta) - \frac{1}{2}\right) + \frac{1 - p^*_\alpha(\delta)}{2}\left(\mu_R - \frac{c}{2} - \frac{1}{2}\right) \tag{83}$$

and

$$\begin{aligned} W_\alpha(\delta) = & \frac{A(1-\rho)}{2} - (1 + p^*_\alpha(\delta))\left(i^*_{R1,\alpha}(\delta) - \frac{1}{2}\right)^2 \\ & - (1 - p^*_\alpha(\delta))\left\{\left(\mu_L + \frac{c}{2} - \frac{1}{2}\right)^2 + c\left[i^*_{R1,\alpha}(\delta) - \left(\mu_L + \frac{c}{2}\right)\right]\right\}. \end{aligned} \tag{84}$$

We observe that $EAP_0(\delta) = EAP(\delta)$ and $W_0(\delta) = W(\delta)$, where $EAP(\delta)$ and $W(\delta)$ correspond to the baseline model and are given in (69) and (72). From (81) and (82), it is obvious that $EAP_\alpha(\delta)$ and $W_\alpha(\delta)$ are constant for

$$\delta \in \left[0, h_\alpha^{-1}\left(\frac{c}{3+\rho}\right)\right] \quad \text{and} \quad \delta \in \left(h_\alpha^{-1}\left(\frac{c}{2}\right), \frac{1}{2}\right]$$

and have a single discontinuity at $\delta = h_\alpha^{-1}\left(\frac{c}{2}\right)$. Moreover, since $h_\alpha(\delta)$ is strictly increasing in δ, $EAP_\alpha(\delta)$ is strictly decreasing in δ and $W_\alpha(\delta)$ is strictly increasing in δ, for

$$\delta \in \left(h_\alpha^{-1}\left(\frac{c}{3+\rho}\right), h_\alpha^{-1}\left(\frac{c}{2}\right)\right].$$

We can now prove statements *(i)-(iv)*. First, statement *(i)* follows from

$$EAP_\alpha\left(\frac{1}{2}\right) - EAP_\alpha(0) = EAP_0\left(\frac{1}{2}\right) - EAP_0(0) < 0 \tag{85}$$

and

$$W_\alpha\left(\frac{1}{2}\right) - W_\alpha(0) = W_0\left(\frac{1}{2}\right) - W_0(0) > 0. \tag{86}$$

The equalities in (85) and (86) hold because, by (81) and (82), $i^*_{R1,\alpha}(0)$, $i^*_{R1,\alpha}\left(\frac{1}{2}\right)$, $p^*_\alpha(0)$ and $p^*_\alpha\left(\frac{1}{2}\right)$ are independent of the value of α. The inequalities hold by (70) and (73). Second, statement *(ii)*, with $\delta^*_\alpha = h_\alpha^{-1}\left(\frac{c}{2}\right)$, follows from

$$EAP_\alpha\left(\frac{1}{2}\right) - EAP_\alpha\left(h_\alpha^{-1}\left(\frac{c}{2}\right)\right) = EAP_0\left(\frac{1}{2}\right) - EAP_0\left(\frac{c}{2}\right) > 0$$

and

$$W_\alpha\left(\frac{1}{2}\right) - W_\alpha\left(h_\alpha^{-1}\left(\frac{c}{2}\right)\right) = W_0\left(\frac{1}{2}\right) - W_0\left(\frac{c}{2}\right) < 0,$$

where the inequalities hold by (71) and (74). Third, since $\delta^*_\alpha = h_\alpha^{-1}\left(\frac{c}{2}\right)$, where

$$h_\alpha^{-1}(x) = \begin{cases} x & \text{if } \alpha = 0, \\ \frac{(4\alpha x - \alpha + 1)^2 - (1-\alpha)^2}{8\alpha} & \text{if } \alpha \in (0,1], \end{cases} \tag{87}$$

δ^*_α is continuous and decreasing in $\alpha \in [0,1]$. Indeed,

$$\lim_{\alpha \to 0} \frac{(4\alpha x - \alpha + 1)^2 - (1-\alpha)^2}{8\alpha} = x,$$

by l'Hôpital's rule, and

$$\frac{dh_\alpha^{-1}(x)}{d\alpha} = x(2x-1) < 0, \tag{88}$$

for all $\alpha \in (0,1)$ and for all $x \in \left(0, \frac{1}{2}\right)$. This completes the proof of *(iii)*. Fourth, recall that $EAP_\alpha(\delta)$ and $W_\alpha(\delta)$ are constant in δ for $\delta \in \left[0, h_\alpha^{-1}\left(\frac{c}{3+\rho}\right)\right]$. Moreover, for $\delta \in \left(h_\alpha^{-1}\left(\frac{c}{3+\rho}\right), h_\alpha^{-1}\left(\frac{c}{2}\right)\right]$, $EAP_\alpha(\delta)$ is strictly decreasing and $W_\alpha(\delta)$ strictly increasing in δ. Since, by (88), $h_\alpha^{-1}\left(\frac{c}{3+\rho}\right)$ and $h_\alpha^{-1}\left(\frac{c}{2}\right)$ are decreasing in α, statement *(iv)* therefore holds if

$$\frac{dEAP_\alpha(\delta)}{d\alpha} < 0 \tag{89}$$

and

$$\frac{dW_\alpha(\delta)}{d\alpha} > 0 \tag{90}$$

hold for all $\alpha \in (0,1)$ and all $\delta \in \left(h_\alpha^{-1}\left(\frac{c}{3+\rho}\right), h_\alpha^{-1}\left(\frac{c}{2}\right)\right]$. Since for $\delta \in \left(h_\alpha^{-1}\left(\frac{c}{3+\rho}\right), h_\alpha^{-1}\left(\frac{c}{2}\right)\right]$, (81), (82), (83), and (84) yield

$$\frac{dEAP_\alpha(\delta)}{d\alpha} = \frac{1+p^*_\alpha(\delta)}{2} \cdot \frac{di^*_{R1,\alpha}(\delta)}{d\alpha} = -(1+p^*_\alpha(\delta)) \cdot \frac{dh_\alpha(\delta)}{d\alpha}$$

and

$$\begin{aligned} \frac{dW_\alpha(\delta)}{d\alpha} &= -2 \cdot (1+p^*_\alpha(\delta)) \cdot \left(i^*_{R1,\alpha}(\delta) - \frac{1}{2}\right) \cdot \frac{di^*_{R1,\alpha}(\delta)}{d\alpha} - (1-p^*_\alpha(\delta)) \cdot c \cdot \frac{di^*_{R1,\alpha}(\delta)}{d\alpha} \\ &= 2 \cdot \underbrace{\left[2 \cdot (1+p^*_\alpha(\delta)) \cdot \left(i^*_{R1,\alpha}(\delta) - \frac{1}{2}\right) + (1-p^*_\alpha(\delta)) \cdot c\right]}_{>0} \cdot \frac{dh_\alpha(\delta)}{d\alpha}, \end{aligned}$$

and since $\frac{dh_\alpha(\delta)}{d\alpha} > 0$ for all $\alpha \in (0,1)$ and all $\delta \in \left(0, \frac{1}{2}\right)$, (89) and (90) hold for all $\alpha \in (0,1)$ and all $\delta \in \left(h_\alpha^{-1}\left(\frac{c}{3+\rho}\right), h_\alpha^{-1}\left(\frac{c}{2}\right)\right]$. It remains to show that $\frac{dh_\alpha(\delta)}{d\alpha} > 0$ does indeed hold. For $\alpha \in (0,1)$,

$$\begin{aligned}
\frac{dh_\alpha(\delta)}{d\alpha} &= \frac{4\alpha \cdot \left[\frac{\alpha-1+4\delta}{\sqrt{(1-\alpha)^2+8\alpha\delta}} + 1\right] - 4 \cdot \left[\sqrt{(1-\alpha)^2+8\alpha\delta} + \alpha - 1\right]}{16\alpha^2} \\
&= \frac{\alpha \cdot [\alpha - 1 + 4\delta] - \left[(1-\alpha)^2 + 8\alpha\delta - \sqrt{(1-\alpha)^2+8\alpha\delta}\right]}{4\alpha^2\sqrt{(1-\alpha)^2+8\alpha\delta}} \\
&= \frac{\alpha(1-4\delta) - 1 + \sqrt{(1-\alpha)^2+8\alpha\delta}}{4\alpha^2\sqrt{(1-\alpha)^2+8\alpha\delta}},
\end{aligned}$$

which is strictly positive if and only if

$$\begin{aligned}
&(1-\alpha)^2 + 8\alpha\delta - [(1-\alpha) + 4\alpha\delta]^2 > 0 \\
\Leftrightarrow &8\alpha\delta - 16\alpha^2\delta^2 - 2(1-\alpha)4\alpha\delta > 0 \\
\Leftrightarrow &8\alpha^2\delta(1-2\delta) > 0.
\end{aligned}$$

The last inequality is satisfied for all $\alpha \in (0,1)$ and all $\delta \in \left(0, \frac{1}{2}\right)$.

□

Appendix B: Analysis of Large Marginal Cost Levels (for online publication only)

In this appendix, we analyze the case where

$$c \geq 2\Pi.$$

This is in contrast to the condition (5) assumed in the main body of the paper. For Proposition 2 to hold for $c \geq 2\Pi$, we need the assumption that, in the second election, every voter who is indifferent between the incumbent and the challenger votes for the incumbent.[29] This assumption is necessary for the following reason: If $i_{k1} \in [\mu_R - \frac{c}{2}, \mu_L + \frac{c}{2}]$, then (†) from the proof of Proposition 2 equals zero for all $i \in [0,1]$. This means that, for $a_k = 0$ and $i_{k1} \in [\mu_R - \frac{c}{2}, \mu_L + \frac{c}{2}]$, all voters are indifferent between k and his/her competitor. Therefore, if, for instance, indifferent voters randomize between both candidates with probability $\frac{1}{2}$, Proposition 2 only holds for $\delta = 0$, since, by Section 3.3, the incumbent wins the election if s/he receives a vote-share of $\frac{1}{2}+\delta$ or larger. With the above assumption regarding indifferent voters in place, we obtain

Proposition 4

Let $\delta \in \left[0, \frac{1}{2}\right]$. Then the following holds:

(i) If $2\Pi \leq c < (3+\rho)\Pi$, then in $t=1$, the incumbent $k \in R$ chooses

$$i^*_{R1}(\delta) = \begin{cases} \mu_R - \frac{c}{2} \cdot \frac{1-\rho}{3+\rho} & \text{if } \delta \in \left[0, \frac{c}{3+\rho}\right], \\ \mu_R + \frac{c}{2} - 2\delta & \text{if } \delta \in \left(\frac{c}{3+\rho}, \Pi\right], \\ \mu_L + \frac{c}{2} & \text{if } \delta \in \left(\Pi, \frac{1}{2}\right] \end{cases}$$

in equilibrium of the game $\mathcal{G}^R$.

(ii) If $(3+\rho)\Pi \leq c < 4\Pi$, then in $t=1$, the incumbent $k \in R$ chooses

$$i^*_{R1}(\delta) = \mu_L + \frac{c}{2}$$

in equilibrium of the game $\mathcal{G}^R$.

(iii) If $c \geq 4\Pi$, then in $t=1$, the incumbent $k \in R$ chooses

$$i^*_{R1}(\delta) = \mu_R$$

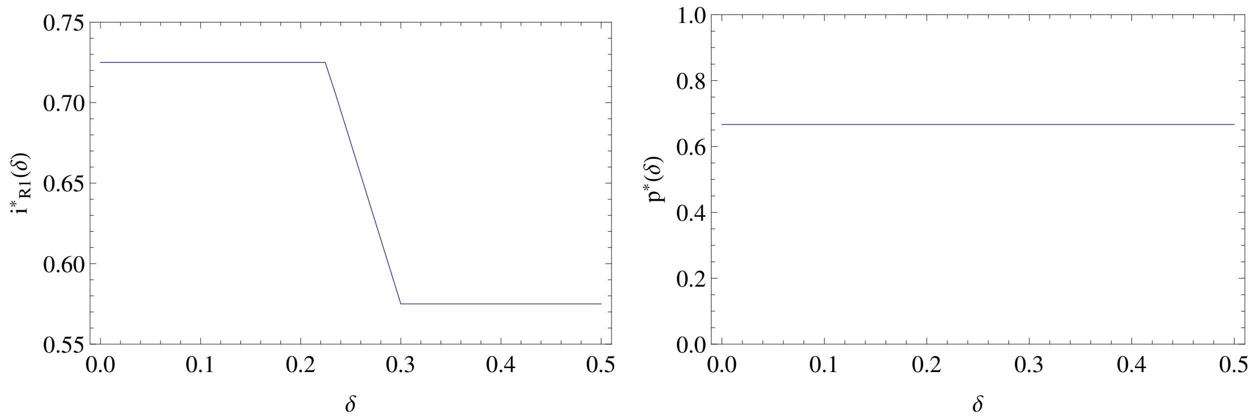
in equilibrium of the game $\mathcal{G}^R$.

[29] This assumption has been made in Section 3.3.

The proof of Proposition 4 is very similar to the proof of Proposition 3.[30] From Proposition 4 we see that, for $c \geq 2\Pi$, $i^*_{R1}(\delta)$ is weakly decreasing in δ. However, there is no unique value of δ that minimizes $i^*_{R1}(\delta)$. For instance, if $c \geq (3+\rho)\Pi$, then $i^*_{R1}(\delta)$ is constant in δ. Due to the assumption that every voter who is indifferent between incumbent and challenger votes for the incumbent, k's re-election probability is equal to $\frac{1+\rho}{2}$ if s/he chooses $i_{k1} \in [\mu_R - \frac{c}{2}, \mu_L + \frac{c}{2}]$. Thus, if $c \geq 2\Pi$ (in contrast to the case where $c < 2\Pi$), the re-election probability in equilibrium is

$$p^*(\delta) = \frac{1+\rho}{2} \tag{91}$$

for all $\delta \in \left[0, \frac{1}{2}\right]$.[31] This is why $i^*_{R1}(\delta)$ is continuous in δ if $c \geq 2\Pi$. The behavior of $i^*_{R1}(\delta)$ and $p^*(\delta)$ for parameter values satisfying $2\Pi \leq c < (3+\rho)\Pi$ is illustrated in Figures 8(a) and 8(b).



(a) Equilibrium policy choice of right-wing policy-maker in $t = 1$.

(b) Re-election probability in equilibrium.

Figure 8: Illustration of the equilibrium analysis in $t = 1$, using the parameter values $\mu_R = 0.8$, $c = 0.75$, $A = 2$, and $\rho = \frac{1}{3}$, satisfying $2\Pi \leq c < (3+\rho)\Pi$.

Finally, from Propositions 4 and 1 it follows that, if $2\Pi \leq c < (3+\rho)\Pi$, then

$$i^*_{R2}(\delta) = i^*_{R1}(\delta), \text{ and} \tag{92}$$

$$i^*_{L2}(\delta) = \mu_L + \frac{c}{2}. \tag{93}$$

Accordingly, if $c \geq (3+\rho)\Pi$, then

$$i^*_{R2}(\delta) = i^*_{L2}(\delta) = i^*_{R1}(\delta). \tag{94}$$

This completes the description of the equilibrium outcomes of the game $\mathcal{G}^R$ for $c \geq 2\Pi$. Analogously to Section 4.3, the equilibrium policy choices of the game $\mathcal{G}^L$ follow immediately, because ideal policy positions of candidates are distributed symmetrically around $\frac{1}{2}$ and candidates' ability distribution is independent of their policy orientation.

[30] The detailed calculations are available upon request.

[31] Recall that, by (10), $p^*(\delta)$ is not constant in δ for $c < 2\Pi$.

In the following, we analyze the dependence of ex-ante policy polarization and welfare on $\delta \in \left[0, \frac{1}{2}\right]$. $EAP(\delta)$ and $W(\delta)$ are defined by (14) and (15) respectively, with $i^*_{R1}(\delta)$ given by Proposition 4, $i^*_{R2}(\delta)$ and $i^*_{L2}(\delta)$ given by (92), (93), and (94), and $p^*(\delta)$ given by (91). Let us now state the counterpart of Theorem 1.

Theorem 4

In equilibrium of $\mathcal{G}$, the following holds:

(i) *Any extra-hurdle $\delta \in \left(0, \frac{1}{2}\right]$ is both weakly W-increasing and weakly P-reducing.*

(ii) *If $2\Pi \leq c < (3+\rho)\Pi$, then*

$$\underset{\delta \in \left[0, \frac{1}{2}\right]}{\operatorname{argmin}} EAP(\delta) = \underset{\delta \in \left[0, \frac{1}{2}\right]}{\operatorname{argmax}} W(\delta) = \left[\Pi, \frac{1}{2}\right].$$

(iii) *If $c \geq (3+\rho)\Pi$, then*

$$\underset{\delta \in \left[0, \frac{1}{2}\right]}{\operatorname{argmin}} EAP(\delta) = \underset{\delta \in \left[0, \frac{1}{2}\right]}{\operatorname{argmax}} W(\delta) = \left[0, \frac{1}{2}\right].$$

Proof

The proof is analogous to the proof of Theorem 1.[32] It is useful, though, to set out the following intermediate results:

First, if $2\Pi \leq c < (3+\rho)\Pi$, then

$$EAP(\delta) = \begin{cases} \frac{(2\mu_R-1)(1+\rho)}{4} & \text{if } \delta \in \left[0, \frac{c}{3+\rho}\right], \\ \frac{(2\mu_R-1)(1+\rho)}{4} + \frac{c-(3+\rho)\delta}{2} & \text{if } \delta \in \left(\frac{c}{3+\rho}, \mu_R - \frac{1}{2}\right], \\ \frac{1-2\mu_R+c}{2} & \text{if } \delta \in \left(\Pi, \frac{1}{2}\right] \end{cases}$$

and

$$W(\delta) = \begin{cases} \frac{A(1-\rho)}{2} - \frac{(2\mu_R-1)^2}{2} + \frac{c^2(1-\rho)}{2(3+\rho)} & \text{if } \delta \in \left[0, \frac{c}{3+\rho}\right], \\ \frac{A(1-\rho)-(1-2\mu_R-c)^2}{2} - 2(3+\rho)\delta^2 + \left[4c + (2\mu_R-1)(3+\rho)\right]\delta & \text{if } \delta \in \left(\frac{c}{3+\rho}, \Pi\right], \\ \frac{A(1-\rho)}{2} - \frac{(1-2\mu_R+c)^2}{2} & \text{if } \delta \in \left(\Pi, \frac{1}{2}\right]. \end{cases}$$

Second, if $(3+\rho)\Pi \leq c < 4\Pi$, then

$$EAP(\delta) = \frac{1-2\mu_R+c}{2}$$

and

$$W(\delta) = \frac{A(1-\rho) - (1-2\mu_R+c)^2}{2}.$$

[32]The detailed calculations are available upon request.

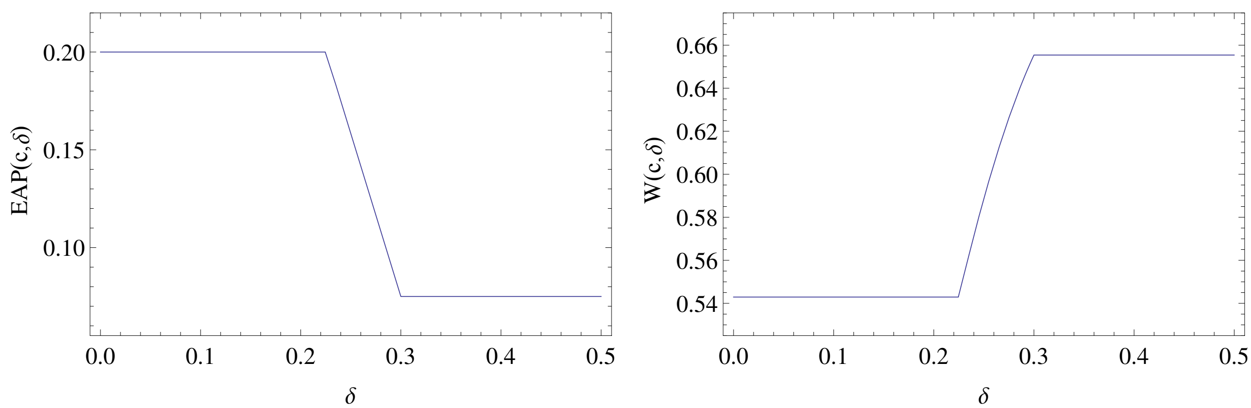
Third, if $c \geq 4\Pi$, then

$$EAP(\delta) = \mu_R - \frac{1}{2}$$

and

$$W(\delta) = \frac{A(1-\rho)}{2} - 2\left(\mu_R - \frac{1}{2}\right)^2.$$

$\square$



(a) Ex-ante policy polarization. (b) Welfare.

Figure 9: Illustration of the equilibrium values of ex-ante policy polarization and welfare, using the parameter values $\mu_R = 0.8$, $c = 0.75$, $A = 2$, and $\rho = \frac{1}{3}$, satisfying $2\Pi \leq c < (3+\rho)\Pi$.

Note that statement *(i)* of Theorem 4 is the same as *(i)* in Theorem 1. That is, compared to the case of $\delta = 0$, the introduction of any non-zero extra-hurdle still weakly improves welfare and weakly reduce ex-ante policy polarization. According to *(ii)* and *(iii)* of Theorem 4, the set of W-optimal extra-hurdles is still identical to the set of P-optimal extra-hurdles. However, in contrast to the main part of the paper, this set is now a continuum, i.e. there is a range of values of δ that are both W- and P-optimal. Theorem 4 is illustrated in Figures 9(a) and 9(b), which show plots of $EAP(\delta)$ and $W(\delta)$, for $\mu_R = 0.8$, $c = 0.75$, $A = 2$, and $\rho = \frac{1}{3}$. These parameter values satisfy $2\Pi \leq c < (3+\rho)\Pi$. The figures show that $EAP(\delta)$ is weakly decreasing and $W(\delta)$ weakly increasing in δ. For $c \geq (3+\rho)\Pi$, ex-ante policy polarization and welfare are constant, because in this case $i^*_{R1}(\delta)$, $i^*_{R2}(\delta)$, $i^*_{L2}(\delta)$ and $p^*(\delta)$ are constant in δ.

Appendix C: Extensions of the Baseline Model (for online publication only)

In this appendix, we investigate several extensions of the baseline model analyzed in the main body of the paper. First, we consider the existence of a status-quo policy in place at the beginning of the first period. Second, we depart from the assumption that costs of change are linear. This includes considering additional fixed costs, on the one hand, and convex costs on the other.[33]

C.1 Initial costs of change

An implicit assumption in the baseline model is that there are no costs of change in $t = 1$. This may occur because there is no status-quo policy or because the significance of costs of change only materializes after the policy in $t = 1$ is chosen. It is thus worth analyzing how our results change if we assume the existence of a status-quo policy in $t = 0$, denoted by $i_0 \in [0, 1]$. Suppose that such a status-quo policy imposes additional costs of change in $t = 1$, given by

$$U^{\tilde{c}}(i_0, i_{k1}) = -\tilde{c} \cdot |i_0 - i_{k1}|,$$

where i_{k1} is the policy chosen by office-holder k in $t = 1$ and $\tilde{c} \in [0, c]$. These costs are added to the lifetime utility of all agents. With methods similar to those used in Sections 4.2 and 5.2, the following theorem can be proved:[34]

Theorem 5

Let the status-quo policy be $i_0 = \frac{1}{2}$. Then,

(i) If $\tilde{c} \in [0, (1 - \rho) \cdot c)$, the statement of Theorem 1 holds.

(ii) If $\tilde{c} \in [(1 - \rho) \cdot c, c]$, the statement of Theorem 1 does not hold always.

The first-period policy choice is still driven by the same two effects as in the case with no status quo. That is, up to some critical value of δ, the office-holder moves closer to the median voter as δ increases to ensure re-election with ability zero. For extra-hurdles above the critical level, s/he cannot ensure re-election with ability zero and thus chooses a more partisan policy. Due to costs of change

[33] The proofs of the results contained in this appendix are available upon request.

[34] The most challenging part of the proof is finding an expression for $i^*_{R1}(\delta)$, but this can be done similarly to the proof of Proposition 3. We point out that we assume $i_0 = \frac{1}{2}$. Finding general results for arbitrary values of i_0 is extremely cumbersome. Nevertheless, a continuity argument shows that the results described in Theorem 5 are robust with respect to small changes of the status-quo policy. A full-scale analysis of the impact of the status-quo policy on policy choices (and thus on polarization and welfare) is left for further research.

in the first period, there is now an additional third effect that induces the first-period office-holder to choose a moderate policy. Indeed, for given $\delta \in \left[0, \frac{1}{2}\right]$, the larger $\tilde{c}$ is, the closer the first-period policy choice is to the median voter's bliss point. The reason is that the office-holder minimizes the cost of change that s/he incurs. Thus, the initial costs of change reinforce the policy-moderating effect of extra-hurdles and, if $\tilde{c}$ is small, leave the maximizer of welfare and the minimizer of ex-ante policy polarization unchanged. This is part *(i)* of Theorem 5.

We are now in a position to discuss the negative statement of Theorem 5. That is, we consider part *(ii)* and then assume that $(1-\rho)\cdot c < \tilde{c} \leq c$. While any extra-hurdle is both weakly W-increasing and weakly P-reducing—so statement *(i)* of Theorem 1 still holds—, uniqueness and existence of an optimal extra-hurdle are no longer guaranteed—so statement *(ii)* of Theorem 1 does not hold. The reason is that, when $\tilde{c}$ is close to c and when δ is above the critical level, the third effect mentioned above is particularly strong. This implies that the W-optimal and the P-optimal extra-hurdles may differ. However, both are still larger than zero, and any W-optimal extra-hurdle is P-reducing. Similarly, any P-optimal δ is W-increasing. If, for instance, $\tilde{c} = c$, then

$$\underset{\delta \in \left[0, \frac{1}{2}\right]}{\operatorname{argmin}}\, EAP(\delta) = \left(\frac{c}{2}, \frac{1}{2}\right]$$

and, if c is smaller than $\frac{2(2\mu_R - 1)}{4+\rho}$,

$$\underset{\delta \in \left[0, \frac{1}{2}\right]}{\operatorname{argmax}}\, W(\delta) = \left\{\frac{c}{2}\right\}.$$

That is, while the sets of W-optimal and P-optimal extra-hurdles are disjoint, all $\delta \in \left(\frac{c}{2}, \frac{1}{2}\right]$ is W-increasing, and $\delta = \frac{c}{2}$ is P-reducing.

C.2 Fixed costs of change

Assume now that in addition to the linear costs of change, any policy shift generates a cost that is independent of the extent of the change. To be more specific, when policies $i_1 \in [0,1]$ and $i_2 \in [0,1]$ are implemented in $t = 1$ and $t = 2$, respectively, all voters and policy-makers incur costs in $t = 2$ equal to

$$U^c(i_1, i_2) = -c \cdot |i_1 - i_2| - K \cdot \mathbb{1}(i_1, i_2),$$

with $K > 0$ and

$$\mathbb{1}(i_1, i_2) = \begin{cases} 1 & \text{if } i_1 \neq i_2, \\ 0 & \text{otherwise.} \end{cases}$$

As in the baseline model, we focus on comparatively small values of c. More precisely, we assume that

$$0 \leq \frac{c}{2} < \Pi - \sqrt{K}. \tag{95}$$

We obtain the following theorem:

Theorem 6

Theorem 1 holds if

$$0 < K < min\left\{\overline{K}_1, \overline{K}_2\right\}, \tag{96}$$

where

$$\overline{K}_1 = \frac{(2\Pi - c)^4}{36} \quad \textit{and} \quad \overline{K}_2 = \frac{2c^2(1-\rho)}{(3-\rho)^2}.$$

The logical steps for proving Theorem 6 are the same as for Theorem 1.[35] The intuition for the result in Theorem 6 is as follows: Fixed costs that are not too large $\left(\text{i.e. below } min\{\overline{K}_1, \overline{K}_2\}\right)$ increase the incumbency advantage of the first-period office-holder. The reason is that the critical extra-hurdle up to which s/he can reach re-election with ability zero is increasing as a function of K. Hence, for any given δ below this critical value, s/he can reach re-election with ability zero by choosing a weakly more polar position compared to the case without fixed costs. This effect merely increases the value of the unique W- and P-optimal extra-hurdle. As a consequence, the results of Theorem 1 hold, but the W- and P-optimal δ^* changes to

$$\delta^* = \frac{c}{2} + \frac{\sqrt{K}(2\Pi - c)}{2(2\Pi - c - \sqrt{K})} > \frac{c}{2},$$

given that (95) and (96) are satisfied. Figure 10 illustrates the constraints on c and K given in (95) and (96).

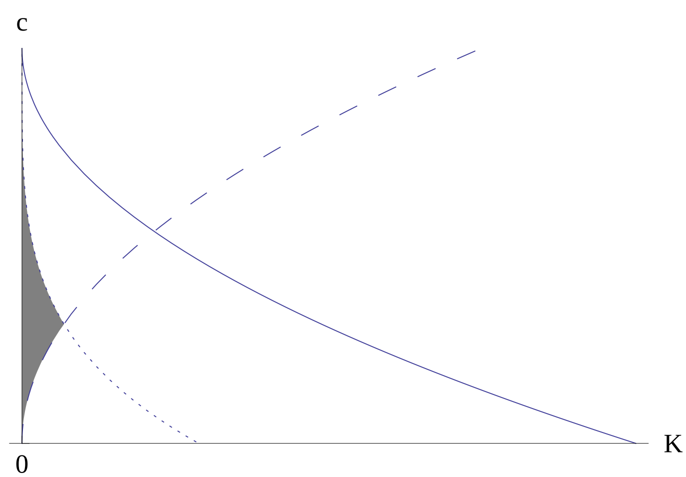


Figure 10: Sketch of the conditions (95) and (96): All pairs (K, c) below the solid curve satisfy (95). If (K, c) lies to the left of the dotted (resp. dashed) curve, then it additionally satisfies $K < \overline{K}_1(c)$ (resp. $K < \overline{K}_2(c)$). The shaded area represents all (K, c) that fulfill (95) and (96).

[35]Details are available upon request.

The area below the solid curve contains all pairs of values (K, c) that satisfy condition (95). The shaded area represents all such pairs that additionally satisfy condition (96). More precisely, if (K, c) lies to the left of the dotted curve, then $K < \overline{K}_1(c)$ is satisfied. Similarly, any pair (K, c) to the left of the dashed curve satisfies $K < \overline{K}_2(c)$.

C.3 Convex costs of change

We next investigate the robustness of our results when the cost function in (3) is replaced by

$$U^c(i_{k1}, i_{k'2}) = -c \cdot |i_{k1} - i_{k'2}|^\eta,$$

with $\eta \in (1, 2]$. On the one hand, $\eta > 1$ implies that costs are strictly convex in the difference between the policies adopted in the two periods. On the other, for $c = 1$, $\eta \leq 2$ implies that the relative utility losses of voters and candidates from policies that differ from their ideal policy positions is higher than the relative increase in the costs of engineering the policy change.

When costs of change are convex, there is no explicit expression for the best response of the second-period office-holder, which is then based on a set of implicit functions. These equations are described in Proposition 5, which shows that, in equilibrium, the office-holder in $t = 2$ responds locally to changes of the policy choice in $t = 1$.[36] This is a feature that cannot be obtained when costs of change are linear. However, as we show below—see Theorem 7—, this more sophisticated property when office-holders change positions is immaterial to the main thrust of our results.

Proposition 5 (Gersbach et al. (2020a))
Let $k \in R$ and $k' \in R \cup L$ be the office-holders in $t = 1$ and $t = 2$, respectively. Let also $c \geq 0$. In $t = 2$, the best response of k' to a policy i_{k1} chosen in the first period is given by

$$i_{k'2}(i_{k1}) = \begin{cases} \underline{i_{\mu_{k'}}} & \text{if } i_{k1} < \mu_{k'}, \text{ where } 2\left(\underline{i_{\mu_{k'}}} - \mu_{k'}\right) + c\eta\left(\underline{i_{\mu_{k'}}} - i_1\right)^{\eta-1} = 0 \text{ and } \underline{i_{\mu_{k'}}} \in (i_1, \mu_{k'}), \\ \mu_{k'} & \text{if } i_{k1} = \mu_{k'}, \\ \overline{i_{\mu_{k'}}} & \text{if } i_{k1} > \mu_{k'}, \text{ where } 2\left(\overline{i_{\mu_{k'}}} - \mu_{k'}\right) - c\eta\left(i_1 - \overline{i_{\mu_{k'}}}\right)^{\eta-1} = 0 \text{ and } \overline{i_{\mu_{k'}}} \in (\mu_{k'}, i_1), \end{cases}$$

where $\mu_{k'}$ is the ideal policy of policy-maker k'.

The following theorem can then be shown with the help of Proposition 5.

Theorem 7
Let $\eta \in (1, 2]$ and $c \geq 0$. Then the following holds:

(i) There exists $\eta^* \in (1, 2)$ such that, if $\eta \in (1, \eta^*)$ and $c < 2\Pi$, Theorem 1 holds.

[36] The proof is of technical nature and relies on the Implicit Function Theorem.

(ii) If $c > 0$, there exists $\delta^{**} \in (0, \Pi)$ such that any extra-hurdle $\delta \in \left(\delta^{**}, \frac{1}{2}\right]$ is P-reducing.

(iii) If $c > 0$ and $\eta = 2$:

(iii.a) Any extra-hurdle $\delta \in \left(0, \frac{1}{2}\right]$ is P-reducing, and $W(\delta)$ is constant for all $\delta \in \left[0, \frac{1}{2}\right]$.

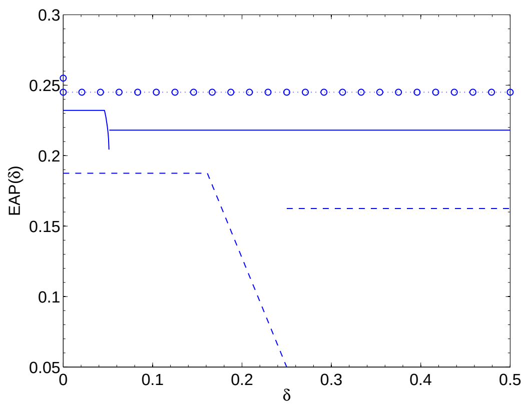
(iii.b) The set of extra-hurdles that are both W- and P-optimal is given by

$$\underset{\delta \in \left[0, \frac{1}{2}\right]}{\operatorname{argmin}} EAP(\delta) \cap \underset{\delta \in \left[0, \frac{1}{2}\right]}{\operatorname{argmax}} W(\delta) = \left(0, \frac{1}{2}\right].$$

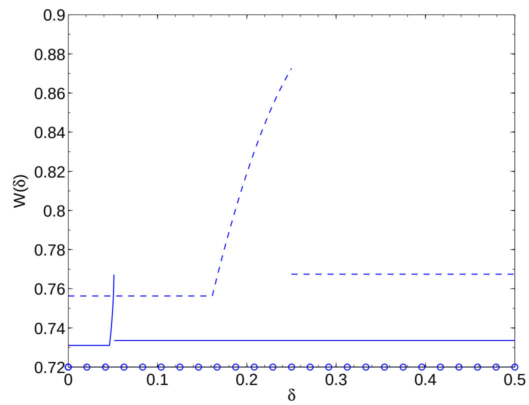
The results in Theorem 7 add to the robustness of Theorem 1. First, as stated in *(i)*, if costs of change are moderately convex, the results of Theorem 1 are unchanged. This is due to the continuity of $EAP(\delta)$ and $W(\delta)$ with respect to η. Second, in the special case where $\eta = 2$, the results of Theorem 1 also hold, except that the set of extra-hurdles that are both W- and P-optimal is not a singleton. This is nonetheless a knife-edged case for $\delta = 0$, since the electorate is then indifferent between selecting either candidate when $\eta = 2$. This property follows from the fact that the marginal utility loss of voters and candidates from a policy that differs from their ideal policy positions coincides with the marginal cost of engineering the policy change. Third, *(ii)* shows that for arbitrary values of $\eta \in (1, 2)$ and provided that $c > 0$, $\delta = 0$ is never optimal in terms of ex-ante policy polarization. The reason is that any sufficiently large extra-hurdle yields lower ex-ante policy polarization than $\delta = 0$. Figures 11(a)–11(d) illustrate the results of Theorem 7.

These graphs show welfare and ex-ante policy polarization for both linear ($\eta = 1$) and convex costs of change ($\eta = 1.5$ and $\eta = 2$).[37] The plots for $\eta = 1.5$ (solid lines) suggest that, for arbitrary $\eta \in (1, 2)$, even stronger results than statement *(ii)* of Theorem 7 might hold. More precisely, the solid plots of Figure 11(b) and Figure 11(d) suggest that, even if there are convex costs of change, $\delta = 0$ cannot be optimal in terms of welfare, because any sufficiently large extra-hurdle is W-increasing. Moreover, Figures 11(a)–11(d) suggest that for $\eta \in (1, 2)$, any non-zero extra-hurdle is both weakly W-increasing and weakly P-reducing. Thus, although for convex costs of change there does not always exist an extra-hurdle that is both W- and P-optimal, there is numerical evidence for the fact that any W-optimal extra-hurdle is P-reducing and any P-optimal extra-hurdle W-increasing. An example of a situation where there is no extra-hurdle that is both W- and P-optimal is given in Figure 11(c) and Figure 11(d) for the case of $\eta = 1.5$.

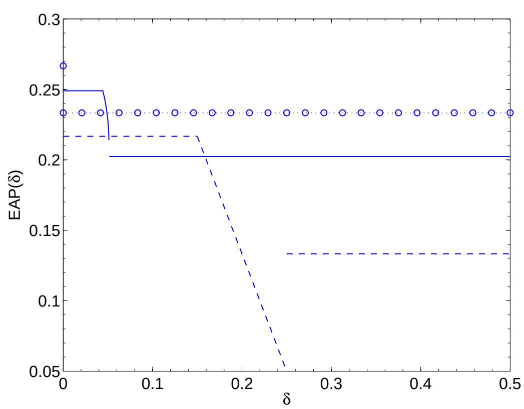
[37] The equilibrium policy choices (and the resulting values of ex-ante policy polarization and welfare) in the case of $\eta = 1.5$ have been computed numerically. Details of the calculations are available upon request.



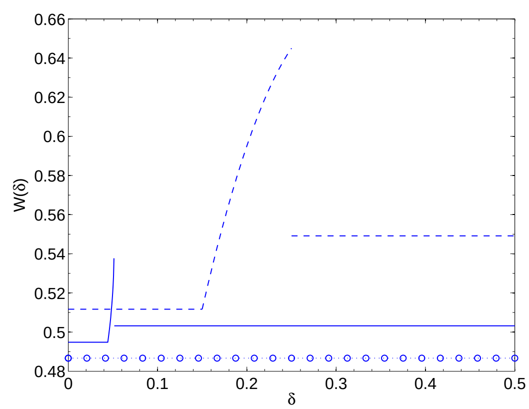
(a) Ex-ante policy polarization with $\rho = 0.1$.



(b) Welfare with $\rho = 0.1$.



(c) Ex-ante policy polarization with $\rho = \frac{1}{3}$.



(d) Welfare with $\rho = \frac{1}{3}$.

Figure 11: Equilibrium values of ex-ante policy polarization and welfare as a function of $\delta \in \left[0, \frac{1}{2}\right]$, using the parameter values $\mu_R = 0.8$, $c = 0.5$, $A = 2$, and $\rho = 0.1$ (resp. $\rho = \frac{1}{3}$). Each figure shows three different values of η: $\eta = 1$ (dashed), $\eta = 1.5$ (solid), and $\eta = 2$ (dotted).