

# Polarization and Policy Design\*

Giovanni Andreottola<sup>†</sup>      Christopher Li<sup>‡</sup>

November 2020

## Abstract

Political polarization has been on the rise in various democracies. One major concern is whether this leads to more divisive policies. The paper addresses this issue in a dynamic model of policy design. We show that increasing electoral polarization may in fact motivate politicians to design policies with less partisan bias. More generally, partisan discrimination is non-monotonic in the level of mass polarization. The prospect of political turnover and the gradual resolution of uncertainty over the effects of the policy are key for the result. We also consider the impact of different electoral systems. Counter to conventional wisdom, a proportional system may lead to more particularistic policymaking than a majoritarian system.

---

\*We would like to thank Tim Feddersen, David Levine, Andrea Mattozzi and Alberto Bennardo for useful comments and suggestions. We are also grateful to the audiences at Princeton, the European University Institute, the Federico II University of Naples, the Virtual Formal Theory Seminar and the OPESS Political Economy Seminar for their attention and stimulating questions.

<sup>†</sup>CSEF and University of Naples Federico II

<sup>‡</sup>Florida State University

# 1 Introduction

Many democracies are experiencing growing political polarization. In the U.S., the partisan divide is at its highest in over a century (McCarty et al., 2016). American voters are more ideologically aligned with their parties than ever before (Fiorina and Abrams, 2008) and also demonstrate greater antipathy and distrust toward the other party and its supporters (Iyengar and Krupenkin, 2018). Regardless of the specific measure of polarization, the consequence on voting behavior is clear: party affiliation has become the predominant factor in voters' decisions, sometimes at the expense of issue and platform considerations (Bartels, 2000; Abramowitz and Webster, 2016).<sup>1</sup>

Many observers view the growing polarization as the biggest threat to democracy today.<sup>2</sup> For example, it has been suggested that political polarization leads to political gridlock and lower legislation quality (Barber and McCarty, 2015; Mann and Ornstein, 2016).<sup>3</sup> More generally, there is widespread concern that a growing partisan polarization can lead to so-called “tribal politics”, where partisan interests are prioritized in policymaking.<sup>4</sup> The intuition is that when voters are highly partisan, it is more difficult and therefore less desirable for politicians to “reach across the aisle”. Indeed, there are instances of governments in polarized societies enacting divisive policies aimed to benefit one group of voters at the expense of the rest: the redistributive policies undertaken under various populist regimes are an example that reflect such dynamics.

---

<sup>1</sup>For example, straight-ticket voting, that is the situation where voters vote for the same party throughout the whole ballot, has been on the steep rise in the last several decades.

<sup>2</sup>A recent study suggests, somewhat worryingly, that some American voters prioritize party loyalty over even fundamental democratic principles (Graham and Svobik 2020).

<sup>3</sup>There is also evidence that a polarized society is more prone to civil conflicts (Esteban and Ray 1994, Montalvo and Reynal-Querol 2005).

<sup>4</sup>Philosophers from Plato to John Stuart Mill consider the “tyranny of majority” as an intrinsic threat to the inclusive and pluralistic ideals of democracy. Social psychologists have long documented the natural tendency for individuals to favor those considered of the same “group” and exclude those of the “out-group” (Tajfel 1970; Chen and Li 2009).

Despite its intuitive appeal, the reasoning above conflicts with some anecdotal evidence, suggesting that the relationship between polarization and policy partisanship is more complex and nuanced. Consider the Affordable Care Act (ACA), a major healthcare reform introduced in the U.S. in 2010. The political climate surrounding the legislation was highly partisan, with Republicans, who were the minority, steadfastly refusing to support the bill or even provide any input to it. Public opinion of the legislation was similarly divided along partisan lines. Nonetheless, the Democrats made several compromises when designing the ACA, despite the political gains from doing so being little if any. The bill incorporated a variety of policy ideas originating from previous Republican healthcare legislation. In fact, in many respects the ACA was less radical than the healthcare reform proposed by the first Clinton administration, despite the growing partisan divide that took place in the intervening years. Finally, as it became clear later on, the fear of many Republicans that they would stand to lose from the reform proved unjustified. By one estimate, as of 2016 68 of the 70 counties most heavily subsidized under the ACA were Republican-leaning (Noam, 2017).

Another interesting fact is that despite the growing antagonism between the Democratic and Republican party, they have in fact converged on some economic issues. A case in point is trade policy. The Republican party has recently shifted from its traditional pro-trade stance to a more protectionist one. The goal is to attract the traditionally Democratic-leaning blue collar workers in manufacturing states. Arguably, the success of this strategy relies on a strong level of partisanship, allowing to retain the support of the segments of the Republican base in favor of free trade.

The discussion so far begs the question: Is “tribal politics” a necessary consequence of mass polarization? Does increasing partisanship among voters translate to more-partisan policies? We explore these questions in a dynamic model of distributive politics and find that, contrary to the pessimistic views of polarization, politicians may be motivated to adopt less-partisan and more-inclusive policies when the electorate becomes more polarized. This occurs even

though politicians are assumed to serve partisan interests. More generally, the effect of polarization on partisan discrimination (i.e. the disparity in policy outcome between the different partisan factions) is non-monotonic.

The dynamic aspect of policymaking is central to our narrative, and it sets our model apart from other models of distributive politics (e.g., Lindbeck and Weibull, 1987; Lizzeri and Persico, 2001). One salient feature of the setup is that policies may be repealed in the future because of political turnover. Another notable assumption is that the full consequences of a policy only become known to voters over time. These considerations shape the policy choice of the incumbent, who serve the interests of their partisan base. Specifically, the incumbent must take into account the policy's durability given the prospect of political turnover. And because of the inter-temporal resolution of uncertainty, the incumbent can take two different approaches to achieving policy durability. One relies on building a majority coalition to stay in power and another relies on making the policy "repeal-proof". These strategies require the incumbent to appeal to different factions of voters. Moreover, the optimal strategy depends on polarization, which shapes the responsiveness of each voter faction to the policy.

The model, while parsimonious, is rich in implications. In addition to partisan discrimination, political turnover and the extent of partisan voting are also non-monotonic in the degree of polarization. Comparative statics with respect to the incumbent's capability to target voters and the efficiency of the policy also yield some interesting insights. In an extension, we show that the proportional system may lead to more partisan discrimination compared to the majoritarian system. This is in contrast to the common perception that a proportional system allows for more inclusive representation and thus curbs particularist policies. In another extension we consider policy design on both the distributive dimension and the efficiency dimension. Somewhat paradoxically, if the incumbent has strong office motivation, he may refrain from choosing the most efficient policy.

## 2 Related Literature

Our model is related thematically to the literature on distributive/pork-barrel politics, though the modeling approaches are quite different. In particular, most of these works are in the Downsian tradition and thus static in nature. Classic works such as Lindbeck and Weibull (1987), Cox and McCubbins (1986) and Dixit and Londregan (1995, 1996) feature office-motivated parties competing by promising particularistic benefits to various voter groups, e.g. moderates vs. partisans. The general take away is that parties try to appeal to whoever is most responsive to benefits at the margin. Among these, the recent paper by Krasa and Polborn (2014) is the closest to our work. They build on the classic models by considering voters who are heterogeneous in two dimensions, economic and cultural, and candidates who are exogenously differentiated. In equilibrium, candidates diverge in their policy positions and adjust their policies in the same direction (e.g., more taxation) when voters become more culturally polarized. Their findings, however, do not have a clear implication in terms of partisan discrimination, and should therefore be considered complementary to our results.

Our observation regarding the impact of majoritarian vs. proportional systems is related to the literature on electoral institutions and the size/composition of government spending (Myerson 1993, Persson and Tabellini 1999, Lizzeri and Persico 2001, Milesi-Ferretti et al. 2002). In particular, Persson and Tabellini (1999) predict higher public goods provision and less private transfers under the proportional system compared to the majoritarian system. Similarly, Lizzeri and Persico (2001) compare the composition of spending under the majoritarian vs. proportional system, and find that the electoral system affects public goods provision but not redistribution.

Various papers have explored policymaking in the shadow of political turnover. Persson and Svensson (1989) and Alesina and Tabellini (1990) take political turnover as exogenous and study its effects on government debt. They show that the incumbent may purposefully

borrow more than is optimal to tie the hands of future governments. Aghion and Bolton (1990) endogenize political turnover and show that fiscally conservative governments may paradoxically take on inefficiently high levels of debt in order to secure electoral victory. In a similar vein, Biais and Perotti (2002) show how incumbents would design privatization policies (e.g. rationing, underpricing of shares) as to encourage public participation and lock in political support. Robinson and Torvik (2005) and Besley and Coate (1998) show that incumbents manipulate spending (e.g. infrastructure) in order to gain electoral advantages. Finally, Battaglini and Harstad (2020) find that governments may deliberately make international agreements weak (i.e., non-binding) in order to shore up electoral support. The effects of polarization have been studied in the context of legislative bargaining, though the focus there is on gridlock and efficiency. Work such as Krehbiel (1998); Brady and Volden (2006); McCarty (2018), and Dziuda and Loeper (2016) show that the polarization of the bargaining players can lead to more gridlock/policy inertia. Austen-Smith et al. (2019) show that bargaining players may settle on inefficient policies because they are easier to repeal in the future. Moreover, this inclination has a non-monotonic relationship with the divergence of preferences among players.

Finally, in modeling policy uncertainty and its temporal resolution, we follow Fernandez and Rodrik (1991). The aggregate outcome under the policy is known but individuals face uncertainty regarding whether they are winners or losers under the policy. Fernandez and Rodrik (1991) take the policy proposals as given and focus on the reversal of public support due to gradual resolution of uncertainty. We incorporate this idea in a richer setting where the incumbent can shape the policy to manipulate public support at various stage of the uncertainty resolution.

### 3 Model

There are two parties, Democrat ( $\mathcal{D}$ ) and Republican ( $\mathcal{R}$ ).<sup>5</sup> There is a unit mass of voters, who are defined by two characteristics  $(\pi, \tau) \in \{D, R\} \times \{\delta, \rho\}$ . We call  $\pi$  partisanship and  $\tau$  targetable trait, which represents some socio-economic or demographic status. As the choice of notation suggests,  $(D, \cdot)$  voters are partisans for party  $\mathcal{D}$  and  $(R, \cdot)$  are partisans for party  $\mathcal{R}$ . Voters are equally divided in terms of partisanship and of the targetable trait (i.e.,  $Pr(D) = Pr(\delta) = \frac{1}{2}$ ). Partisanship and targetable trait are correlated, with  $Pr(\delta|D) = Pr(\rho|R) = q > \frac{1}{2}$ . The parameter  $q$  measures the extent of partisan sorting on the basis of observable traits like socio-economic status or geography.

The game proceeds for two periods, with date indexed by  $t \in \{1, 2\}$ . One of the parties is the incumbent at  $t = 1$ , which we take to be the Democrat for simplicity. The incumbent has the opportunity to adopt a reform. Under the status quo, all voters receive a payoff normalized to 0. Under the reform, on the other hand, some voters would be winners while others would be losers. We assume that at the end of the game, a winner obtains a payoff of 1 whereas losers obtains a payoff of  $-1$ . Now, the aggregate share of winners under the reform is known to be  $w > 0$ , which can be interpreted as the efficiency of the reform. There is, however, uncertainty about the reform at the individual level. Specifically, as in Fernandez and Rodrik (1991), voters only know the probability of being a winner in the immediate aftermath of the reform; their status as winners or losers is revealed only at  $t = 2$ .

The incumbent can design the reform to statistically discriminate voters based on their targetable traits. Formally, let  $w_\tau$  denote the probability that a  $(\cdot, \tau)$  voter is a winner. The incumbent can adopt any reform  $(w_\delta, w_\rho) \in [0, 1] \times [0, 1]$  subject to the feasibility constraint:

$$\frac{w_\delta + w_\rho}{2} = w. \tag{1}$$

---

<sup>5</sup>We use these names purely for exposition purposes.

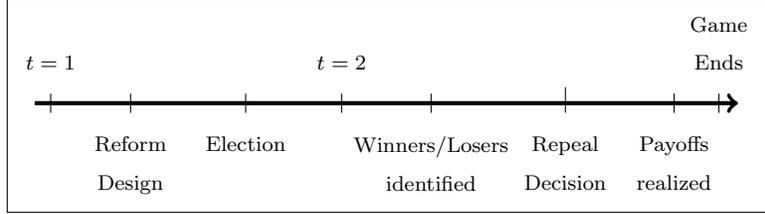


Figure 1: Game timing

In other words, the design of the reform corresponds to a lottery conditional on targetable traits. Voters observe the incumbent's choice of reform and therefore form correct beliefs about their probability of being a winner. Note that the incumbent cannot directly target voters based on partisanship, although the correlation between targetable traits and partisanship allows him to do so indirectly. This is reasonable given the legal restrictions on policies that explicitly discriminate voters on the basis of partisanship. More generally, it would be difficult for parties to perfectly identify voters' partisanship.

The incumbent may choose not to reform, in which case the status quo payoff is realized and the game ends. Otherwise, there is an election between the two parties, which is decided by majority rule, with a fair-coin toss to break ties. The winner of the election becomes the decision maker in period 2. If there has been a reform, the identity of winners and losers is revealed at  $t = 2$ . However, before voters obtain their payoffs, the party in office decides whether or not to repeal the reform.<sup>6</sup> If the reform is repealed, then all voters get the status quo payoff of 0. Otherwise, voters obtain their payoffs according to their status of winners or losers under the reform. Figure 1 summarizes the timing of the game.

Voters are forward looking, in that they anticipate the decisions of the two parties at  $t = 2$ . They vote sincerely, i.e. based on the expected utility differential between parties, but with a bias toward the party they are affiliated to. Formally, let  $u_{\pi,\tau}^j$  be the expected payoff for

<sup>6</sup>Notice that we assume that no payoffs are realized between the reform design stage and the repeal decision. This simplifies the analysis, but it is not crucial for our results, especially if the first round of payoffs is small compared to the second one, for example because the payoffs realized after the repeal decision represent the net present value of the stream of payoffs from the reform.

$(\pi, \tau)$  voters if party  $j$  wins.<sup>7</sup> Then,  $\mathcal{D}$  partisans (i.e.  $(D, \tau)$  voters) vote for  $\mathcal{D}$  if and only if

$$u_{D,\tau}^{\mathcal{D}} + v \geq u_{D,\tau}^{\mathcal{R}},$$

where  $v > 0$ . An analogous decision criterion applies for  $\mathcal{R}$  partisans. In other words, voters are loyal to their party unless the other party offers a net utility gain greater than  $v$ .<sup>8</sup> The parameter  $v$  measures the extent of party loyalty in voting, which, as we argued in the initial paragraph, directly reflects the intensity of partisanship/polarization.<sup>9</sup> Thus it is reasonable to take  $v$  as a proxy measure for polarization. In fact, measuring polarization by its behavioral implication allows us to be agnostic about the source of polarization. For example, one may think of polarization as the result the ideological alignment between voters and their parties on some unmodeled policy dimension, or alternatively attribute polarization to innate psychological factors associated with group identity.

We assume that parties serve partisan interests in that they seek to maximize the expected utility of their median party member.<sup>10</sup> A microfoundation for this type of party objective is that politicians are fully office motivated but accountable towards party members, for example due to primary elections.<sup>11</sup> Given the nature of the reform design and the inter-temporal resolution of uncertainty, the “identity” of the median partisan differs in the two periods. At  $t = 1$ , voters’ expected utility is a function of their targetable trait. Given the positive correlation between partisanship and targetable trait, the median member of party  $\mathcal{D}$ , the incumbent, is a  $(D, \delta)$  voter. Thus, party  $\mathcal{D}$ ’s objective in period 1 is to maximize the payoff

---

<sup>7</sup>Specifically,  $u_{\pi,\tau}^j$  is zero if  $j$  is expected to repeal the reform, and it is  $2w_\tau - 1$  otherwise.

<sup>8</sup>For simplicity, we assume that  $v$  is the same for all voters. However, we could relax this assumption and allow  $v$  to vary across different groups of voters without qualitatively changing the results.

<sup>9</sup>Party loyalty in voting has long been a standard measure of polarization of the legislators (see Barber and McCarty 2015 for a discussion).

<sup>10</sup>Given the assumption on the payoffs, this is also equivalent to maximizing aggregate welfare of the partisans.

<sup>11</sup>Indeed, the decision resulting from the objective is equivalent to the Condorcet winner of a “referendum” held by the partisans of the incumbent party.

of  $(D, \delta)$  voters. voters.<sup>12</sup> In period 2, the party in office decides whether to repeal the reform, which impacts voters' payoffs based on their status as winners or losers. Maximizing the payoff of its median member means that the party repeals the reform if and only if a majority of its members are losers.

## 4 Analysis

In this section we first discuss the different types of reforms that can arise in equilibrium. We then explore the comparative statics of reform choice with respect to polarization and its implications for partisan discrimination, which constitute the main result of our paper. The section concludes with a brief discussion of other findings of interest of the model.

### 4.1 Three “Focal” Reforms

Suppose the Democrat decides to proceed with the reform in period 1: how would it structure the reform? As discussed above, its objective is to maximize the expected utility of  $(D, \delta)$  voters, which is a majority among Democrat partisans. One obvious strategy for the Democrat is to set  $w_\delta$  as high as possible, which we call the *No-Compromise* reform.<sup>13</sup> However, by myopically maximizing  $w_\delta$ , the Democrat may alienate voters with  $\tau = \rho$ , who constitute the majority of the Republican base. This can result in the repeal of the reform if the Republican comes into power. Indeed, the Republican will want to repeal the No-Compromise reform as long as  $w < 1 - \frac{1}{4q} \equiv \bar{w}$ . In this case, the Democrat must balance between setting favorable

---

<sup>12</sup>To be precise,  $\mathcal{D}$  maximizes  $\mathbb{E}[u_{D,\delta}^j(w_\delta, w_\rho)]$  where we incorporate the reform explicitly as input and the expectation is taken with respect to the identity of the incumbent at  $t = 2$ . Note that parties do not take  $v$  into account in their objective, since we do not treat  $v$  as instrumental utility, but as a warm glow from voting for own party. Nevertheless, the results would not change qualitatively even if we incorporate  $v$  in the parties' objective.

<sup>13</sup>Note that because of the feasibility constraint 1, it is without loss to refer to the reform by either  $w_\delta$  or  $w_\rho$ .

terms for its partisans and ensuring the durability of the reform.<sup>14</sup>

Given the structure of the game, the reform is durable when either (i) the opposition (the Republican in this case) does not repeal it when in power or (ii) the incumbent builds a majority coalition and wins the election.<sup>15</sup> We call the best reform (from the perspective of the median Democrat partisan) that satisfies (i) the *Repeal-Proof* reform, and the best reform that satisfies (ii) the *Election-Winning* reform. Notice that if  $w$  is too low, then the reform cannot be made durable no matter the design. The lower bound on  $w$  that we need in order to have a non-trivial reform design problem is  $\frac{1}{2}$ .<sup>16</sup> Coincidentally, this condition also implies that the incumbent strictly prefers to adopt some version of the reform over the status quo. From now on, the analysis takes as given the following parameter restrictions, unless otherwise stated:

$$\frac{1}{2} < w < 1 - \frac{1}{4q} \quad (2)$$

The following Lemma summarizes our observations so far.

**Lemma 1.** *Assume that condition (2) holds. The incumbent chooses reform over status quo, and the optimal reform takes one of three forms:*

- (*No-Compromise*):  $w_\delta = 1$
- (*Election-Winning*):  $w_\delta = \min \left\{ 2w - \frac{1-v}{2}, 1 \right\}$
- (*Repeal-Proof*):  $w_\delta = \frac{wq-1/4}{q-1/2}$

It should be noted that the Election-Winning reform may coincide with the No-Compromise

---

<sup>14</sup>If this condition is violated, then the incumbent would trivially set  $w_\delta = 1$ . The reform will be not be repealed by either party at  $t = 2$ .

<sup>15</sup>This claim takes for granted the fact that the incumbent will not repeal if it wins the election. This is indeed the case given the restrictions on  $w$  that we shall impose.

<sup>16</sup>When  $w < \frac{1}{2}$ , there is no alternative for the incumbent besides setting  $w_\delta = 2w$  (i.e., the No-Compromise reform in Lemma 1 below).

reform under some parameter values. To avoid confusion, we reserve the term No-Compromise for when such policy is *not* election-winning.

The endogenous variable  $w_\delta$  is a natural measure of the degree of *partisan discrimination* under the reform. In particular, given the feasibility constraint and the correlation between partisanship and the targetable trait, the greater the  $w_\delta$ , the larger the gap between the share of winners among Democrat partisans and among Republican partisans.<sup>17</sup> Based on our discussion thus far, several useful facts can be derived.

**Fact 1.** *The No-Compromise reform ranks the highest in terms of partisan discrimination, followed by the Election-Winning reform, and then the Repeal-Proof reform.*

The fact that the No-Compromise reform is the most discriminatory is fairly obvious. The intuition for why the Election-Winning reform is more discriminatory than the Repeal-Proof reform is instead more subtle. It follows from the observation that the Election-Winning reform requires the parties to act differently at  $t = 2$ , since otherwise there is no reason for voters to vote across party line. Thus, a necessary condition for the Election-Winning reform is that the Republican repeals the reform, while the Democrat upholds it. This means that under the Election-Winning reform, less than a majority of Republican partisans are winners. On the other hand, a reform is repeal-proof precisely when a majority of Republican partisans are winners.

**Fact 2.** *By definition, the Election-Winning reform results in an outright win by the incumbent. The Repeal-Proof and the No-Compromise reforms result in a tied election with voters voting for their own respective parties.*

Under the No-Compromise reform, the Republican would repeal the reform if it wins. So long as polarization is not too high, voters split among the targetable trait, with  $(\cdot, \delta)$  voters voting

---

<sup>17</sup> $w_\delta$  more directly captures the discrimination between voters with different targetable traits.

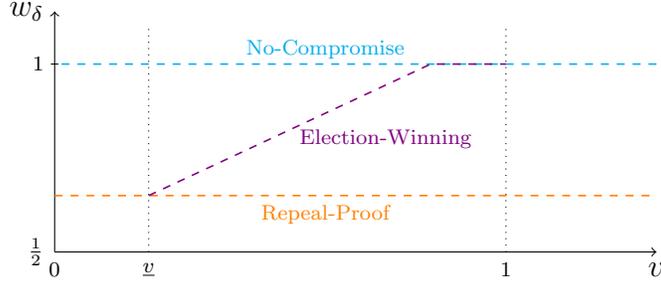


Figure 2: Three focal policies, as a function of  $v$

for the Democrat and  $(\cdot, \rho)$  for the Republican. If polarization is sufficiently high, then voters split along partisan lines. Either case results in a tied election. The Repeal-Proof reform also leads to a tied election but for a different reason: by making the reform repeal-proof, there is no differentiation between the parties on instrumental grounds, hence partisanship is the sole determinant of the vote.

The main result presented in the next section concerns the relationship between partisan discrimination and polarization. Therefore it is useful to consider the three focal reforms as functions of  $v$ .

**Fact 3.** *The No-Compromise and Repeal-Proof reforms are independent of  $v$ . The Election-Winning reform is feasible if and only if  $v$  is in the interval  $[\underline{v}, 1]$ , and  $w_\delta$  as induced by the reform is increasing in  $v$  in this interval.*

Fact 3 follows from Lemma 1, and is illustrated in Figure 2. It is obvious that the No-Compromise reform should be independent of  $v$  as the incumbent myopically maximizes  $w_\delta$ .<sup>18</sup> Similarly, the Repeal-Proof reform is independent of  $v$  since it is pinned down by the condition that the Republican is indifferent to repeal at  $t = 2$ . On the contrary, partisanship does have an effect on the Election-Winning reform since it influences voting. Specifically, it shapes how responsive the various types of voters are to the benefits and harms of the policy.

<sup>18</sup>Our assumption that  $w > \frac{1}{2}$  implies that the incumbent can set  $w_\delta = 1$ .

In order for the Democrat to build a majority coalition to win the election, it needs the support of some Republican partisans while maintaining the support of its base. The easiest group of Republican voters to persuade are those with  $\tau = \delta$  since these voters already stand to benefit from the Democrat inclination to favor  $(D, \delta)$  voters. At the same time, the Democrat cannot tilt the policy too much lest in favor of  $(\cdot, \delta)$  voters lest  $(D, \rho)$  voters defect. It turns out that when it exists, the Election-Winning reform is pinned down by the constraint that the  $(D, \rho)$  voters do not defect. As partisanship increases,  $(D, \rho)$  voters are less willing to defect and therefore the Democrat can set a larger  $w_\delta$ . To see why the Election-Winning reform is feasible only within a sub-interval of  $v$ , note first that when  $v$  is too high, it is impossible to sway  $(R, \delta)$  voters. Conversely, as  $v$  becomes small, the Democrat must correspondingly lower  $w_\delta$  in order to prevent  $(D, \rho)$  voters from defecting. At a certain point,  $w_\delta$  is low enough (and  $w_\rho$  high enough) such that the reform becomes repeal-proof, which as discussed above implies that it is no longer possible to sway  $(R, \delta)$  voters.

## 4.2 Result: Partisan Discrimination and Polarization

The main result of the paper is Proposition 1 below. It shows that polarization has a non-monotonic effect on partisan discrimination.

**Proposition 1.** *If  $q \leq \frac{1}{6-8w} \equiv \bar{q}$ , then in equilibrium, partisan discrimination is inverse U-shaped in the level of polarization. Specifically, there is a threshold  $\underline{v} < 1$  such that the incumbent chooses:*

- *the Repeal-Proof reform for  $v \leq \underline{v}$  and  $v \geq 1$ .*
- *the Election-Winning reform for  $v \in (\underline{v}, 1)$*

*If  $q > \bar{q}$ , then in equilibrium partisan discrimination is U-shaped in the level of polarization. Specifically, there is a threshold  $\tilde{v}$  where  $\underline{v} < \tilde{v} < 1$ , such that the incumbent chooses:*

- *the No-Compromise policy for  $v \leq \tilde{v}$  and  $v \geq 1$*
- *the Election-Winning policy for  $v \in (\tilde{v}, 1)$*

The proposition is illustrated in Figure 3a and 3b. To see the intuition behind the result, notice first that the Election-Winning reform, whenever feasible, is preferred to the Repeal-Proof reform. Recall that the Democrat faces a trade-off between maximizing the welfare of its partisans and ensuring the durability of the reform. Both the Election-Winning and the Repeal-Proof reform are ways to reduce the chance of the reform being overturned, but it turns out that the former is less costly for the Democrat, and thus preferred. As alluded in the previous section, by definition the Repeal-Proof reform requires a weak majority of Republican partisans being winners. The Election-Winning reform, on the other hand, rests on the fact that the Republican will overturn the reform at  $t = 2$ , that is, the share of winners within the Republican base is less than one half.

The Democrat also prefers the Election-Winning reform to the No-Compromise one depending on the degree of polarization. Recall that when the Democrat seeks to build a majority coalition, the binding constraint is to prevent defections by  $(D, \rho)$  voters. When partisanship (i.e.,  $v$ ) is low, building such a coalition is costly, as it requires relatively high  $w_\rho$  to ensure  $(D, \rho)$  voters do not defect. This cost decreases as  $v$  increases to the point where the incumbent need not appease  $(D, \rho)$  voters to maintain a majority coalition. Thus the incumbent prefers the Election-Winning reform (when feasible) to the No-Compromise reform when there is a relatively high degree of polarization.

For low and high levels of polarization, the Election-Winning reform is not available, and the incumbent must choose between the No-Compromise and Repeal-Proof reforms. Similarly to what happens with the Election-Winning reform, making the reform repeal-proof comes at a cost to the Democrat. When this cost is too high, the incumbent finds it optimal to take his chances with the No-Compromise reform; otherwise, the Repeal-Proof reform is optimal.

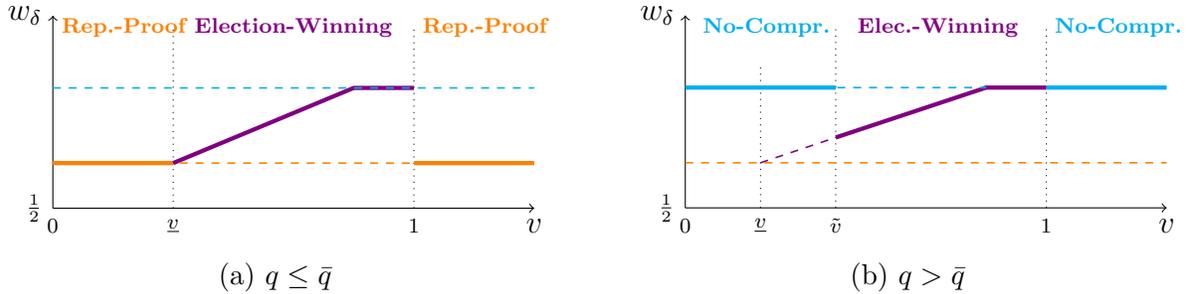


Figure 3: Equilibrium Reform

Now, the cost is inversely proportional to the extent of partisan sorting, as measured by  $q$ . If  $q$  is low it is less costly to ensure repeal-proofness, since the  $(R, \delta)$  voters, who stand to gain from the Democrat’s reform, make up for a larger portion of the Republican base. As a result, there is a threshold on  $q$  below which the Repeal-Proof policy becomes optimal.<sup>19</sup>

### 4.3 Discussion

Given the stylized nature of the model, the reader may question the robustness of our findings. We have already pointed out the (in)significance of various modeling assumptions in the footnotes in Section 3. Here we address the generalizability of our results further. One notable feature of the model is its symmetry. For example, the electorate is split evenly along both partisanship and the targetable trait. This leads to tied elections in several scenarios, which may seem knife-edge. However, if one allows for some electoral uncertainty (more on this below), then asymmetries in the distribution of voter characteristics would not change our results qualitatively. Finally, it should be noted that the symmetric tie-breaking rule is not crucial to our insight; all that is required is that the incumbent loses with some probability when the election is tied.

The model can also be generalized to allow for asymmetries in the gain/loss for the winners/losers, as well as to allow for the intensity of “party loyalty”,  $v$ , to vary with the

<sup>19</sup>To be precise, the threshold  $\bar{q}$  may not exist if  $w$  is large.

targetable trait. Finally, our insights are robust to the presence of electoral uncertainty and parties being office-motivated. These extensions are explored in more detail in Sections 5 and 6, since they also give rise to some additional insights that are interesting in their own right. An element that is crucial for our results is the gradual resolution of policy uncertainty. The implications of this assumption are two-fold. First, uncertainty gives the incumbent the leverage to influence election outcomes through policy design. If the identity of winners and losers were known at the time of the election, the outcome of the election would then depend solely on the intensity of partisanship, and not on the nature of the reform.<sup>20</sup>

The second implication of the inter-temporal resolution of uncertainty is that the identity of the “median” partisan changes across the two periods. In particular, at  $t = 1$ , the “median” partisan is characterized by her targetable trait, and this induces the incumbent to privilege voters based on such trait. At  $t = 2$ , instead, the “median” partisan is characterized by her status as winner or loser. Interestingly, this may give rise to a phenomenon where the attitude of a party towards a policy changes over time, which echos the central insight of Fernandez and Rodrik (1991). In the context of our model, the Republican opposes the Repeal-Proof reform initially but then does not have the incentive to repeal it in period 2. This may help rationalize the puzzling observation of opposition parties who are adversarial to the incumbent’s policy initiatives – and may in fact win office on the promise of undoing such policies – but who then lack the political will or support to go through with their promise once in office. This was the case for the ACA: the Republicans swept into House majority in 2010, and later into the White House, partly on the promise to repeal the ACA. Yet, once in power (what is more, with a unified government), their various attempts to repeal ACA failed.

The gradual resolution of uncertainty is undeniably a salient feature of policymaking. Rarely

---

<sup>20</sup>As a matter of fact, voters would either vote based on their status as winners/losers, or based on their partisan affiliation.

are the full consequences of a policy known immediately after it is adopted. This feature has been somewhat under-emphasized in the literature and our model demonstrates that this can have important implications for policymaking in a dynamic environment.

#### 4.4 Additional Comparative Statics

In addition to our main result, the model yields a rich variety of comparative statics, some of which may be empirically testable. First, as a straightforward corollary of Proposition 1, political turnover and partisan voting are non-monotonic in the degree of polarization.

**Corollary 1.** *The incumbent’s reelection probability is inverse U-shaped in  $v$ . If  $q \leq \frac{1}{6-8w}$ , the share of voters who vote according to their party affiliation is U-shaped in  $v$ .*

How the voting pattern varies with respect to the different “focal” policies warrants some discussion. Under the Repeal-Proof reform, voting is split along partisan line. The same holds for the No-Compromise reform when polarization is high, whereas when polarization is low, voting is split along the targetable trait. Finally, the Election-Winning reform leads to a mixture of partisan and trait-based voting, with some Republican partisans voting for the Democrat.

Proposition 1 also has interesting implications on how partisan sorting, measured by  $q$ , affects partisan discrimination. Notably, for high and low levels of polarization,  $\bar{q}$  is the threshold at which the incumbent switches from the No-Compromise reform to the Repeal-Proof reform. This gives rise to the following result.

**Corollary 2.** *Partisan discrimination is higher when  $q > \bar{q}$  than when  $q \leq \bar{q}$  (strictly when  $v < \tilde{v}$  and  $v > 1$ ).*

This observation is in line with the impression that in societies that are polarized along cleavages that are highly targetable, e.g. race or income, policies are often marked with

extreme bias and discrimination. On the contrary, in societies where political polarization is driven by cultural differences, as arguably the case in the U.S., policies may still remain relatively inclusive.

Lemma 1 and Proposition 1 also give rise to some interesting comparative statics with respect to the efficiency of the reform,  $w$ .

**Corollary 3.** *There exists a threshold  $\bar{w}$  such that for  $v < \tilde{v}$  and  $v > 1$ , partisan discrimination is highest but constant when  $w < \bar{w}$ . At  $w = \bar{w}$ , partisan discrimination displays a downward jump, after which it is again increasing in  $w$ .*

This result shows that partisan discrimination may be non-monotonic in the efficiency of the reform. Thus, the “best” reform from a utilitarian perspective may not be the best from an equity stand point.

## 5 Proportional System

So far, we assumed that the party with the majority of votes wins outright. It is standard in the literature to interpret this assumption as the majoritarian or winner-take-all system (see Lizzeri and Persico (2001) for example). Another electoral institution common to democracies is the proportional system, in which power is split based on the share of votes earned in the election.

In this section we explore the outcomes of our model under the proportional system and compare them to that under the majoritarian system. In modeling a proportional system, we follow Lizzeri and Persico (2001) and assume that the probability of a party winning the election is equal to its vote share.<sup>21</sup> Clearly, this has a “probabilistic voting” flavor and

---

<sup>21</sup>See the Appendix for a formal definition.

therefore this exercise can also be seen as an evaluation of the robustness of our results to the presence of electoral uncertainty.

The two conclusions from this exercise, drawn from Proposition 2 below, are that (i) the observation that polarization has a non-monotonic effect on partisan discrimination still holds qualitatively and that (ii) the proportional system may reduce or worsen partisan discrimination.

**Proposition 2.** *If  $q \leq \frac{1}{6-8w} = \bar{q}$ , then the Repeal-Proof reform is chosen for a broader range of values of  $v$  compared to the majoritarian benchmark. Specifically, the incumbent chooses*

- *the Repeal-Proof reform for  $v \leq \underline{v}_{prop}$ , where  $\underline{v}_{prop} > \underline{v}$ , and for  $v \geq 1$ .*
- *the Election-Winning reform<sup>22</sup> for  $v \in (\underline{v}_{prop}, 1)$ .*

*If  $q > \bar{q}$ , then the No-Compromise reform is chosen for a broader range of values of  $v$  compared to the majoritarian benchmark. Specifically, the incumbent chooses*

- *the No-Compromise reform for  $v \leq \tilde{v}_{prop}$ , where  $\tilde{v}_{prop} > \tilde{v}$ , and for  $v \geq 1$*
- *the Election-Winning reform for  $v \in (\tilde{v}_{prop}, 1)$ .*

The result follows from the observation that the proportional system does not affect the incumbent’s calculus when it comes to the No-Compromise and Repeal-Proof reforms, since in either case the electorate is split, resulting in a toss-up election as before. The “Election-Winning” reform, on the other hand, no longer guarantees the incumbent’s reelection.<sup>23</sup> This reduces the incumbent’s incentives to choose the Election-Winning reform. Nonetheless,

---

<sup>22</sup>Strictly speaking this is a misnomer given that the incumbent cannot win the election with probability 1. This caveat aside, the incumbent indeed chooses the “Election-Winning” reform as defined in Lemma 1, which maximizes the share of votes won in the election.

<sup>23</sup>It does still maximize the probability of reelection.

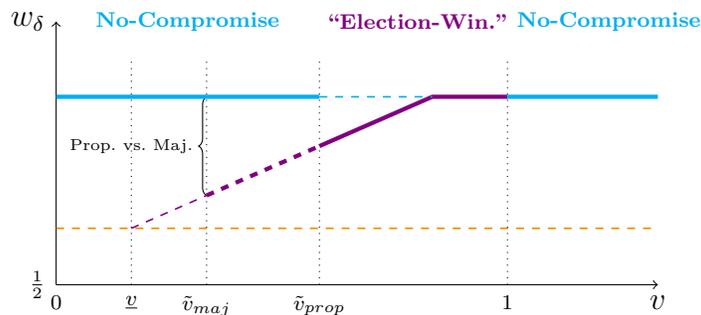


Figure 4: Proportional vs Majoritarian in Scenario 2

polarization still has a non-monotonic effect on partisan discrimination as in the benchmark case.

When  $q \leq \bar{q}$ , the proportional system reduces partisan discrimination relative to the majoritarian system, as the Repeal-Proof reform is chosen over a larger set of parameters. This is in line with the general wisdom that the proportional system generates more egalitarian and less particularistic outcomes. However, when  $q > \bar{q}$ , the proportional system leads to more partisan discrimination relative to the majoritarian system (see Figure 4). This is in contrast to the conventional wisdom and lessons from the literature. For example, Persson and Tabellini (1999) and Lizzeri and Persico (2001) suggest that the proportional system leads to greater provision of public goods and/or less targeted transfers.

## 6 Office Motivation and Endogenous $w$

In the baseline model, the parties' objective is to maximize the payoffs of their median partisan supporters. It is also natural to consider the case where parties are directly office motivated.<sup>24</sup> We explore this possibility in this section. First, we show that our results are robust to incorporating office motivation. Moreover, allowing for office motivation gives rise

<sup>24</sup>It should be noted that the incumbent does care about winning office in the baseline model, albeit indirectly.

to the interesting observation that, unlike in the baseline model, the incumbent can be worse off when designing more efficient reforms (i.e., higher  $w$ ).

Formally, we assume that the objective function of the incumbent (the Democrat) in period 1 is:

$$\mathbb{E}[u_{D,\delta}^j + \mathbb{1}\{reelection\}r], \quad (3)$$

where  $\mathbb{1}$  is the indicator function and  $r$  represents level of the office rents. Note that the baseline model is a special case with  $r = 0$ . The parties' objective in period 2 remains unchanged. The first result is that the main insight concerning the comparative statics of reform choice with respect to polarization still holds.

**Proposition 3.** *If  $q \leq \bar{q}$ , the incumbent's choice of reform in equilibrium is the same as in the benchmark. If  $q > \bar{q}$ , the Election-Winning reform is chosen for a larger range of  $v$  compared to the benchmark. Specifically, the incumbent chooses:*

- *the No-Compromise policy for  $v \leq \tilde{v}_{off}$ , where  $\tilde{v}_{off} < \tilde{v}$ , and for  $v \geq 1$*
- *the Election-Winning policy for  $v \in (\tilde{v}_{off}, 1)$*

*Additionally, for sufficiently large  $r$ ,  $\tilde{v}_{off} = \underline{v}$ , i.e. the Election-Winning policy is chosen whenever available.*

The intuition for the result is straightforward. Being office motivated makes the incumbent more inclined towards the Election-Winning reform. This results in no change in behavior compared to the benchmark when  $q \leq \bar{q}$ , since the incumbent already chooses the Election-Winning reform whenever possible. For  $q > \bar{q}$ , instead, with office motivation, the incumbent chooses the Election-Winning reform over the No-Compromise one for a strictly larger set of values of  $v$ .

One can show that in the benchmark setting, if the incumbent could choose efficiency of the reform, i.e.,  $w$ , prior to the allocative decision, he would choose the maximal  $w$ . However, this is no longer necessarily the case when the incumbent is directly office motivated.

**Proposition 4.** *If the incumbent’s office motivation is sufficiently large (i.e.  $r$  is sufficiently high) and  $q \leq \bar{q}$ , then there exist  $w$  and  $w' > w$  such that for some values of  $v$  the incumbent prefers  $w$  over  $w'$ .*

When  $q \leq \bar{q}$ , the incumbent switches from the Repeal-Proof reform to the Election-Winning reform as  $v$  crosses  $\underline{v}$ . Suppose for instance  $v = \underline{v} + \epsilon$ : in this case, the Election-Winning reform is very close to the Repeal-Proof reform i.e., there is almost a majority of winners among the Republican partisans. As  $w$  increases, this has the paradoxical effect of making the Election-Winning reform infeasible. Since the “surplus budget” must be allocated based only on the targetable trait, the outcome is a majority of Republican partisans being winners of the reform regardless of how the extra resources are allocated. Consequently this no longer allows the Democrat to differentiate itself from the Republican on instrumental grounds and this leads to a toss-up election. This is clearly undesirable for the Democrat if it has strong office motivation.

## 7 Conclusion

Virtually all policies result in winners and losers: as a result, policymaking is a balancing act between serving narrow partisan interests and building broader political coalitions. In this paper we explore how politicians address this trade-off when facing a polarized electorate. We show, somewhat surprisingly, that the increasing partisan divide need not result – as the conventional wisdom about “tribal politics” suggests – in more partisan policymaking. Politicians, even those under partisan pressure, may choose more inclusive policies when

the electorate becomes more polarized. Another contribution of our model is to highlight the role played by the gradual resolution of policy uncertainty, which has been somewhat underappreciated in the existing literature. Investigating the implications of this feature for policymaking in a dynamic environment, especially in conjunction with political turnover, could be a fruitful subject of future research.

Our model is parsimonious yet delivers a variety of findings. In addition to the main result, the equilibrium characterization has implications in terms of how political turnover and partisan voting are affected by voters' polarization. The model also sheds light on the comparative impact of different electoral systems on partisan discrimination. In opposition to conventional wisdom, we find that the proportional system may exacerbate political particularism compared to a majoritarian system. We also show that relaxing the "resource constraint" faced by politicians may not improve outcomes from the perspective of equity.

## References

- Abramowitz, A. and S. Webster (2016). The rise of negative partisanship and the nationalization of us elections in the 21st century. *Electoral Studies* 41, 12–22.
- Aghion, P. and P. Bolton (1990). Government domestic debt and the risk of default: a political-economic model of the strategic role of debt. *Public debt management: theory and history* 315.
- Alesina, A. and G. Tabellini (1990). A positive theory of fiscal deficits and government debt. *The Review of Economic Studies* 57(3), 403–414.
- Austen-Smith, D., W. Dziuda, B. Harstad, and A. Loeper (2019). Gridlock and inefficient policy instruments. *Theoretical Economics* 14(4), 1483–1534.
- Barber, M. and N. McCarty (2015). Causes and consequences of polarization. *Political negotiation: A handbook* 37, 39–43.
- Bartels, L. M. (2000). Partisanship and voting behavior, 1952-1996. *American Journal of Political Science*, 35–50.
- Battaglini, M. and B. Harstad (2020). The political economy of weak treaties. *Journal of Political Economy* 128(2), 544–590.
- Besley, T. and S. Coate (1998). Sources of inefficiency in a representative democracy: a dynamic analysis. *American Economic Review*, 139–156.
- Biais, B. and E. Perotti (2002). Machiavellian privatization. *American Economic Review* 92(1), 240–258.
- Brady, D. W. and C. Volden (2006). *Revolving gridlock: Politics and policy from Jimmy Carter to George W.* Westview Press.

- Chen, Y. and S. X. Li (2009). Group identity and social preferences. *American Economic Review* 99(1), 431–57.
- Cox, G. W. and M. D. McCubbins (1986). Electoral politics as a redistributive game. *The Journal of Politics* 48(2), 370–389.
- Dixit, A. and J. Londregan (1995). Redistributive politics and economic efficiency. *American political science Review*, 856–866.
- Dixit, A. and J. Londregan (1996). The determinants of success of special interests in redistributive politics. *the Journal of Politics* 58(4), 1132–1155.
- Dziuda, W. and A. Loeper (2016). Dynamic collective choice with endogenous status quo. *Journal of Political Economy* 124(4), 1148–1186.
- Esteban, J.-M. and D. Ray (1994). On the measurement of polarization. *Econometrica: Journal of the Econometric Society*, 819–851.
- Fernandez, R. and D. Rodrik (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *The American economic review*, 1146–1155.
- Fiorina, M. and S. Abrams (2008). Political polarization in the american public. *Annual Review of Political Science* 11, 563–588.
- Graham, M. H. and M. W. Svobik (2020). Democracy in america? partisanship, polarization, and the robustness of support for democracy in the united states. *American Political Science Review* 114(2), 392–409.
- Iyengar, S. and M. Krupenkin (2018). The strengthening of partisan affect. *Political Psychology* 39, 201–218.

- Krasa, S. and M. Polborn (2014). Social ideology and taxes in a differentiated candidates framework. *American Economic Review* 104(1), 308–22.
- Krehbiel, K. (1998). *Pivotal politics: A theory of US lawmaking*. University of Chicago Press.
- Lindbeck, A. and J. W. Weibull (1987). Balanced-budget redistribution as the outcome of political competition. *Public choice* 52(3), 273–297.
- Lizzeri, A. and N. Persico (2001). The provision of public goods under alternative electoral incentives. *American Economic Review* 91(1), 225–239.
- Mann, T. E. and N. J. Ornstein (2016). *It's even worse than it looks: How the American constitutional system collided with the new politics of extremism*. Basic Books.
- McCarty, N. (2018). Pivotal politics, partisan polarization, and policy predictability. *The Journal of Politics* 80(3), 1076–1081.
- McCarty, N., K. T. Poole, and H. Rosenthal (2016). *Polarized America: The dance of ideology and unequal riches*. mit Press.
- Milesi-Ferretti, G. M., R. Perotti, and M. Rostagno (2002). Electoral systems and public spending. *The quarterly journal of economics* 117(2), 609–657.
- Montalvo, J. G. and M. Reynal-Querol (2005). Ethnic polarization, potential conflict, and civil wars. *American economic review* 95(3), 796–816.
- Myerson, R. B. (1993). Incentives to cultivate favored minorities under alternative electoral systems. *American Political Science Review*, 856–869.
- Noam, L. (2017). Trump voters would be among the biggest losers in republicans' obamacare replacement plan. *Los Angeles Times*.

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

Persson, T. and G. Tabellini (1999). The size and scope of government:: Comparative politics with rational politicians. *European economic review* 43(4-6), 699–735.

Robinson, J. A. and R. Torvik (2005). White elephants. *Journal of public economics* 89(2-3), 197–210.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American* 223(5), 96–103.

# A Additional Model Details and Proofs

## Derivation of Upper Bound on $w$

Suppose  $w \geq \frac{1}{2}$ . Setting  $w_\delta = 1$  would be the reform design the Democratic party would choose if it could single-handedly determine the outcome. However, such a policy only leaves the residual  $2w - 1$  to the  $\tau = \rho$  group, which forms a majority in the Republican party. The share of winners in the Republican constituency, that is among voters with  $\pi = R$ , is therefore equal to  $(1 - q)(1) + q(2w - 1)$ . In the second period, if the Republican party takes office it repeals the reform as long as the share of winners in its constituency is smaller than one half. This requires to solve the following inequality:

$$(1 - q) + q(2w - 1) \leq \frac{1}{2} \tag{4}$$

which then yields condition  $w < 1 - \frac{1}{4q}$ .

In other words, for high enough values of  $w$  and/or low enough values of  $q$ , the reform design problem is straightforward: the Democratic party can set the highest possible winning probability for its majority group and not face the possibility of repeal. That is why we focus on the more interesting parameter space in which  $w < 1 - \frac{1}{4q}$ .

## Voting Behavior

Voters are forward-looking, hence they anticipate the decision that each party takes at  $t = 2$  if it wins the election. Given a reform design  $(w_\delta, w_\rho)$ , therefore, their expected utility conditional on their social trait  $\tau$  and on the anticipated decision of party  $j \in \{\mathcal{D}, \mathcal{R}\}$  is:

$$u_\tau^j(w_\delta, w_\rho) = \begin{cases} 2w_\tau - 1, & \text{if } j \text{ upholds} \\ 0, & \text{if } j \text{ repeals} \end{cases} \tag{5}$$

The voting decision depends on  $u_\tau^j(w_\delta, w_\rho)$  and on the level of partisanship  $v$ . In particular, a voter with social trait  $\tau$  and partisanship  $\pi$  votes for the Democratic party as long as the following condition is satisfied:

$$\begin{aligned} u_\tau^{\mathcal{D}}(w_\delta, w_\rho) + v &\geq u_\tau^{\mathcal{R}}(w_\delta, w_\rho), & \text{if } \pi = D \\ u_\tau^{\mathcal{D}}(w_\delta, w_\rho) &\geq u_\tau^{\mathcal{R}}(w_\delta, w_\rho) + v, & \text{if } \pi = R \end{aligned} \tag{6}$$

As a consequence of this, a voter anticipating that both parties will take the same decision has  $u_\tau^{\mathcal{D}}(w_\delta, w_\rho) = u_\tau^{\mathcal{R}}(w_\delta, w_\rho)$  and will vote according to her partisan affiliation.

Notice that if voters vote based on their partisan affiliation, the election results in a tie, since  $Pr[\pi = D] = \frac{1}{2}$ , that is exactly one half of the population has each type of partisanship.

The same happens if voters vote according to their social trait  $\tau$ , which also divides the population in half. Finally, notice that since  $u_\tau^j(w_\delta, w_\rho) \in [0, 1]$ , for  $v > 1$  all voters vote according to their partisan affiliation  $\pi$ .

### Repeal-Proof Policy

A policy is repeal-proof if it satisfies the following two conditions. The first is:

$$(1 - q)w_\delta^{rp} + qw_\rho^{rp} \geq \frac{1}{2} \tag{7}$$

which implies that at least a half of Republican partisans are winners of the reform. This prevents repeal if the Republican party wins the election. The second condition is:

$$qw_\delta^{rp} + (1 - q)w_\rho^{rp} \geq \frac{1}{2} \tag{8}$$

which states that at least one half of Democrats have to be winners of the reform, in order to prevent repeal if the Democratic party wins the election. From condition (7) and condition

(8) we get  $w \geq \frac{1}{2}$  as a necessary condition for the Repeal-Proof reform to exist.

Since the objective of the Democratic party is to maximize the expected payoff of  $\tau = \delta$  voters, it is optimal to let condition (7) hold with equality. Then, solving the feasibility constraint (1) and condition (7) results in the following probabilities for citizens with  $\tau = d$  and  $\tau = r$  to be winners under the Repeal-Proof reform:

$$w_{\delta}^{rp} = \frac{1}{2} \frac{4wq - 1}{2q - 1} \quad (9)$$

and

$$w_{\rho}^{rp} = \frac{1}{2} \frac{1 - 4w(1 - q)}{2q - 1} \quad (10)$$

Finally, notice from expression (10) that the value of  $w_{\rho}^{rp}$  is decreasing in  $w$  and increasing in  $q$ , meaning that the larger efficiency and the lower sorting, the more the Repeal-Proof policy favors the  $\delta$  group.

### **Election-Winning Policy**

The Election-Winning policy needs to satisfy the following two conditions, which follow from equation (6). The first is:

$$w_{\rho}^{ew} \geq \frac{1 - v}{2} \quad (11)$$

which makes sure  $(D, \rho)$  voters support the Democratic party in the election. In order to get the support of  $(R, \delta)$  voters, instead, it is necessary that:

$$w_{\delta}^{ew} \geq \frac{1 + v}{2} \quad (12)$$

Summing these two inequalities we get, as a necessary condition for the Election-Winning policy to exist, that  $w \geq \frac{1}{2}$ , just like for the Repeal-Proof policy. Moreover, as a consequence

of the feasibility constraint (1) and the fact that  $w \geq \frac{1}{2}$ , notice that as long as either constraints (11) or (12) hold with equality, the other one is satisfied, too. Given that the Democratic party aims to maximize the expected utility of the  $\delta$  group, it always finds it optimal to let constraint (11) hold with equality. For this reason, the Election-Winning policy is pinned down by the indifference of the  $(D, \rho)$  voters, and hence it is downward sloping in  $v$ . From condition (12) it is also clear that when  $v > 1$ , the Election-Winning policy cannot be achieved. The reason is that it would require  $w_\delta^{ew} > 1$  to convince  $(R, \delta)$  voters to switch to the Democratic party.<sup>25</sup>

The second crucial requirement of the Election-Winning policy is that it is not repeal-proof. In other words, the Election-Winning policy cannot satisfy condition (7).<sup>26</sup> The reason is that if that were the case, voters would expect Republicans not to repeal the policy if they win the election. However, the optimal response in this case would be to vote according to partisanship. If this happens, the election results in a tie, contradicting the definition of Election-Winning policy. Mathematically, the Election-Winning policy is not repeal-proof as long as:

$$w_\rho^{ew} = \frac{1-v}{2} \leq w_\rho^{rp}. \quad (13)$$

Solving this inequality for  $v$  using expression (10) yields the expression for  $\underline{v}$ :

$$v \geq \frac{2(1-q)(2w-1)}{2q-1} \equiv \underline{v} \quad (14)$$

The intuition for this result is that if  $v$  is small, buying the loyalty of  $(D, \rho)$  voters requires a large value of  $w_\rho^{ew} = \frac{1-v}{2}$ , which gets close to  $\frac{1}{2}$  as  $v$  goes to 0. If  $w_\rho$  grows too much, however, at some point condition (7) is satisfied. Hence, the policy becomes repeal-proof and it cannot be election-winning.

---

<sup>25</sup>In fact, given our assumption that if indifferent, voters vote according to their ideological affiliation, at  $v = 1$  the Election-Winning policy is also unattainable.

<sup>26</sup>Notice that, on the other hand, condition (8) has to be satisfied.

Finally, notice that for  $v$  sufficiently close to 1, the Election-Winning policy coincides with the No-Compromise. In order to see this, consider  $w_\rho^{ew} = \frac{1-v}{2}$ . As  $v$  increases, this quantity decreases. Since at  $v = 1$  it holds that  $w_\rho^{ew} = 0$ , there is some value of  $v$  such that  $\frac{1-v}{2} = 2w - 1$ . Solving this yields the threshold  $\bar{v} = 3 - 4w$ . For  $v \in [3 - 4w, 1)$ , the Election-Winning policy coincides with the No-Compromise, with the crucial difference that it allows the Democratic party to win the election for sure.

### Choice Between Repeal-Proof and No-Compromise

**Lemma A1.** *Suppose that  $v < 3 - 4w$ , so that the Election-Winning policy does not coincide with the No-Compromise, or that  $v \geq 1$ , so that voting follows partisan affiliations. Moreover, consider  $w \leq 1 - \frac{1}{4q}$ , so that the No-Compromise policy is not repeal-proof. Between the Repeal-Proof and the No-Compromise policy, the Democratic party prefers the former as long as condition (15) is satisfied.*

*Proof.* Remember that the decision of the Democratic party reflects the preferences of the  $\delta$  group. Therefore, the payoff from the No-Compromise policy is equal to  $\frac{1}{2}$ . The payoff from the Repeal-Proof policy, on the other hand, is equal to  $2w_\delta^{rp} - 1$ . Solving the inequality  $2w_\delta^{rp} - 1 \geq \frac{1}{2}$  yields that the Repeal-Proof policy is preferred as long as  $w_\delta^{rp} \geq \frac{3}{4}$ . Using expression (10) then yields the following condition:

$$w \geq \frac{3}{4} - \frac{1}{8q} \tag{15}$$

□

**Definition A1.** *We denote as Scenario 1 the parameter values satisfying condition (15). Analogously, we denote as Scenario 2 the parameter values such that  $\frac{1}{2} \leq w \leq \frac{3}{4} - \frac{1}{8q}$ .*

**Lemma A2.** *Consider the choice between the No-Compromise policy and the Election-Winning policy. The No-Compromise policy is preferred as long as  $v \leq \tilde{v}$ , where  $\tilde{v} = \frac{5}{2} - 4w$ .*

*Proof.* The payoff from the No-Compromise policy is  $\frac{1}{2}$ . The payoff from the Election-Winning policy is instead  $2\left(2w - \frac{1-v}{2}\right) - 1$ . Comparing the two gives the threshold  $\tilde{v} = \frac{5}{2} - 4w$ . For  $v \leq \tilde{v}$ , the Democratic party chooses the No-Compromise policy. Notice that  $\tilde{v} < \bar{v}$ . Moreover,  $\tilde{v} > \underline{v}$  if and only if (15) does not hold, that is as long as we are in Scenario 2.  $\square$

### **Proof of Proposition 1**

*Proof.* Consider Scenario 1 first. Following Lemma A1, the No-Compromise policy is never chosen, unless it coincides with the Election-Winning. Therefore, we are left with the choice between the Repeal-Proof and the Election-Winning. Since one of the requirements of the Election-Winning policy is to not be repeal-proof (see condition (13)), the Election-Winning policy, when it exists, is more beneficial than the Repeal-Proof for the  $\delta$  group. Given that neither the Repeal-Proof nor the Election-Winning get overturned in the second period, the Election-Winning dominates the Repeal-Proof, whenever it exists. Hence, in Scenario 1 the Election-Winning is chosen for  $v \in (\underline{v}, 1)$ , and the Repeal-Proof is chosen otherwise.

Consider now Scenario 2. Following Lemma A1, the Repeal-Proof policy is never chosen. Therefore, we are left with the choice between the No-Compromise and the Election-Winning. For  $v \geq 1$ , the election winning does not exist and therefore the reform chosen is the No-Compromise. For  $v < 1$ , instead, following Lemma A2, the No-Compromise policy is chosen for  $v \leq \tilde{v}$ , whereas the Election-Winning is chosen for  $v \in (\tilde{v}, 1)$ .

$\square$

## Welfare of group $\rho$ voters

When commenting on the meaning of the results summarized in Proposition 1, we consider  $w_\delta - w_\rho$  as an index of policy divisiveness: hence, the larger  $w_\rho$ , the lower the divisiveness of the policy designed by the Democratic party. Notice that a slightly different characterization would involve considering the welfare of the  $\rho$  group instead of  $w_\rho$ . When the policy outcome is the Repeal-Proof policy or the Election-Winning policy, these two measures of divisiveness coincide, since the policy is never repealed and the  $\rho$  group welfare is  $2w_\rho - 1$ , that is an affine transformation of  $w_\rho$ . However, when the No-Compromise policy is chosen, there is a  $\frac{1}{2}$  probability of repeal<sup>27</sup>. In that case, therefore, the expected welfare of  $\rho$  voters is  $\frac{1}{2}(2w_\rho - 1)$ , which accounts for the probability of achieving the status quo payoff of 0 following a repeal. Therefore, comparing welfare under the Election-Winning and the No-Compromise policy is not the same as comparing the respective values of  $w_\rho$ .

As a consequence of this, notice that for values of  $v$  such that the Election-Winning policy is close to the No-Compromise one, going from the No-Compromise policy to the Election-Winning policy decreases the welfare of  $\rho$  voters, despite the higher  $w_\rho$  under the Election-Winning policy. The reason is that whereas under the No-Compromise policy there is a probability of repeal, under the Election-Winning policy this is not the case: hence, the  $\rho$  group is stuck with a policy which is better than the No-Compromise one, but still worse than the status quo payoff.

Nevertheless, there always exists an interval of values of  $v$  such that the  $\rho$  group has a higher welfare under the Election-Winning policy compared to the No-Compromise policy. In particular, comparing the payoff for  $\rho$  voters from the No-Compromise policy, which is equal to  $\frac{1}{2}[2(2w - 1) - 1]$ , and the payoff from the Election-Winning policy, which is equal to

---

<sup>27</sup>Unless of course the No-Compromise and Election-Winning policy coincide.

$2^{\frac{1-v}{2}} - 1 = -v$ , we obtain the following threshold:

$$v \leq \frac{3 - 4w}{2} \quad (16)$$

Imposing  $\frac{3-4w}{2} \geq \underline{v}$  we then find the following condition:

$$w \leq \frac{q}{2} + \frac{1}{4} \quad (17)$$

It can be checked that this quantity is larger than  $\frac{3}{4} - \frac{1}{8q}$ , implying that in Scenario 2, condition (17) is always satisfied. This means that in Scenario 2, there always exists a range of values of  $v$  such that switching from the No-Compromise policy to the Election-Winning policy is welfare improving. Intuitively, these values of  $v$  are close to  $\underline{v}$ , such that the Election-Winning policy is close to the Repeal-Proof policy.

## Proof of Proposition 2

*Proof.* The payoff from the No-Compromise policy is still equal to  $\frac{1}{2}$ , and also the payoff from the Repeal-Proof policy is unchanged at  $2w_{\delta}^{rp} - 1$ . Therefore, the results of Lemma A1 still hold. Concerning the Election-Winning policy, instead, the probability for the Democratic party to win is now equal to  $\frac{1+(1-q)}{2} \in (\frac{1}{2}, \frac{3}{4})$ . Hence, the payoff from the Election-Winning policy is  $\frac{2-q}{2} [2(2w - \frac{1-v}{2}) - 1]$ .

For Scenario 1, comparing the payoffs from the Repeal-Proof and the Election-Winning policy yields that the Repeal-Proof policy is preferred for:

$$v \leq \frac{2(2w - 1)(2q^2 - 3q + 2)}{(2 - q)(2q - 1)} \equiv \underline{v}_{prop} \quad (18)$$

and it can be checked that  $\underline{v}_{prop} > \underline{v}$  since  $\underline{v}_{prop} = \underline{v} \frac{2q^2 - 3q + 2}{(2 - q)(1 - q)}$ , where  $\frac{2q^2 - 3q + 2}{(2 - q)(1 - q)} > 1$ .

For Scenario 2, comparing the payoffs from the No-Compromise and the Election-Winning policy yields that the No-Compromise policy is preferred for:

$$v \leq \frac{5 - 2q}{2 - q} - 4w \equiv \tilde{v}_{prop} \quad (19)$$

Notice that since  $\frac{5-2q}{2-q} > \frac{5}{2}$ , we have that  $\tilde{v}_{prop} > \tilde{v} = \frac{5}{2} - 4w$ . □

### Proof of Proposition 3

*Proof.* Consider Scenario 1 first and consider the results of Proposition 1. Since the Election-Winning policy is always chosen whenever it is available, the Democratic party already maximizes the probability of winning the election. Therefore, introducing an additional office rents term  $r$  in the objective of the reform designer changes nothing.

Consider now Scenario 2. Here, the Democratic party sometimes chooses the No-Compromise policy despite the availability of the Election-Winning one. Given office motivation, the Democratic party prefers the No-Compromise over the Election-Winning reform if:

$$\frac{1 + r}{2} > 2w - \frac{1 - v}{2} + r \quad (20)$$

which can be rearranged to:

$$v < \frac{5}{2} - 4w - \frac{r}{2} \equiv \tilde{v}_{off} \quad (21)$$

It is immediate to see that  $\tilde{v}_{off} < \tilde{v}$ . Moreover, by solving for  $\tilde{v}_{off} < \underline{v}$  we obtain:

$$r > \frac{2q(3 - 4w) - 1}{2q - 1} \quad (22)$$

By condition (15),  $\frac{2q(3-4w)-1}{2q-1} > 0$  as long as we are in Scenario 2. Therefore, for  $r > \bar{r} \equiv \frac{2q(3-4w)-1}{2q-1}$ ,  $\tilde{v}_{off} = \underline{v}$ , i.e. the Election-Winning reform is always chosen if available. □

## Proof of Proposition 4

*Proof.* Consider a reform designer who can choose the value of  $w$  without being able to commit to a given reform design.

In Scenario 2, increasing  $w$  increases  $w_\delta^{ew}$  and shifts  $v_{off}$  downwards. Hence, increasing  $w$  within Scenario 2 weakly increases the chances of the Democratic party to win the election, as well as the share of winners in the  $\delta$  group (in fact, of both groups). As a result, there are no distortions in the choice of  $w$  by the reform designer.

Consider now Scenario 1. Increasing  $w$  decreases  $w_\rho^{rp}$ , while  $w_\rho^{ew} = \frac{1-v}{2}$  does not depend on  $w$ . Therefore, the threshold  $\underline{v}$  increases in  $w$ . This means that by increasing  $w$ , the optimal reform might switch from being the Election-Winning to being the Repeal-Proof. In particular, consider the choice between  $w$  and  $w'$ , where  $w' > w$ : if  $\underline{v}(w) < v < \underline{v}(w')$ , then under  $w$  the policy chosen is the Election-Winning, whereas under  $w'$  it is the Repeal-Proof. For small changes in  $w$  relative to the value of  $r$ , therefore, an office motivated reform designer prefers the less efficient reform. Mathematically, the reform designer prefers the less efficient reform as long as  $w' - w = dw < \frac{2q-1}{2q} \frac{r}{2}$ .

Notice that for sufficiently high  $r$ , a reform designer who can choose any  $w \in [0, 1]$  will choose the highest  $w$  such that the policy design is Election-Winning. Given  $v$ , this is the value of  $w$  denoted by  $w^*$  such that  $\underline{v}(w^*) = v$ . □