Dynamic Collective Choice with Endogenous Status Quo*

Wioletta Dziuda† Antoine Loeper‡(Job market paper)

November 14, 2010

Abstract

This paper analyzes an ongoing bargaining situation in which i) preferences evolve over time, ii) the interests of individuals are not perfectly aligned, and iii) the previous agreement becomes the next status quo and determines the payoffs until a new agreement is reached. We show that the endogeneity of the status quo exacerbates the players' conflict of interest and decreases the responsiveness of the bargaining outcome to the environment. Players with arbitrarily similar preferences can behave as if their interests were highly discordant. When players become very patient, the endogeneity of the status quo can bring the negotiations to a complete gridlock.

Under mild regularity conditions, fixing the status quo throughout the game via an automatic sunset provision improves welfare. The detrimental effect of the endogeneity of the status quo can also be mitigated by concentrating decision rights, for instance, by lowering the supermajority requirement.

JEL Classification Numbers: C73, D72, D78

Keywords: Dynamic voting, endogenous status quo, supermajority, partisanship, polarization, policy inertia, sunset provision.

*Previously circulated under the title "Ongoing Negotiation with an Endogenous Status Quo." We thank David Austen-Smith, David Baron, Daniel Diermeier, Bard Harstad, and the participants of the seminars at Bonn University, Leuven University, Northwestern University, Nottingham University, Rice University, University of Chicago, and Paris School of Economics.

†Kellogg School of Management, Northwestern University. Email: wdziuda@northwestern.edu

‡Kellogg School of Management, Northwestern University. Email: a-loeper@northwestern.edu

1 Introduction

This paper analyzes an ongoing collective decision problem in which i) there are shocks to the environment that affect individual preferences, and hence call for renegotiation of the past agreement, ii) the interests of individuals are not perfectly aligned, and iii) the previous agreement becomes the next status quo and determines the payoffs until a new agreement is reached.

Negotiations in a changing environment with an endogenous default are at the center of many economically relevant situations, one prominent example being legislative bargaining. Legislators' preferences reflect heterogeneous ideologies (or constituencies) and shocks, such as business cycles, technological improvements, or the vagaries of public opinion. For example, during a recession, generous public spending may be favored by all parties to stimulate short-term economic growth and employment. Conversely, when facing investors' skepticism, all parties may prefer to curb spending to convince the market that the public debt is under control. In normal times, however, legislators may genuinely disagree on these issues. Hence, a changing environment calls for renegotiations. At the same time, many budgetary policies are continuing in nature: the enacted policy continues in effect unless further legislative action is taken.[1] Recognizing this, fiscal conservatives may be reluctant to increase public spending during a recession, fearing that their liberal counterparts will oppose a return to fiscal discipline when the economy improves.

This fundamental trade-off between responding to the current environment and securing a favorable bargaining position for the future plays a key role in many dynamic bargaining settings. In the legislative sphere, besides fiscal policy over the business cycle, many ideologically charged issues such as taxation, immigration, minimum wage, pensions, or civil liberties are also subject to recurring shocks (e.g., demographic transitions, inflation pressures, or terrorist threats) and the policies that address them are typically continuing in nature.[2] Likewise, prior agreements determine the default option in a host of non legisla-

[1] In the U.S., the budget process distinguishes between two expenditure categories: discretionary spending and direct spending. The former requires an annual appropriation bill while the latter is continuing in nature. Direct spending consists mainly of entitlement programs such as Social Security benefits, Medicare and Medicaid. It has been the larger of the two categories since the 1990s. See Weaver (1985, 1988), Hird (1991) and Lowi (1969) on the ongoing nature of public policies in the U.S.

[2] Sunset provisions are rather the exception than the norm, as laws and regulations are meant to be permanent. They are somewhat more frequently used in taxation. For instance, the U.S. Earned Income Tax Credit and its subsequent expansions in 1986, 1990, 1993, and 2001 did not have a sunset provision, but the "Vietnam tax surcharge" of The Revenue and Expenditure Control Act of 1968 had a two-year sunset clause. More recently, the "Bush tax cuts" of the Economic Growth and Tax Relief Reconciliation Act of 2001 and the Jobs and Growth Tax Relief Reconciliation Act of 2003 had a ten-year sunset clause. Another prominent exception is the U.K. income tax which is repealed and voted on again every year.

tive bargaining contexts such as monetary policy,[3] labor contracts, financial contracts, and international treaties (such as those of the WTO, or the GATT).

Despite its pervasiveness, the institution of the endogenous status quo in a changing environment has received little attention in the literature. This is likely due to the complexity of the dynamics of these games. This paper builds an analytically tractable model which isolates the incentives generated by the endogeneity of the status quo in a transparent way. Despite its simplicity, the model generates rich dynamics but delivers clear results on the responsiveness of the bargaining outcome to the environment. It also provides a tractable framework to study the efficiency of alternative legislative rules such as sunset provisions.

In the basic model, two players engage in a finite sequence of collective choices over two alternatives. In each period, the environment changes and affects players' preferences. At the beginning of each period one alternative—called the current status quo—is in place. If both players agree to move away from the status quo, the new alternative is implemented. Otherwise, the status quo remains in place. In both cases, the implemented alternative determines the players' payoffs in this period and becomes the status quo in the next period. The preferences are independently and identically distributed over time, but can be correlated across players. Players are forward-looking and discount the future.

The endogeneity of the status quo introduces a *dynamic linkage* between otherwise independent bargaining periods, and this distorts players' voting behavior. In the unique equilibrium, in each period a player votes for a given alternative if and only if her current payoff difference between this alternative and the other one exceeds a certain, time-dependent threshold. Since each player is willing to sacrifice her current payoff to secure a favorable status quo, the voting thresholds are typically different from zero. This means that each player's vote is biased in favor of one alternative.

We show that this bias is given by the expected preferences *conditional on disagreement*. This means that the bias of a given player does not depend solely on her preferences, but on how her preferences conflict with those of her opponent. The players' willingness to ignore their actual preference realizations and follow instead interests defined in opposition to the rival reminds us of what is commonly referred to as *partisanship*, and we use this term to denote the voting threshold of each player.

The central finding of this paper is that by inducing partisanship, the endogenous status quo exacerbates the players' conflict of interest and decreases the responsiveness of the bargaining outcome to the environment. We show that players with arbitrarily similar preference

[3] In the U.S., the interest rates are negotiated within the Federal Open Market Committee and remain in place until the committee agrees to change them according to its internal voting rule. See Riboni and Ruge-Murcia (2008) for more on the role of the status quo in monetary policy institutions.

distributions can behave as if their interests were highly discordant. Moreover, if players are patient enough, the endogeneity of the status quo can bring the negotiations to a complete gridlock in which the enacted policy is completely unresponsive to the environment.

Our results extend to an $N-$player game in which i) individual preferences are parameterized by a constant, heterogeneous ideological type, ii) preferences are subject to a common shock, and iii) the status quo stays in place unless at least M players vote for the other alternative. As in the two-player case, the endogeneity of the status quo makes all players partisan. In addition, in this setting partisanship unambiguously decreases the probability that M players agree, increasing thereby status quo inertia. Partisanship and status quo inertia increase with patience, the bargaining horizon, and ideological polarization. Partisanship also increases with the supermajority requirement M. This last finding means that a higher supermajority increases status quo inertia not only because more players need to agree, but also because it exacerbates the players' conflict of interests.

Partisanship results in the Pareto worse alternative being implemented with positive probability in every period, which certainly has a detrimental impact on welfare. However, by bundling the players' vote on today's policy and tomorrow's status quo, the endogenous status quo prompts the players to express the intensity of their preference realizations, which could be socially beneficial. To assess the welfare effect of partisanship, we compare our game to the repetition of a static bargaining protocol in which one alternative is exogenously designated as the status quo. Under this protocol, players vote only on the basis of their current preferences in each period. Our model delivers a somewhat negative result: under mild regularity conditions, an exogenous status quo achieves a higher level of social welfare than an endogenous one.

In legislative decision making, designating one policy as a fixed default can be interpreted as an automatic sunset provision. Sunset provisions have usually been advocated to improve parliamentary control of executive agencies, to evaluate the efficiency of new laws, or to impose a time limit on temporary measures. The rationale advanced by this paper has a more strategic flavor: sunset provisions sever the link between today's agreement and tomorrow's status quo, which mitigates polarization and makes policies more responsive to the environment.

Partisanship can also be mitigated by concentrating decision rights, because if fewer players have to agree, the status quo is less likely to be a binding constraint. In the two-player case, this can be done by vesting one player with dictatorial power. We show that in some cases, partisanship is so detrimental that both players are willing to give up their veto power and let their opponent be the dictator. In the $N-$player extension of the model, if preferences are symmetrically distributed along the ideological spectrum, decreasing the supermajority requirement is socially beneficial.

The paper is organized as follows. Section 2 describes the model. Section 3 derives the equilibrium. In section 4, we characterize the dynamics of partisanship. Section 5 extends the model to N players. Section 6 contains the welfare analysis and discusses the implication of our model for sunset provisions and concentrating decision power. Section 7 introduces intertemporal correlation of preferences and allows shocks to differ in their persistency. Section 8 concludes. All proofs not provided in the main text can be found in the appendix.

1.1 Related Literature

The formal literature on ongoing legislation with endogenous status quo started with the seminal paper of Baron (1996).[4] His model has been extended to various multidimensional settings by Baron and Herron (2003), Kalandrakis (2004, 2007), Cho (2005), Fong (2006), Bernheim et al. (2006), Diermeier and Fong (2007a), Baron, Diermeier, and Fong (2007), and Battaglini and Palfrey (2007).[5] These models, however, consider static environments: policies evolve over time not because preferences change, but because the set of actions available to each player varies across voting stages. They focus on the dynamics of the proposal power under different institutional rules. We abstract away from the distributional issue of the proposal power and focus instead on the efficiency of the policy-making process and its responsiveness to economic and political shocks.[6]

Battaglini and Coate (2007, 2008) study the inefficiency of a dynamic legislative bargaining model of public finance. In their papers, the status quo is fixed and the dynamic linkage is the accumulation of the public good or debt, which affects the relative returns of pork-barrel programs. In their model, the availability of targeted public spending leads legislatures to pass inefficient budgets and be present-biased, more so the lower the supermajority requirement, while in our model, the continuing nature of policies lead voters to be future-biased, more so the larger the supermajority requirement.

Even though dynamic bargaining with an endogenous status quo and evolving preferences is at the center of many economically relevant situations, the existing literature on this topic is scarce. This may be a consequence of the relative intractability of these games.

[4] Epple and Riordan (1987) study a similar model but consider nonstationary equilibria. The principle of an evolving status quo was first introduced in a cooperative bargaining literature by Kalai (1977).

[5] The models of Bernheim et al. (2006) and Diermeier and Fong (2007a) are originally cast in a single policy period, but they can be extended or interpreted as dynamic legislative bargaining games.

[6] Because most of these models consider the division of a pie of exogenous size or single-peaked preferences, equilibrium outcomes are always efficient in a static sense and can be inefficient in a dynamic sense only when citizens are sufficiently risk-averse. In contrast, when preferences vary as in our model, equilibrium outcomes are typically Pareto inefficient independently of risk aversion.

As Romer and Rosenthal (1978) showed in a static setup with single-peaked preferences, the induced preferences over the status quo are typically not convex, which makes the multiperiod extension technically hard to analyze. With a continuum of alternatives and an infinite horizon, Markov equilibrium existence is not guaranteed even under standard preference specifications.[7] To the best of our knowledge, only Diermeier and Fong (2007b), Riboni and Ruge-Murcia (2008) and Duggan and Kalandrakis (2009) make progress on this front. Adding noise to the status quo, Duggan and Kalandrakis (2009) establish the existence of an equilibrium. The generality of their model does not allow an analytical equilibrium characterization, so they resort instead to numerical methods. Riboni and Ruge-Murcia (2008) analyze a game with quadratic utility functions and a finite state space. They analytically solve a two-period two-state example, but use numerical solutions for the general model. Diermeier and Fong (2007b) analyze a two-period three-state model with a richer institutional framework. Our paper differs from these contributions in that we simplify the space of alternatives, but fully characterize the policy dynamics for any preference distributions and any bargaining horizon. Our institutionally sparse model allows us to isolate the effect of the endogeneity of the status quo in a transparent way.

Montagnes (2010) looks at a two-period financial contracting environment in which the current contract serves as the default option in future negotiations. He shows that both contracting parties may prefer to commit ex ante to ceding a future decision power to one of them. Such a commitment breaks the dynamic linkage and avoids inefficiencies in the initial contract.

Fernandez and Rodrik (1991) and Alesina and Drazen (1991) have emphasized that the distributional uncertainty of policy reforms can lead to status quo inertia. In our model, status quo inertia would also arise in an environment without uncertainty but with evolving preferences.

Our results on policy responsiveness are related to the political economy literature on growth and on the dynamics of welfare policies (Glomm and Ravikumar 1995; Krussell and Rios-Rull 1996,1999, Coate and Morris 1999; Saint Paul and Verdier 1997; Benabou 2000; Saint Paul 2001; Hassler et al. 2003, 2005). These models emphasize the effect of the current policy on private investment decisions, which in turn affect the policy preferences of voters in future periods and thus generate policy persistence. In contrast, in our paper, the current policy does not affect future preferences, but inertia emerges because today's policy affect future bargaining positions.

Finally, Casella (2005) shows that linking voting decisions across time allows voters to express their preference intensity, which can be socially beneficial. Our results suggests

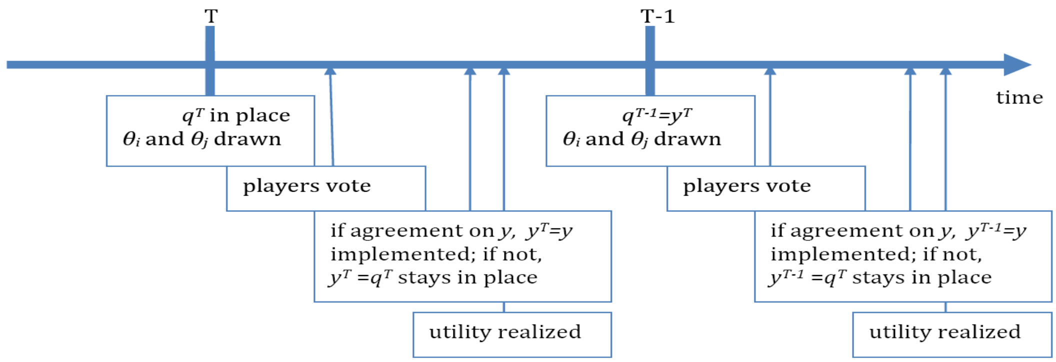
[7] See, e.g., Kalandrakis (2004b, 2007) or Duggan and Kalandrakis (2009) for more on this issue.

that the intertemporal trade-off induced by the endogeneity of the status quo, despite the pervasiveness of this institution, is not an efficient way to elicit preference intensity. Barbera and Jackson (2010) let ex ante identical voters choose the group decision rule after having learned their first period preferences. As in our framework, bundling the current and the future decision rules generate inefficiencies. But since the dynamic linkage is only between the first and the subsequent periods, sufficiently patient players always select the optimal voting rule.

2 The Basic Model

Two players, i and j, are in a relationship that lasts for T periods. For notational simplicity, we adopt the convention that $t = 1$ denotes the last period. In other words, t measures the number of periods remaining in the game. In each period t, players adopt one of two alternatives, $y^t \in \{-1, 1\}$. The utility of player $k \in \{i, j\}$ in period t depends on the alternative adopted in period t only, and is given by $u\left(\theta_k^t, y^t\right) = \theta_k^t y^t$. We refer to θ_k^t as player k's *current preference*. Players discount future payoffs with the same factor $\delta \in (0, 1)$. Throughout the paper, for any individual parameter p_k, the bold symbol $\boldsymbol{p}$ will refer to the vector (p_i, p_j).

The game proceeds as follows. Each period starts with one alternative in place. We call this alternative *the status quo* in period t and denote it by q^t. At the beginning of each period, the preferences of both players $\boldsymbol{\theta}^t$ are drawn from a joint distribution which is integrable, i.i.d. over time, but the preferences can be arbitrarily correlated across players. We shall assume that the distribution of $\boldsymbol{\theta}^t$ has full support and admits a probability density function f with mean $\bar{\boldsymbol{\theta}}$. After players observe $\boldsymbol{\theta}^t$, they vote on which alternative to adopt in period t. If both players vote for the same alternative, this alternative is implemented. If they disagree, the status quo q^t stays in place. The implemented alternative y^t, be it the new agreement or the status quo q^t, determines the payoff in period t and becomes the status quo for the next period q^{t-1}. Hence, each period is an independent social choice problem but the endogeneity of the status quo introduces a strategic linkage between bargaining periods. We denote the game that lasts T periods and begins with status quo q by $\Gamma_{T,q}^{en}$. The following diagram summarizes the model.



To isolate the effect of the endogeneity of the status quo on equilibrium behavior and welfare, we shall compare $\Gamma^{en}_{T,q}$ to the game $\Gamma^{ex}_{T,q}$ in which the status quo is exogenously fixed at q in each of the T bargaining periods.

As usual with voting games, subgame perfection in the strict sense is not enough to eliminate pathological equilibria such as both players always voting for the status quo. Therefore, we restrict our attention to stage undominated equilibria (henceforth equilibria) as defined in Baron and Kalai (1993). This solution concept, standard in dynamic voting games, basically amounts to assuming that in every period, players cast their votes as if they were pivotal.

A few comments on the model are in order. First, we analyze a two-player game with a unanimity requirement, but we show in section 5 that our results extend to an $N-$player game with a (super) majoritarian approval rule. Second, the stationarity of the preference distribution is a simplifying assumption which is meant to capture the recurring nature of shocks (e.g., economic cycles, demographic transitions, public opinion swings, or national security threats) that affect issues such as taxation, public spending, immigration, or civil liberties. We relax this assumption in section 7. Third, restricting attention to two alternatives allows us to abstract away from the details of the stage game and the issue of proposal power.[8] It thereby allows us to isolate the effect of the endogeneity of the status quo on the efficiency and responsiveness of the bargaining outcomes to the environment in a transparent way. Fourth, what players know about each other's preferences is immaterial.[9]

Finally, we look at a $T-$period game, as opposed to $\Gamma^{en}_{2,q}$ or $\Gamma^{en}_{\infty,q}$ (where $\Gamma^{en}_{\infty,q}$ denotes the corresponding infinite-horizon game), for the following reasons. First, as we will see, the

[8] With two alternatives, many static bargaining protocols are equivalent. In particular, equilibrium outcomes are the same whether players vote simultaneously or sequentially, if they make take-it-or-leave-it offers, or if we allow for n rounds of bargaining within each period with either a random or alternating proposer.

[9] As we shall see, even if $\boldsymbol{\theta}^t$ is common knowledge at the beginning of each period t, the equilibrium strategy of each player depends only on her own preference realization.

equilibrium behavior in some period $T > 2$ may differ qualitatively from the equilibrium behavior in a two-stage game. Hence, restricting attention to $\Gamma^{en}_{2,q}$ may not be informative about players' behavior when the bargaining horizon is longer. Second, the equilibrium of $\Gamma^{en}_{T,q}$ as T approaches infinity may not be a Markov equilibrium of $\Gamma^{en}_{\infty,q}$. So restricting attention to the Markov equilibria of $\Gamma^{en}_{\infty,q}$ could miss interesting equilibrium dynamics. However, when the equilibrium converges, our results can be generalized to $\Gamma^{en}_{\infty,q}$, and we shall state the generalizations as we move along.

3 Equilibrium Analysis

3.1 The Equilibrium

Proposition 1 shows that the equilibria of the games $\left(\Gamma^{en}_{T,q}\right)_{T\in\mathbb{N},q\in\{-1,1\}}$ are characterized by a unique sequence of voting thresholds $\left(\mathbf{c}^t\right)_{t\geq 1}$.

Proposition 1 $\Gamma^{en}_{T,q}$ *admits a unique (up to a zero-measure subset of preference realizations) equilibrium which is characterized by a sequence of thresholds* $\left(\mathbf{c}^t\right)_{t=1}^{T}$. *In period* t, *player* $k \in \{i, j\}$ *votes for* 1 *if* $\theta^t_k > c^t_k$ *and for* -1 *if* $\theta^t_k < c^t_k$. *For all* t, $\mathbf{c}^t$ *is independent of* T *and* q, *and is defined recursively by* $\boldsymbol{c}^1 = (0, 0)$ *and by* $\boldsymbol{c}^{t+1} = \boldsymbol{H}\left(\boldsymbol{c}^t\right)$ *where*

$$H_k\left(\boldsymbol{c}\right) = \delta\left(\int_{-\infty}^{c_j}\int_{c_i}^{\infty}\left(c_k - \theta_k\right) f\left(\boldsymbol{\theta}\right) d\theta_i d\theta_j + \int_{c_j}^{\infty}\int_{-\infty}^{c_i}\left(c_k - \theta_k\right) f\left(\boldsymbol{\theta}\right) d\theta_i d\theta_j\right). \quad (1)$$

The proof of proposition 1 is relatively simple and helps build intuition for the rest of the analysis. Observe first that the policy implemented in the last period $t = 1$ determines only the current payoff in that period. Therefore, in that period each player votes according to her current preferences, so $\boldsymbol{c}^1 = (0, 0)$.[10]

Suppose now that the proposition holds for all $T \leq T'$. For all $t \leq T'+1$, the continuation game after period t's outcome y^t is simply Γ^{en}_{t-1,y^t}. From the induction hypothesis, this game has a unique equilibrium, and we denote its value for player k by $V^{t-1}_k\left(y^t\right)$. Since player $k \in \{i, j\}$ votes as if she were pivotal, in every period $t \leq T' + 1$, she votes for 1 if

$$\theta^t_k + \delta V^{t-1}_k(1) > -\theta^t_k + \delta V^{t-1}_k(-1) \Leftrightarrow \theta^t_k > \frac{\delta}{2}\left(V^{t-1}_k(-1) - V^{t-1}_k(1)\right),$$

[10]The uniqueness of the thresholds comes from the full support assumption.

and for -1 if the reverse inequality holds. Therefore, she uses a threshold strategy with the threshold being

$$c_k^t = \frac{\delta}{2}\left(V_k^{t-1}(-1) - V_k^{t-1}(1)\right). \tag{2}$$

This shows that $\Gamma_{T'+1,q}^{en}$ has a unique equilibrium. Observe from (2) that the sign of c_k^t determines whether player k prefers the next period's status quo to be 1 or -1, and its absolute value measures the intensity of this preference.

The following observation is the key step to understanding the function $\boldsymbol{H}$ in (1): the status quo in a given period t matters only when players vote for opposite alternatives. From what precedes, this happens when players' preferences θ_i^t and θ_j^t are on opposite sides of their respective thresholds c_i^t and c_j^t. Therefore, from (2) we have

$$c_k^{T'+1} = \frac{\delta}{2}\left(\begin{array}{c} \int_{-\infty}^{c_j^{T'}}\int_{c_i^{T'}}^{\infty}\left(-\theta_k + \delta V_k^{T'-1}(-1) - \left(\theta_k + \delta V_k^{T'-1}(1)\right)\right) f(\boldsymbol{\theta})\, d\theta_i d\theta_j \\ + \int_{c_j^{T'}}^{\infty}\int_{-\infty}^{c_i^{T'}}\left(-\theta_k + \delta V_k^{T'-1}(-1) - \left(\theta_k + \delta V_k^{T'-1}(1)\right)\right) f(\boldsymbol{\theta})\, d\theta_i d\theta_j \end{array}\right).$$

Substituting (2) inside the integral, we get the recursive relation $\boldsymbol{c}^{T'+1} = \boldsymbol{H}\left(\boldsymbol{c}^{T'}\right)$, which proves proposition 1. The above expression implies that the voting behavior of each player is given by her expected intertemporal preferences (i.e., $\theta_k^t - c_k^t$) in the next period *conditional on disagreement* (i.e. $\theta_i^t - c_i^t$ and $\theta_j^t - c_j^t$ of opposite sign).

3.2 Partisanship

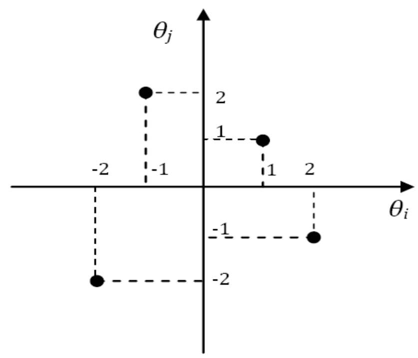
The following is a straightforward but important corollary of proposition 1:

Corollary 1 *In equilibrium, players use nonzero thresholds (i.e., $c_i^t = c_j^t = 0$ for all t) if and only if*

$$\int_0^{\infty}\int_{-\infty}^{0} \boldsymbol{\theta} f(\boldsymbol{\theta})\, d\theta_i d\theta_j + \int_{-\infty}^{0}\int_0^{\infty} \boldsymbol{\theta} f(\boldsymbol{\theta})\, d\theta_i d\theta_j = (0,0). \tag{3}$$

Condition (3) shows that players use nonzero thresholds if, conditional on disagreement, at least one player is ex-ante not indifferent between the two alternatives. Condition (3) is satisfied only in special cases. For instance, if the distribution of $\boldsymbol{\theta}$ is bivariate normal, condition (3) holds if and only if $\bar{\boldsymbol{\theta}} = (0,0)$ (see figure in section 4). The following example shows that it can be violated even if the preference distribution is symmetric (i.e., for all (θ_i, θ_j), $f(\theta_i, \theta_j) = f(\theta_j, \theta_i)$) and the marginal distribution of the preference of each player is symmetric around zero.

Example 1 *Each of the four points in the figure below represents the mean of a smooth and symmetric density function that has full support and a vanishing variance. Agents'* *preferences are drawn from each possible joint distribution with equal probability. Observe that the marginal distribution for each player is symmetric around zero; hence, ex ante, each player is indifferent between the two alternatives. However, conditional on the current preferences disagreeing, that is, conditional on θ_i and θ_j being of opposite sign, each player is equally likely to prefer 1 and -1, but her preferences for 1 are more intense.*[11]



Corollary 1 shows that the voting thresholds are typically different from zero, which means that players are willing to sacrifice their current payoff to secure a better bargaining position in the future. It is instructive to compare this situation with the game with an exogenous status quo $\Gamma^{ex}_{T,q}$. With the exogenous status quo, each period can be considered in isolation; so in the unique equilibrium of $\Gamma^{ex}_{T,q}$ the players use constant threshold strategies $(0,0)$. Hence, by comparing $\Gamma^{en}_{T,q}$ with $\Gamma^{ex}_{T,q}$, we see that the endogeneity of the status quo introduces a bias in the equilibrium behavior: each player k in $\Gamma^{en}_{T,q}$ behaves as if the status quo was exogenous but her payoff θ^t_k was shifted by $-c^t_k$.

As proposition 1 shows, these voting biases are determined by the players' expected preference conditional on disagreement. This implies that the bias of a given player does not depend solely on her preferences, but on how her preferences conflict with those of her opponent. The players' willingness to ignore their actual preference realizations and follow instead interests defined in opposition to the rival reminds us of what is commonly referred to as *partisanship.* This resemblance will be emphasized in section 5, in which we show that when players are ranked on an ideological spectrum, the direction of the bias of a given player depends only on her relative ideological position. Therefore, in the sequel we use the term partisanship to refer to the voting threshold of each player.

Definition *The partisanship of player $k \in \{i,j\}$ in period t is c^t_k.*

[11] To see this, observe that conditional on θ_i and θ_j having opposite sign, the players' type are drawn from the "top/left" distribution with mean $(-1,2)$ or the "bottom/right" distribution with mean $(2,-1)$.

Because of partisanship, Pareto dominated alternatives are implemented with positive probability. For example, if $c_i^t < 0$, then with positive probability $\theta_j^t < 0$ and $c_i^t < \theta_i^t < 0$, in which case player i vetoes the Pareto optimal alternative -1 when $q^t = 1$. If additionally $c_j^t < 0$, then with positive probability, $c_j^t < \theta_j^t < 0$, and the Pareto dominated alternative $y^t = 1$ garners unanimous approval.

4 The dynamics of partisanship

Proposition 1 implies that partisanship evolves over time. In this section, we analyze the dynamics of partisanship and its limit $\lim_{T\to\infty} \mathbf{c}^T$ as the bargaining horizon T increases, which we denote by $\boldsymbol{c}^\infty$ when it exists.

Before we proceed, let us establish some notation. First, let $(\leq, \geq)$ denote the partial order on $\mathbb{R}^2$ defined by $\boldsymbol{c}(\leq, \geq)\boldsymbol{c}'$ if $c_i \leq c_i'$ and $c_j \geq c_j'$; $(<, >)$ and $(\geq, \leq)$ are defined similarly. Second, it will prove convenient to partition of the set of preference distributions as follows.

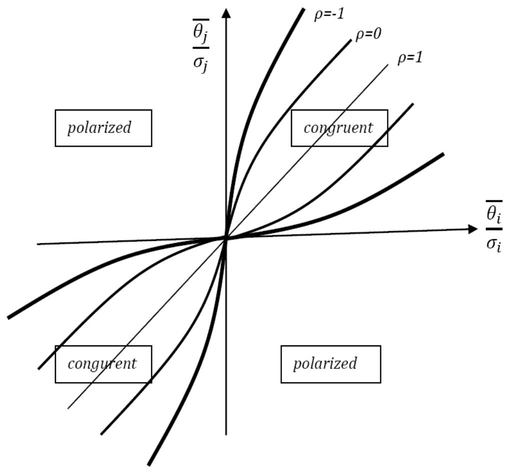
Definition *A preference distribution is congruent if $H_i(0,0)$ and $H_j(0,0)$ are of the same sign. It is polarized if they are of opposite sign.*

From the definition of $\boldsymbol{H}$ (see (1)), a preference distribution is congruent if, conditional on disagreement, each player's expected preference favors the same alternative. Otherwise, it is polarized. Stated differently, if players had the choice between playing the game $\Gamma_{T,q}^{ex}$ with an exogenous status quo $q = 1$ or $q = -1$, congruent players would have the same preferences over q, while polarized player would disagree.

The following figure illustrates this partition for normal bivariate distributions for $\rho = -1$ (the thickest curve), $\rho = 0$, and $\rho = 1$ (the thinnest curve).[12] Observe that for a given correlation, the preference distribution is congruent only if the means are sufficiently close to each other. In particular, the preference distribution is always polarized when the means are of opposite sign, or the preferences are perfectly correlated. To see the latter, suppose for simplicity that $\sigma_i = \sigma_j = 1$ and $\bar{\theta}_i > \bar{\theta}_j$. As ρ approaches 1, then $\theta_i > \theta_j$ almost surely. Therefore, conditional on disagreement, player i prefers alternative 1 while player j prefers -1. This shows in particular that players with very similar preference distributions can be

[12] This partition is formally characterized in example 3 in the appendix.

polarized.



4.1 Polarized preferences

From proposition 1, $\boldsymbol{c}^2 = \boldsymbol{H}(0,0)$, so if the preference distribution is polarized, in the penultimate period of $\Gamma_{T,q}^{en}$ players are partisan for different alternatives. The following proposition shows that in this case, players will remain biased in favor of different alternatives in all periods. Moreover, their degree of partisanship increases with the bargaining horizon.

Proposition 2 *If the preference distribution is polarized with* $\boldsymbol{H}(0,0)(\leq,\geq)(0,0)$*, then*

- *a) for all* t*,* $c_i^{t+1} < c_i^t < 0 < c_j^t < c_j^{t+1}$*;*
- *b)* $\boldsymbol{c}^\infty$ *exists and is the least fixed point of* $\boldsymbol{H}$ *on* $\mathbb{R}_- \times \mathbb{R}+$ *for the order* $(\leq,\geq)$*.*

Proof. From lemma 1 (see the appendix), $\frac{\partial H_i}{\partial c_i} \geq 0$ and $\frac{\partial H_i}{\partial c_j} \leq 0$, so $\boldsymbol{H}(\boldsymbol{c})$ is monotone in the order $(\leq,\geq)$. Since $\boldsymbol{H}(0,0)(\leq,\geq)(0,0)$, it follows that $\boldsymbol{H}(\mathbb{R}_- \times \mathbb{R}_+) \subset \mathbb{R}_- \times \mathbb{R}_+$. From (1), for all $k \in \{i,j\}$,

$$|H_k(\boldsymbol{c})| \leq \delta\left(\int_0^\infty \int_{-\infty}^0 (|\theta_k - c_k|)\, f(\boldsymbol{\theta})\, d\,\boldsymbol{\theta} + \int_{-\infty}^0 \int_0^\infty (|\theta_k - c_k|)\, f(\boldsymbol{\theta})\, d\boldsymbol{\theta}\right) \leq \delta\left(E(|\theta_k|) + |c_k|\right).$$

This implies that $\left|H_k\left(-\frac{\delta E(|\theta_i|)}{1-\delta}, \frac{\delta E(|\theta_j|)}{1-\delta}\right)\right| \leq \frac{\delta E(|\theta_k|)}{1-\delta}$. Hence, we have shown that $\boldsymbol{H}(L) \subset L$,where $L = \left[-\frac{E(|\theta_i|)}{1-\delta}, 0\right] \times \left[0, \frac{E(|\theta_j|)}{1-\delta}\right]$. Since L is a complete lattice for $(\leq,\geq)$, Tarski's theorem implies that $\boldsymbol{H}$ has a least fixed point $\boldsymbol{c}^f$ on L (and so on $\mathbb{R}_- \times \mathbb{R}_+$) in the order $(\leq,\geq)$.

Since $\boldsymbol{c}^1 = (0,0)$, it follows that $\boldsymbol{c}^f(\leq,\geq)\boldsymbol{c}^1$, and by monotonicity of $\boldsymbol{H}$, $\boldsymbol{c}^f(\leq,\geq)\boldsymbol{c}^2$. By assumption, $\boldsymbol{H}(0,0)(\leq,\geq)(0,0)$, so $\boldsymbol{c}^2(\leq,\geq)\boldsymbol{c}^1$. Hence, $\boldsymbol{c}^f(\leq,\geq)\boldsymbol{c}^{t+1}(\leq,\geq)\boldsymbol{c}^t(\leq,\geq)(0,0)$ holds for $t = 1$. Suppose that it holds for some $t \geq 1$; by applying $\boldsymbol{H}$ and using the fact that

$\boldsymbol{H}(0,0)(\leq,\geq)(0,0)$, the same inequalities hold for $t+1$. This shows by induction that $\boldsymbol{c}^t$ is increasing in t for the order $(\leq,\geq)$. Since $\boldsymbol{c}^t$ is bounded above by $\boldsymbol{c}^f$ in the order $(\leq,\geq)$, it has a limit, and the limit must be a fixed point of $\boldsymbol{H}$ no greater than $\boldsymbol{c}^f$ in the order $(\leq,\geq)$. Therefore, it must be $\boldsymbol{c}^f$. ■

Since partisanship measures the willingness of players to ignore their current preferences in order to secure a favorable status quo, proposition 2 implies that the responsiveness of the equilibrium behavior to the preference realizations decreases as the bargaining horizon increases. To understand the intuition for this result, consider the case in which player i prefers status quo 1 and player j prefers status quo -1 in the last period $t=1$. In the penultimate period $t=2$, the players' partisanship increases the probability that player i votes for 1 and player j votes for -1. Since such disagreement becomes more likely, securing each player's most preferred status quo becomes more important for $t=2$ than for $t=1$. The higher probability of such disagreement makes securing each player's most preferred status quo more important for $t=2$ than for $t=1$. As a result, players are more partisan in $t=3$. Hence, with polarized preferences, partisanship feeds on itself by increasing the probability of disagreement, and thus the importance of the status quo.

Technically, the proof of proposition 2 uses the monotonicity of H in the partial order $(\leq,\geq)$. This order structure, together with the recursive nature of the equilibrium, allows us to derive monotone comparative statics in the main primitives of the problem.

Proposition 3 *If the distribution of $\boldsymbol{\theta}$ is polarized with $\boldsymbol{H}(0,0)(\leq,\geq)(0,0)$, then*

a) partisanship increases with patience: for all t, $\mathbf{c}^t$ is increasing in δ in the order $(\leq,\geq)$;

b) partisanship increases with the polarization of preferences: For any $\boldsymbol{m}\in\mathbb{R}^2$, let $(\boldsymbol{c}^t(\boldsymbol{m}))_{t\geq 1}$ be the equilibrium voting thresholds for the preference distribution of $\boldsymbol{\theta}+\boldsymbol{m}$. Then $\mathbf{c}^t(\boldsymbol{m})$ is decreasing in $\boldsymbol{m}$ in the order $(\leq,\geq)$.

Proof. Part (a): If $\boldsymbol{H}(0,0)(\leq,\geq)(0,0)$ for some δ^o, the same inequality holds for all δ. Since $\boldsymbol{H}$ is monotonic in the order $(\leq,\geq)$, then for all $\boldsymbol{c}\in\mathbb{R}_-\times\mathbb{R}_+$, $\boldsymbol{H}(\boldsymbol{c})(\leq,\geq)(0,0)$ and from (1), $\frac{\partial H_i(\boldsymbol{c})}{\partial\delta}=\frac{H_i(\boldsymbol{c})}{\delta}\leq 0$ and $\frac{\partial H_j(\boldsymbol{c})}{\partial\delta}=\frac{H_j(\boldsymbol{c})}{\delta}\geq 0$. Lemma 2 (see the appendix) with $p=\delta$, $P=]0,1[$ and $C=\mathbb{R}_-\times\mathbb{R}_+$ completes the argument.

Part (b): Lemma 1 implies that for all $k\neq k'$, $\frac{\partial H_k(\boldsymbol{c})}{\partial m_k}\geq 0$ and $\frac{\partial H_k(\boldsymbol{c})}{\partial m_{k'}}\leq 0$. Lemma 2 for $p=\boldsymbol{m}$, $P=C=\mathbb{R}^2$, and $\succeq=(\leq,\geq)$ completes the argument. ■

The intuition for part (a) is that when players trade off the adequacy of the policy to the current environment versus securing a favorable status quo for tomorrow, more patient players put more weight on the latter and thus are more partisan. As for part (b), the preferences of more polarized players are more likely to disagree, which makes the status quo more important and thus increases partisanship. This, in turn, increases the likelihood of a disagreement. Hence, the endogeneity of the status quo magnifies the conflict of interest between players.[13]

The following proposition shows that when players become patient, the magnitude of their partisanship can lead to complete gridlock: if players are sufficiently polarized, even if their preferences agree with positive probability in every period, they almost always disagree in equilibrium.

Proposition 4 *For any $\boldsymbol{m} \in \mathbb{R}^2$, let $\left(\boldsymbol{c}^t\left(\boldsymbol{m}\right)\right)_{t\geq 1}$ be the equilibrium voting thresholds for the preference distribution of $\boldsymbol{\theta} + \boldsymbol{m}$. There exists $\boldsymbol{m}^o \in \mathbb{R}^2$ such that for all $\boldsymbol{m}\left(\geq, \leq\right)\boldsymbol{m}^o$, $\lim_{\delta\to 1} \boldsymbol{c}^\infty\left(\boldsymbol{m}\right) = \left(-\infty, +\infty\right)$.*

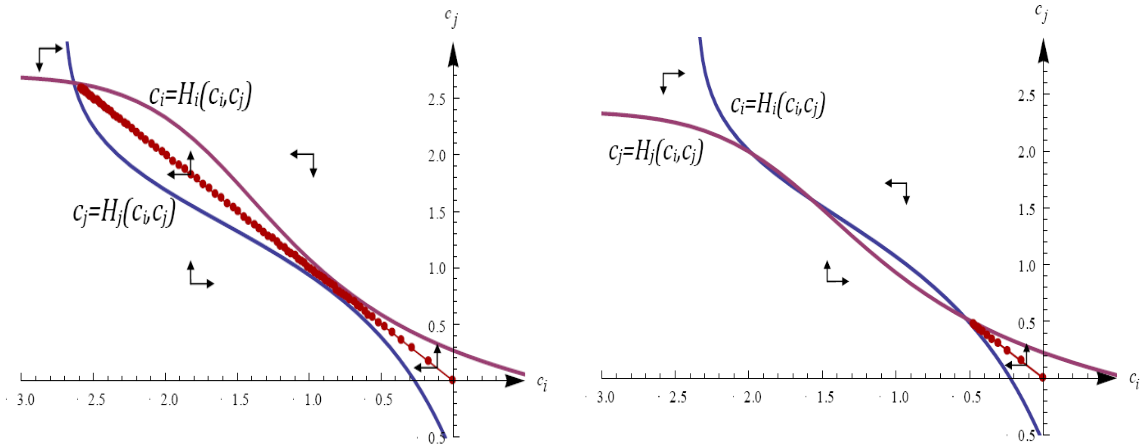
Observe that this result is not a mechanical consequence of increasing patience because today's status quo has direct consequences only on tomorrow's status quo, and today and tomorrow carry the same weight for patient players. What drives the completely unresponsive behavior of patient players is the vicious cycle in which patience increases partisanship, which itself increases the life expectancy of the status quo, which in turn increases the impact of patience on partisanship.[14]

We conclude this section with an example that illustrates the above propositions. The following figure represents the phase diagram of the function $\boldsymbol{H}$ and the equilibrium dynamics when the preference distribution is bivariate normal with mean $\bar{\boldsymbol{\theta}} = (\mu, -\mu)$ for two different values of μ. The dots in each panel represent the equilibrium sequence of thresholds. The arrows at any point $\boldsymbol{c}$ represent the direction of the vector $\boldsymbol{H}\left(\boldsymbol{c}\right) - \boldsymbol{c}$ and the two curves show the locus of the points $\boldsymbol{c}$ such that if players use voting thresholds $\boldsymbol{c}$ in period t, one player will use the same threshold in period $t+1$. The intersections of these two curves are the fixed points of $\boldsymbol{H}$. As shown in proposition 2, if the bargaining horizon is long enough, partisanship starts near the smallest fixed point of $\boldsymbol{H}$ in the northwest quadrant,

[13] It turns out, however, that partisanship is not monotone in the degree of correlation. On the one hand, more correlated preferences are less likely to disagree. On the other hand, as correlation increases, the nature of the disagreement is more predictable: conditionally on disagreement, one player almost always prefers -1, and the other player almost always prefers 1. Which effect dominates depends on the preference distribution.

[14] The results of this section extend to the infinite horizon game $\Gamma^{en}_{\infty,q}$ as follows: the set of Markov equilibria of $\Gamma^{en}_{\infty,q}$ is the set of constant threshold strategies $\boldsymbol{c}$ where $\boldsymbol{c}$ is a fixed point of $\boldsymbol{H}$. The comparative statics results in proposition 3 can be extended to the greatest and least Markov equilibria of $\Gamma^{en}_{\infty,q}$ in the order $(\leq, \geq)$. Indeed, as shown in the proof of propositions 2 and 3, $\boldsymbol{H}\left(\boldsymbol{c}\right)$ is monotonic in $\boldsymbol{c}$, δ and $\boldsymbol{m}$. Tarski's fixed point theorem and theorem 4 in Villas-Boas 1997 complete the argument. Moreover, in the proof of proposition 4, we show that for $\boldsymbol{m}\left(\leq, \geq\right)\boldsymbol{m}^o$, all fixed points of $\boldsymbol{H}^{\boldsymbol{m}}$ diverge to $(-\infty, +\infty)$ as $\delta \to 1$. This shows that for $\boldsymbol{m}\left(\leq, \geq\right)\boldsymbol{m}^o$, all Markov equilibria of $\Gamma^{en}_{\infty,q}$ diverge to $(-\infty, +\infty)$ as $\delta \to 1$.

and decreases towards the origin as the game progresses.



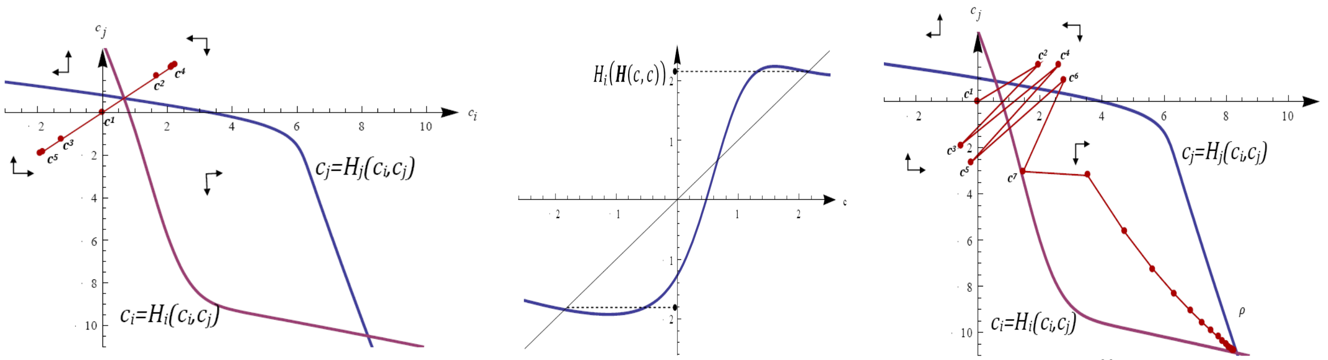
In the left panel, $\boldsymbol{H}$ has a unique fixed point, while in the right panel, $\boldsymbol{H}$ has three fixed points. The pictures show that $\boldsymbol{c}^{\infty}$ may not be continuous in the primitives of the problem. However, consistent with proposition 3, partisanship increases in the preference polarization μ. Setting δ equal to 1, one can show numerically that for $\mu \geq 0.22$, $\boldsymbol{H}$ has no fixed point. This means that, as established in the proof of proposition 4, there is complete gridlock as players become very patient, even though for $\mu = 0.22$ players preferences agree with probability 0.64.

4.2 Congruent preferences

Unlike in the polarized case, partisanship for congruent preference distributions may exhibit complex and even chaotic dynamics. To understand why, note that if a preference distribution is congruent, since $\boldsymbol{c}^2 = \boldsymbol{H}(0,0)$, in the penultimate period $t = 2$ of $\Gamma^{en}_{T,q}$, players are partisan for the same alternative. Suppose that this alternative is 1. This means that player j is more likely to vote for 1 in $t = 2$; therefore, player i has less incentive to secure the status quo $q = 1$ for the period $t = 2$; in other words, c_i^3 is decreasing in c_j^2. As a result, player i may be less partisan for 1 in period $t = 3$, or may even be partisan for -1. Hence, partisanship may not vary monotonically over time.

The example below magnifies the negative effect of c_j^{t-1} on c_i^t by considering a preference distribution with fat tails. It shows that when extreme preferences are sufficiently likely, partisanship can cycle over time and arbitrarily similar players can become strongly partisan for different alternatives. Intuitively, fat tails make c_i^t very negatively sensitive to c_j^{t-1} because when extreme preferences for either alternative are very likely, each player wants to secure as a status quo the alternative for which the opponent is less likely to vote in the next period.[15]

Example 2 *Preferences θ_i and θ_j are independent, and for $k \in \{i, j\}$, θ_k is drawn from $N\left(\bar{\theta}_k, 1\right)$ with probability $\frac{1}{2}$ and from $N\left(\bar{\theta}_k, 1000\right)$ with the remaining probability. Panel A in the figure below shows the phase diagram of $\boldsymbol{H}$ and the dynamics of partisanship for $\bar{\theta}_i = \bar{\theta}_j = 1$. Given that players are ex ante identical, they have the same voting thresholds. The unique fixed point $\mathbf{c}^f$ of $c \to H_i(c, c)$ is unstable since $\frac{\partial[H_i(c,c)]}{\partial c}\left(\mathbf{c}^f\right) < -1$. Panel B graphs $H_i(\boldsymbol{H}(c, c))$ and shows that $c \to H_i(\boldsymbol{H}(c, c))$ has two additional (and stable) fixed points around $\mathbf{c}^f$. Hence, although $\mathbf{c}^t$ does not converge, it oscillates between the least and the greatest fixed point of $c \to H_i(\boldsymbol{H}(c, c))$. Panel C shows the phase diagram of $\boldsymbol{H}$ and the dynamics of partisanship when we slightly perturb the previous, symmetric example by setting $\bar{\theta}_i = 1$ and $\bar{\theta}_j = 1.1$. The voting thresholds do not converge to the almost symmetric (unstable) fixed point of $\boldsymbol{H}$, but to the (stable) least fixed point of $\boldsymbol{H}$ for the order $(\leq, \geq)$.*



Panel A *Panel B* *Panel C*

Qualitatively, example 2 resonates with our results on polarized preference distribution in that it shows that the endogeneity of the status quo magnifies the conflict of interest between players: it induces arbitrarily similar players to behave as if their interests were highly discordant.

The discussion above suggests that in the congruent case the conditions under which partisanship increases monotonically over time, oscillates or exhibit discontinuities depend on the details of the preference distribution. Since we believe that the case of polarized preference distribution is empirically more relevant, we do not elaborate on the characterization

[15]It should be noted that the only role of fat tails is to make the discontinuity of $\boldsymbol{c}^\infty$ with respect to the preference distribution more salient. Even with standard distributions such as bivariate normal, $\boldsymbol{c}^\infty$ can be discontinuous (see the example in subsection 4.1) and congruent players may end up partisan for different alternatives as $T \to \infty$.

of these conditions in the interest of brevity.[16]

5 The case of N players

Since one of the motivating examples for our analysis is legislative bargaining, it seems important to look at the case of more than two players and majoritarian approval rules. In this section we show that under sensible assumptions on the preference distribution, our results readily extend to an $N-$player game—denoted by $\Gamma^{en}_{T,q,N}$—which differs from $\Gamma^{en}_{T,q}$ in that in every period, an alternative replaces the current status quo if at least M players vote for it, where $N/2 < M \leq N$. We assume that the preferences of player n in period t are given by

$$\theta_n \left(\varepsilon^t\right) = v_n + \varepsilon^t, \tag{4}$$

where $v_1 < .. < v_N$ parameterize the ideology of each voter, and ε^t is a common shock which is i.i.d. over time, has full support, and admits a p.d.f. f. This condition placed on preferences essentially amounts to ruling out preference reversal among voters.

Proposition 5 *When $M > \frac{N+1}{2}$, the unique equilibrium of $\Gamma^{en}_{T,q,N}$ is characterized by a sequence of voting thresholds $(c^t_1, .., c^t_N)_{t=1..T}$ given by $\boldsymbol{c}^1 = (0, .., 0)$ and $\boldsymbol{c}^{t+1} = \boldsymbol{H}\left(\boldsymbol{c}^t\right)$, where for all n,*

$$\boldsymbol{H}_n\left(\boldsymbol{c}\right) = \delta \int_{c_M - v_M}^{c_{N-M+1} - v_{N-M+1}} \left(c_n - \theta_n\left(\varepsilon\right)\right) f\left(\varepsilon\right) d\varepsilon. \tag{5}$$

In any period t, the voting thresholds are ordered with the players' ideology: $c^t_n \geq c^t_{n+1}$ for any $n = 1, ...N-1$. The thresholds of the M^{th} most right-wing and left-wing players have opposite sign: $c^t_M < 0 < c^t_{N-M+1}$. When $M = \frac{N+1}{2}$, all players are nonpartisan in all periods.

As in the two-player case, the requirement that players vote as if they were pivotal uniquely determines the equilibrium behavior, and implies that all players are partisan. The proof of proposition 5 proceeds by showing that since θ^t_n is always increasing in n, partisanship is also monotonic in n, and hence the M^{th} most right-wing and the M^{th} most left-wing voters are always pivotal. Hence, the game $\Gamma^{en}_{T,q,N}$ boils down to the $2-$player game $\Gamma^{en}_{T,q}$ in which $\boldsymbol{\theta}^t = \left(\theta^t_M, \theta^t_{N-M+1}\right)$.

Since θ^t_M is always greater than θ^t_{N-M+1}, the preference distribution of the decisive players is polarized, which implies that those players exhibit opposite partisanship. This means that the direction of the voting bias of a given decisive player is not determined by her ideological

[16]In the previous version of this paper, we derived sufficient conditions for partisanship to have a stable direction over time, to converge and to vary monotonically with the main primitives of the problem. These results are available from the authors upon request.

inclination for one or the other alternative but instead by her *relative* position as compared to that of the other decisive player. The players' willingness to ignore their actual preference realizations and base their vote instead on their relative ideological positions is what lead us to refer to the voting thresholds as partisanship.

To understand the bounds of the integral in (5), observe that from (4) the pivotal voters disagree when

$$\varepsilon^t \in \left[c_M^t - v_M, c_{N-M+1}^t - v_{N-M+1}\right]. \tag{6}$$

Note that under the exogenous status quo, players would not be partisan, so the set of shock realizations for which the status quo would remain in place would simply be $[-v_M, -v_{N-M+1}]$, and since $c_M^t < 0 < c_{N-M+1}^t$, this set is strictly included in (6). Hence, the endogeneity of the status quo increases the probability of disagreement.

Moreover, with the endogenous status quo, the disagreement region given in (6) is strictly increasing (in the inclusion sense) in $\left(c_M^t, c_{N-M+1}^t\right)$ in the order $(\leq, \geq)$. This means that under the preference specification (4), for a given ideology profile $\boldsymbol{v}$, the probability of disagreement increases in partisanship. Together with propositions 2 and 3, this shows the following:

Corollary 2 *The probability of disagreement—and thus the probability of status quo inertia—increases with bargaining horizon T and patience δ.*

Proposition 4 further implies that if the ideological difference between the pivotal players is large enough, then as players become very patient, players disregard the common shock and vote purely along ideological lines, which results in complete gridlock. With symmetric shocks, one can further show that gridlock can arise even with a modest degree of ideological polarization:

Corollary 3 *If ε is symmetrically distributed around 0 with standard deviation σ_ε, $\left(c_M^t, c_{N-M+1}^t\right)$ tends to $(-\infty, +\infty)$ whenever $(v_M, v_{N-M+1})\,(\geq, \leq)\,(\sigma_\varepsilon, -\sigma_\varepsilon)$.*

Finally, the following proposition states that the supermajority requirement exacerbates the effect of the endogeneity of the status quo on status quo inertia.

Proposition 6 *Let $\boldsymbol{c}^t(M)$ be the equilibrium voting thresholds in period t when the supermajority requirement is M. Then for all $M < M'$,*

$$c_{M'}^t(M') \leq c_{M'}^t(M) \leq 0 \leq c_{N-M'+1}^t(M) \leq c_{N-M'+1}^t(M'). \tag{7}$$

Propositions 5 and 6 imply that a higher supermajority requirement increases the probability of disagreement for three distinct reasons, each of which corresponds to an occurrence of M in each of the two bounds $c_M^t(M) - v_M$ and $c_{N-M+1}^t(M) - v_{N-M+1}$ of the disagreement region (6). First, increasing M to M' changes the identity, and hence the ideology v_M and v_{N-M+1} of the decisive voters, which expands the disagreement region (6). Second, as shown in proposition 5, for a given supermajority, voters M' and $N - M' + 1$ are more partisan than voters M and $N - M + 1$, which also increases (6). And finally, proposition 6 implies that the supermajoritarian requirement has a direct effect on the partisanship of voters M' and $N - M' + 1$. The reason is that voters M' and $N - M' + 1$ are more likely to disagree than voters M and $N - M + 1$. So making the former rater than the latter decisive increases the importance of the status quo and thus partisanship, increasing (6) even further. The second and third effect would not occur with an exogenous status quo. This shows that when the status quo is endogenous, an increase in the supermajority increases status quo inertia not only because more voters have to agree, but also because it exacerbates the ideological polarization of decisive voters.

Note that under simple majority rule (i.e., $M = \frac{N+1}{2}$), the median voter is pivotal in every period and thus always follows her current preferences. Hence, what drives partisanship and the inertia of collective decision making in $\Gamma_{T,q,N}^{en}$ is not the endogeneity of the status quo per se, but its combination with a supermajority requirement, as only then is the status quo strategically relevant.

6 Welfare effect of partisanship

The welfare effect of the endogenous status quo is ambiguous. On the one hand, the equilibrium analysis shows that the endogeneity of the status quo is detrimental to the responsiveness of the equilibrium to the environment. On the other hand, by bundling the players' vote on today's policy and tomorrow's status quo, the endogenous status quo prompts the players to express the intensity of their preference realizations, which could be socially beneficial.

To assess the welfare effect of the endogeneity of the status quo, we compare our game to bargaining protocols which break the dynamic linkage between periods. In the two-player game $\Gamma_{T,q}^{en}$, the dynamic linkage results from the interplay of two elements. First, the endogeneity of the status quo constrains players to bundle their vote on today's policy and tomorrow's status quo; second, having more than one player decisive renders the status quo strategically relevant. Designating one alternative as a fixed status quo throughout the game (as in $\Gamma_{T,q}^{ex}$), or concentrating the voting rights in the hands of one of the players (dictatorship in the two-player game or simple majority rule in the $N-$player game) severs the link between bargaining periods and therefore eliminates partisanship.[17]

6.1 Fixed status quo and sunset provisions

In this section, we compare the game with an endogenous status quo $\Gamma^{en}_{T,q}$ to the game with an exogenous status quo $\Gamma^{ex}_{T,q}$. We denote by $W\left(\Gamma^{en}_{T,q}\right)$ and $W\left(\Gamma^{ex}_{T,q}\right)$ the corresponding expected level of utilitarian welfare in the equilibrium. Since $W\left(\Gamma^{en}_{T,q}\right)$ depends on T, we compare $W\left(\Gamma^{en}_{T,q}\right)$ to $W\left(\Gamma^{ex}_{T,q}\right)$ as T approaches infinity.[18] The following definition will be useful in characterizing the welfare properties of $\Gamma^{en}_{T,q}$:

Definition *A preference distribution is uniformly ordered if the distribution of $\theta_i - \theta_j$ conditional on $\theta_i + \theta_j$ is such that for all $b > 0$ and all $a \in [-b, b]$,*

$$f\left(\theta_i - \theta_j = -b|\theta_i + \theta_j\right) \leq f\left(\theta_i - \theta_j = a|\theta_i + \theta_j\right). \tag{8}$$

Condition (8) means that one player is more likely to be to the right of the other player conditional on any value of the average preferences $\theta_i + \theta_j$.[19]

The next proposition shows that fixing the status quo is generally socially beneficial.

Proposition 7 *If $\left(\boldsymbol{c}^t\right)_{t\geq 1}$ converges to some $\boldsymbol{c}^\infty$ such that $c_i^\infty + c_i^\infty \neq 0$,[20] and (i) c_i^∞ and c_j^∞ have the same sign, or (ii) the preference distribution is uniformly ordered, then there exists $q' \in \{-1, 1\}$ such that for all q, $W\left(\Gamma^{en}_{T,q}\right) < W\left(\Gamma^{ex}_{T,q'}\right)$ for T sufficiently large. Moreover, in case (i), $\Gamma^{ex}_{T,q'}$ ex-ante Pareto dominates $\Gamma^{en}_{T,q}$ for T sufficiently large.*

Roughly speaking, condition (i) corresponds to congruent preference distributions for which players' behavior remains stable over time, while condition (ii) refers to either polarized preferences, or congruent preferences for which players end up being partisan for different alternatives (as in example 2).

To understand the intuition for proposition 7, note that partisanship affects welfare in three different ways. First, as argued earlier, partisan players enact Pareto inefficient

[17] Actually, dictatorship and a fixed status quo are the only static games with two alternatives: one can show that for a one-period interaction, the set of (Pareto optimal) incentive-compatible direct mechanisms without transfers is the convex hull of a fixed status quo $q = 1$, a fixed status quo $q = -1$, the dictatorship of player i, and the dictatorship of player j. From the revelation principle, the Bayesian Nash equilibrium of any game form can be replicated by one of these direct mechanisms.

[18] The results of this section also apply to any Markov perfect equilibria of $\Gamma^{en}_{\infty,q}$.

[19] In particular, this condition is satisfied if the preference distribution is uniform on $[a, b] \times [a', b']$ with $a > a'$ and $b > b'$; bivariate normal with $\sigma_i = \sigma_j$ and $\bar{\theta}_i > \bar{\theta}_j$; or if $\theta_i > \theta_j$ with probability one, as in $\Gamma^{en}_{T,q,N}$.

[20] If $c_i^\infty + c_i^\infty = 0$, then one can easily see from the proof of proposition 7 that any exogenous status quo dominates the endogenous status quo, but in a weak sense: for all $q, q' \in \{-1, 1\}$, $\lim_{T\to\infty}\left(W\left(\Gamma^{ex}_{T,q'}\right)\right) \geq \lim_{T\to\infty}\left(W\left(\Gamma^{en}_{T,q}\right)\right)$. The inequality is strict in particular whenever $W\left(\Gamma^{en}_{T,1}\right) \neq W\left(\Gamma^{en}_{T,-1}\right)$.

policies with positive probability, which does not happen under a fixed status quo. Second, partisanship affects the distribution of the status quo. In the polarized case, players' attempts at affecting the status quo offset each other but increase the probability of disagreement. This is socially detrimental and can be avoided with a fixed status quo. In the congruent case, players sacrifice their current payoff to shift the distribution of the status quo in the same, mutually favorable direction. However, designating the preferred status quo of both players as a fixed status quo has the same effect, but does not distort their voting behavior.

Third, a partisan player, while voting for her preferred status quo, may defer to her opponent's preferences. This may be socially beneficial if the opponent's preferences are relatively more intense. Condition (8) makes this type of preference reversal sufficiently unlikely so that the two aforementioned detrimental effects of partisanship dominate. In the appendix (example 4), we construct a polarized preference distribution which violates condition (8) and for which an endogenous status quo is socially better than any static bargaining protocol, and any fixed status quo in particular.

Proposition 7 provides support for automatic sunset provisions. A sunset provision is a clause that repeals a law, a tax change, or a regulation after a specific date, unless further legislative action is taken. Hence, it severs the link between the current agreement and the future status quo by automatically reverting the policies to a pre-specified default. Therefore, a sunset clause is essentially equivalent to a fixed status quo. In the U.S., sunset provisions were used extensively at the state level in the 1970s and 1980s (e.g., in Texas, Colorado, and Alabama), primarily to contain the multiplication of executive agencies and regulations. They were seldom used at the federal level prior to 2000, but have become increasingly frequent since then. For example, many of the provisions of the USA Patriot Act of 2001 had a four-year sunset clause, while the tax cuts authorized in the Economic Growth and Tax Relief Reconciliation Act of 2001 and the Jobs and Growth Tax Relief Reconciliation Act of 2003 had a ten- and five-year sunset clauses, respectively.

The standard rationale for sunset clauses is threefold: to improve parliamentary control of executive agencies through periodic reviews, to evaluate the efficiency of new laws (in order to avoid inefficient laws, unintended consequences, or loopholes), and to impose a time limit on regulations designed to deal with a temporary issue.[21] The argument advanced by our model does not rest on a delegation problem, policy learning, or one-time events, but has instead a more strategic underpinning. We show that sunset provisions sever the link between today's agreement and tomorrow's status quo, which eliminates partisanship and

[21] The idea of sunset clauses as a tool for legislatures to monitor and evaluate regulatory agencies was pervasive in the state sunset laws introduced in the 1970s and 1980s (see Kearney 1990). A historical example of a sunset provision on a policy meant to be temporary is the two-year limit on the "Vietnam tax surcharge" authorized by the Revenue and Expenditure Control Act of 1968.

makes policies more responsive to the environment.

We conclude this section by raising the question whether, in light of our results, sunset clauses could emerge endogenously. One can shed light on this issue by considering the following extension of our model: in every period the players first vote on whether the policy should be subject to a sunset clause or whether it should become the future status quo, and only then they vote on the policy. Although a detailed analysis of this game is beyond the scope of this paper, a quick inspection reveals that both sunsets and endogenous status quo would occur in equilibrium. To see that, consider a situation in which for the last period player i prefers the status quo 1 and player j prefers the status quo -1. Suppose that in the penultimate period the status quo is -1, and both players prefer policy 1. If the dominant force in player j's intertemporal trade-off is her preference for status quo -1 for the last period, player i needs to attach a sunset provision to the penultimate policy to convince player j to approve the Pareto optimal policy 1. On the other hand, if the dominant force in player j's intertemporal trade-off is her preference for policy 1 in the penultimate period, she will vote for 1 even without the sunset provision. Anticipating this, player i will veto the sunset provision to assure that 1 will also become the last period's status quo. Hence, despite its inefficiency, the endogenous status quo can prevail.

6.2 Concentrating decision rights

Concentrating decision rights unambiguously reduces partisanship, but its effect on welfare depends on the details of the preference distribution.[22] However, by strengthening the uniform order requirement of proposition 7, we have the following:

Corollary 4 *If $\theta_i > \theta_j$ with probability 1, then for all q, there exists $k \in \{i, j\}$ such that if $\Gamma^d_{T,k}$ denotes the T period game in which k is a dictator, then for all q, $W\left(\Gamma^{en}_{T,q}\right) < W\left(\Gamma^{d}_{T,k}\right)$ for T sufficiently large.*

The intuition for corollary 4 is simple: if $\theta_i > \theta_j$ with probability 1, then preferences are uniformly ordered. Moreover, in case of disagreement, player i always prefers 1 while player j always prefers -1; hence, $\Gamma^d_{T,k}$ is equivalent to $\Gamma^{ex}_{T,q}$ where q is the status quo most preferred by player k. The result then follows from proposition 7.

More surprisingly, one can find examples in which both players prefer to give up their veto power and let their opponent be the dictator. In the appendix, we show (see example 5) that if preferences are perfectly correlated and normally distributed with opposite means

[22] For instance, for the preference distribution constructed in example 4 (see the appendix), the endogenous status quo is socially better than a dictatorship of any player.

and if players are sufficiently patient, then the dictatorship of any player is Pareto preferred to $\Gamma^{en}_{T,q}$ for some intermediate values of the means. That is, if the means are far apart enough to lead to complete gridlock (see proposition 4), but close enough so that θ_i and θ_j agree sufficiently often, then both players prefer to vest the opponent with dictatorial power rather than face a totally unresponsive policy.

In the $N-$player game $\Gamma^{en}_{T,q,N}$, concentrating decision rights can be achieved by decreasing the supermajority requirement of the approval rule. In particular, under simple majority the pivotal player is de facto a dictator, and hence all players become nonpartisan. From proposition 6, we know that under any other majority rule the players are partisan, but decreasing M reduces partisanship. The next proposition shows that when the preference distribution is symmetric, a lower supermajority increases the welfare of all players.

Proposition 8 *If $\boldsymbol{v}$ and ε^t are symmetrically distributed around zero, then for all T and q, $W\left(\Gamma^{en}_{T,q,N}\right)$ decreases in the supermajority requirement M.*

To see the intuition for proposition 8 it is helpful to consider the extreme case of simple majority rule. From proposition 7, an exogenous status quo improves the utilitarian welfare of the pivotal players. Simple majority rule, or equivalently dictatorship of the median player, further improves social welfare because whenever the pivotal players M and $N-M+1$ agree, they also agree with the median voter. When they disagree, however, the policy preferred by the median voter is more responsive to the average preferences of the pivotal players than a fixed status quo. Observe that simple majority rule is also more likely to leave both sides of the ideological spectrum better off than a fixed status quo because it distributes the benefits from breaking the dynamic linkage more evenly across the ideological spectrum.

The symmetry assumption in proposition 8 is only needed to guarantee that the utilitarian welfare of pivotal players is representative of the utilitarian welfare of all players. If the symmetry assumption is violated, for example, in that the preferences of the median voter are the same as the preferences of the M^{th} most right-wing voter, moving from a supermajority M to simple majority rule is equivalent to awarding dictatorship to voter M. This may not be beneficial to left-wing voters, and may decrease social welfare if the distribution of ideology is skewed to the left.

Most Western parliamentary democracies require some form of supermajority in their legislative process. For instance, in the U.S. Senate, senators have the right to limitlessly debate (filibuster), with debate ending only when sixty senators vote in favor of so doing. As a result, in order for controversial measures to pass the Senate, a supermajority of sixty votes, rather than a simple majority, is required. Our results suggest that supermajoritarian requirements, when used to enact continuing policies, have a negative welfare effect.

These results contrast with the literature on majoritarian incentives with distributive policies. As Buchanan and Tullock (1962) and Riker (1962) have argued, majoritarian rules allows the concentration of benefits and the collectivization of costs and thus lead to the adoption of inefficient pork-barrel programs, more so the lower the supermajority requirement. Ferejohn, Fiorina, and McKelvey (1987) and Baron (1991) first formalized this prediction in models of legislative bargaining. Our results contrast with these analyses in that we rule out targeted spending programs—which are a negligible fraction of the U.S. federal budget—and focus instead on entitlement programs and other continuing policies.

7 Intertemporal Correlation

In many applications, certain shocks to the environment are likely to have a persistent effect on preferences, while others are transient. Moreover, economic agents often receive signals about shocks' persistence. For instance, economic indicators inform policy makers about the duration of booms and recessions. In this section, we relax the assumption of the stationarity of the environment and analyze how players' behavior depends on the persistence of the current situation and the volatility of the environment.

In each period t, $\pi^t \in [0,1]$ measures the persistence of the current environment: with probability π^t the preferences and their persistence remain unchanged for the next period; with probability $1-\pi^t$, the preferences and the persistence parameter are redrawn. Formally:

$$\left(\boldsymbol{\theta}^{t-1}, \pi^{t-1}\right) = \begin{cases} \left(\boldsymbol{\theta}^{t}, \pi^{t}\right) \text{ with probability } \pi^t \\ \boldsymbol{\theta}^{t-1} \sim f\left(.\right) \text{ and } \pi^{t-1} \sim g\left(.\right) \text{ with probability } 1-\pi^t, \end{cases}$$

where $g\left(\pi\right)$ is a p.d.f. with support on $[0,1]$. A realization of $(\boldsymbol{\theta}, \pi)$ is called *a phase*.

We denote by $\Gamma^{en}_{L,q}$ the game that lasts L phases and starts with status quo q. As before, we adopt the convention that $l = 1$ denotes the last phase. As the next proposition shows, in equilibrium, the voting behavior is constant within each phase, and thus depends on l but not on t. For this reason, we index all variables by l instead of t. As in the basic model, the equilibrium behavior depends neither on L nor on the initial status quo q.

Proposition 9 $\Gamma^{en}_{L,q}$ *has a unique (up to a zero measure subset of preference realizations) equilibrium, in which players use threshold strategies which are constant within each phase. In each phase l, the voting thresholds depend on the preference persistence π^l and are given by $\boldsymbol{c}^l\left(\pi^l\right) = \left(1-\pi^l\right)\hat{\boldsymbol{c}}^l$, with the sequence $\left(\hat{\boldsymbol{c}}^l\right)_{l\geq 1}$ defined recursively by $\hat{\boldsymbol{c}}^1 = (0,0)$ and $\hat{\boldsymbol{c}}^{l+1} = \boldsymbol{G}\left(\hat{\boldsymbol{c}}^l\right)$, where*

$$\boldsymbol{G} : \hat{\boldsymbol{c}} \rightarrow \int_0^1 \frac{\boldsymbol{H}\left(\left(1-\pi\right)\hat{\boldsymbol{c}}\right)}{1-\pi\delta} g\left(\pi\right) d\pi. \tag{9}$$

Note that although players' partisanship in each phase l is a function of π^l, the sequence of functions $\left(\boldsymbol{c}^l(.)\right)_{l\geq 1}$ is completely characterized by the sequence of scalars $\left(\hat{\boldsymbol{c}}^l\right)_{l\geq 1}$. Since $\hat{\boldsymbol{c}} \to \boldsymbol{G}(\hat{\boldsymbol{c}})$ is monotonic in the order $(\leq,\geq)$ and since the sign of $G_k(0,0)$ is the same as the sign of $H_k(0,0)$, using the same definition for polarization and congruence one can easily check that the results stated in proposition 2, 3 and 4 on the sequence of thresholds $(\boldsymbol{c}^t)_{t\geq 1}$ hold for the the sequence $\left(\hat{\boldsymbol{c}}^l\right)_{l\geq 1}$. That is, for a given persistence, partisanship is increasing with the bargaining horizon and patience, and can lead to complete gridlock if the preferences are sufficiently polarized.[23]

Proposition 9 implies that for a given distribution of persistence, players are more partisan in transient states (small π). The next proposition compares partisanship across environments with different volatility. It shows that for a given realization of persistence, patient players are more partisan in an environment in which persistent shocks are more likely.

Proposition 10 *Let $\boldsymbol{\theta}$ be strictly polarized, i.e., $\boldsymbol{H}(0,0)(<,>)(0,0)$, and let g and g' be such that $\lim_{\pi\to 1^-} g'(\pi) > \lim_{\pi\to 1^-} g(\pi)$, then there exists $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$ and all l,*

$$\boldsymbol{c}^l(\pi, g')(<,>)\,\boldsymbol{c}^l(\pi, g)(<,>)(0,0).$$

Proof. There exists $\epsilon > 0$ such that $\inf_{\pi\in[1-\epsilon,1]} g'(\pi) > \sup_{\pi\in[1-\epsilon,1]} g(\pi)$. As shown in the proof of proposition 2, $\boldsymbol{H}$ is bounded, so there exists $A \in \mathbb{R}$ such that for all $\delta \in [0,1]$ and all $\hat{\boldsymbol{c}} \in \mathbb{R}^2$, $\left|\int_0^{1-\epsilon} \frac{H_j((1-\pi)\hat{\boldsymbol{c}})}{1-\pi\delta} h(\pi)\,d\pi\right| \leq A$, and likewise for $h \in g'$. Since $\boldsymbol{H}$ is monotone in the order $(\leq,\geq)$, for all $\hat{\boldsymbol{c}} \in \mathbb{R}_- \times \mathbb{R}_+$,

$$\int_{1-\epsilon}^{1} \frac{H_j((1-\pi)\,\hat{\boldsymbol{c}})}{1-\pi\delta} g'(\pi)\,d\pi - \int_{1-\epsilon}^{1} \frac{H_j((1-\pi)\,\hat{\boldsymbol{c}})}{1-\pi\delta} g(\pi)\,d\pi \geq \int_{1-\epsilon}^{1} \frac{H_j(0,0)}{1-\pi\delta}\left(g'(\pi) - g(\pi)\right) d\pi.$$

This shows that for all $\hat{\boldsymbol{c}} \in \mathbb{R}_- \times \mathbb{R}_+$,

$$\boldsymbol{G}(\hat{\boldsymbol{c}}, g') - \boldsymbol{G}(\hat{\boldsymbol{c}}, g) \geq H_j(0,0)\left(\inf_{\pi\in[1-\epsilon,1]} g'(\pi) - \sup_{\pi\in[1-\epsilon,1]} g(\pi)\right)\int_{1-\epsilon}^{1} \frac{1}{1-\pi\delta}\,d\pi - A.$$

Since $\int_{1-\epsilon}^{1} \frac{1}{1-\pi\delta}\,d\pi \to \infty$ as $\delta \to 1$ and since $G_j(0,0) > 0$ implies $H_j(0,0) > 0$, it follows that there exists $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$ and all $\hat{\boldsymbol{c}} \in \mathbb{R}_- \times \mathbb{R}_+$, $\boldsymbol{G}^{g'}(\hat{\boldsymbol{c}}) > \boldsymbol{G}^{g}(\hat{\boldsymbol{c}})$. Lemma 2 (see the appendix) for $\boldsymbol{H} = \boldsymbol{G}$, $P = \{g, g'\}$ and $C_p = \mathbb{R}_- \times \mathbb{R}_+$ completes the argument. ■

The result of proposition 10 is not straightforward. To see that, note that increasing the probability of persistent shocks has two effects. On the one hand, a disagreement in

[23] As in the stationary case, whenever $\hat{\boldsymbol{c}}^l$ converge as l approaches ∞, the constant cut-off strategy defined by $\boldsymbol{c}^\infty(\pi) = (1-\pi)\,\hat{\boldsymbol{c}}^\infty$ is a Markov perfect equilibrium of the infinite horizon game $\Gamma^{en}_{L=\infty,q}$.

the next phase is more likely to be long-lived, which drives today's partisanship up. On the other hand, from proposition 9 partisanship is low in persistent states; hence, since the next phase is more likely to be highly persistent, players are likely to be less partisan in that phase. This decreases the probability of future disagreement, which drives today's partisanship down. The proof of proposition 10 shows that as players become very patient, the second effect is bounded while the first effect is not, so the latter dominates the former.

8 Conclusion

Negotiations in a changing environment with an endogenous default option are at the center of many economically relevant situations. They present the negotiating parties with a fundamental trade-off between responding to the current environment and securing a favorable bargaining position for the future. In this paper, we show that this trade-off has a detrimental impact on the efficiency of agreements and their responsiveness to political and economic shocks. Bundling the vote on today's policy and tomorrow's status quo exacerbates the players' conflict of interest and increases the probability of a disagreement, which in turn increases status quo inertia. Even if some agreements are commonly known to be mutually beneficial, they may not be adopted.

Our paper sheds light on the effect of some important rules governing legislative institutions: we provide a new argument in favor of sunset provisions and we show that a supermajority requirement exacerbates the detrimental impact of an endogenous default on the responsiveness of the policies to the environment.

This parsimonious model lends itself to many extensions. First, adding transfers—interpreted as pork-barrel spending—to the $N-$player model would allow us to analyze the trade-off between their positive role as a lubricant for passing efficient policies and the perverse incentives they generate to concentrate benefits and collectivize cost. Second, by enriching the policy space one could analyze whether the evolving environment can make inefficient compromises persistent. Third, in many situations, implemented policies affect the future state of the economy, which introduces an additional dynamic linkage. For example, an expansionary fiscal policy increases public debt, leading all players to adopt a more fiscally conservative stand in the future. Technically, this amounts to introducing a state variable in the model. And finally, one could introduce elections to see whether strategic delegation would exacerbate or mitigate the partisanship of the legislature.

9 Appendix

Throughout the appendix, Φ and ϕ will denote the c.d.f. and the p.d.f. of the normal distribution (univariate or bivariate depending on the context) with mean 0 (or $(0,0)$) and standard deviation 1 (or $(1,1)$).

Lemma 1 *Let $\boldsymbol{H}$ be defined by (1), then*

$$\frac{\partial H_i}{\partial c_i} = \delta \left(\int_{c_j}^{\infty} \int_{-\infty}^{c_i} f(\boldsymbol{\theta})\, d\theta_i d\theta_j + \int_{-\infty}^{c_j} \int_{c_i}^{\infty} f(\boldsymbol{\theta})\, d\theta_i d\theta_j \right) > 0 \text{ and } \frac{\partial H_i}{\partial c_j} = -\delta \int_{-\infty}^{+\infty} |\theta_i - c_i|\, f(\theta_i, c_j)\, d\theta_i < 0.$$

Let $\boldsymbol{\theta}$ be a preference distribution and for all $\boldsymbol{m} \in \mathbb{R}^2$, let $\boldsymbol{H}^{\boldsymbol{m}}$ be map defined in (1) for the preference distribution $\boldsymbol{\theta} + \boldsymbol{m}$, then $\frac{\partial H_i^{\boldsymbol{m}}}{\partial \pi_i} = -\frac{\partial H_i^{\boldsymbol{m}}}{\partial c_i}$ and $\frac{\partial H_i^{\boldsymbol{m}}}{\partial \pi_j} = -\frac{\partial H_i^{\boldsymbol{m}}}{\partial c_j}$.If the preference distribution is bivariate normal with mean $\left(\bar{\theta}_i, \bar{\theta}_j\right)$, standard deviation $(1,1)$, and correlation ρ, then

$$\frac{\partial H_i}{\partial \rho} = \delta \left(1 - 2\Phi\left(\frac{c_i - \rho c_j - \left(\bar{\theta}_i - \rho \bar{\theta}_j\right)}{\sqrt{1-\rho^2}} \right) \right) \phi\left(c_j - \bar{\theta}_j\right).$$

Proof. The expressions for $\frac{\partial H_i}{\partial c_i}$, $\frac{\partial H_i}{\partial c_j}$, $\frac{\partial H_i^{\boldsymbol{m}}}{\partial \pi_i}$ and $\frac{\partial H_i^{\boldsymbol{m}}}{\partial \pi_j}$ are obtained using the Leibnitz integral rule on (1). If $\phi_{\rho,\bar{\theta}_i,\bar{\theta}_j}$ is the p.d.f. of the bivariate normal with mean $\left(\bar{\theta}_i, \bar{\theta}_j\right)$, variance $(1,1)$, and correlation ρ, simple calculus yields that $\frac{\partial^2 \phi_{\rho,\bar{\theta}_i,\bar{\theta}_j}}{\partial \theta_i \partial \theta_j} = \frac{\partial \phi_{\rho,\bar{\theta}_i,\bar{\theta}_j}}{\partial \rho}$. Substituting in $\frac{\partial H_i(\boldsymbol{c})}{\partial \rho}$ and integrating by parts, we get

$$\begin{aligned}
\frac{1}{\delta}\frac{\partial H_i(\boldsymbol{c})}{\partial \rho} &= \textstyle\int_{c_j}^{\infty} \int_{-\infty}^{c_i} (c_i - \theta_i) \frac{\partial^2 \phi_{\rho,\bar{\theta}_i,\bar{\theta}_j}}{\partial \theta_i \partial \theta_j}(\theta_i, \theta_j)\, d\theta_i d\theta_j + \int_{-\infty}^{c_j} \int_{c_i}^{\infty} (c_i - \theta_i) \frac{\partial^2 \phi_{\rho,\bar{\theta}_i,\bar{\theta}_j}}{\partial \theta_i \partial \theta_j}(\theta_i, \theta_j)\, d\theta_i d\theta_j \\
&= \textstyle\int_{-\infty}^{c_i} (\theta_i - c_i) \frac{\partial \phi_{\rho,\bar{\theta}_i,\bar{\theta}_j}}{\partial \theta_i}(\theta_i, c_j)\, d\theta_i - \int_{c_i}^{\infty} (\theta_i - c_i) \frac{\partial \phi_{\rho,\bar{\theta}_i,\bar{\theta}_j}}{\partial \theta_i}(\theta_i, c_j)\, d\theta_i \\
&= \left[(\theta_i - c_i)\, \phi_{\rho,\bar{\theta}_i,\bar{\theta}_j}(\theta_i, c_j)\right]_{-\infty}^{c_i} - \int_{-\infty}^{c_i} \phi_{\rho,\bar{\theta}_i,\bar{\theta}_j}(\theta_i, c_j)\, d\theta_i - \left[(\theta_i - c_i)\, \phi_{\rho,\bar{\theta}_i,\bar{\theta}_j}(\theta_i, c_j)\right]_{c_i}^{\infty} \\
&\quad + \textstyle\int_{c_i}^{\infty} \phi_{\rho,\bar{\theta}_i,\bar{\theta}_j}(\theta_i, c_j)\, d\theta_i = \left(1 - 2\Phi\left(\frac{c_i - \rho c_j - \left(\bar{\theta}_i - \rho\bar{\theta}_j\right)}{\sqrt{1-\rho^2}} \right) \right) \phi\left(c_j - \bar{\theta}_j\right).
\end{aligned}$$

■

Example 3 *If the distribution of $\boldsymbol{\theta}$ is bivariate normal with means $\left(\bar{\theta}_i, \bar{\theta}_j\right)$, variances $\left(\sigma_i^2, \sigma_j^2\right)$, and correlation ρ, then the preference distribution is congruent iff $\left|\frac{\bar{\theta}_i}{\sigma_i} - \frac{\bar{\theta}_j}{\sigma_j}\right| \leq a\left(\frac{\bar{\theta}_i}{\sigma_i} + \frac{\bar{\theta}_j}{\sigma_j}; \rho\right)$, where a is a function which is positive whenever $\frac{\bar{\theta}_i}{\sigma_i} + \frac{\bar{\theta}_j}{\sigma_j} \neq 0$ and $\rho < 1$, decreasing in ρ, and tends to 0 when $\rho \to 1$.*

Proof. Let $G\left(\bar{\theta}_i, \bar{\theta}_j, \rho\right) = -\frac{1}{\delta}\boldsymbol{H}(0,0)$ where $\boldsymbol{H}$ is defined by (1) for a bivariate normal preference distribution with mean $\left(\bar{\theta}_i, \bar{\theta}_j\right)$, standard deviation $(1,1)$, and correlation ρ.[24] From lemma 1,

$$\frac{\partial G_i}{\partial \bar{\theta}_i} > 0 \text{ and } \frac{\partial G_i}{\partial \bar{\theta}_j} < 0. \tag{10}$$

The curves of the function $\bar{\theta}_i \to \mu_i\left(\bar{\theta}_i, \rho\right)$ and $\bar{\theta}_i \to \mu_j\left(\bar{\theta}_i, \rho\right)$ in the first figure of section 4 that delineates congruent and polarized distribution are characterized by the equation $G_i\left(\bar{\theta}_i, \mu_i\left(\bar{\theta}_i, \rho\right), \rho\right) = 0$ and $G_j\left(\bar{\theta}_i, \mu_j\left(\bar{\theta}_i, \rho\right), \rho\right) = 0$, respectively. The following steps prove their properties.

Step 1: $\mu_i\left(\bar{\theta}_i, \rho\right)$ and $\mu_j\left(\bar{\theta}_i, \rho\right)$ are well defined and are symmetric of each other with respect to the 45 degree line $\bar{\theta}_i = \bar{\theta}_j$. Fix $\bar{\theta}_i$ and ρ. Using the change of variables $y = \theta_j - \bar{\theta}_j$ in (1), we get

$$\begin{aligned}
\lim_{\bar{\theta}_j \to -\infty} G_i\left(\bar{\theta}_i, \bar{\theta}_j, \rho\right) &= \lim_{\bar{\theta}_j \to -\infty}\left(\int_{-\bar{\theta}_j}^{\infty}\int_{-\infty}^{0} \theta_i \phi\left(\theta_i - \bar{\theta}_i, y\right) d\theta_i dy + \int_{-\infty}^{-\bar{\theta}_j}\int_{0}^{\infty} \theta_i \phi\left(\theta_i - \bar{\theta}_i, y\right) d\theta_i dy\right) \\
&= E\left(\theta_i | \theta_i > 0\right) \Pr\left(\theta_i > 0\right) > 0.
\end{aligned}$$

Analogously, $\lim_{\bar{\theta}_j \to \infty} G_i\left(\bar{\theta}_i, \bar{\theta}_j, \rho\right) = E\left(\theta_i | \theta_i < 0\right) \Pr\left(\theta_i < 0\right)$, which is negative. Hence, $\mu_i\left(\bar{\theta}_i, \rho\right)$ exists. From (10), $\frac{\partial G_i}{\partial \bar{\theta}_j} < 0$ so $\mu_i\left(\bar{\theta}_i, \rho\right)$ is unique. The proof for $\mu_j\left(\bar{\theta}_i, \rho\right)$ is identical. The symmetry comes from the fact that $G_j\left(\bar{\theta}_i, \bar{\theta}_j, \rho\right) = 0$ implies $G_i\left(\bar{\theta}_j, \bar{\theta}_i, \rho\right) = 0$.

Step 2: The preference distribution with parameters $\left(\bar{\theta}_i, \bar{\theta}_j, \rho\right)$ is congruent iff

$$\bar{\theta}_j \in \left(\min\left\{\mu_i\left(\bar{\theta}_i, \rho\right), \mu_j\left(\bar{\theta}_i, \rho\right)\right\}, \max\left\{\mu_i\left(\bar{\theta}_i, \rho\right), \mu_j\left(\bar{\theta}_i, \rho\right)\right\}\right).$$

This is because (10) implies that $G_k\left(\bar{\theta}_i, \bar{\theta}_j, \rho\right) \geq 0$ iff $\bar{\theta}_i \geq \mu_k\left(\bar{\theta}_i, \rho\right)$.

Step 3: $\mu_i\left(\bar{\theta}_i, \rho\right)$ and $\mu_j\left(\bar{\theta}_i, \rho\right)$ are increasing in $\bar{\theta}_i$ and have the same sign as $\bar{\theta}_i$. The implicit function theorem together with (10) implies $\frac{\partial \mu_i\left(\bar{\theta}_i, \rho\right)}{\partial \bar{\theta}_i} > 0$ and $\frac{\partial \mu_j\left(\bar{\theta}_i, \rho\right)}{\partial \bar{\theta}_i} > 0$. To complete the argument, observe that $\mu_i(0, \rho) = \mu_j(0, \rho) = 0$.

Step 4: $\mu_i\left(\bar{\theta}_i, \rho\right)$ and $\mu_j\left(\bar{\theta}_i, \rho\right)$ cross the 45 degree line $\bar{\theta}_i = \bar{\theta}_j$ only at $\bar{\theta}_i = 0$, from below for $\mu_i\left(\bar{\theta}_i, \rho\right)$ and from above for $\mu_j\left(\bar{\theta}_i, \rho\right)$. By symmetry, $G_i\left(\bar{\theta}_i, \bar{\theta}_j, \rho\right) = G_j\left(\bar{\theta}_i, \bar{\theta}_j, \rho\right)$ only if $\bar{\theta}_i = \bar{\theta}_j = \mu$ for some μ. For all $\mu > 0$,

$$\begin{aligned}
&G_i(\mu, \mu, \rho) + G_j(\mu, \mu, \rho) - \left(G_i(-\mu, -\mu, \rho) + G_j(-\mu, -\mu, \rho)\right) = \\
&= \int_0^{\infty}\int_{-\infty}^{0}\left(\theta_i + \theta_j\right)\left(\phi\left(\theta_i - \mu, \theta_j - \mu\right) - \phi\left(\theta_i + \mu, \theta_j + \mu\right)\right) d\theta_i d\theta_j.
\end{aligned} \tag{11}$$

[24] Normalizing standard deviations is w.l.o.g. because the sign of $H_k(0,0)$ when $\boldsymbol{\theta}$ is bivariate normal with mean $\left(\bar{\theta}_i, \bar{\theta}_j\right)$, standard deviation $\left(\sigma_i, \sigma_j\right)$ and correlation ρ, is the same when $\boldsymbol{\theta}$ is bivariate normal with mean $\left(\frac{\bar{\theta}_i}{\sigma_i}, \frac{\bar{\theta}_j}{\sigma_j}\right)$, standard deviation $(1,1)$ and correlation ρ.

Simple algebra shows that $(\phi(\theta_i - \mu, \theta_j - \mu) - \phi(\theta_i + \mu, \theta_j + \mu))$ has the same sign as $\theta_i + \theta_j$, which implies that (11) is positive. By symmetry, $G_i(\mu, \mu, \rho) = -G_i(-\mu, -\mu, \rho)$. Therefore, $G_i(\mu, \mu, \rho)$ has the same sign as μ. Together with (10), this shows that $\mu_i\left(\bar{\theta}_i, \rho\right) > \bar{\theta}_i > \mu_j\left(\bar{\theta}_i, \rho\right)$ for $\bar{\theta}_i > 0$ and $\mu_i\left(\bar{\theta}_i, \rho\right) < \bar{\theta}_i < \mu_j\left(\bar{\theta}_i, \rho\right)$ for $\bar{\theta}_i < 0$.

Step 5: $\mu_i\left(\bar{\theta}_i, \rho\right)$ and $\mu_j\left(\bar{\theta}_i, \rho\right)$ tend towards the 45 degree line as $\rho \to 1$. To see this observe that $\lim_{\rho \to 1} G_i\left(\bar{\theta}_i, \bar{\theta}_j, \rho\right) = \int_{\min\left(0, \bar{\theta}_i - \bar{\theta}_j\right)}^{\max\left(0, \bar{\theta}_i - \bar{\theta}_j\right)} \theta_i f(\theta_i)\, d\theta_i$, which is 0 only if $\bar{\theta}_i = \bar{\theta}_j$.

Step 6: $\mu_i\left(\bar{\theta}_i, \rho\right)$ and $\mu_j\left(\bar{\theta}_i, \rho\right)$ get closer to the 45 degree line as $\rho \to 1$. To show this, we will show that for $\bar{\theta}_i > 0$, $\frac{d\mu_i\left(\bar{\theta}_i, \rho\right)}{d\rho} < 0$ and $\frac{d\mu_i\left(\bar{\theta}_i, \rho\right)}{d\rho} > 0$. The case $\bar{\theta}_i < 0$ is identical. Using the implicit function theorem and lemma 1,

$$\frac{\partial \mu_i}{\partial \rho}\left(\bar{\theta}_i, \rho\right) = -\frac{\frac{\partial G_i}{\partial \rho}\left(\bar{\theta}_i, \mu_i\left(\bar{\theta}_i, \rho\right), \rho\right)}{\frac{\partial G_i}{\partial \bar{\theta}_j}\left(\bar{\theta}_i, \mu_i\left(\bar{\theta}_i, \rho\right), \rho\right)} = -\frac{\left(2\Phi\left(\frac{\rho\mu_i\left(\bar{\theta}_i, \rho\right) - \bar{\theta}_i}{\sqrt{1-\rho^2}}\right) - 1\right)\phi\left(-\bar{\theta}_j\right)}{\frac{\partial G_i}{\partial \bar{\theta}_j}\left(\bar{\theta}_i, \mu_i\left(\bar{\theta}_i, \rho\right), \rho\right)}.$$

Together with (10), this shows that $\frac{\partial \mu_i}{\partial \rho}\left(\bar{\theta}_i, \rho\right)$ has the same sign as $\rho\mu_i\left(\bar{\theta}_i, \rho\right) - \bar{\theta}_i$. Since $\bar{\theta}_i > 0$, by Step 3, $\mu_i\left(\bar{\theta}_i, \rho\right) > 0$. Clearly, $\rho\mu_i\left(\bar{\theta}_i, \rho\right) - \bar{\theta}_i < 0$ for $\rho < 0$. Assume that there exists $\rho_0 > 0$ such that $\mu^i\left(\bar{\theta}_i, \rho_0\right)\rho_0 - \bar{\theta}_i > 0$. From what precedes, $\frac{\partial \mu_i}{\partial \rho}\left(\bar{\theta}_i, \rho_0\right) > 0$, and hence $\frac{\partial\left(\mu^i\left(\bar{\theta}_i, \rho_o\right)\rho_o - \bar{\theta}_i\right)}{\partial \rho_o} > 0$. Therefore, for all $\rho \geq \rho_0$, $\rho\mu_i\left(\bar{\theta}_i, \rho\right) - \bar{\theta}_i > 0$, and hence $\frac{\partial \mu_i}{\partial \rho}\left(\bar{\theta}_i, \rho\right) > 0$. But by Step 4, $\mu_i\left(\bar{\theta}_i, \rho_0\right) > \bar{\theta}_i$, and by Step 5, $\lim_{\rho \to 1} \mu_i\left(\bar{\theta}_i, \rho\right) = \bar{\theta}_i$; a contradiction.

The fact that preferences are congruent if and only if $\left|\bar{\theta}_i - \bar{\theta}_j\right| \leq a\left(\bar{\theta}_i + \bar{\theta}_j; \rho\right)$ for some function $a(.)$ comes from Step 1 and 2 and the symmetry of $\mu_i\left(\bar{\theta}_i, \rho\right)$ and $\mu_j\left(\bar{\theta}_i, \rho\right)$ with respect to 45 degree line. Step 5 and 6 imply that a is decreasing in ρ and tends to 0 when $\rho \to 1$. Step 4 imply that a is positive whenever $\bar{\theta}_i + \bar{\theta}_j \neq 0$ and 0 otherwise. ■

Lemma 2 *If $p \in P$ parametrizes the map $\boldsymbol{H}^p : \mathbb{R}^2 \to \mathbb{R}^2$, $\succeq$ is a partial order on P, and $C \subset \mathbb{R}^2$ such that for all $p \in P$, $(0,0) \in C$, $\boldsymbol{H}^p(C) \subset C$ and for all $p, p' \in P$, $p' \succeq p$ implies $\boldsymbol{H}^{p'}(\boldsymbol{c})(\leq, \geq)\boldsymbol{H}^p(\boldsymbol{c})$ for all $c \in C$. Then for all t, $\boldsymbol{c}^t(p)$ (as defined in proposition 1 using $\boldsymbol{H}^p$) is weakly increasing (in the order $(\leq, \geq)$) in p (in the order $\succeq$).*

Proof. Since $\boldsymbol{c}^1(p) = (0,0) \in C$, by induction, for all t, $\boldsymbol{c}^t(p) \in C$. Let $p' \succ p$. Obviously, $\boldsymbol{c}^1(p')(\leq, \geq)\boldsymbol{c}^1(p)$. Suppose that $\boldsymbol{c}^t(p')(\leq, \geq)\boldsymbol{c}^t(p)$ for some $t \geq 1$, then

$$\boldsymbol{c}^{t+1}(p') = \boldsymbol{H}^{p'}\left(\boldsymbol{c}^t(p')\right)(\leq, \geq)\boldsymbol{H}^p\left(\boldsymbol{c}^t(p')\right)(\leq, \geq)\boldsymbol{H}^p\left(\boldsymbol{c}^t(p)\right) = \boldsymbol{c}^{t+1}(p).$$

By induction, $\boldsymbol{c}^t(p')(\leq, \geq)\boldsymbol{c}^t(p)$ for all t. ■

Proof of proposition 4. Let $\boldsymbol{\theta}$ be a preference distribution. For all $\boldsymbol{m} \in \mathbb{R}^2$ and $\delta \in [0,1]$, $\boldsymbol{H}^{\boldsymbol{m},\delta}$ will refer to the function $\boldsymbol{H}$ defined in (1) for the preference distribution

$\boldsymbol{\theta}+\boldsymbol{m}$, and $(\boldsymbol{c}_t(\boldsymbol{m},\delta))_{t\in\mathbb{R}}$ to the corresponding equilibrium voting thresholds. In this proof, for all $k\in\{i,j\}$, f_k denotes the p.d.f. of θ_k.

Step 1: There exists $\boldsymbol{m}^o\in\mathbb{R}^2$ such that for all $\boldsymbol{m}(\geq,\leq)\boldsymbol{m}^o$, $\boldsymbol{\theta}+\boldsymbol{m}$ is polarized. From (1), as $\boldsymbol{m}\to(+\infty,-\infty)$, $\boldsymbol{H}^{\boldsymbol{m},\delta}(0,0)+E(\boldsymbol{\theta}+\boldsymbol{m})\to(0,0)$, so $\boldsymbol{H}^{\boldsymbol{m},\delta}(0,0)(\leq,\geq)(0,0)$.

Step 2: There exists $\boldsymbol{m}^o\in\mathbb{R}^2$ such that for all $\boldsymbol{m}(\geq,\leq)\boldsymbol{m}^o$, $\boldsymbol{H}^{\boldsymbol{m},1}$ has no fixed point. Let $(\boldsymbol{m}^n)_{n\geq1}$ be such that $\boldsymbol{m}^n\to(+\infty,-\infty)$. From (1), for all $\boldsymbol{c}$,

$$\begin{aligned}H_i^{\boldsymbol{m}^n,1}(\boldsymbol{c})-c_i=&-\int_{-\infty}^{c_j-m_j^n}\int_{-\infty}^{c_i-m_i^n}c_if(\boldsymbol{\theta})\,d\theta_id\theta_j-\int_{c_j-m_j^n}^{\infty}\int_{c_i-m_i^n}^{\infty}c_if(\boldsymbol{\theta})\,d\theta_id\theta_j \quad (12)\\&-\int_{-\infty}^{c_j-m_j^n}\int_{c_i-m_i^n}^{\infty}(m_i^n+\theta_i)f(\boldsymbol{\theta})\,d\theta_id\theta_j-\int_{c_j-m_j^n}^{\infty}\int_{-\infty}^{c_i-m_i^n}(m_i^n+\theta_i)f(\boldsymbol{\theta})\,d\theta_id\theta_j.\end{aligned}$$

Let us call $A^n(\boldsymbol{c})$, $B^n(\boldsymbol{c})$, $C^n(\boldsymbol{c})$ and $D^n(\boldsymbol{c})$ the four integrals in the order they appear on the right-hand side of (12). Observe first that if g is the p.d.f. of an integrable, real random variable, $\int_{-\infty}^{x}|xg(u)|\,du\to0$ as $x\to-\infty$. Therefore

$$\begin{aligned}\sup_{c_i\leq0}|A^n(\boldsymbol{c})|&\leq\sup_{c_i\leq0}\int_{-\infty}^{c_i-m_i^n}|c_i|f_i(\theta_i)\,d\theta_i\to_{n\to\infty}0, \quad (13)\\ \sup_{c_i\leq0}|B^n(\boldsymbol{c})|&\leq\int_{c_j-m_j^n}^{\infty}|c_i|f_j(\theta_j)\,d\theta_j,\\ \sup_{c_i\leq0}|D^n(\boldsymbol{c})|&\leq\int_{-\infty}^{-m_i^n}(|m_i^n|+|\theta_i|)f_i(\theta_i)\,d\theta_i\to_{n\to\infty}0.\end{aligned}$$

Moreover, for all $M>0$, there exists N such that for all $n\geq N$, $m_i^n>M$. So for all $n\geq N$,

$$\inf_{c_i\leq0}|C^n(\boldsymbol{c})|\geq\sup_{c_i\leq0}M\int_{-\infty}^{c_j-m_j^n}\int_{c_i-m_i^n}^{\infty}f(\boldsymbol{\theta})\,d\theta_id\theta_j-E(|\theta_i|)\to_{n\to\infty}M-E(|\theta_i|).$$

This shows that $\inf_{c_i\leq0}|C^n(\boldsymbol{c})|$ is unbounded as $n\to\infty$. If step 2 does not hold, then there exists a sequence $\boldsymbol{m}^n\to(+\infty,-\infty)$ such that for all $n\geq1$, $\boldsymbol{H}^{\boldsymbol{m}^n,1}$ has a fixed point $\boldsymbol{c}^n$. Substituting $H_i^n(\boldsymbol{c}^n)=c_i^n$ in (12), we obtain $A^n(\boldsymbol{c}^n)+B^n(\boldsymbol{c}^n)+C^n(\boldsymbol{c}^n)+D^n(\boldsymbol{c}^n)=0$. From step 1, we can assume that for all n, $\boldsymbol{\theta}+\boldsymbol{m}^n$ is polarized, so $c_i^n\leq0$. As shown above, $C^n(\boldsymbol{c}^n)$ is unbounded, $A^n(\boldsymbol{c}^n)$ and $D^n(\boldsymbol{c}^n)$ tend to 0 as $n\to\infty$, so $B^n(\boldsymbol{c}^n)$ must be unbounded. Since $c_j^n\geq0$, $c_j^n-m_j^n\to+\infty$, so if $B^n(\boldsymbol{c}^n)$ is unbounded, $c_j^n-m_j^n=o(|c_i^n|)$. Since $c_j^n\geq0$ and $m_j^n\leq0$ for n large, this implies $c_j^n=o(|c_i^n|)$. Likewise, $H_j^n(\boldsymbol{c}^n)=c_j^n$ implies that $c_i^n=o\left(\left|c_j^n\right|\right)$, a contradiction.

Step 3: There exists $\boldsymbol{m}^o\in\mathbb{R}^2$ such that for all $\boldsymbol{m}(\geq,\leq)\boldsymbol{m}^o$, $\boldsymbol{c}^\infty(\boldsymbol{m},\delta)$ diverges as $\delta\to1$. From step 1, we can assume that for all $\boldsymbol{m}(\geq,\leq)\boldsymbol{m}^o$, $\boldsymbol{\theta}+\boldsymbol{m}$ is polarized. From proposition

2, $\boldsymbol{c}^{\infty}(\boldsymbol{m},\delta)$ exists for all $\delta\in(0,1)$ and from proposition 3, $\boldsymbol{c}^{\infty}(\boldsymbol{m},\delta)$ is increasing in δ in the order $(\leq,\geq)$. Therefore, if $\boldsymbol{c}^{\infty}(\boldsymbol{m},\delta)$ does not diverge as $\delta\to 1$, $\boldsymbol{c}^{\infty}(\boldsymbol{m},\delta)\to\boldsymbol{c}^{\infty}(\boldsymbol{m},1)$ for some finite $\boldsymbol{c}^{\infty}(\boldsymbol{m},1)$ as $\delta\to 1$. By continuity of $\boldsymbol{H}^{\boldsymbol{m},\delta}$ in δ, $\boldsymbol{c}^{\infty}(\boldsymbol{m},1)$ must a fixed point of $\boldsymbol{H}^{\boldsymbol{m},1}$. Step 2 completes the argument.

Step 4: There exists $\boldsymbol{m}^{o}\in\mathbb{R}^{2}$ *such that for all* $\boldsymbol{m}(\geq,\leq)\boldsymbol{m}^{o}$, *there is no solution* c_i *to* $\lim_{c_j\to+\infty}H_i^{\boldsymbol{m},1}(c_i,c_j)=c_i$. Let $(\boldsymbol{m}^n)_{n\geq 1}$ be such that $\boldsymbol{m}^n\to(+\infty,-\infty)$. From (1)

$$\lim_{c_j\to+\infty}H_i^{\boldsymbol{m}^n,1}(c_i,c_j)-c_i=-\int_{-\infty}^{c_i-m_i^n}c_i f_i(\theta_i)\,d\theta_i-\int_{c_i-m_i^n}^{\infty}(m_i^n+\theta_i)\,f_i(\theta_i)\,d\theta_i, \tag{14}$$

for all c_i. Let us call $A^n(c_i)$ and $C^n(c_i)$ the two integrals in their order they appear on the right-hand side of (14). As argued in the proof of step 2 (see equation (13)), $\sup_{c_i\leq 0}|C^n(c_i)|\to_{n\to\infty}0$. Moreover,

$$\sup_{c_i\leq 0}|A^n(c_i)+m_i^n|\leq\int_{-\infty}^{-m_i^n}|m_i^n|\,f_i(\theta_i)\,d\theta_i+E(|\theta_i|)\to_{n\to\infty}E(|\theta_i|)\,. \tag{15}$$

Therefore, $\sup_{c_i\leq 0}|A^n(c_i)+m_i^n|$ is bounded and since $m_i^n\to-\infty$, $\inf_{c_i\leq 0}|A^n(c_i)|$ must be bounded away from 0. If step 4 does not hold, then there exists $(\boldsymbol{m}^n)_{n\in\mathbb{N}}$ and $(c_i^n)_{n\in\mathbb{N}}$ such that $\boldsymbol{m}^n\to(+\infty,-\infty)$ and for all $n\geq 1$, $c_i^n=\lim_{c_j\to+\infty}H_i^{\boldsymbol{m}^n,1}(c_i^n,c_j)$. Equation (14) implies then that $|A^n(c_i^n)|=|C^n(c_i^n)|$, which is impossible since $A^n(c_i^n)$ diverges and $C^n(c_i^n)\to 0$.

Step 5: There exists $\boldsymbol{m}^{o}\in\mathbb{R}^{2}$ *such that for all* $\boldsymbol{m}(\geq,\leq)\boldsymbol{m}^{o}$, *both* $c_i^{\infty}(\boldsymbol{m},\delta)$ *and* $c_j^{\infty}(\boldsymbol{m},\delta)$ *diverge as* $\delta\to 1$. Let $(\boldsymbol{m}^n)_{n\geq 1}$ be such that $\boldsymbol{m}^n\to(+\infty,-\infty)$. From step 2, there exist N such that either $c_i^{\infty}(\boldsymbol{m}^n,\delta)\to-\infty$ for all $n\geq N$ or $c_j^{\infty}(\boldsymbol{m}^n,\delta)\to\infty$ for all $n\geq N$. Suppose the latter to fix ideas. From step 1, we can assume that for all n, $\boldsymbol{\theta}+\boldsymbol{m}^n$ is polarized. So if $c_i^{\infty}(\boldsymbol{m}^n,\delta)$ does not diverge as $\delta\to 1$, it must have a finite limit $c_i^{\infty}(\boldsymbol{m},1)$. A standard continuity argument implies that necessarily, $\lim_{c_j\to+\infty}H_i^{\boldsymbol{m}^n,1}(c_i^{\infty}(\boldsymbol{m}^n,1),c_j)=c_i^{\infty}(\boldsymbol{m},1)$. Step 4 concludes the proof. ■

Proof of Proposition 5. We proceed by induction. The case $T=1$ is straightforward. Suppose the proposition holds up to some T. For all $t\leq T+1$, the continuation game after period t's outcome y^t is simply $\Gamma^{en}_{t-1,y^t,N}$. From the induction hypothesis, it has a unique equilibrium, and we denote its value $\boldsymbol{V}^{t-1}(y^t)$. Since player n votes as if she is pivotal, in every period $t\leq T+1$, her threshold strategy must be $c_n^t=\frac{\delta}{2}\left(V_n^{t-1}(-1)-V_n^{t-1}(1)\right)$. In particular, $\Gamma^{en}_{T+1,q,N}$ has a unique equilibrium. From the induction hypothesis, c_n^t is decreasing in n, and since $\theta_n(\varepsilon)$ is increasing in n for all ε, so the status quo stays in place if only if players M and $N-M+1$ disagree. Therefore,

$$V_n^t(1) - V_n^t(-1) = \delta \int_{c_M^t - v_M}^{c_{N-M+1}^t - v_{N-M+1}} \left(2\theta_n(\varepsilon) + \delta\left(V_n^{t-1}(1) - V_n^{t-1}(-1)\right)\right) f(\varepsilon)\, d\varepsilon. \quad (16)$$

Substituting the thresholds on both sides of (16), we get $\boldsymbol{c}^{t+1} = H_n(\boldsymbol{c}^t)$. The latter equation with the fact that c_n^t is decreasing in n, implies that c_n^{t+1} is also decreasing in n. When $M = \frac{N+1}{2}$, the bounds of the integral in (16) coincide, so $\boldsymbol{c}^t = (0, .., 0)$ for all t. ■

Proof of Corollary 3. Suppose first that $v_M = -v_{N-M+1} = v > 0$. Since the distribution of ε is symmetric, for all t, $c_M^t = -c_{N-M+1}^t$. From proposition 5, the sequence $\left(c_{N-M+1}^t\right)_{t\geq 1}$ is given recursively by $c_{N-M+1}^1 = 0$ and $c_{N-M+1}^{t+1} = H\left(c_{N-M+1}^t\right)$ where

$$H(c) = \int_{-c-v}^{c+v} (c + v - \varepsilon) f(\varepsilon)\, d\varepsilon = (c+v)(2F(c+v) - 1). \quad (17)$$

Therefore, from the proof of proposition 4, to show that $\left(c_M^\infty, c_{N-M+1}^\infty\right) \to (-\infty, +\infty)$ as $\delta \to 1$ for all $(v_M, v_{N-M+1})(\geq, \leq)(\sigma_\varepsilon, \sigma_\varepsilon)$, it suffices to show that H as defined in (17) has no fixed point on $\mathbb{R}^+$ for $v \geq \sigma_\varepsilon$ and $\delta = 1$. Using the symmetry of ε and Holder inequality, for all $c > 0$, $H(c) - c = -2\int_{c+v}^\infty (c+v) f(u)\, du + v > -2\int_0^\infty u f(u)\, du + v > -\sigma_\varepsilon + v$. ■

Proof of Proposition 6. All but the first and the last inequalities are proven in corollary 5. Let $\boldsymbol{H}^M$ denote the map defined in (5) for the supermajority M; i.e., $H_n^M(\boldsymbol{c}^t(M)) = \delta\int_{c_M^t - v_M}^{c_{N-M+1}^t - v_{N-M+1}} (c_n^t - v_n - \varepsilon) f(\varepsilon)\, d\varepsilon$. We will prove the remaining inequalities by induction. For $t = 1$, all players use zero threshold; therefore, the inequalities hold. Assume they hold up to some t; i.e.,

$$c_{N-M'+1}^t(M') \geq c_{N-M'+1}^t(M) \text{ and } c_{M'}^t(M') \leq c_{M'}^t(M). \quad (18)$$

Since $c_{M'}^{t+1}(M') = H_{M'}^{M'}(\boldsymbol{c}^t(M'))$ and $c_{M'}^{t+1}(M) = H_{M'}^M(\boldsymbol{c}^t(M))$, to prove $c_{M'}^{t+1}(M') \leq c_{M'}^{t+1}(M)$, it suffices to show that $H_{M'}^{M'}(\boldsymbol{c}^t(M')) \leq H_{M'}^M(\boldsymbol{c}^t(M))$. Corollary 5 and (18) imply $c_{M'}^t(M') \leq c_M^t(M)$. Since $v_{M'} > v_M$, $c_{M'}^t(M') - v_{M'} \leq c_M^t(M) - v_M$. A symmetric argument shows that $c_{N-M'+1}^t(M') - v_{N-M'+1} \geq c_{N-M+1}^t(M) - v_{N-M+1}$. Therefore,

$$H_{M'}^{M'}\left(\boldsymbol{c}^t(M')\right) - H_{M'}^M\left(\boldsymbol{c}^t(M)\right) = \delta \begin{pmatrix} \int_{c_{M'}^t(M') - v_{M'}}^{c_{N-M'+1}^t(M') - v_{N-M'+1}} (c_{M'}^t(M') - v_{M'} - \varepsilon) f(\varepsilon)\, d\varepsilon \\ - \int_{c_M^t(M) - v_M}^{c_{N-M+1}^t(M) - v_{N-M+1}} (c_{M'}^t(M) - v_{M'} - \varepsilon) f(\varepsilon)\, d\varepsilon \end{pmatrix}$$

$$\leq \delta \begin{pmatrix} \int_{c^t_{M'}(M')-v_{M'}}^{c^t_{N-M'+1}(M')-v_{N-M'+1}} \left(c^t_{M'}(M') - v_{M'} - \varepsilon\right) f(\varepsilon)\, d\varepsilon \\ -\int_{c^t_{M}(M)-v_{M}}^{c^t_{N-M+1}(M)-v_{N-M+1}} \left(c^t_{M'}(M') - v_{M'} - \varepsilon\right) f(\varepsilon)\, d\varepsilon \end{pmatrix}$$

$$= \delta \begin{pmatrix} \int_{c^t_{M'}(M')-v_{M'}}^{c^t_{M}(M)-v_{M}} \left(c^t_{M'}(M') - v_{M'} - \varepsilon\right) f(\varepsilon)\, d\varepsilon \\ +\int_{c^t_{N-M+1}(M)-v_{N-M+1}}^{c^t_{N-M'+1}(M')-v_{N-M'+1}} \left(c^t_{M'}(M') - v_{M'} - \varepsilon\right) f(\varepsilon)\, d\varepsilon \end{pmatrix} \leq 0,$$

where the first step follows from (18). By a symmetric argument $c^t_{N-M'+1}(M) \leq c^t_{N-M'+1}(M')$.
■

Proof of proposition 7. Let $\boldsymbol{U}\left(\Gamma^{en}_{T,q}\right)$ and $\boldsymbol{U}\left(\Gamma^{ex}_{T,q}\right)$ be the equilibrium payoff profiles (normalized per period) of $\Gamma^{en}_{T,q}$ and $\Gamma^{ex}_{T,q}$, respectively. Let $D(\boldsymbol{c})$ denote the event that $\theta_i - c_i$ and $\theta_j - c_j$ are of opposite sign. Since $\boldsymbol{c}^\infty$ is a fixed point of $\boldsymbol{H}$, equation (1) implies that

$$E(\theta_i + \theta_j | D(\boldsymbol{c}^\infty)) = -\frac{1 - \delta \Pr(D(\boldsymbol{c}^\infty))}{\delta \Pr(D(\boldsymbol{c}^\infty))}\left(c^\infty_i + c^\infty_j\right). \tag{19}$$

Let $A^+(\boldsymbol{c})$ and $A^-(\boldsymbol{c})$ denote the event that $\theta_i - c_i$ and $\theta_j - c_j$ are both positive and both negative, respectively, and let $\pi_t(T)$ be the ex-ante probability that the status quo is 1 at the beginning of period t in $\Gamma^{en}_{T,q}$. Since c^t converges, for all $a \geq 0$, $\lim_{T\to\infty} \pi_{T-a}(T)$ exists. Hence $\frac{\sum_{t=1}^T \delta^t \pi_{T-t}(T)}{\sum_{t=1}^T \delta^t}$ converges to some π as $T \to \infty$. Likewise, since for all t, $\lim_{T\to\infty} \boldsymbol{c}^{T-t} = \boldsymbol{c}^\infty$,

$$\begin{aligned} &U_i\left(\Gamma^{en}_{T,q}\right) = \tfrac{1}{\sum_{t=1}^T \delta^t} \textstyle\sum_{t=1}^T \delta^t \left(2\pi_{T-t}(T) - 1\right) E\left(\theta_i | D\left(\boldsymbol{c}^{T-t}\right)\right) \Pr\left(D\left(\boldsymbol{c}^{T-t}\right)\right) \\ &+ \tfrac{1}{\sum_{t=1}^T \delta^t} \textstyle\sum_{t=1}^T \delta^t \left[E\left(\theta_i | A^+\left(\boldsymbol{c}^{T-t}\right)\right) \Pr\left(A^+\left(\boldsymbol{c}^{T-t}\right)\right) - E\left(\theta_i | A^-\left(\boldsymbol{c}^{T-t}\right)\right) \Pr\left(A^-\left(\boldsymbol{c}^{T-t}\right)\right)\right] \\ &\to_{T\to\infty} (2\pi - 1) E(\theta_i | D(\boldsymbol{c}^\infty)) \Pr(D(\boldsymbol{c}^\infty)) + E(\theta_i | A^+(\boldsymbol{c}^\infty)) \Pr(A^+(\boldsymbol{c}^\infty)) - E(\theta_i | A^-(\boldsymbol{c}^\infty)) \Pr(A^-(\boldsymbol{c}^\infty)). \end{aligned} \tag{20}$$

Case (ii): If c^∞_i and c^∞_j have the same sign, then case (i) applies. If they have opposite signs, the uniform order condition implies that $c^\infty_i \leq 0 \leq c^\infty_j$.[25] Suppose first that $c^\infty_i + c^\infty_j < 0$. Equation 19 implies then that $E(\theta_i + \theta_j | D(\boldsymbol{c}^\infty)) > 0$. Since $\boldsymbol{\theta}$ has full support, $2\pi - 1 < 1$, this implies that

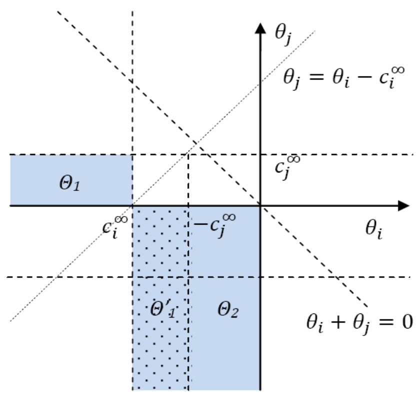
$$\begin{aligned} &\lim_{T\to\infty} W\left(\Gamma^{en}_{T,q}\right) < E(\theta_i + \theta_j | D(\boldsymbol{c}^\infty)) \Pr(D(\boldsymbol{c}^\infty)) \\ &+ E(\theta_i + \theta_j | A^+(\boldsymbol{c}^\infty)) \Pr(A^+(\boldsymbol{c}^\infty)) - E(\theta_i + \theta_j | A^-(\boldsymbol{c}^\infty)) \Pr(A^-(\boldsymbol{c}^\infty)). \end{aligned} \tag{21}$$

[25] For any $c \in \mathbb{R}$, $H_j(c,c) - H_i(c,c) = \int_{\theta_i \geq c}\int_{\theta_j \leq c} (\theta_i - \theta_j) f(\boldsymbol{\theta})\, d\boldsymbol{\theta} + \int_{\theta_i \leq c}\int_{\theta_j \geq c} (\theta_i - \theta_j) f(\boldsymbol{\theta})\, d\boldsymbol{\theta}$, and the right-hand side is positive by the uniform order condition. Since $\frac{\partial H_i}{\partial c_i} \geq 0$, $\frac{\partial H_j}{\partial c_i} \leq 0$ (see lemma 1), and $H_j(c,c) \geq H_i(c,c)$, for all $c_i < c_j$, $H_i(c_i, c_j) \leq H_i(c_j, c_j) \leq H_j(c_j, c_j) \leq H_j(c_i, c_j)$. Therefore, a simple induction argument shows that for all t, $c^t_i \leq c^t_j$ and $c^\infty_i \leq c^\infty_j$.

With an exogenous status quo $q=1$,

$$\begin{aligned}\lim_{T\to\infty} W\left(\Gamma^{ex}_{T,1}\right) = E\left(\theta_i+\theta_j|D\left(\mathbf{0}\right)\right)\Pr\left(D\left(\mathbf{0}\right)\right)\\ +E\left(\theta_i+\theta_j|A^+\left(\mathbf{0}\right)\right)\Pr\left(A^+\left(\mathbf{0}\right)\right)-E\left(\theta_i+\theta_j|A^-\left(\mathbf{0}\right)\right)\Pr\left(A^-\left(\mathbf{0}\right)\right).\end{aligned} \tag{22}$$

Comparing the right-hand sides of (21) and (22), we see that the two expectations differ only on the events $\Theta_1=\left\{\boldsymbol{\theta}:\theta_i\leq c_i^\infty,\theta_j\in\left[0,c_j^\infty\right]\right\}$ and $\Theta_2=\{\boldsymbol{\theta}:\theta_j\leq 0,\theta_i\in[c_i^\infty,0]\}$ (see the figure below).



Therefore,

$$\lim_{T\to\infty}\left(W\left(\Gamma^{ex}_{T,1}\right)-W\left(\Gamma^{en}_{T,q}\right)\right)>2E\left(\theta_i+\theta_j|\Theta_1\right)\Pr\left(\Theta_1\right)-2E\left(\theta_i+\theta_j|\Theta_2\right)\Pr\left(\Theta_2\right). \tag{23}$$

Let Θ_1' be the symmetric of Θ_1 with respect to the line $\{\boldsymbol{\theta}:\theta_j=\theta_i+c_i^\infty\}$, as in the figure above. Since $c_i^\infty+c_j^\infty<0$, $\Theta_1'\subset\Theta_2$, and for all $\boldsymbol{\theta}\in\Theta_1\cup\Theta_2$, $\theta_i+\theta_j<0$. Let $\boldsymbol{\theta}'\in\Theta_1'$ be the symmetric of $\boldsymbol{\theta}\in\Theta_1$. Since the preference distribution is uniformly ordered one can see from the figure above that $f\left(\boldsymbol{\theta}'\right)\geq f\left(\boldsymbol{\theta}\right)$. Therefore,

$$E\left(\theta_i+\theta_j|\Theta_2\right)\Pr\left(\Theta_2\right)\leq E\left(\theta_i+\theta_j|\Theta_1'\right)\Pr\left(\Theta_1'\right)\leq E\left(\theta_i+\theta_j|\Theta_1\right)\Pr\left(\Theta_1\right).$$

Together with (23), the inequality above shows that $W\left(\Gamma^{ex}_{T,1}\right)>W\left(\Gamma^{en}_{T,q}\right)$ for T sufficiently large. A symmetric argument shows that if $c_i^\infty+c_j^\infty>0$, $W\left(\Gamma^{ex}_{T,-1}\right)>W\left(\Gamma^{en}_{T,q}\right)$ for T sufficiently large.

Case (i): Suppose w.l.o.g. that c_i^∞ and c_j^∞ are both nonnegative. Since $\boldsymbol{c}^\infty$ is a fixed point of $\boldsymbol{H}$, equation (1) implies that $E\left(\theta_i|D\left(\boldsymbol{c}^\infty\right)\right)\leq 0$. Substituting this in (20), we get

$$\begin{aligned}\lim_{T\to\infty} U_i\left(\Gamma^{en}_{T,q}\right)\leq -E\left(\theta_i|D\left(\boldsymbol{c}^\infty\right)\right)\Pr\left(D\left(\boldsymbol{c}^\infty\right)\right)\\ +E(\theta_i|A^+(\boldsymbol{c}^\infty))\Pr(A^+(\boldsymbol{c}^\infty))-E(\theta_i|A^-(\boldsymbol{c}^\infty))\Pr(A^-(\boldsymbol{c}^\infty)).\end{aligned} \tag{24}$$

With an exogenous status quo $q = -1$,

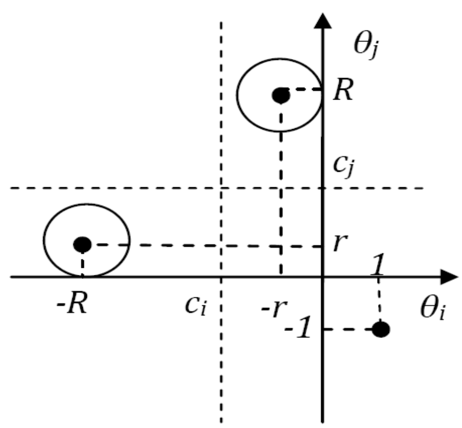
$$U_i\left(\Gamma^{ex}_{T,-1}\right) = -E\left(\theta_i | D\left(\mathbf{0}\right)\right)\Pr\left(D\left(\mathbf{0}\right)\right) + E\left(\theta_i | A^+\left(\mathbf{0}\right)\right)\Pr\left(A^+\left(\mathbf{0}\right)\right) - E\left(\theta_i | A^-\left(\mathbf{0}\right)\right)\Pr\left(A^-\left(\mathbf{0}\right)\right). \tag{25}$$

Comparing the right-hand sides of (24) and (25), we see that the two expectations differ only on the events $\Theta_1 = \left\{\boldsymbol{\theta} : \theta_i > 0, \theta_j \in \left(0, c_j^\infty\right)\right\}$ and $\Theta_2 = \left\{\boldsymbol{\theta} : \theta_j \geq c_j^\infty, \theta_i \in \left(0, c_i^\infty\right)\right\}$. Hence,

$$\lim_{T\to\infty}\left(U_i\left(\Gamma^{ex}_{T,-1}\right) - U_i\left(\Gamma^{en}_{T,q}\right)\right) \geq 2E\left(\theta_i|\Theta_1\right)\Pr\left(\Theta_1\right) + 2E\left(\theta_i|\Theta_2\right)\Pr\left(\Theta_2\right).$$

Since $\boldsymbol{c}^\infty \neq (0,0)$, $\Theta_1 \cup \Theta_2$ has positive measure. Moreover, $\theta_i > 0$ for all $\boldsymbol{\theta} \in \Theta_1 \cup \Theta_2$, so $\lim_{T\to\infty} U_i\left(\Gamma^{ex}_{T,-1}\right) > \lim_{T\to\infty} U_i\left(\Gamma^{en}_{T,q}\right)$. Hence, we have shown that whenever $\boldsymbol{c}^\infty \neq (0,0)$, there exists q such that for T sufficiently large, $U_k\left(\Gamma^{ex}_{T,-1}\right) > U_k\left(\Gamma^{en}_{T,q}\right)$ for all $k \in \{i,j\}$. ■

Example 4 *Let $0 < r < R$ and consider the following preference distribution: with probability $1/2$, θ_i and θ_j are independently distributed with $\theta_i \sim N(1,\varepsilon)$ and $\theta_j \sim N(-1,\varepsilon)$; with probability $1/4$, $\boldsymbol{\theta}$ is uniformly distributed on a circle of center $(-R,r)$ and radius r; with probability $1/4$, $\boldsymbol{\theta}$ is uniformly distributed on a circle of center $(-r,R)$ and radius r. The corresponding p.d.f. are denoted f^1, f^2 and f^3, respectively. Conditional on (static) disagreement and on the preferences being drawn from f^1, player i prefers 1 and player j prefers -1. However, with probability $1/2$ a preference reversal happens: under f^2 and f^3, conditional on (static) disagreement, player i prefers -1 and player j prefers 1.*



Let Γ^s_T be a T repetition of any static mechanism. One can show that for all (r,R) such that

$$\frac{1}{2-\delta} + r < R < 2 - \left(\frac{8}{\delta} + 1\right) r, \tag{26}$$

the equilibrium $\boldsymbol{c}^t$ are such that for all $t \geq 2$, the disagreement region $\left\{\boldsymbol{\theta} : \left(\theta_i - c_i^t\right)\left(\theta_j - c_j^t\right) \leq 0\right\}$

does not intersect the support of f_2 and f_3. Let Γ_T^s be the T repetition of any static mechanism. In the game Γ_T^s, nonpartisan players almost always disagree as $\varepsilon \to 0$. However, in the game $\Gamma_{T,q}^{en}$, conditionally on the preferences being drawn from f_2 or f_3, partisan players vote unanimously for the socially optimal policy. Formally, one can show that

$$\lim_{\varepsilon \to 0, T \to \infty} \left(W\left(\Gamma_{T,q}^{en}\right) - W\left(\Gamma_T^s\right)\right) = \frac{1}{4}\left(R + r\right). \tag{27}$$

If $\delta \in (0,1)$, $R \in \left[\frac{1}{2-\delta}, 2\right]$ and r is sufficiently small, (26) is satisfied and (27) is positive.

Proof of example 4. In what follows, we let $\varepsilon \to 0$ and omit to mention "for ε sufficiently small" for brevity. We first show that for all r and R that satisfy (26), $\boldsymbol{c}^\infty$ is independent of r and R. Let $\boldsymbol{H}$ be the map defined in (1) with p.d.f. $f = \frac{2f^1+f^2+f^3}{4}$.By construction,

$$\boldsymbol{c}^2 = \boldsymbol{H}\left(\boldsymbol{0}\right) \to \delta\left(-\frac{1}{2} + \frac{R+r}{4}, \frac{1}{2} - \frac{R+r}{4}\right). \tag{28}$$

Simple algebra shows that $\frac{2\delta+(4-\delta)r}{4+\delta} < \frac{1}{2-\delta} + r$ so (26) implies $\frac{2\delta+(4-\delta)r}{4+\delta} < R < 2 - \left(\frac{8}{\delta}+1\right)r$. Substituting these inequalities in (28), we get that for $t = 2$

$$\left(-R+r, R-r\right)\left(\leq,\geq\right)\boldsymbol{c}^t\left(\leq,\geq\right)\left(-2r, 2r\right). \tag{29}$$

Hence, only f^1 has positive weight on the disagreement region when using voting thresholds $\boldsymbol{c}^2$, so $\boldsymbol{H}\left(\boldsymbol{c}^2\right)$ can be evaluated using $f = \frac{f^1}{2}$. Moreover, from proposition 2, $\boldsymbol{c}^t$ is increasing in t in the order $(\leq,\geq)$. So for all $t \geq 2$, $\left(c_i^t, c_j^t\right)(\leq,\geq)(-2r,2r)$. If we show furthermore that $(-R-r, R+r)(\leq,\geq)(\boldsymbol{c}^\infty)$, then we can conclude that (29) holds for all t, which implies that $\boldsymbol{c}^\infty$ does not depend on r and R and can be computed via $\boldsymbol{H}$ using the measure $f = \frac{f^1}{2}$.

From what precedes, to prove the first point, it suffices to show that the fixed points $\boldsymbol{c}^f$ of $\boldsymbol{H}$ evaluated using the p.d.f. $f = \frac{f^1}{2}$ (or equivalently the p.d.f. f^1 and $\delta/2$ instead of δ) satisfy $(-R+r, R-r)(\leq,\geq)\boldsymbol{c}^f$. As shown in the proof of proposition 2,

$$\boldsymbol{c}^f\left(\geq,\leq\right)\left(\frac{\int|\theta_i| f\left(\boldsymbol{\theta}\right)d\boldsymbol{\theta}}{1-\delta/2}, \frac{\int|\theta_j| f\left(\boldsymbol{\theta}\right)d\boldsymbol{\theta}}{1-\delta/2}\right) \to \left(\frac{-1}{2-\delta}, \frac{1}{2-\delta}\right).$$

To conclude the proof, observe that (26) implies that $\frac{1}{2-\delta} \leq R - r$.

We now prove (27) for all r and R satisfying (26). Using the symmetry of the distribution and the notations of proposition 7, for all $c \in \mathbb{R}$, $E\left(\theta_i + \theta_j | D\left(-c, c\right)\right) = 0$. Since any static bargaining protocol implements policy 1 on $A^+(\mathbf{0})$ and -1 on $A^-(\mathbf{0})$, for all q,

$$\begin{aligned} W\left(\Gamma^{en}_{T\to\infty,q}\right) &= E\left(\theta_i+\theta_j|A^+(\boldsymbol{c}^\infty)\right)\Pr\left(A^+(\boldsymbol{c}^\infty)\right) - E\left(\theta_i+\theta_j|A^-(\boldsymbol{c}^\infty)\right)\Pr\left(A^-(\boldsymbol{c}^\infty)\right), \\ W\left(\Gamma^s_T\right) &= E\left(\theta_i+\theta_j|A^+(\mathbf{0})\right)\Pr\left(A^+(\mathbf{0})\right) - E\left(\theta_i+\theta_j|A^-(\mathbf{0})\right)\Pr\left(A^-(\mathbf{0})\right). \end{aligned}$$

Comparing the two expressions above, if we define $\Theta_i = \left\{\boldsymbol{\theta} : \theta_i \le c_i^\infty \text{ or } \theta_i \ge 0, \theta_j \in \left[0, c_j^\infty\right]\right\}$ and $\Theta_j = \left\{\boldsymbol{\theta} : \theta_i \in \left[c_i^\infty, 0\right], \theta_j \le 0 \text{ or } \theta_j \ge c_j^\infty\right\}$, we get

$$W\left(\Gamma^{en}_{T\to\infty,q}\right) - W\left(\Gamma^s_T\right) = -E\left(\theta_i+\theta_j|\Theta_i\right)\Pr\left(\Theta_i\right) + E\left(\theta_i+\theta_j|\Theta_j\right)\Pr\left(\Theta_j\right).$$

The latter equality implies (27) because as $\varepsilon \to 0$, f_1 has a vanishing weight on $\Theta_i \cup \Theta_j$, while f^2 has all its weight on Θ_i and f^3 has all its weight on Θ_j. ■

Example 5 (when dictatorship of j is better for i than the endogenous status quo) *Let $\theta_j = \mu + \varepsilon$ and $\theta_i = -\mu + \varepsilon$, with $\mu > 0$ and $\varepsilon \sim N(0,1)$. We will show that there exists some $\bar{\delta} < 0$ such that for all $\delta > \bar{\delta}$ and $\mu \in (0.4, 0.5)$ player i prefers $\Gamma^d_{T,j}$ to $\Gamma^{en}_{T,q}$ for T sufficiently large.*

Proof. By symmetry, for all t, $c_j^t = -c_i^t \doteq c^t$. The preference distribution is polarized, so from proposition 2, $\boldsymbol{c}^\infty$ exists.

Let us first show that for all $\mu > 0.4$, $c^\infty \to \infty$ as $\delta \to 1$. Using the symmetry of the normal distribution, at $\delta = 1$, $\boldsymbol{c}^\infty = \boldsymbol{H}(\boldsymbol{c}^\infty)$ can be rewritten as

$$\left(2\Phi\left(c^\infty+\mu\right)-1\right)\left(c^\infty+\mu\right) = c^\infty \Leftrightarrow 2\left(\Phi\left(c^\infty+\mu\right)-1\right)\left(c^\infty+\mu\right) = \mu. \tag{30}$$

By plotting $x \to 2(\Phi(x)-1)x$, one can see that it is bounded by 0.4, so for $\mu > 0.4$, (30) has no solution. As shown in the proof of proposition 4, this implies that $c^\infty \to \infty$ as $\delta \to 1$.

Using the notations and the argument in the proof of proposition 7 (see equation (20)),

$$\begin{aligned} &U_i\left(\Gamma^{en}_{T,q}\right) - U_i\left(\Gamma^d_{T,j}\right) = \\ &= \frac{1}{\sum_{t=1}^T \delta^t}\sum_{t=1}^T \delta^t \left[\begin{array}{c} 2\pi_{T-t}(T)\,E\left(\theta_i|\theta_i\in[-c^t,\mu]\right)\Pr\left(\theta_i\in[-c^t,\mu]\right) \\ +2\left(\pi_{T-t}(T)-1\right)E\left(\theta_i|\theta_i\in[\mu,\mu+c^t]\right)\Pr\left(\theta_i\in[\mu,\mu+c^t]\right) \end{array}\right] \to_{T\to\infty} \\ &2\pi E\left(\theta_i|\theta_i\in[-c^\infty,2\mu]\right)\Pr\left(\theta_i\in[-c^\infty,2\mu]\right) + 2(\pi-1)E\left(\theta_i|\theta_i\in[2\mu,2\mu+c^\infty]\right)\Pr\left(\theta_i\in[2\mu,2\mu+c^\infty]\right). \end{aligned}$$

Since $E\left(\theta_i|\theta_i\in[2\mu,2\mu+c^\infty]\right) \ge 0$,

$$\lim_{T\to\infty} U_i\left(\Gamma^{en}_{T,q}\right) - U_i\left(\Gamma^d_{T,j}\right) \le 2E\left(\theta_i|\theta_i\in[-c^\infty,2\mu]\right)\Pr\left(\theta_i\in[-c^\infty,2\mu]\right) = 2\int_{-c^\infty}^{2\mu}\theta_i\phi\left(\theta_i-\mu\right)d\theta_i. \tag{31}$$

For all $\mu > 0.4$, $c^{\infty} \to \infty$ as $\delta \to 1$, so using Mill's ratio formula, the right-hand side of (31) tends to $2\left(\mu\Phi\left(\mu\right) - \phi\left(\mu\right)\right)$ as $\delta \to 1$. By plotting $\mu \to 2\left(\mu\Phi\left(\mu\right) - \phi\left(\mu\right)\right)$, we see that it is increasing for $\mu \geq 0$ and negative at $\mu = 0.5$. Hence, we have shown that for all $\mu \in (0.4, 0.5)$, there exists $\overline{\delta}$, such that for all $\delta \geq \overline{\delta}$, for T sufficiently large $U_i\left(\Gamma^{en}_{T,q}\right) < U_i\left(\Gamma^{d}_{T,j}\right)$. ■

Proof of proposition 8. Let $\left(\boldsymbol{c}^t\left(M\right)\right)_{t\leq T}$ be the equilibrium thresholds in $\Gamma^{en}_{T,q,N}$ with supermajority M. Let $U^t_n\left(\left(\overline{\boldsymbol{c}}^t\left(M\right)\right)_{t\geq 1}\right)$ be the expected utility of player n in period t of $\Gamma^{en}_{T,q,N}$. If $\pi\left(t, T, q\right)$ denotes the probability that the status quo is 1 at the beginning of period t, if we denote $e^t_M = c^t_M\left(M\right) - v_M$ and $e^t_{N-M+1} = c^t_{N-M+1}\left(M\right) - v_{N-M+1}$, then

$$\begin{aligned} U^t_n\left(\left(\overline{\boldsymbol{c}}^t\left(M\right)\right)_{t\geq 1}\right) &= \left(2\pi\left(t,T,q\right) - 1\right) E\left(\theta_n | e_M \leq \varepsilon \leq e_{N-M+1}\right) \Pr\left(e_M \leq \varepsilon \leq e_{N-M+1}\right) \\ &\quad + E\left(\theta_n | \varepsilon \geq e_{N-M+1}\right) \Pr\left(\varepsilon \geq e_{N-M+1}\right) - E\left(\theta_n | \varepsilon \leq e_M\right) \Pr\left(\varepsilon \leq e_M\right). \end{aligned} \tag{32}$$

By symmetry, for all t, $e^t_{N-M+1} = -e^t_M$, for all ε, $\theta_n\left(\varepsilon\right) + \theta_{N-n+1}\left(\varepsilon\right) = 2\varepsilon$ and for all $e \in \mathbb{R}$, $E\left(\theta_n + \theta_{N-n+1} | -e \leq \varepsilon \leq e\right) = 0$. So (32) implies

$$\left(U^t_n + U^t_{N-n+1}\right)\left(\left(\overline{\boldsymbol{c}}^t\left(M\right)\right)_{t\geq 1}\right) = \int_{e^t_{N-M+1}}^{\infty} 2\varepsilon f\left(\varepsilon\right) d\varepsilon - \int_{-\infty}^{e^t_M} 2\varepsilon f\left(\varepsilon\right) d\varepsilon \tag{33}$$

From Propositions 5 and 6, for all t, e^t_{N-M+1} is positive and increasing in M while e^t_M is negative and decreasing in M. Therefore, (33) is decreasing in M. ■

The following lemma will be used in the proof of proposition 9.

Lemma 3 *Let $\Gamma_q\left(\pi, \boldsymbol{V}\left(1\right), \boldsymbol{V}\left(-1\right), \boldsymbol{\theta}\right)$ be game in which in the first period players play $\Gamma^{en}_{1,q}$, and in the second period, with probability $\left(1-\pi\right)$ the game stops and players get terminal payoff $\boldsymbol{V}\left(q\right)$, and with probability π, players are called to play $\Gamma_y\left(\pi, \boldsymbol{V}\left(1\right), \boldsymbol{V}\left(-1\right), \boldsymbol{\theta}\right)$, where y is outcome of the first period game $\Gamma^{en}_{1,q}$. For almost all $\boldsymbol{\theta}$, $\Gamma_q\left(\pi, \boldsymbol{V}\left(1\right), \boldsymbol{V}\left(-1\right), \boldsymbol{\theta}\right)$ has a unique equilibrium. In each period, player use constant threshold strategies $\boldsymbol{c} = \frac{\delta(1-\pi)}{2}\left(\boldsymbol{V}\left(-1\right) - \boldsymbol{V}\left(1\right)\right)$. If c_i and c_j have different signs, the value of $\Gamma_q\left(\pi, \boldsymbol{V}\left(1\right), \boldsymbol{V}\left(-1\right), \boldsymbol{\theta}\right)$ for player k is $\frac{\theta_k q + \delta(1-\pi)V_k(q)}{1-\delta\pi}$.*

Proof. Notice that the game is strategically equivalent to the infinite horizon game in which, in each period t, player k gets $\theta_k y^t + \delta\left(1-\pi\right) V_k\left(y^t\right)$ with certainty, where y^t is the outcome of the vote in period t, and players use the discount factor $\delta\pi$. Therefore, the set of possible intertemporal payoff from $\Gamma_q\left(\pi, \boldsymbol{V}\left(1\right), \boldsymbol{V}\left(-1\right), \boldsymbol{\theta}\right)$ is included in $\left[\frac{-\theta_k + \delta(1-\pi)V_k(-1)}{1-\delta\pi}, \frac{\theta_k + \delta(1-\pi)V_k(1)}{1-\delta\pi}\right]$. Assume for concreteness that $\theta_k + \delta\left(1-\pi\right) V_k\left(1\right) > -\theta_k + \delta\left(1-\pi\right) V_k\left(-1\right)$. In any subgame where 1 is the status quo, player k can secure his highest possible intertemporal payoff by always voting for 1, so the outcome must be 1 in all subsequent periods. Therefore, in any period t, stage undomination implies that voter k must vote for 1. Likewise, he will vote for -1 when the reverse inequality holds. Since $\boldsymbol{\theta}$ has full support, this uniquely determines the voting threshold. ■

Proof of proposition 9. We proceed by induction. Proposition 9 for $L=1$ follows from lemma 3 for $\boldsymbol{V}(1)=\boldsymbol{V}(-1)=(0,0)$. Suppose it holds up to some T. Stage undomination implies that in the first phase of $\Gamma^{en}_{L+1,q}$, players use the same strategy as in the equilibrium of $\Gamma_q\left(\pi^{L+1},\boldsymbol{V}\left(\Gamma^{en}_{L,1}\right),\boldsymbol{V}\left(\Gamma^{en}_{L,-1}\right),\boldsymbol{\theta}^{L+1}\right)$ (see the notations of lemma 3) where $\boldsymbol{V}\left(\Gamma^{en}_{L,q}\right)$ is the continuation payoff of $\Gamma^{en}_{L,q}$, uniquely defined by the induction hypothesis. Lemma 3 implies then that Γ^{L+1}_q has a unique equilibrium. Moreover, for all $l\le L$, the voting threshold in phase $l+1$ are constant and given by

$$\boldsymbol{c}^{l+1}\left(\pi^{l+1}\right)=\delta\left(1-\pi^{l+1}\right)\frac{\boldsymbol{V}\left(\Gamma^{en}_{l,-1}\right)-\boldsymbol{V}\left(\Gamma^{en}_{l,1}\right)}{2}. \tag{34}$$

Since the status quo matters only when players vote for different alternatives, using lemma 3, we obtain

$$\begin{aligned}
&V_k\left(\Gamma^{en}_{l,-1}\right)-V_k\left(\Gamma^{en}_{l,1}\right) \qquad (35)\\
&=\int_0^1\left(\begin{array}{c}\int_{c^l_j(\pi)}^{\infty}\int_{-\infty}^{c^l_i(\pi)}\left(\begin{array}{c}V_k\left(\Gamma_{-1}\left(\pi,\boldsymbol{V}\left(\Gamma^{en}_{l-1,-1}\right),\boldsymbol{V}\left(\Gamma^{en}_{l-1,1}\right),\boldsymbol{\theta}\right)\right)\\-V_k\left(\Gamma_{1}\left(\pi,\boldsymbol{V}\left(\Gamma^{en}_{l-1,-1}\right),\boldsymbol{V}\left(\Gamma^{en}_{l-1,1}\right),\boldsymbol{\theta}\right)\right)\end{array}\right)f(\boldsymbol{\theta})\,d\theta_i d\theta_j\\+\int_{-\infty}^{c^l_j(\pi)}\int_{c^l_i(\pi)}^{\infty}\left(\begin{array}{c}V_k\left(\Gamma_{-1}\left(\pi,\boldsymbol{V}\left(\Gamma^{en}_{l-1,-1}\right),\boldsymbol{V}\left(\Gamma^{en}_{l-1,1}\right),\boldsymbol{\theta}\right)\right)\\-V_k\left(\Gamma_{1}\left(\pi,\boldsymbol{V}\left(\Gamma^{en}_{l-1,-1}\right),\boldsymbol{V}\left(\Gamma^{en}_{l-1,1}\right),\boldsymbol{\theta}\right)\right)\end{array}\right)f(\boldsymbol{\theta})\,d\theta_i d\theta_j\end{array}\right)g(\pi)\,d\pi\\
&=\int_0^1\left(\begin{array}{c}\int_{c^l_j(\pi)}^{\infty}\int_{-\infty}^{c^l_i(\pi)}\left(\frac{-2\theta_k+\delta(1-\pi)\left(V_k\left(\Gamma^{en}_{l-1,-1}\right)-V_k\left(\Gamma^{en}_{l-1,1}\right)\right)}{1-\delta\pi}\right)f(\boldsymbol{\theta})\,d\theta_i d\theta_j\\+\int_{-\infty}^{c^l_j(\pi)}\int_{c^l_i(\pi)}^{\infty}\left(\frac{-2\theta_k+\delta(1-\pi)\left(V_k\left(\Gamma^{en}_{l-1,-1}\right)-V_k\left(\Gamma^{en}_{l-1,1}\right)\right)}{1-\delta\pi}\right)f(\boldsymbol{\theta})\,d\theta_i d\theta_j\end{array}\right)g(\pi)\,d\pi.
\end{aligned}$$

Substituting (34) in both sides of (35), we get

$$c^{l+1}_k\left(\pi^{l+1}\right)=\left(1-\pi^{l+1}\right)\int_0^1\frac{H_k\left(\boldsymbol{c}^l(\pi)\right)}{1-\pi\delta}g(\pi)\,d\pi, \tag{36}$$

which shows that for all l, $\frac{c^l_k(\pi)}{(1-\pi)}$ is constant in π. Substituting $\hat{c}^l_k=\frac{c^l_k(\pi)}{(1-\pi)}$ in both sides of (36), we obtain $\hat{\boldsymbol{c}}^{l+1}=\boldsymbol{G}\left(\hat{\boldsymbol{c}}^l\right)$. ■

References

[1] Alesina, A. and A. Drazen, 1991. Why Are Stabilizations Delayed? American Economic Review, 81(5), pp. 1170-88.

[2] Barbera, S. and M. Jackson, 2010. The Role of a Status Quo and Compromise in Dynamic Voting. Unpublished manuscript.

[3] Baron, D., 1991. Majoritarian Incentives, Pork Barrel Programs, and Procedural Control. American Journal of Political Science, 35(1), pp. 57–90.

[4] Baron, D., 1996. A Dynamic Theory of Collective Goods Programs. The American Political Science Review, 90 (2), pp. 316-330

[5] Baron, D., D. Diermeier, and P. Fong, 2007. Policy Dynamics and Inefficiency in a Parliamentary Democracy with Proportional Representation. Unpublished manuscript.

[6] Baron, D. and M. Herron, 2003. A Dynamic Model of Multidimensional Collective Choice. Computational Models of Political Economy, Ken Kollman, John H. Miller, and Scott E. Page, eds., MIT Press, 2003, pp. 13-47.

[7] Baron D. and E. Kalai, 1993. The Simplest Equilibrium of a Majority-Rule Division Game. Journal of Economic Theory, 61(2), pp. 290-301.

[8] Battaglini, M. and S. Coate, 2007. Inefficiency in Legislative Policymaking: A Dynamic Analysis. American Economic Review, 97(1), pp. 118-149.

[9] Battaglini, M. and S. Coate, 2008. A Dynamic Theory of Public Spending, Taxation, and Debt. American Economic Review, 98(1), pp. 201–236.

[10] Battaglini, M. and T. Palfrey, 2007. The Dynamics of Distributive Politics. Unpublished manuscript.

[11] Benabou, R., 2000. Unequal Societies: Income Distribution and the Social Contract. American Economic Review, 90(1), pp. 96–129.

[12] Bernheim, D., A. Rangel, and L. Rayo, 2006. The Power of the Last Word in Legislative Policy Making. Econometrica, 74(5), pp. 1161-90.

[13] Buchanan, J., and G. Tullock, 1962. The Calculus of Consent. Ann Arbor: University of Michigan Press.

[14] Casella, A., 2005. Storable votes. Games and Economic Behavior 51, pp. 391–419.

[15] Cho, S., 2005. A Dynamic Model of Parliamentary Democracy. Unpublished manuscript.

[16] Coate, S. and S. Morris, 1999. Policy Persistence. American Economic Review, 89(5), pp. 1327-1336.

[17] Diermeier, D., and P. Fong, 2007a. Legislative Bargaining with Reconsideration. Unpublished manuscript.

[18] Diermeier, D. and P. Fong, 2007b. Policy Persistence in Multi-Party Parliamentary Democracies. Unpublished manuscript.

[19] Duggan, J. and T. Kalandrakis, 2009. Dynamic Legislative Policy Making. Unpublished manuscript.

[20] Epple, D. and M. Riordan, 1987. Cooperation and Punishment Under Repeated Majority Voting. Public Choice, 55, pp. 41-73.

[21] Ferejohn, J. A., M. P. Fiorina, and R. D. McKelvey, 1987. Sophisticated Voting and Agenda Independence in the Distributive Politics Setting. American Journal of Political Science, 31, pp.169-93.

[22] Fernandez, R. and D. Rodrik, 1991. Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty. American Economic Review, 81(5), pp. 1146-55.

[23] Fong, P., 2006. Dynamics of Government and Policy Choice. Unpublished manuscript.

[24] Glomm, G. and B. Ravikumar, Endogenous Public Policy and Multiple Equilibria. European Journal of Political Economy, December 1995, 11(4), pp. 653–62.

[25] Hassler, J., J. V. Rodriguez Mora, K. Storesletten, and F. Zilibotti, 2001. The Survival of the Welfare State. American Economic Review, 93, pp. 87-112.

[26] Hassler J., P. Krusell, K. Storesletten, and F. Zilibotti, 2005. The Dynamics of Government. Journal of Monetary Economics, 52, pp. 1331-1358.

[27] Hird, J. A., 1991. The Political Economy of Pork: Project Selection at the U.S. Army Corps of Engineers. American Political Science Review,85, pp. 430-56.

[28] Kalai, E., 1977. Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons. Econometrica, 45 (7), pp. 1623-1630.

[29] Kalandrakis, T., 2004. A three-player dynamic majoritarian bargaining game. Journal of Economic Theory, 116, 294–322.

[30] Kalandrakis, T., 2006. Proposal Rights and Political Power. American Journal of Political Science, 50(2), pp. 441-448.

[31] Kalandrakis, T., 2007. Majority Rule Dynamics with Endogenous Status Quo. Unpublished manuscript.

[32] Kearney, R., 1990. Sunset: A Survey and Analysis of the State Experience. Public Administration Review, 50(1), pp. 49-57.

[33] Krusell, P. and J. V. Rios-Rull, 1996. Vested Interests in a Positive Theory of Stagnation and Growth. Review of Economic Studies, 63(2), pp. 301–329.

[34] Krusell, P. and J. V. Rios-Rull, 1999. On the Size of the U.S. Government: Political Economy in the Neoclassical Growth Model." American Economic Review, 89(5), pp. 1156–81.

[35] Montagnes, P., 2010. Voting Rules and Contractual Defaults. Unpublished manuscript.

[36] Lowi, T.J., 1969. The End of Liberalism: Ideology, Policy, and the Crisis of Public Authority. Norton.

[37] Riboni, A. and F. Ruge-Murcia, 2008. The Dynamic (In)efficiency of Monetary Policy by Committee. Journal of Money, Credit, and Banking, 40, pp. 1001-1032.

[38] Riker, W. H., 1962. The Theory of Political Coalitions. New Haven: Yale University Press.

[39] Romer, T. and H. Rosenthal, 1978. Political Resource Allocation, Controlled Agendas, and the Status Quo. Public Choice, 33, pp. 27-43.

[40] Saint Paul, G., 2001. The Dynamics of Exclusion and Fiscal Conservatism. Review of Economic Dynamics, 4(2), pp. 275–302.

[41] Saint Paul, G. and T. Verdier, 1997. Power, Distributive Conflicts, and Multiple Growth Paths. Journal of Economic Growth, 2(2), pp. 155–68.

[42] Villas-Boas, J.M., 1997. Comparative Statics of Fixed Points. Journal of Economic Theory, 73, pp. 183-198.

[43] Weaver, R. K., 1985. Controlling Entitlements. The New Directions in American Politics, eds. John E. Chub and Paul E. Peterson. Washington, D.C.: Brookings Institution.

[44] Weaver, R. K., 1988. Automatic Government: The Politics of Indexation. Washington, D.C.: Brookings Institution.