INSTITUTO DE ECONOMÍA

DOCUMENTO de TRABAJO

536
2020

How Does Court Stability Affect Legal Stability?

Álvaro Bustos y Nuno Garoupa.

www.economia.puc.cl • ISSN (edición impresa) 0716-7334 • ISSN (edición electrónica) 0717-7593

HOW DOES COURT STABILITY AFFECT LEGAL STABILITY?

ÁLVARO BUSTOS
Pontificia Universidad Católica de Chile, School of Management

NUNO GAROUPA
George Mason University, Scalia Law School•

February, 2020

Abstract

The formation of coalitions in a court has attracted the attention of political scientists and legal economists. The question we address in this article is the extent to which coalition stability impacts the law. We consider a model where a court has two judicial coalitions, majority and minority. However, they may change their relative influence over time. We show that, while both sides have a preferred legal policy and want their standard to become law, the two coalitions may compromise on not changing the standard, thus maintaining the status quo, because of majority uncertainty in the future. One important implication from our article is that less certainty concerning the future (in terms of majority and minority) could actually make the law more stable in the present (since the standard is unchanged). In addition, we prove that moderate standards are more likely to endure the passage of time when compared to extreme standards.

JEL: K10, K40, H0

Key Words: Political Coalition; Legal Standard; Changing Legal Standard; Legal Stability.

• ngaroup@gmu.edu; alvaro.bustos@uc.cl. We are grateful to Lewis Kornhauser, Giri Parameswaran, Jide Nzelibe, Mauricio Bugarin, Abe Wickelgren, Michael Livermore, Michael Gilbert and Lawle-2019/Mexico City participants for very useful comments. Constance L. McKinnon provided superb research assistance. The usual disclaimers apply.

1. INTRODUCTION

Looking around the world, we know that some courts have a more stable composition than others, depending on appointment rules and party system (Garoupa and Ginsburg, 2015). For example, due to the different political cycles in the United States, since the early 1950s the composition of the US Supreme Court, in terms of Republican and Democrat appointees, is to some extent balanced and somehow predictable. However, in Brazil, where the US model of judicial appointment was imported a long time ago (Oliveira and Garoupa, 2011), the composition of the top court reflects significant political swings, in part due to mandatory retirement at 70 (because a President serving two full terms has the ability to appoint several justices, almost the entire bench). In Argentina, there have been periods of great change (for example, after the democratic transition) and periods of long stability (in the last decade) reflecting cycles in partisanship (González Bertomeu et al, 2016; Helmke, 2005). Across European constitutional courts, we also find some interesting variations.[1]

In this article we address the extent to which coalition stability impacts the law. We consider a model where a court has two judicial coalitions, majority and minority. However, they may change their relative influence over time. Let us call these coalitions Blue and Red. Blue could be majority in the present and minority in the future; the opposite applies to Red. We show that, while both sides have a preferred legal policy and want their standard to become law, the two coalitions may compromise on not changing the standard, thus preserving the status quo, because of majority uncertainty in the future. If one coalition

[1] Germany has practiced a quota system for many decades (the same institutional mechanism exists in countries like Austria and Portugal) – the main political parties have a fixed number of seats and appointments do not reflect short-term parliamentary balance of power. In other countries, such as France, Italy or Spain, constitutional courts have less stable political composition since it displays changes in the legislative and executive branches of government (Garoupa and Ginsburg, 2015).

polarizes and imposes its own standard, the other side will change the standard in the opposite direction given the opportunity (hence, inducing an excessively costly “standards-race” to both sides). However, if the majority keeps the original standard, the other side is also likely to keep the original standard (if given a chance to decide the standard in the future).[2] All of these effects depend on different parameters such as the costs from amending the law, stability of coalitions, predictability and so on. In particular, results significantly reflect the position of the original standard vis-à-vis the standards favored by both coalitions. Reminiscent of the logic revealed by the literature that studies US Supreme Court nominations and confirmations (Moranski and Shipan, 1999; Bustos and Jacobi, 2014), predictions vary depending on whether coalitions are ideologically aligned or opposed.

One important implication from our article is that uncertainty concerning the future (in terms of majority and minority) could actually make the law more stable (since the standard is unchanged). We uncover a pattern somehow similar to a Rawlsian veil of ignorance narrative – when both coalitions are afraid of possible repercussions in the future from swinging positions (behind the veil of ignorance), they are disciplined in the present and compromise by avoiding significant changes in the law, thus preventing a costly “standards-race”. The argument is also related to the political insurance theory in the context of judicial independence (Vanberg, 2015) – the possibility of being a minority in the future convinces the current majority to protect judicial independence (in our model, keeping the standard untouched) at the cost of foregoing their preferred policy.

There are some commonalities between our results and the seminal article by

[2] In other words, the initial majority coalition of a given court will see its expected extra payoff as associated with a new standard substantially affected by the probability of remaining as majority in the future.

Stephenson (2003). The latter explains judicial independence review as the outcome of a competitive political system, moderate judicial doctrine and a degree of forward looking risk aversion. Although our model is about courts, and not the political balance within the executive and legislative branches, the resulting suggestions are aligned. Political competition (reflected in the composition of the court) and uncertainty about the future is what induces moderate judicial doctrines. Our theory predicts that, when compared to extreme standards, moderate standards are more likely to be preserved over time. Extreme standards will be changed by one of the coalitions whenever it gets the chance. In this circumstance, a costly standard race is a risk that the coalition is willing to endure because the cost of keeping the extreme standard is greater. A moderate standard generates exactly the opposite incentives, which are shared by both coalitions.

In our article, we abstain from other possible relevant considerations. For example, we do not consider dynamic formation of coalitions (they are always Blue or Red with no possibility for a future Yellow). We also do not distinguish between salient cases and less important cases, thus not addressing the possibility of adjudication trade between coalitions within the same time period, because we focus on strategies across periods for a given legal standard (presumably important enough to merit attention by both coalitions). Still, we suggest that our main insights matter for the debate on the evolution of the common law (Gennaioli and Shleifer, 2007; Niblett et al, 2010; Parameswaran, 2018; Ponzetto and Fernandez, 2008). In the context of that literature, we show that the role of particular judicial preferences (or possible judicial biases) is affected by uncertainty concerning court composition. In fact, the present varying preferences of the court could actually deter change, therefore diluting the so-called evolutionary adaptability of common law (as opposed to the result suggested by Gennaioli and Shleifer, 2007) when there is significant uncertainty concerning the prevailing future preferences of the court.

Another area related to our article is the political economy of horizontal stare decisis (Cameron et al, 2019; Kornhauser, 1989; O'Hara, 1992; Rasmusen, 1994). We show that horizontal stares decisis happens because both coalitions are too uncertain about the future and prefer to defer to the current standard rather than risking an unwelcome change in the future**.**

We start with a short literature review in section 2. Section 3 introduces the model. The solutions to the model are presented in section 4. In section 5, we generalize our results. We discuss the results in section 6. Final remarks are addressed in section 7.

2. LITERATURE REVIEW

By now, there is a vast theoretical and empirical literature on judicial behavior, from the attitudinal model to multiple variations on the strategic model (Bergara et al, 2003; Epstein and Knight, 1998; Epstein et al, 2013; Maltzman et al, 2000; Segal and Spaeth, 2002; Spiller and Tiller, 1996). The theoretical models explain how judges make decisions in collegial courts. Starting with a rational theory of what judges maximize (Posner 1993, 2004 & 2010), the different approaches in the literature emphasize various relevant aspects in order to explain judicial decision-making. They differ in the possible determinants of judicial behavior, from ideology and other political variables to gender, age, religion, professional and social background. Interactions with further actors have been also considered: politicians, lobbyists, regulators, media, lawyers and law audiences, and so on (Segal, 1997; Segal and Westerland, 2005). The empirical literature has broadly provided validity about the theoretical predictions while not immune to controversy (Epstein and Jacobi, 2010; Cameron and Kornhauser, 2017; Segal et al, 2011).

There is considerably less work at the aggregate level, that is, explaining court behavior as a complex institution rather than the sum of individual judges (some exceptions include Baker and Mezzetti, 2012; Bustos and Jacobi, 2019; Callander and Clark, 2017; Landes and Posner, 1975; Ramseyer and Rasmusen, 1997). Most literature focuses on explaining the role of courts in promoting and protecting judicial independence as well as the interaction of the judiciary with the other two branches of government, legislative and executive (Landes and Posner, 1975). At this level, the empirical literature documents the impact of judicial independence and court performance on economic growth and other macroeconomic variables (Hanssen, 2004; Helmke and Rosenbluth, 2009; La Porta et al, 2008).

Coalition formation in a court has attracted attention from political scientists and legal economists given the historical dynamics of the US Supreme Court. The main goals of the literature have been to explain why judges join or configure certain coalitions (Cameron and Kornhauser, 2010; Epstein et al, 2007; Kornhauser, 1992a & 1992b; Segal et al, 1995), the role of the median judge (Clark, 2008; Martin et al, 2005; Segal and Cover, 1989) or coalition survival (Caminker, 1999; Geli and Spiller, 1990 & 1992; Segal 1997). However, the way in which the possibility of changes in court composition affect or shape behavior and legal policy has not been debated extensively. Particularly, at the meso level (that is, above the micro level of individual judges, but below the macro level of courts), work aimed at clarifying how coalitions interact and affect legal policy is limited (Kornhauser, 1995; Wahlbeck, 1997).

One recent article is particularly close to ours (Cameron et al, 2019), and focuses on horizontal stare decisis and model a heterogeneous bench. In that model, judges have different preferences for legal standards. These differences not only generate a non-cooperative equilibrium in the implementation of the standards, but also create opportunities for gains from trade. Such gains are maximal when all judges agree on applying a common standard. That is, partial stare decisis emerges as a coordinating mechanism that benefits judges from allowing full trade.

Apart from the applications (we are more interested in the stability of the law as a function of the court's political stability than in explaining partial stare decisis as a commitment device) and context (we model legal stability as the outcome of a non-cooperative game rather than gains in trade from cooperation), there are, however, a few additional important differences. Although both models discuss static and dynamic features, our result is centrally dynamic (sequential moves game) while theirs is static (simultaneous moves game that breaks with a prisoners dilemma). We explicitly consider a cost of changing the standard which adds to the cost of facing a suboptimal standard (implicit costs of trade in Cameron et al, 2019). We rely on uncertainty about future influence (political power changes exogenously) to explain choices rather than using the multiplicity of possible cases to trade on possible consensus. Still, we emphasize a similar insight – against common wisdom, that judicial polarization might provide stability rather than undermine it.

3. THE MODEL

A two-period court exercising constitutional review (simply "Court") has the option to change the legal standard at the beginning of each period. We denote $s_0 \in [0,1]$ as the standard initially faced by the Court and $s_t \in \{s_0, b, r\}$ with $t \in \{1,2\}$ as the standard set by the Court in periods 1 and 2 respectively. In Section 5, among other extensions, we discuss our results when $s_t \in [0,1]$.[3]

Political coalitions and the standard setting process

The Court is only conformed by justices that belong to either a Red or a Blue coalition. While Red has ideology $r \in [0,1]$, Blue has ideology $b \in [0,1]$. Each period one of the political coalitions is majority and the other coalition is minority. Without loss of generality, we impose that $r > b$, where 1 is most conservative and 0 is most liberal. While we know that Blue is the majority in the first period, it remains the majority in the second period only with probability x exogenously given.[4]

The majority coalition can freely change the legal standard in the first period, the minority coalition cannot stop that process. That is, in the first period, Blue sets s_1. Instead, in the second period, two possibilities arise. Under one possibility, Blue keeps the majority, with probability x, in which case the standard is set to be $s_2 = s_1$, that is, Blue cannot make additional changes in the second period (hence Blue has a first-move advantage in period one that cannot be delayed to period two). Alternatively, Red is able to gain a majority in the court, with probability 1-x, and two scenarios can happen. If $s_1 = s_0$, Red sets s_2 with no restriction, that is, either keeping s_0 or switching to r. However, if $s_1 = b$, Red necessarily sets $s_2 = r$; in other words, if Blue imposes its standard in period one, Red will impose its

[3] Usually justices cannot make decisions supported by any point in the ideological spectrum. The reason is that judicial decisions are restricted by legal doctrines used by justices to support their decisions and legal doctrines are typically discrete ideological constructs. See Bustos and Tiller (2019) for an extensive discussion of these ideas in the context of the decision made by the Court in *Sebelius*, 567 U.S. 519 (2012).

[4] This probability captures expectations of changes in the composition of the Court conditional on the political and electoral system.

standard in period two.[5] The cost of changing the standard is $c > 0$ regardless of whether the majority is Blue or Red.[6] Figure 1 shows the timing of the game.

<<Insert Figure 1 about here>>

Pay-offs

The only two players of the game are the Red and Blue coalitions. Without loss of generality, we call $U_i(s) = 1 - (i - s)^2$ the benefit obtained by coalition $i \in \{r, b\}$ when the Court sets standard s such that $U_i(s)$ is concave with $\frac{\partial U_i(i)}{\partial s} = 0$ and $U_r(r) = U_b(b) = 1$. The payoff of each coalition is maximal when the standard imposed by the Court is the coalition's standard. In addition to the benefit that each coalition gets every period because of the state of the law, the majority coalition faces cost c for changing the standard at the beginning of the period — note, however, that the majority coalition might decide not to change the standard. As usual, we denote the discount factor as δ. For parsimony, we introduce the following notation

$$\Delta \equiv U_b(b) - U_b(r) = U_r(r) - U_r(b) = (r - b)^2 \text{ (measures polarization across coalitions)}$$

$$\Delta_i(s) \equiv U_i(i) - U_i(s) = (i - s)^2 \text{ (measures disposition for each coalition)}$$

We define the scenarios: "opposed ideologies" and "aligned ideologies" as

$$\text{opposed ideologies} \equiv b < s_0 < r$$

[5] Note that the model is equivalent to a general setting in which coalitions always choose to keep the standard or set their optimal standard but the cost is restricted to intermediate values. In section 5 we specify those restrictions and show that our main results hold under the general setting without the restrictions.

[6] The cost of changing the standard is the same regardless of which coalition is in the majority because we interpret c to be the political cost of changing the law. Main results do not change if we assume that the cost in the first period is a constant but a random variable distributed according to a certain distribution in the second period. The mathematical analysis becomes more cumbersome if the cost is different in both periods, but it does not affect the qualitative results we derive with the basic model.

$$\text{aligned ideologies} \equiv b < r < s_0 \text{ or } s_0 < b < r$$

We also introduce the following definitions for consideration in later sections of the article:

Maximum Stability: No coalition changes the initial standard;

Maximum Instability: Both coalitions change any standard that is not their optimum.

4. MAIN RESULTS

In this section, we answer two central questions: When does the first period majority change the initial standard? And, under which conditions does the standard achieve maximum stability (hereafter, MS)? As for the first question, we show that when c takes intermediate values, Blue sets b only when its probability to remain the majority in the second period is large enough. Indeed, when s_0 takes extreme values, Blue sets b only when Blue's benefits associated with a Blue second period majority that keeps b, are big enough. Instead, when s_0 takes intermediate values, Blue sets b only when Blue´s loses associated with a Red second period majority that sets r, are small enough. In both cases, Blue´s decision to change the standard is only triggered when x is large enough. Moreover, in the case that s_0 takes intermediate values, Blue´s decision aims to avoid a costly standard-race. If Blue sets b, it exposes itself to a Red second period majority that moves the standard all the way to r.

As for the second question, we show that in addition to the trivial case in which MS takes place when c is very large, MS also takes place when c, x and s_0 take intermediate values. If x is big then Blue eventually changes the standard, and if x is small then Red eventually changes the standard. Only for intermediate values of x do both coalitions keep s_0 because doing the opposite leads to a costly standard-race. Even more relevant because of its policy implications, MS is more likely to take place (there is a greater set of values of c

that support a MS region) for intermediate values of s_0 than for extreme values of s_0. That is, moderate legal standards are the most stable standards. Intuitively, when s_0 is very close to b (r) then Red (Blue) always changes the standard, ergo MS requires an intermediate value of s_0. Finally, we also show that the region of MS tends to get smaller the more periods into the future are internalized by the coalitions when making their decisions.

Here, we only discuss the case in which coalitions have opposing ideologies. In Section 4 we identify commonalities and differences when coalitions are ideologically aligned.

4.1. WHEN DOES BLUE CHANGE THE STANDARD?

In the context of the game described in section 3, a second period Red facing $s_1 = s_0$ sets $s_2 = r$ if and only if $c < (r - s_0)^2$. With that in mind, we characterize the decision made by Blue in the first period. If Blue sets $s_1 = b$ then it gets payoff

$$U_b(b) = \underbrace{1 - c}_{Blue\ sets\ b} + \delta \left[\underbrace{x}_{Blue\ keeps\ b} + \underbrace{(1 - x)(1 - (r - b)^2)}_{Red\ sets\ r} \right]$$

Instead, Blue´s payoff when it keeps $s_1 = s_0$ depends on whether c is greater or smaller than $(r - s_0)^2$. If $c < (r - s_0)^2$, then Blue gets payoff

$$U_b(s_0) = \underbrace{1 - (s_0 - b)^2}_{Blue\ keeps\ s_0} + \delta \left[\underbrace{x(1 - (s_0 - b)^2)}_{Blue\ keeps\ s_0} + \underbrace{(1 - x)(1 - (r - b)^2)}_{Red\ sets\ r} \right]$$

Hence, Blue sets b in the first period only if the cost of doing so is smaller than the cost faced because the standard is s_0 in both periods. That is, Blue sets b if

$$c < (1 + \delta x)(s_0 - b)^2 \leftrightarrow x > \frac{c - (s_0 - b)^2}{\delta (s_0 - b)^2} \qquad (1)$$

Note that Blue´s decision does not depend on the scenario in which Red is majority in the second period because regardless of whether Blue sets b or keeps s_0, Red always sets r.

If $c > (r - s_0)^2$ then Blue gets the following payoff if it keeps $s_1 = s_0$

$$U_b(s_0) = \underbrace{1 - (s_0 - b)^2}_{Blue\ keeps\ s_0} + \delta \left[\underbrace{x(1 - (s_0 - b)^2)}_{Blue\ keeps\ s_0} + \underbrace{(1 - x)(1 - (s_0 - b)^2)}_{Red\ keeps\ s_0} \right]$$

As before, Blue sets b if the cost of doing so is smaller than the cost faced because the standard is suboptimal s_0, but this time this cost is mitigated by the avoidance of the cost Blue would face because a second period Red sets r in response to $s_1 = b$. That is, Blue sets b if

$$c < (1 + \delta)(s_0 - b)^2 - \delta(1 - x)(r - b)^2$$

$$\leftrightarrow x > \frac{c - (1 + \delta)(s_0 - b)^2 + \delta(r - b)^2}{\delta(r - b)^2} \qquad (2)$$

In both cases, as clarified by expressions (1) and (2), Blue changes the standard only if the probability of remaining the majority coalition at the second period is above a certain threshold. In other words, *the higher the probability that Blue maintains control of the Court, the more likely it will change the standard at the first period*. The previous results are summarized by Proposition 1.

PROPOSITION 1 (Blue's decisions at t = 1 with opposed coalitions):

For all values of c there exists $x^*(c)$ *with* $\frac{\partial x^*(c)}{\partial c} \geq 0$ *such that Blue sets*

$$s_1 = \begin{cases} b \text{ if } x \geq x^*(c) \\ s_0 \text{ if } x < x^*(c) \end{cases}$$

Proof: See previous mathematical derivations.

Because a priori we do not know whether $(1+\delta)(s_0 - b)^2$ is larger than $(r - s_0)^2$ or $(r-b)^2$, figures 2 and 3 distinguish two scenarios, one for small values of s_0 and the other for intermediate values of s_0, in which Blue sets b.[7]

<<Insert Figures 2&3 about here>>

The figures allow us to uncover two regularities. First, an increment in s_0 increases the likelihood that Blue changes standard s_0 (the set of values of c that support this strategy increases) because Blue becomes less interested in keeping a standard that is farther away from b. Second, an increment in r (increment in polarization) weakly decreases the number of scenarios in which Blue sets b. While a change in r is irrelevant in (1), a change in r makes (2) more difficult to hold because (2) captures the fact that Red sets r if Blue sets b.

4.2. MAXIMUM STABILITY AND MAXIMUM INSTABILITY

The previous discussion was exclusively centered on the decisions Blue makes in the first period. An additional central concern of this article is to determine the conditions under which the legal standard is never (always) changed by neither (both) Blue or Red and which values of x and s_0 make those realities more likely.

There is *maximum stability* if and only if $c > (r - s_0)^2$ and

[7] Without loss of generality and for sake of exposition, we assume that $s_0 < \frac{r+b}{2}$.

$$x < \frac{c - (1+\delta)(s_0 - b)^2 + \delta(r-b)^2}{\delta(r-b)^2} \qquad (3)$$

Maximum stability requires the cost to be high enough to deter Red from changing the standard in the second period, and the probability of Blue being a majority in the second period to be low enough to deter Blue from engaging in attrition with Red, which is the consequence if Blue sets b in the first period. Figures 4 and 5 show the regions of MS when s_0 takes small and intermediate values respectively.

<< Insert Figures 4&5 about here>>

When s_0 is small (figure 4) MS only takes place when the cost is large enough — expressions (1) and (2) become irrelevant for the identification of the region of MS because $(r - s_0)^2 > (1+\delta)(s_0 - b)^2$. Basically, whenever Blue keeps s_0 so does Red, but not the other way around. In other words, (3) does not bind. On the other side, when s_0 is not so small (figure 5) MS takes place not only when the cost is large enough but also when the cost takes intermediate values and, in addition, (3) holds. This distinction of cases has an important implication. The size of the region of MS first increases with s_0 but eventually starts decreasing. *That is, it is more likely that the standard is preserved over time when the standard is moderate than when it is extreme.* [8]

We can also define the conditions for *maximum instability*. This is the case where Blue sets b in period one and Red sets r in period two, given the chance. In this context, it

[8] This finding is equivalent to note that the area of MS is concave in $s_0 \in [b, r]$. Indeed, when $s_0 < \frac{r}{1+\sqrt{1+\delta}} + \frac{\sqrt{1+\delta}}{1+\sqrt{1+\delta}} b$ then $MS(s_0) = Cte - (r - s_0)^2$ which increases with s_0. But when $s_0 \geq \frac{r}{1+\sqrt{1+\delta}} + \frac{\sqrt{1+\delta}}{1+\sqrt{1+\delta}} b$ then $MS(s_0) = Cte - \left((r - s_0)^2 + \frac{((1+\delta)(s_0-b)^2-(r-s_0)^2)^2}{2\delta(r-b)^2}\right)$ which is increasing at $s_0 = \frac{r}{1+\sqrt{1+\delta}} + \frac{\sqrt{1+\delta}}{1+\sqrt{1+\delta}} b$, it is decreasing at $s_0 = r$ and its second derivative is strictly decreasing in $s_0 \in \left[\frac{r}{1+\sqrt{1+\delta}} + \frac{\sqrt{1+\delta}}{1+\sqrt{1+\delta}} b, r\right]$.

has to be that $c < (r - s_0)^2$ and expression (1) or (2) are satisfied. When the cost of changing is close to zero, trivially, we have *maximum instability*. This result also occurs when the cost takes intermediate values and the probability that Blue maintains the majority in the second period is large enough. Under these conditions, Blue is willing to set b because a costly standard race is not likely enough.

Because in a 2-Period model the implications of the decisions made by Red cannot be completely appreciated — Red only makes a decision in the last period of the game — we then derive general properties of the MS region in the context of a T-Period model.

4.3. MAXIMUM STABILITY WITHIN A T PERIOD MODEL

In a T-Period model, Blue and Red keep s_0 at $t = T$ if the next two conditions hold:

Cond $\boldsymbol{MS_T^r(T)}$: $c > (r - s_0)^2$

Cond $\boldsymbol{MS_T^b(T)}$: $c > (s_0 - b)^2$

In addition, if Cond MS_T^r and Cond MS_T^b hold — we evidently have to restrict the discussion to the subgames in which both coalitions preserve s_0 in the future — then at $t = T - 1$ both coalitions preserve s_0 when:

Cond $\boldsymbol{MS_{T-1}^r(T)}$: $c > \max\{(r - s_0)^2, (1 + \delta)(r - s_0)^2 - \delta x(r - b)^2\}$

Cond $\boldsymbol{MS_{T-1}^b(T)}$: $c > \max\{(s_0 - b)^2, (1 + \delta)(s_0 - b)^2 - \delta(1 - x)(r - b)^2\}$

While previously we derived the condition that leads a second period Blue to keep s_0, which is the same condition that leads a first period Blue to keep s_0 in a 2-Period model, we are missing the condition under which a second period Red keeps s_0. That happens only if

$$1 - (r - s_0)^2 - \delta[1 - (r - s_0)^2] > 1 - c - \delta[x(1 - (r - b)^2) + (1 - x)]$$

$$\leftrightarrow c > (1 + \delta)(r - s_0)^2 - \delta x(r - b)^2$$

This is equivalent to (2) but from the perspective of the Red coalition. That is, Red keeps s_0 when the cost of changing the standard is larger than the cost of facing a suboptimal standard mitigated by the threat that Blue sets standard b in case that Red sets r.

We can follow the same steps to derive the conditions that assure maximum stability at period t. Those are:

Cond $\boldsymbol{MS_t^r(T)}$: $c > \max\left\{(r-s_0)^2, \left\{\frac{(r-s_0)^2\sum_{i=0}^{N}\delta^i - \delta x(r-b)^2\sum_{i=0}^{N-1}\delta^i}{1+(1-x)x\delta\left(\sum_{i=0}^{N-1}\delta^i - 1\right)}\right\}_{N\in\{1,..,T-t\}}\right\}$

Cond $\boldsymbol{MS_t^b(T)}$: $c > \max\left\{(s_0-b)^2, \left\{\frac{(s_0-b)^2\sum_{i=0}^{N}\delta^i - \delta(1-x)(r-b)^2\sum_{i=0}^{N-1}\delta^i}{1+(1-x)x\delta\left(\sum_{i=0}^{N-1}\delta^i - 1\right)}\right\}_{N\in\{1,\dots,T-t\}}\right\}$

The numerators of these conditions capture the same two effects that were present in the first models. These effects are the cost of having a suboptimal standard and the mitigation for avoiding a costly standard race. The denominators refer to the paths in the sequence of decisions made by the coalitions in which the standard is changed. All the expressions that do not appear in Cond MS_2^r and Cond MS_1^b are missing because they appear both in the sequence in which the coalition keeps s_0 or changes it and ergo they cancel out.

Proposition 2 not only summarizes the required conditions for the two coalitions to always decide to keep the initial standard during T periods, it also states three important properties associated to the region of MS. While we formally prove these properties in the Appendix, we discuss their intuitions and implications in the following section.

PROPOSITION 2 (Region of MS in a T-Period Model):

Maximum Stability takes place when simultaneously **Cond** $\boldsymbol{MS_2^r(T)}$ and **Cond** $\boldsymbol{MS_1^b(T)}$. *In addition, the region of MS defined in the* (x, c) *plane satisfies the next properties:*

i. *The size of the region strictly decreases with* T.

ii. *The size of the region is strictly concave with* s_0.

iii. *For a given* c, *Cond* MS_2^r *defines a lower bound* $\underline{x}(c)$ *and Cond* MS_1^b *defines an upper bound* $\overline{x}(c)$ *such that MS only takes place if* $x \in [\underline{x}(c), \overline{x}(c)]$.

Proof: See the Appendix.

The more periods the coalitions consider at the moment of making their decisions, the smaller the region of maximum stability. With more periods into the future, there are more subgames in which one of the coalitions changes the standard.[9] That observation is graphically captured by figure 6.

<<Insert Figure 6 about here>>

The concavity of the MS region follows after we note that the size of the region always increases for a marginal increment in s_0 when s_0 is small enough, because in that case $MS_1^b(T)$ is not binding and $MS_2^r(T)$ becomes a looser restriction. On the other extreme, the size of the region always decreases for a marginal increment in s_0 when it is large enough, because in that case $MS_2^r(T)$ is not binding and $MS_1^b(T)$ becomes a tighter restriction. The proof presented in the Appendix concludes after we note that $MS_1^b(T)$ and $MS_2^r(T)$ are binding and define a region of MS that is concave in s_0 when s_0 takes intermediate values.[10] This concavity implies that *moderate initial standards are the most likely to endure the pass of time*. As we mentioned before, extremely conservative (liberal) values of s_0 will be

[9] The reduction in the region of MS with an increment in the number of periods follows after we note that $\frac{(r-s_0)^2\sum_{i=0}^{N}\delta^i-\delta x(r-b)^2\sum_{i=0}^{N-1}\delta^i}{1+(1-x)x\delta\left(\sum_{i=0}^{N-1}\delta^i-1\right)}$ is stricter than $\frac{(r-s_0)^2\sum_{i=0}^{N'}\delta^i-\delta x(r-b)^2\sum_{i=0}^{N'-1}\delta^i}{1+(1-x)x\delta\left(\sum_{i=0}^{N'-1}\delta^i-1\right)}$ when $N' > N$. The same happens for $\frac{(s_0-b)^2\sum_{i=0}^{N}\delta^i-\delta(1-x)(r-b)^2\sum_{i=0}^{N-1}\delta^i}{1+(1-x)x\delta\left(\sum_{i=0}^{N-1}\delta^i-1\right)}$ when compared with $\frac{(s_0-b)^2\sum_{i=0}^{N'}\delta^i-\delta(1-x)(r-b)^2\sum_{i=0}^{N'-1}\delta^i}{1+(1-x)x\delta\left(\sum_{i=0}^{N'-1}\delta^i-1\right)}$.

[10] The critical values for s_0 are $\frac{1}{1+\sqrt{\frac{1-\delta^T}{1-\delta}}}r+\frac{\sqrt{\frac{1-\delta^T}{1-\delta}}}{1+\sqrt{\frac{1-\delta^T}{1-\delta}}}b$ and $\frac{\sqrt{\frac{1-\delta^{T-1}}{1-\delta}}}{1+\sqrt{\frac{1-\delta^{T-1}}{1-\delta}}}r+\frac{1}{1+\sqrt{\frac{1-\delta^{T-1}}{1-\delta}}}b$.

changed by Blue (Red) in the future, but moderate values of s_0 are more likely to be preserved by the coalitions as a way to avoid a costly standard race between them.

Finally, *maximum stability takes place only when there is enough uncertainty as to which coalition will be the majority in the future.* When x is large then eventually Blue decides to change the standard. The same happens for Red when x is low. But when x takes an intermediate value then no coalition sets its preferred standard because of the associated risk in starting a standard race. Mathematically, the cost bound defined by $MS_2^r(T)$ is decreasing in x while the cost bound defined by $MS_1^b(T)$ is increasing in x. This result implies that MS_2^r defines a lower bound for x and MS_1^b defines an upper bound for x, such that *MS* only takes place when x takes values within that intermediate range.[11] Note that $MS_2^r(T)$ tells us that Red is more willing to keep s_0 the smaller is b. Specifically, $MS_2^r(T)$ relaxes when b is reduced. Also $MS_1^b(T)$ tells us that Blue is more willing to keep s_0 the greater is r since $MS_2^r(T)$ tightens when r increases. Both effects capture the efforts that the coalitions make to avoid a costly standard race, including preserving the initial standard.

5. EXTENSIONS

Here we discuss the generality and robustness of our main results. We show that the link between ideological stability and legal instability is remarkably general as it holds under a general formulation of our base model. We also uncover that Proposition 1 might not hold when coalitions are ideologically aligned. We end this section by exploring the sensitivity of the results to the specification of the cost of a standard change.

[11] We show that the cost boundary defined by $MS_2^r(T)$ is decreasing in x only when $x < 1/2$, while the cost boundary defined by $MS_1^b(T)$ is increasing in x only when $x > 1/2$. Both are the relevant ranges of values for x.

5.1. GENERAL MODEL

Consider a general version of the model we have discussed so far. That is, in the second period, Blue sets s_2 with probability x and Red sets s_2 with probability $1 - x$ with no restrictions on possible strategies. Figure 7 summarizes the new timing of the game.

<<Insert Figure 7 about here>>

Note that our base model is a particular case of this general model when these two conditions are satisfied:

Cond 1: $c \in [(s_0 - b)^2, (r - b)^2]$.

Cond 2: $(s_0 - b)^2 < (r - s_0)^2 \leftrightarrow s_0 < \frac{r+b}{2}$.

The first condition implies that first period Blue sets $s_1 = b$ or keeps $s_1 = s_0$ but second period Blue always keep s_1 (because $c > (s_0 - b)^2$). Instead, second period Red always changes $s_1 = b$ to $s_1 = r$ (because $c < (r - b)^2$). The second condition implies that second period Red might keep or change standard $s_1 = s_0$ conditional on the value of c.

Clearly, under this more general specification of the model, Proposition 1 becomes Proposition 1G. Under Proposition 1G Blue follows the strategy identified in Proposition 1 when c takes intermediate values but Blue follows a strategy that does not depend on the value of x when the cost of a change takes extreme values. Indeed, Blue follows a bang-bang strategy in which it never (always) changes the standard when c is too big (small), but when c takes intermediate values Blue sets b only if the probability to stay as majority is large enough. As we uncovered in the basic model, the latter captures Blue's incentives to avoid a costly standard race, which is triggered when it sets b in the first period.

PROPOSITION 1G (Blue's decisions at t = 1 with opposed coalitions and discrete s_t):

There exists $\underline{c}$, $\overline{c}$ and $x^*(c)$ when $c \in [\underline{c}, \overline{c}]$ with $\frac{\partial x^*(c)}{\partial c} \geq 0$ such that Blue sets

$$s_1 = \begin{cases} b \text{ when } c < \underline{c} \\ \begin{cases} b \text{ if } x \geq x^*(c) \\ s_0 \text{ if } x < x^*(c) \end{cases} \text{when } c \in [\underline{c}, \overline{c}] \\ s_0 \text{ when } c > \overline{c} \end{cases}$$

Proof: See the Appendix.

Proposition 1G allows us to emphasize that an increment in polarization (increment in r) reduces Blue's incentives to set b. However that is only true when s_0 takes intermediate values. Figures 8 and 9, associated with the scenarios in which s_0 take intermediate values, show that Blue is less willing to set b when r increases. These are the scenarios in which the threat of a costly standard race centrally determines Blue's decisions.

<<Insert Figures 8 & 9 about here>>

Extreme values of s_0 do not generate the previous adjustment in Blue's strategy because in those cases the strategies followed by a second period Red are the same, regardless of whether Blue sets b or keeps s_0 in the first period.

5.2. CONTINUOUS MODEL

We carried out the previous analysis under the assumption that s_t can only take values within a restricted set of options. Proposition 1C summarizes Blue's strategy when neither Cond 1 nor Cond 2 hold and, s_t can take any value in the ideological interval.

PROPOSITION 1C (Blue's decision at t = 1 with opposed coalitions and continuous s_t):

When $s_0 \in [b, r]$ there exists $\underline{c}$, $\overline{c}$ and $x^*(c; s_0), x_1^*(c; s_0), x_2^*(c; s_0)$ when $c \in [\underline{c}, \overline{c}]$ with $\frac{\partial x^*(c; s_0)}{\partial c} > 0$ and $x_1^*(c; s_0), x_2^*(c; s_0)$ concave functions such that Blue sets

$$
s_1 = \begin{cases} b \text{ when } c < \underline{c} \\ \begin{cases} b \text{ if } x > \max\{x^*(c;s_0), x_2^*(c;s_0)\} \\ \begin{cases} \begin{cases} s_0 \text{ if } x \in \begin{bmatrix} \min\{x_1^*(c;s_0), x_2^*(c;s_0)\}, \\ \max\{x^*(c;s_0), x_2^*(c;s_0)\} \end{bmatrix} \\ r - \sqrt{c} \text{ if } x < \min\{x_1^*(c;s_0), x_2^*(c;s_0)\} \end{cases} \text{if } s_0 < \dfrac{r+b}{2} \\ \begin{cases} r - \sqrt{c} \text{ if } x \in \begin{bmatrix} \min\{x_1^*(c;s_0), x^*\ (c;s_0)\}, \\ \max\{x^*(c;s_0), x_2^*(c;s_0)\} \end{bmatrix} \\ s_0 \text{ if } x < \min\{x_1^*(c;s_0), x^*\ (c;s_0)\} \end{cases} \text{if } s_0 > \dfrac{r+b}{2} \end{cases} \end{cases} \text{ when } c \in [\underline{c}, \overline{c}] \\ s_0 \text{ when } c > \overline{c} \end{cases}
$$

Proof: See the Appendix.

A direct comparison of propositions 1G and 1C tells us that in the most general formulation (1C), Blue still follows a bang-bang strategy when c takes extreme values and follows a strategy in which more ideological stability (larger x) implies more legal instability (smaller s_1) when c takes intermediate values. Still, we should emphasize that the notion of legal stability in the continuous model (standard stays close to the coalition ideology) is not exactly the same as in the discrete model (standard is unchanged).[12]

Proposition 1C summarizes two main differences with the analysis and results presented in Proposition 1G. First, Blue optimally fixes a standard that belongs to the set $\{b, r - \sqrt{c}, s_0\}$. For Blue, it might be optimal to set $s_1 = r - \sqrt{c}$ because $r - \sqrt{c}$ convinces a second period Red not to set r in the second period given that

$$1 - c = 1 - (r - s_1)^2 \rightarrow s_1 = r - \sqrt{c}$$

Hence, even when $s_1 = b$ is Blue's first best standard, in the case of a second period Red majority, Blue ends up much worse with $s_2 = r$ than with $s_2 = r - \sqrt{c}$. On the other hand,

[12] While in the discrete model legal instability could be simply captured by $Var(s_t)$, in the continuous model that statistic does not seem to be enough because the direction in the adjustment of s_t is relevant. $Var(s_t - s_{t-1})$ seems to be a better instrument to capture instability in the continuous model.

even when $s_1 = s_0$ saves Blue cost c, it might also imply that in the second period a Red majority sets $s_2 = r$ and leaves Blue in a worse position than if the standard is $r - \sqrt{c}$. Evidently, only for certain parameters of the model, Blue prefers to set $r - \sqrt{c}$ instead of b and s_0.[13]

Second, when c takes intermediate values then Blue faces a richer set of decisions than to be "aggressive" (sets b) or to be "accommodating" (keeps s_0). In the context of the continuous model, Blue might be "aggressive", meaning that it sets b, "semi-accommodating", it sets $r - \sqrt{c} \in [b, s_0]$, "accommodating", it keeps s_0, or "concessive", it sets $r - \sqrt{c} \in [s_0, r]$.[14] Beyond the specifics, the strategy followed by Blue tells us that, for opposed coalitions, *the more likely it is that the Court is stable (same coalition is in majority in both periods) the closer to the ideology of the majority coalition is the standard set by that coalition in the first period.*

The proof of Proposition 1C is more involved than in the case of 1 and 1G because Blue's decisions require three, as opposed to one, comparison of utilities given Blue's decisions. Mathematical details can be found in the Appendix.

---

[13] Blue prefers at least one of $\{b, r - \sqrt{c}, s_0\}$ to any other value of s_1. To see that, consider that $b < r - \sqrt{c} < s_0$, then all $s_1 \in [b, r - \sqrt{c}]$ are at least dominated by b because we are in a region in which a second period Red majority changes the standard. Hence, Blue gets a maximum payoff in the first period and as a second period majority if the standard is b and not s_1. In addition, all $s_1 \in [r - \sqrt{c}, s_0]$ are at least dominated by $r - \sqrt{c}$ because we are in a region in which a second period Red majority does not change the standard, hence Blue gets a maximum payoff in the first period and as a second period majority if the standard is $r - \sqrt{c}$ and not s_1. The same steps can be used to prove that all $s_1 \notin \{b, r - \sqrt{c}, s_0\}$ are dominated when $b < s_0 < r - \sqrt{c}$.

[14] Note that $r - \sqrt{c} < s_0$ when $c > \Delta_r(s_0)$ but $r - \sqrt{c} > s_0$ when $c < \Delta_r(s_0)$. That is why, in Proposition 1G, we make a distinction when s_0 is larger or smaller than $(r + b)/2$.

5.3. ALIGNED IDEOLOGIES

Until now, we have concentrated our discussion in the case in which coalitions have opposite ideologies. In many ways this is the most relevant scenario because, many higher courts of the world are dominated by coalitions with opposed ideological preferences. Consider, however, the case of aligned ideologies. Although it is still true that Blue always (never) changes the standard when the cost is small (large enough), results are somehow distinct when the cost has an intermediate value. For example, *if the initial standard is to the right of both Blue and Red ideal points then the probability of Blue remaining as majority in the second period becomes irrelevant to the strategy followed by Blue in the first period*. The reason is because the payoff that Blue gets in the second period (when it sets b) always dominates the second period payoff when Blue keeps s_0. Hence the only scenario in which Blue keeps the standard is when the cost of change is greater than the benefit of not facing the suboptimal standard s_0 in both periods, which is $(1+\delta)\Delta_b(s_0)$ and that boundary does not depend on x.

In order to see the previous scenario more clearly, assume that $b < r < s_0$ and $c \in [\Delta_r(s_0), \Delta_b(s_0)]$ then if Blue keeps s_0 it gets

$$U_b(s_0) = \underbrace{1-\Delta_b(s_0)}_{Blue\ keeps\ s_0} + \delta\left(\underbrace{x(1-c)}_{Blue\ sets\ b} + \underbrace{(1-x)(1-\Delta_b(s_0))}_{Red\ keeps\ s_0}\right)$$

But if it sets b it gets

$$U_b(b) = \underbrace{1-c}_{Blue\ sets\ b} + \delta\left(\underbrace{x}_{Blue\ keeps\ b} + \underbrace{(1-x)}_{Red\ keeps\ b}\right)$$

Consequently, $U_b(b)$ is greater than $U_b(s_0)$ for all values of x because $U_b(b)$ dominates $U_b(s_0)$, vis-à-vis, in both periods. Moreover, when $c > \Delta_b(s_0)$ the corresponding pay-offs become

$$U_b(s_0) = \underbrace{1 - \Delta_b(s_0)}_{Blue\ keeps\ s_0} + \delta\left(\underbrace{x\Delta_b(s_0)}_{Blue\ sets\ b} + \underbrace{(1-x)(1-\Delta_b(s_0))}_{Red\ keeps\ s_0}\right) = (1+\delta)(1-\Delta_b(s_0))$$

and

$$U_b(b) = \underbrace{1 - c}_{Blue\ sets\ b} + \delta\left(\underbrace{x}_{Blue\ keeps\ b} + \underbrace{(1-x)}_{Red\ keeps\ b}\right) = 1 + \delta - c,$$

respectively. Accordingly, $U_b(b)$ dominates $U_b(s_0)$ only when $c < (1+\delta)\Delta_b(s_0)$. Proposition 3, formally proven in the Appendix, states that the previous logic always hold as long as $b < r < s_0$.

PROPOSITION 3 (Blue's decisions at t = 1 with aligned coalitions):

When $b < r < s_0$ then Blue sets

$$s_1 = \begin{cases} b & \text{when } c \leq (1+\delta)\Delta_b(s_0) \\ s_0 & \text{when } c > (1+\delta)\Delta_b(s_0) \end{cases}$$

Proof: See the Appendix.

In the Appendix, we also discuss the case in which the initial standard is located to the left of both Blue and Red ideal points. There we prove that *if the initial standard is to the left of both Blue and Red ideal points then the first period majority coalition follows a strategy in which it adjusts the legal standard if x is large enough or follows a strategy in which it adjusts the legal standard if x is small enough*. The second strategy is possible when ideologies are aligned (but not when they are opposed) because, under alignment, Blue's future expected benefit due to imposing its preferred ideology can be greater if Red, and not Blue, is majority in the second period. This approach is not possible when the coalitions have opposed ideologies because Blue is always better off when it remains in the majority if it chooses to change the standard in the first period.

Although the last result emphasizes the relevance of the coalition ideology relative to the initial standard, we still believe that results derived in the case where coalitions are ideologically opposed are the most relevant ones. Namely their practical value in explaining judicial politics.

5.4. COST FUNCTION

Our formal model assumes a fixed cost imposed on the majority when the standard is changed. We argue that this is a reasonable assumption. First, since the legal change is imposed by the majority without (explicit) bargaining with the minority (a passive party in each period of time), the direct cost is borne by the active party (in terms of public opinion, general criticism or workload). Notice, however, in the base model developed in section 4, the minority does face a reduction in its payoff as reflected by the distance between its disposition and the new standard (imposed by the majority). Thus, there is a cost to the minority, just not a direct cost from changing the law. Second, at least in the base model, since the choices are discrete (original standard, Blue's disposition, Red's disposition), a proportional cost function adds little insight while complicating the results in unnecessary ways according to our view.

6. DISCUSSION

The main insight from our model is that uncertainty about future court preferences might deter legal change in the present. A court dominated by the same party leaning produces stable law in the sense that the prevalent majority pushes its standard when costs are minimal and there is no need for future change (assuming everything else stays the same). However, in a court where majoritarian coalitions alternate, we could also have stable law in the sense that the initial standard is unchanged because each side is deterred by the threat of triggering a costly "standard-race" (notice that, in our base model, technically the standards are unchanged because the first majority has strategically set out the standard such that the rival group does not find it worthwhile to change it). The opposite situation, legal instability, takes place when majoritarian coalitions endure or the costs of adjustment are sufficiently low to make a potential "standard-race" a reasonable risk to take (in our reduced-form base model, we have a "standard-race" when Blue picks *b* in period one and Red picks *r* in period two).

We suggest that this insight has possible policy applications. The canonical parameters are the expectations coalitions have concerning the future and the relative costs of adjustment, both inevitably tainted with hindsight biases when we discuss specific illustrations. Still, one general application is in explaining why increasing court polarization (Bartels, 2015; Clark, 2008; Epstein et al, 2007; Gooch, 2015) may not bring about legal change. If polarization is associated with increasing the relative costs borne by each side when a "standard-race" takes place, then it actually enhances rather than diminishes legal stability. In fact, observed ideological swings, for example, from a "liberal court" (Warren court; Powe, 2002) to more "conservative courts" (Burger and Rehnquist courts; Lindquist and Solberg, 2007) might not produce as much change as the attitudinal model would suggest, precisely because of deterrence induced by the possibility of triggering costly legal instability. Yet, once a new ideological majority has a perception of significant durability, we should expect to see legal change, consistent with long periods of a “liberal court” or a “conservative court” (Tushnet, 2005).[15] Similarly, evidence about US Supreme Court precedent seems to support our modeling approach - as predicted by our model, less radical precedents are more likely to survive.[16]

Turning to anecdotal evidence and historical debates, our model might provide additional reasoning to opposing accounts by Leuchtenburg (1995) and Cushman (1998) concerning the constitutional change in the 1930s (New Deal) – that an internal threat of costly instability provides the analytical framework for a continuous gradual change rather than a full switch story.

An additional angle of analysis is the prevalence of confrontational mood versus a more harmonious working environment. For example, Crawford Greenburg (2007) provides two useful examples – Lee v. Weisman (1992) and Lawrence v. Texas (2003). Writing for the majority in both cases, Justice Kennedy explains how he had to ponder departing from precedent -in the latter case, Bowers v. Hardwick (1986) established the precedent- and find the appropriate language to search for compromise as much as possible in order to avoid future turmoil. Although, in both cases, Justice Kennedy could not ultimately avoid a split,

[15] A few critics of the modern trends in the US Supreme Court have voiced/argued that confrontational polarization changes the law too far and too often, thus documenting that a “standard-race” is a real possibility. For example, McCloskey (2010) argues that when pragmatic compromise is replaced by “negative capability”, the resistance to push a logical extreme is broken and legal standards become uncertain. According to the author, this happened in the New Deal period as well as in the “liberal court” of the 1950s and 1960s because “judges of the Supreme Court itself did not know the history of the bench they occupied, or had failed to understand it”.

[16] In their study of 6,363 precedents, from 1946 to 2001, Hansford and Spriggs (2006) find the survival of a precedent seems to be helped by ideological proximity and vitality (as measured by more prior positive citations and less negative citations). As vitality increases, ideological proximity matters more. However, focusing on majority interpretation of law, the same authors conclude that only ideological distance seems to matter (the majority seems to use precedents closer to the median ideology) while vitality seems unimportant.

he tried to suppress hostile language while engaging with the (minority) conservative wing of the court (Crawford Greenberg, 2007, at 36-63 and at 139-163). Eventually, both cases set new precedents that have largely been stable.[17]

Our model could also be understood as explaining why minimalism in constitutional review is likely to prevail over fundamentalism or extremism (visionaries), to paraphrase Sunstein (2005, 2009). Minimalism (either conservative or liberal) is defined as proceeding step by step without unbalancing the law, preserving the past, with slow change and no legal revolution. Conversely, visionaries (either conservative or liberal) seek big and sudden change (either back to the original meaning or moving forward in the progressive agenda). In fact, our model explains why minimalism might emerge even when both sides are polarized by visionaries. In a way, visionaries are afraid of each other and might converge on minimalism as the best possible solution to legal debates.

In Europe, the relationship between constitutional courts and supreme courts has not been uniform. Some countries exhibit a tense relationship (Italy and Spain quite remarkably, France more recently), whereas other countries have a reasonably peaceful engagement (Germany comes to mind). Different incentives might explain these differences (Garoupa and Ginsburg, 2015). Our results provide an additional useful reasoning – we could imagine that the situation is similar to our model where both "majority" (constitutional court) and "minority" (supreme court) know their relative strength in the future (constitutional court always prevails over supreme court). However, in countries with more unstable composition

[17] Another author, looking at different personalities and temperaments within the same ideological area (Rosen, 2007), argues that justices more inclined to compromise or less confrontational might have more influence in the US Supreme Court due to their ability to build stable coalitions. Specifically, on the liberal side, he compares Black and Douglas; on the conservative side, he compares Rehnquist and Scalia.

(such as France, Italy or Spain), the preference of the constitutional court could change over time, therefore triggering more conflicts with the “minority”. In countries with a more stable composition (such as Germany, Austria or Portugal), the preferences of the constitutional court are less likely to change over time, hence allowing some form of settlement with the “minority” or an inclination for gradual change.

7. CONCLUSIONS

In this article, we present a discussion of legal stability based on the game played by coalitions in a court of law. Changing the law has two aspects. There is a direct cost from setting a new standard, therefore limiting legal volatility to a certain degree. However, there is an indirect cost – a change of law in present could trigger a reaction from the opposing coalition in the future. Therefore, a current majority could be deterred from shaping the law more actively if the future is too uncertain.

In courts where composition is stable and the dynamics of majority and minority is largely unchanged, we expect the majority to adjust the law to their preferences at some point. However, a court with strong and uncertain majority and minority dynamics can produce two very different results. If the direct costs of changing the law are somehow relevant, the coalitions may settle on keeping the status quo, thus not reforming the law, and we have *maximum stability*. In this case, judicial polarization, for example, is consistent with stable law. However, when the direct costs of changing the law are moot, we may observe huge swings in the law reflecting *maximum instability*.

Our article points out that the relationship between court composition and legal stability could be more nuanced than anticipated. Important ideological shifts in composition could have less of an impact on the law precisely because the different coalitions are deterred by the possibility of triggering a costly "standard-race" that harms both sides with excessive costs.

REFERENCES

Baker, Scott and Claudio Mezzetti. 2012. A Theory of Rational Jurisprudence, Journal of Political Economy 120: 513-551.

Bartels, Brandon L. 2015. The Sources and Consequences of Polarization in the US Supreme Court, in American Gridlock: The Sources, Character, and Impact of Political Polarization (James A. Thurber and Antoine Yoshinaka, eds), New York, NY: Cambridge University Press.

Bergara, Mario, Barak D. Richman and Pablo T. Spiller. 2003. Modeling Supreme Court Strategic Decision Making: The Congressional Constraint, Legislative Studies Quarterly 28 (2): 247-280.

Bustos, Álvaro and Tonja Jacobi. 2014. Strategic Judicial Preference Revelation, Journal of Law and Economics, 57 (1): 113-137.

Bustos, Álvaro and Tonja Jacobi. 2019. Judicial Choice among Cases for Certiorari, Supreme Court Economic Review, forthcoming,

Callander, Steven and Tom S. Clark. 2017. Precedent and Doctrine in a Complicated World, American Political Science Review 111(1):184-203.

Cameron, Charles M. and Lewis A. Kornhauser. 2010. Modeling Collegial Courts III: Adjudication Equilibria, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2153785

Cameron, Charles M. and Lewis A. Kornhauser. 2017. Theorizing the U.S. Supreme Court, Oxford Research Encyclopedia of Politics.

Cameron, Charles M., Lewis A. Kornhauser and Giri Parameswaran. 2019. Stare Decisis and Judicial Log-Rolls: A Gains-from-Trade Model, Rand Economic Journal, forthcoming.

Caminker, Evan H. 1999. Sincere and Strategic Voting Norms on Multimember Courts, Michigan Law Review 97(8): 2297-2380.

Clark, Tom S. 2008. Measuring Ideological Polarization on the United States Supreme Court, Political Research Quarterly 62 (1): 146-157.

Crawford Greenburg, Jan. 2007. Supreme Conflict: The Inside Story of the Struggle for Control of the United States Supreme Court, NY: Penguin Group.

Cushman, Barry. 1998. Rethinking the New Deal Court: The Structure of a Constitutional Revolution. New York, NY: Oxford University Press.

Epstein, Lee and Tonja Jacobi. 2010. The Strategic Analysis of Judicial Decisions, Annual Review of Law and Social Science 6: 341–58.

Epstein, Lee and Jack Knight. 1998. The Choices Justices Make. Washington, D.C.: C.Q. Press.

Epstein, Lee, William M. Landes, and Richard A. Posner. 2013. The Behavior of Federal Judges: A Theoretical and Empirical Study of Rational Choice. Boston: Harvard University Press.

Epstein, Lee, Andrew D. Martin, Kevin M. Quinn, and Jeffrey A. Segal. 2007. Ideological Drift among Supreme Court Justices: Who, When, and How Important, Northwestern University Law Review 101: 1483-1541.

Garoupa, Nuno and Tom Ginsburg. 2015. Judicial Reputation: A Comparative Theory. Chicago, Illinois: University of Chicago Press.

Gennaioli, Nicola and Andrei Shleifer. 2007. The Evolution of Common Law, Journal of Political Economy 115: 43-68.

Gely, Rafael and Pablo T. Spiller. 1990. A Rational Choice Theory of Supreme Court Statutory Decisions with Applications to the State Farm and Grove City Cases, Journal of Law, Economics, and Organization 6:263-300.

Gely, Rafael and Pablo T. Spiller. 1992. The Political Economy of Supreme Court Constitutional Decisions: The Case of Roosevelt's Court-Packing Plan, International Review of Law and Economics 12(1):45-67.

González Bertomeu, Juan, Lucia Dalla Pellegrina, and Nuno Garoupa. 2016. Estimating Judicial Ideal Points in Latin America: The Case of Argentina, Review of Law and Economics 13, https://doi.org/10.1515/rle-2015-0040.

Gooch, Donald M. 2015. Ideological Polarization on the Supreme Court: Trends in the Court's Institutional Environment and Across Regimes, 1937-2008, American Politics Research 43, 999-1040

Hansford, Thomas G. and James F. Spriggs II. 2006. The Politics of Precedent on the US Supreme Court, Princeton NJ: Princeton University Press.

Hanssen, F. Andrew Hanssen. 2004. There a Politically Optimal Level of Judicial Independence?, American Economic Review 94, 712-729.

Helmke, Gretchen. 2005. Courts Under Constraints: Judges, Generals, and Presidents in Argentina, New York: Cambridge University Press.

Helmke, Gretchen and Frances Rosenbluth. 2009. Regimes and the Rule of Law: Judicial Independence in Comparative Perspective, Annual Review of Political Science 12, 345-366.

Kornhauser, Lewis A. 1989. An Economic Perspective on Stare Decisis, Chicago-Kent Law Review 65, 63-92.

Kornhauser, Lewis A. 1992a. Modeling Collegial Courts I: Path-Dependence, International Review of Law and Economics 12(2):169-185.

Kornhauser, Lewis A. 1992b. Modeling Collegial Courts. II. Legal Doctrine, Journal of Law, Economics, and Organization 8(3):441-470.

Kornhauser, Lewis A. 1995. Adjudication by a Resource-Constrained Team: Hierarchy and Precedent in a Judicial System, Southern California Law Review 68:1605-1629.

Landes, William M. and Richard A. Posner. 1975. The Independent Judiciary in an Interest-Group Perspective, Journal of Law and Economics 18: 875-901.

La Porta, Rafael, Florencio Lopez-de-Silanes and Andrei Shleifer. 2008. The Economic Consequences of Legal Origins, Journal of Economic Literature 46 (2): 285-322.

Leuchtenburg, William E. 1995. The Supreme Court Reborn: The Constitutional Revolution in the Age of Roosevelt, New York, NY: Oxford University Press.

Lindquist, S. A. and R. S. Solberg. 2007. Judicial Review by the Burger and Rehnquist Courts: Explaining Justices' Responses to Constitutional Challenges, Political Research Quarterly 60, 71-90.

Maltzman, Forrest, James F. Spriggs II, and Paul J. Wahlbeck. 2000. Crafting Law on the Supreme Court: The Collegial Game. New York: Cambridge University Press.

Martin, Andrew D., Kevin M. Quinn, and Lee Epstein. 2005. The Median Justice on the United States Supreme Court, North Carolina Law Review 83: 1275-1322.

McCloskey, Robert G. 2010. The American Supreme Court, 5th edition, Chicago: University of Chicago Press.

Moraski, Bryon and Charles Shipan. 1999. The Politics of Supreme Court Nominations: A Theory of Institutional Constraints and Choices, American Journal of Political Science 43 (4): 1069-1095.

Niblett, Anthony, Richard A. Posner, and Andrei Shleifer. 2010. The Evolution of a Legal Rule, Journal of Legal Studies 39: 325-358.

O'Hara, Erin. 1993. Social Constraint or Implicit Collusion? -Toward a Game Theoretic Analysis, Seton Hall Law Review 24, 736-778.

Oliveira, Maria Angela and Nuno Garoupa. 2011. Choosing Judges in Brazil: Reassessing Legal Transplants from the United States, American Journal of Comparative Law 59 (2): 529-561.

Parameswaran, Giri. 2018. Endogenous Cases and the Evolution of the Common Law, Rand Journal of Economics 49, 781-818.

Ponzetto, Giacomo A. M. and Patricio A. Fernandez. 2008. Case Law versus Statute Law: An Evolutionary Comparison, Journal of Legal Studies 37: 379-430.

Posner, Richard A. 1993. What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does), Supreme Court Economic Review 3(1):1-41.

Posner, Richard A. 2004. Judicial Behavior and Performance: An Economic Approach, Florida State University Law Review 32: 1259-1279.

Posner, Richard A. 2010. How Judges Think. Cambridge, Mass: Harvard Univ. Press.

Powe. Lucas A. 2002. The Warren Court and American Politics. Cambridge, Mass: Belknap Press.

Ramseyer, J. Mark and Eric B. Rasmusen. 1997. Judicial Independence in a Civil Law Regime: The Evidence from Japan, Journal of Law, Economics and Organization 13: 259-286.

Rasmusen, Eric B. 1994. Judicial Legitimacy as a Repeated Game, Journal of Law, Economics and Organization 10, 63-83.

Rosen, Jeffrey. 2007. The Supreme Court: The Personalities and Rivalries that Defined America, NY: Holt.

Segal, Jeffrey A. 1997. Separation-of-Powers Games in the Positive Theory of Congress and Courts, American Political Science Review 91(1):28-44.

Segal, Jeffrey A. and Albert D. Cover. 1989. Ideological Values and the Votes of U.S. Supreme Court Justices, American Political Science Review 83(2):557-565.

Segal, Jeffrey A. and Chad Westerland. 2005. The Supreme Court, Congress, and Judicial Review, North Carolina Law Review 83(5):1323-1352.

Segal, Jeffrey A., Chad Westerland and Stefanie A. Lindquist. 2011. Congress, the Supreme Court, and Judicial Review: Testing a Constitutional Separation of Powers Model, American Journal of Political Science 55(1):89-104.

Segal, Jeffrey A. and Harold J. Spaeth. 2002. The Supreme Court and the Attitudinal Model Revisited. New York: Cambridge University Press.

Segal, Jeffrey A., Lee Epstein, Charles M. Cameron, and Harold J. Spaeth. 1995. Ideological Values and the Votes of U.S. Supreme Court Justices Revisited, Journal of Politics 57(3):812-823.

Spiller, Pablo T. and Emerson H. Tiller. 1996. Invitations to Override: Congressional Reversals of Supreme Court Decisions, International Review of Law and Economics 16 (4): 503–521.

Stephenson, Matthew C. 2003. "When the Devil Turns . . . ": The Political Foundations of Independent Judicial Review, Journal of Legal Studies 32, 59-88.

Sunstein, Cass R. 2005. Radicals in Robes: Why Extreme Right-Wing Courts are Wrong for America, Cambridge MA: Basic Books.

Sunstein ,Cass R. 2009. A Constitution of Many Minds, Why the Founding Document Doesn't Mean What It Meant Before, Princeton: Princeton University Press.

Tushnet, Mark. 2005. A Court Divided: The Rehnquist Court and the Future of Constitutional Law. New York City, NY: WW Norton& Company.

Vanberg, Georg. 2015. Constitutional Courts in Comparative Perspective: A Theoretical Assessment, Annual Review of Political Science 18, 167-185.

Walhbeck, Paul J. 1997. The Life of the Law: Judicial Politics and Legal Change, Journal of Politics 59: 778-802.

APPENDIX WITH PROOFS

Proof Proposition 1: We focus only in the case in which $s_0 < \frac{b+r}{2}$ which implies $\Delta_b(s_0) < \Delta_r(s_0)$. The proof when $s_0 < \frac{b+r}{2}$ is analogous because, at t = 2, Blue never changes s_0 and Red always change b. We have to distinguish two cases: $c < \Delta_r(s_0)$; and $c \geq \Delta_r(s_0)$.
When $c < \Delta_r(s_0)$ then (remember that Blue cannot change s_0 in the second period)

$$U_b(s_0) = 1 - (s_0 - b)^2 + \delta\big(x(1 - (s_0 - b)^2) + (1 - x)(1 - (r - b)^2)\big)$$

and

$$U_b(b) = 1 - c + \delta(x + (1 - x)(1 - (r - b)^2))$$

$$\rightarrow U_b(b) > U_b(s_0) \leftrightarrow x > \frac{c - (s_0 - b)^2}{\delta(s_0 - b)^2} = x^*(c)$$

When $c \geq \Delta_r(s_0)$ then (remember that Red always change b)

$$U_b(s_0) = 1 - (s_0 - b)^2 + \delta\big(x(1 - (s_0 - b)^2) + (1 - x)(1 - (s_0 - b)^2)\big)$$

and

$$U_b(b) = 1 - c + \delta(x + (1 - x)(1 - (r - b)^2))$$

$$\rightarrow U_b(b) > U_b(s_0) \leftrightarrow x > 1 - \frac{(1 + \delta)(s_0 - b)^2 - c}{\delta(r - b)^2} = x^*(c)$$

In both cases it is direct that $x^*(c)$ is increasing in c. **End of Proof.**

Proof of Proposition 2: We first use mathematical induction to prove that

Cond $\boldsymbol{MS_t^r(T)}$: $c > \max\left\{(r - s_0)^2, \left\{\frac{(r-s_0)^2 \sum_{i=0}^{N} \delta^i - \delta x (r-b)^2 \sum_{i=0}^{N-1} \delta^i}{1 + (1-x)x\delta\left(\sum_{i=0}^{N-1} \delta^i - 1\right)}\right\}_{N \in \{1,..,T-t\}}\right\}$

Cond $\boldsymbol{MS_t^b(T)}$: $c > \max\left\{(s_0 - b)^2, \left\{\frac{(s_0-b)^2 \sum_{i=0}^{N} \delta^i - \delta(1-x)(r-b)^2 \sum_{i=0}^{N-1} \delta^i}{1 + (1-x)x\delta\left(\sum_{i=0}^{N-1} \delta^i - 1\right)}\right\}_{N \in \{1,...,T-t\}}\right\}$

In the main text we already proved that when $T = 2$ then s_0 is unchanged when

$$c > (r - s_0)^2;\ c > (s_0 - b)^2; c > (1 + \delta)(s_0 - b)^2 - \delta(1 - x)(r - b)^2$$

We accept that for T periods MS takes place when Cond $MS_2^r(T)$ and Cond $MS_1^b(T)$ are true. We prove that for $T + 1$ periods MS takes place when Cond $MS_2^r(T + 1)$ and Cond $MS_1^b(T + 1)$ are true. To do that we only need to prove that at t = 2 Red doesn´t change the standard if

$$c > \frac{(r - s_0)^2 \sum_{i=0}^{T-2} \delta^i - x(r - b)^2 \sum_{i=0}^{T-3} \delta^{i+1}}{1 + (1 - x)x\delta\left(\sum_{i=0}^{T-3} \delta^i - 1\right)}$$

And at t = 1 Blue does not change the standard if

$$c > \frac{(s_0 - b)^2 \sum_{i=0}^{T-1} \delta^i - (1-x)(r-b)^2 \sum_{i=0}^{T-2} \delta^{i+1}}{1 + (1-x)x\delta(\sum_{i=0}^{T-2} \delta^i - 1)}$$

knowing that in future neither Blue or Red change s_0 and both coalitions set their preferred ideology if the standard has the opposed value. Red faces a net present cost for not being at its optimal of $(r - s_0)^2 \sum_{i=0}^{T-2} \delta^i$ if it keeps s_0 at t = 2. That expression has to be compared to the costs Red faces if it sets r at t = 2. These are two costs. First, the net present cost of changing the standard in the future, which is $c(1 + x(1-x)(\delta^2 + \cdots + \delta^{T-2}))$. Second, the net present cost because the standard is b and not r which is $x(r-b)^2(\delta + \cdots + \delta^{T-2})$. Hence Red does not change s_0 when

$$(r - s_0)^2 \sum_{i=0}^{T-2} \delta^i < \left(1 + (1-x)x \sum_{i=2}^{T-2} \delta^i\right) c + x(r-b)^2 \sum_{i=1}^{T-2} \delta^i$$

which gives us the inequality we wanted to retrieve. Following analogous steps we conclude that Blue keeps the initial standard at t = 1 if and only if

$$(s_0 - b)^2 \sum_{i=0}^{T-1} \delta^i < \left(1 + (1-x)x \sum_{i=2}^{T-1} \delta^i\right) c + (1-x)(r-b)^2 \sum_{i=1}^{T-1} \delta^i$$

Before proving i.-iii., note that the set of values of Cond $MS_2^r(T)$ and the set of values of Cond $MS_1^b(T)$ that define MS take place only when $x < 1/2$ and only when $x > 1/2$ respectively. The reason is that $MS_2^r(T) = MS_1^b(T)$ when $x = 1/2$ and $s_0 = (r+b)/2$.

In order to prove i. it is enough to prove that for $MS_1^b(T)$

$$\left.\frac{(s_0 - b)^2 \sum_{i=0}^{N} \delta^i - \delta(1-x)(r-b)^2 \sum_{i=0}^{N-1} \delta^i}{1 + (1-x)x\delta(\sum_{i=0}^{N-1} \delta^i - 1)}\right|_{x=1} < \left.\frac{(s_0 - b)^2 \sum_{i=0}^{N+1} \delta^i - \delta(1-x)(r-b)^2 \sum_{i=0}^{N} \delta^i}{1 + (1-x)x\delta(\sum_{i=0}^{N} \delta^i - 1)}\right|_{x=1}$$

$$\left.\frac{(s_0 - b)^2 \sum_{i=0}^{N} \delta^i - \delta(1-x)(r-b)^2 \sum_{i=0}^{N-1} \delta^i}{1 + (1-x)x\delta(\sum_{i=0}^{N-1} \delta^i - 1)}\right|_{x=0} > \left.\frac{(s_0 - b)^2 \sum_{i=0}^{N+1} \delta^i - \delta(1-x)(r-b)^2 \sum_{i=0}^{N} \delta^i}{1 + (1-x)x\delta(\sum_{i=0}^{N} \delta^i - 1)}\right|_{x=0}$$

And for $MS_2^r(T)$

$$\left.\frac{(r - s_0)^2 \sum_{i=0}^{N} \delta^i - \delta x(r-b)^2 \sum_{i=0}^{N-1} \delta^i}{1 + (1-x)x\delta(\sum_{i=0}^{N-1} \delta^i - 1)}\right|_{x=1} < \left.\frac{(r - s_0)^2 \sum_{i=0}^{N+1} \delta^i - \delta x(r-b)^2 \sum_{i=0}^{N} \delta^i}{1 + (1-x)x\delta(\sum_{i=0}^{N} \delta^i - 1)}\right|_{x=1}$$

$$\left.\frac{(r-s_0)^2\sum_{i=0}^{N}\delta^i-\delta x(r-b)^2\sum_{i=0}^{N-1}\delta^i}{1+(1-x)x\delta(\sum_{i=0}^{N-1}\delta^i-1)}\right|_{x=0} < \left.\frac{(r-s_0)^2\sum_{i=0}^{N+1}\delta^i-\delta x(r-b)^2\sum_{i=0}^{N}\delta^i}{1+(1-x)x\delta(\sum_{i=0}^{N}\delta^i-1)}\right|_{x=0}$$

But the previous inequalities are equivalent to

$$0<\delta;\ (r-b)^2>(s_0-b)^2$$
$$(r-b)^2>(r-s_0)^2; 0<\delta$$

Which evidently hold. **In order to prove ii.** it is enough to prove that for each N the region

$$\text{Red: } c>\max\left\{(r-s_0)^2, \frac{(r-s_0)^2\sum_{i=0}^{N}\delta^i-\delta x(r-b)^2\sum_{i=0}^{N-1}\delta^i}{1+(1-x)x\delta(\sum_{i=0}^{N-1}\delta^i-1)}\right\} \qquad (A1)$$

$$\text{Blue: } c>\max\left\{(s_0-b)^2, \frac{(s_0-b)^2\sum_{i=0}^{N}\delta^i-\delta(1-x)(r-b)^2\sum_{i=0}^{N-1}\delta^i}{1+(1-x)x\delta(\sum_{i=0}^{N-1}\delta^i-1)}\right\} \qquad (A2)$$

is concave in s_0. Note that when $s_0<(r+b)/2$ then $(s_0-b)<(r-s_0)$ but vice versa when $s_0>(r+b)/2$. Because $Cond^B(x)=\frac{(s_0-b)^2\sum_{i=0}^{N}\delta^i-\delta(1-x)(r-b)^2\sum_{i=0}^{N-1}\delta^i}{1+(1-x)x\delta(\sum_{i=0}^{N-1}\delta^i-1)}$ is increasing with x then as long as $(s_0-b)^2\sum_{i=0}^{N}\delta^i\le(r-s_0)^2 \leftrightarrow s_0\le\frac{r}{1+\sqrt{\sum_{i=0}^{N}\delta^i}}+\frac{\sqrt{\sum_{i=0}^{N}\delta^i}b}{1+\sqrt{\sum_{i=0}^{N}\delta^i}}$ then only (A1) binds and an increment in s_0 relaxes (A1) and hence increases MS. On the other side, because $Cond^R(x)=\frac{(r-s_0)^2\sum_{i=0}^{N}\delta^i-\delta x(r-b)^2\sum_{i=0}^{N-1}\delta^i}{1+(1-x)x\delta(\sum_{i=0}^{N-1}\delta^i-1)}$ is decreasing with x then as long as $(r-s_0)^2\sum_{i=0}^{N}\delta^i\le(s_0-b)^2 \leftrightarrow s_0\ge\frac{\sqrt{\sum_{i=0}^{N}\delta^i}r}{1+\sqrt{\sum_{i=0}^{N}\delta^i}}+\frac{b}{1+\sqrt{\sum_{i=0}^{N}\delta^i}}$ then only (A2) binds and an increment in s_0 restricts (A2) and hence decreases MS. We only have to prove that MS is concave when $s_0\in\left[\frac{r}{1+\sqrt{\sum_{i=0}^{N}\delta^i}}+\frac{\sqrt{\sum_{i=0}^{N}\delta^i}b}{1+\sqrt{\sum_{i=0}^{N}\delta^i}}, \frac{\sqrt{\sum_{i=0}^{N}\delta^i}r}{1+\sqrt{\sum_{i=0}^{N}\delta^i}}+\frac{b}{1+\sqrt{\sum_{i=0}^{N}\delta^i}}\right]$.[18] We first define MS

$$MS=Cte-(r-s_0)^2 - \left[\int_{x^*}^{1}(Cond^B(x)-(r-s_0)^2)\,dx-\int_{0}^{x^{**}}(Cond^R(x)-(r-s_0)^2)\,dx\right]$$

In which

$$Cond^B(x^*)=(r-s_0)^2; Cond^R(x^{**})=(r-s_0)^2 \rightarrow x^*>x^{**} \leftrightarrow (s_0-b)<(r-s_0)$$

If we focus in the case $(s_0-b)<(r-s_0)$ which implies $x^{**}>1/2$ then

[18] Note that $\frac{1}{1+\sqrt{\sum_{i=0}^{N}\delta^i}}=1-\frac{\sqrt{\sum_{i=0}^{N}\delta^i}}{1+\sqrt{\sum_{i=0}^{N}\delta^i}}\le\frac{1}{2}$.

$$\frac{\partial MS}{\partial s_0} = 2(r-s_0) - 2(s_0-b)\left(\sum_{i=0}^{N}\delta^i - \frac{\sum_{i=0}^{N}\delta^i}{1+(1-x^*)x^*\delta(\sum_{i=0}^{N-1}\delta^i - 1)}\right)$$
$$- 2(r-s_0)(1-x^*) - (Cond^B(x^*) - (r-s_0)^2)\frac{\partial x^*}{\partial s_0}$$
$$+ \left[-2(r-s_0)\left(\frac{\sum_{i=0}^{N}\delta^i}{1+(1-x^{**})x^{**}\delta(\sum_{i=0}^{N-1}\delta^i - 1)} - \sum_{i=0}^{N}\delta^i\right)\right.$$
$$\left. + 2(r-s_0)x^{**} + (Cond^R(x^{**}) - (r-s_0)^2)\frac{\partial x^{**}}{\partial s_0}\right]$$
$$= \underbrace{2(r-s_0)(x^*+x^{**})}_{+} + \underbrace{2\sum_{i=0}^{N}\delta^i\left(\begin{array}{c}(r-s_0)\left(1-\frac{1}{1+(1-x^{**})x^{**}\delta(\sum_{i=0}^{N-1}\delta^i-1)}\right) - \\ (s_0-b)\left(1-\frac{1}{1+(1-x^{*})x^{*}\delta(\sum_{i=0}^{N-1}\delta^i-1)}\right)\end{array}\right)}_{+}$$

The proof is completed after we note that in the case that $(s_0 - b) > (r - s_0)$, which implies $x^* < 1/2$, the derivative $\frac{\partial MS}{\partial s_0}$ is negative when s_0 is close enough to r but it is positive when s_0 is close enough to $(r+b)/2$. That shows concavity.

In order to prove iii. we need to prove that $MS_2^r(T)$ is decreasing in x when $x < 1/2$ and $MS_1^b(T)$ is increasing in x when $x > 1/2$. Indeed is the case because

$$\frac{\partial MS_2^r(T)}{\partial x} = -\frac{\begin{bmatrix}(r-s_0)^2\sum_{i=0}^{N}\delta^i\left(\left(\sum_{i=0}^{N-1}\delta^i-1\right)\delta(1-2x)\right)\\ +\delta(r-b)^2\sum_{i=0}^{N-1}\delta^i\left(1+x^2\left(\sum_{i=0}^{N-1}\delta^i-1\right)\delta\right)\end{bmatrix}}{\left(1+(1-x)x\delta(\sum_{i=0}^{N-1}\delta^i-1)\right)^2} < 0$$

$$\frac{\partial MS_1^b(T)}{\partial x} == \frac{\begin{bmatrix}-(s_0-b)^2\sum_{i=0}^{N}\delta^i\left(\left(\sum_{i=0}^{N-1}\delta^i-1\right)\delta(1-2x)\right)\\ +\delta(r-b)^2\sum_{i=0}^{N-1}\delta^i\left(1+(1-x)^2\left(\sum_{i=0}^{N-1}\delta^i-1\right)\delta\right)\end{bmatrix}}{\left(1+(1-x)x\delta(\sum_{i=0}^{N-1}\delta^i-1)\right)^2} > 0$$

End of Proof.

Proof Proposition 1G: We are interested in comparing the pay-offs obtained by Blue when that coalition sets $s_1 = b$ or keeps $s_1 = s_0$ (coalitions are opposed). Those pay-offs are:

$$U_B(b) = \begin{cases}1 - c + \delta\left(x + (1-x)(1-\Delta)\right) & \text{if } c < \Delta \\ 1 + \delta - c & \text{if } c \geq \Delta\end{cases}$$

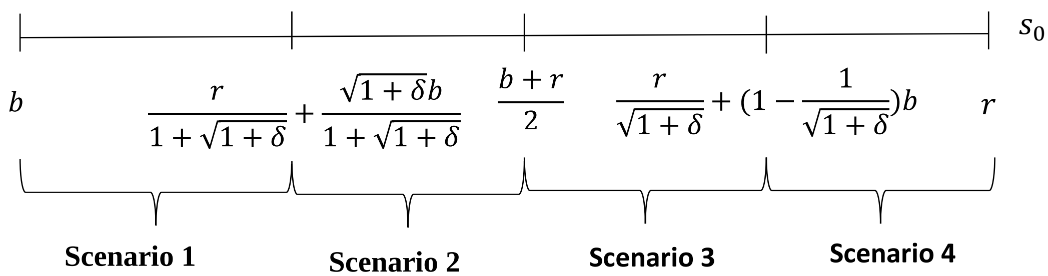
In addition, when $s_0 \in \left[b, \frac{r+b}{2}\right]$

$$U_B(s_0) = \begin{cases} 1-\Delta_b(s_0)+\delta((1-c)x+(1-x)(1-\Delta)) \text{ if } c<\Delta_b(s_0) \\ 1-\Delta_b(s_0)+\delta\big(x(1-\Delta_b(s_0))+(1-x)(1-\Delta)\big) \text{ if } c\in[\Delta_b(s_0),\Delta_r(s_0)] \\ (1-\Delta_b(s_0))(1+\delta) \quad \text{if } c>\Delta_r(s_0) \end{cases}$$

But when $s_0 \in \left[\frac{r+b}{2}, r\right]$

$$U_B(s_0) = \begin{cases} 1-\Delta_b(s_0)+\delta((1-c)x+(1-x)(1-\Delta)) \text{ if } c<\Delta_r(s_0) \\ 1-\Delta_b(s_0)+\delta\big((1-c)x+(1-x)(1-\Delta_b(s_0))\big) \text{ if } c\in[\Delta_r(s_0),\Delta_b(s_0)] \\ (1-\Delta_b(s_0))(1+\delta) \quad \text{if } c>\Delta_b(s_0) \end{cases}$$

To compare these expressions we discuss four scenarios that depend on the value of s_0.



Scenario 1: $s_0 \in \left[b, \frac{r}{1+\sqrt{1+\delta}} + \left(1-\frac{1}{1+\sqrt{1+\delta}}\right)b\right] \to (s_0-b)^2 < (1+\delta)(s_0-b)^2 < (r-s_0)^2 < (r-b)^2$

If $c<(s_0-b)^2$: b is a dominant strategy for Blue and r is a dominant strategy for Red

If $c\in[(s_0-b)^2,(1+\delta)(s_0-b)^2]$: Blue sets b only if x is large enough. Regardless whether Blue sets b or keeps s_0 then Red sets r. Then Blue sets b iff:

$$\begin{aligned}\big(1-c+\delta(1-(r-b)^2)\big)(1-x)+(1-c+\delta)x \\ > \big((1-(s_0-b)^2)+\delta(1-(r-b)^2)\big)(1-x) \\ +(1-(s_0-b)^2)(1+\delta)x\end{aligned}$$

$$\leftrightarrow -c+\delta x > -(s_0-b)^2+\delta x-(s_0-b)^2\delta x$$

$$\leftrightarrow c<(s_0-b)^2(1+\delta x)$$

Which is true only if large enough. To see that, note that the inequality always hold when $x=1$but it is never true when $x=0$.

If $c\in[(1+\delta)(s_0-b)^2,(r-s_0)^2]$: Regardless whether Blue sets b or keeps s_0 then Red sets r. Then Blue sets b iff:

$$\begin{aligned}\left(1-c+\delta(1-(r-b)^2)\right)(1-x)+(1-c+\delta)x \\ > \left((1-(s_0-b)^2)+\delta(1-(r-b)^2)\right)(1-x) \\ +(1-(s_0-b)^2)(1+\delta)x\end{aligned}$$

$$\leftrightarrow -c+\delta x > -(s_0-b)^2+\delta x-(s_0-b)^2\delta x$$

$$\leftrightarrow c < (s_0-b)^2(1+\delta x)$$

Which is never true. Hence Blue keeps s_0 which is also kept by a Blue second period majority and a Red second period majority changes it to r.

If $c \in [(r-s_0)^2,(r-b)^2]$: Both coalitions keep the standard because the cost of a change is too high.

If $c > (r-b)^2$: Both coalitions keep the standard because the cost of a change is too high.

Scenario 2: $s_0 \in \left[\frac{r}{1+\sqrt{1+\delta}}+\left(1-\frac{1}{1+\sqrt{1+\delta}}\right)b,\frac{r+b}{2}\right] \rightarrow (s_0-b)^2 < (r-s_0)^2 < (1+\delta)(s_0-b)^2 < (r-b)^2$

If $c < (s_0-b)^2$: b is a dominant strategy for Blue and r is a dominant strategy for Red

If $c \in [(s_0-b)^2,(r-s_0)^2]$: Blue sets b only if x is large enough. Indeed, regardless whether Blue sets b or keeps s_0 then Red sets r. Then Blue sets b iff:

$$\begin{aligned}\left(1-c+\delta(1-(r-b)^2)\right)(1-x)+(1-c+\delta)x \\ > \left((1-(s_0-b)^2)+\delta(1-(r-b)^2)\right)(1-x) \\ +(1-(s_0-b)^2)(1+\delta)x\end{aligned}$$

$$\leftrightarrow -c+\delta x > -(s_0-b)^2+\delta x-(s_0-b)^2\delta x$$

$$\leftrightarrow c < (s_0-b)^2(1+\delta x)$$

Which is true only if x is large enough. To see that, note that the inequality always hold when $x=1$ but it is never true when $x=0$.

If $c \in [(r-s_0)^2,(1+\delta)(s_0-b)^2]$: Once more Blue sets b only if x is large enough. If Blue sets b then Red sets r and if B keeps s_0 then Red keeps that as well, then Blue sets b iff:

$$\begin{aligned}\left(1-c+\delta(1-(r-b)^2)\right)(1-x)+(1-c+\delta)x \\ > (1-(s_0-b)^2)(1+\delta)(1-x)+(1-(s_0-b)^2)(1+\delta)x \\ =(1-(s_0-b)^2)(1+\delta)\end{aligned}$$

$$\leftrightarrow -c-(r-b)^2(1-x) > -(s_0-b)^2(1+\delta)$$

$$\leftrightarrow c < (s_0-b)^2(1+\delta)-\delta(r-b)^2(1-x)$$

Which is true if and only if x is large enough. To see that note that the inequality always hold when $x=1$ but it is never true when $x=0$ because

$$(s_0-b)^2(1+\delta)-\delta(r-b)^2 < (s_0-b)^2 \leftrightarrow (s_0-b)^2 < (r-b)^2$$

If $c \in [(1+\delta)(s_0 - b)^2, (r-b)^2]$: Both coalitions keep the standard because the cost of a change is too high.

If $c > (r-b)^2$: Both coalitions keep the standard because the cost of a change is too high.

Scenario 3: $s_0 \in \left[\frac{r+b}{2}, \frac{r}{\sqrt{1+\delta}} + \left(1 - \frac{1}{\sqrt{1+\delta}}\right) b\right] \rightarrow (r-s_0)^2 < (s_0-b)^2 < (1+\delta)(s_0 - b)^2 < (r-b)^2$

If $c < (r-s_0)^2$: b is dominant strategy for Blue and b is a dominant strategy for Red.

If $c \in [(r-s_0)^2, (s_0-b)^2]$: Blue decides b at t = 1 only if x is large enough. To see that:

$$1 - c + \delta\big(x + (1-x)(1-(r-b)^2)\big) > 1 - (s_0-b)^2 + \delta((1-c)x + (1-x)(1-(s_0-b)^2))$$

$$\leftrightarrow c < \frac{(s_0-b)^2 - \delta[(r-b)^2 - (s_0-b)^2](1-x)}{1-\delta x}$$

Which always hold when $x = 1$ because the inequality becomes $c < (s_0 - b)^2$ and only holds when c is small enough when $x = 0$. More specifically it only holds when

$$c < (s_0-b)^2(1+\delta) - \delta(r-b)^2$$

Note that we know that $(s_0-b)^2(1+\delta) - \delta(r-b)^2 < (s_0-b)^2$ but $(s_0-b)^2(1+\delta) - \delta(r-b)^2 > (r-s_0)^2$ only when s_0 is large enough.

If $c \in [(s_0-b)^2, (1+\delta)(s_0-b)^2]$: Once more Blue sets b only if x is large enough. If Blue sets b then Red sets r and if B keeps s_0 then Red keeps that as well, then Blue plays b iff:

$$\begin{aligned}\big(1 - c + \delta(1-(r-b)^2)\big)(1-x) + (1-c+\delta)x &> (1-(s_0-b)^2)(1+\delta)(1-x) + (1-(s_0-b)^2)(1+\delta)x \\ &= (1-(s_0-b)^2)(1+\delta)\end{aligned}$$

$$\leftrightarrow -c - (r-b)^2(1-x) > -(s_0-b)^2(1+\delta)$$

$$\leftrightarrow c < (s_0-b)^2(1+\delta) - \delta(r-b)^2(1-x)$$

Which is true if and only if x is large enough. To see that note that the inequality always hold when $x = 1$ but it is never true when $x = 0$ because

$$(s_0-b)^2(1+\delta) - (r-b)^2 < (s_0-b)^2 \leftrightarrow (s_0-b)^2 < (r-b)^2$$

If $c \in [(1+\delta)(s_0-b)^2, (r-b)^2]$: Both coalitions keep the standard because the cost of a change is too high. To see that

$$1 - c + \delta\big(x + (1-x)(1-(r-b)^2)\big) < (1-(s_0-b)^2)(1+\delta)$$

$$\leftrightarrow c > (s_0-b)^2(1+\delta) - \delta(1-x)(r-b)^2$$

Which is always true when $c \in [(1+\delta)(s_0-b)^2, (r-b)^2]$.

If $c > (r-b)^2$: Both coalitions keep the standard because the cost of a change is too high.

Scenario 4: $s_0 \in \left[\frac{r}{\sqrt{1+\delta}} + \left(1-\frac{1}{\sqrt{1+\delta}}\right)b, r\right] \rightarrow (r-s_0)^2 < (s_0-b)^2 < (r-b)^2 < (1+\delta)(s_0-b)^2$

If $c < (r-s_0)^2$: Knowing that Red always decide r at t = 2 then Blue always set b. Formally, it is direct that

$$1-c+\delta\big(x+(1-x)(1-(r-b)^2)\big) > 1-(s_0-b)^2+\delta((1-c)x+(1-x)(1-(r-b)^2))$$

Because $c < (s_0-b)^2$ and $1-c < 1$.

If $c \in [(r-s_0)^2, (s_0-b)^2]$: Blue decides b at t = 1 only if x is large enough. To see that:

$$1-c+\delta\big(x+(1-x)(1-(r-b)^2)\big) > 1-(s_0-b)^2+\delta((1-c)x+(1-x)(1-(s_0-b)^2))$$

$$\leftrightarrow c < \frac{(s_0-b)^2-\delta[(r-b)^2-(s_0-b)^2](1-x)}{1-\delta x}$$

Which always hold when $x=1$ because the inequality becomes $c < (s_0-b)^2$ and only holds when c is small enough when $x=0$. More specifically it only holds when

$$c < (s_0-b)^2(1+\delta)-\delta(r-b)^2$$

Note that we know that $(s_0-b)^2(1+\delta)-\delta(r-b)^2 < (s_0-b)^2$ but $(s_0-b)^2(1+\delta)-\delta(r-b)^2 > (r-s_0)^2$ only when s_0 is large enough.

If $c \in [(s_0-b)^2, (r-b)^2]$: Once more Blue sets b only if x is large enough. If Blue sets b at t = 1 then Red sets r at t = 2 and if B keeps s_0 at t = 1 then Red also keeps that, then Blue sets b iff:

$$1-c+\delta\big(x+(1-x)(1-(r-b)^2)\big) > (1-(s_0-b)^2)(1+\delta)$$

$$\leftrightarrow -c-\delta(r-b)^2(1-x) > -(s_0-b)^2(1+\delta)$$

$$\leftrightarrow c < (s_0-b)^2(1+\delta)-\delta(r-b)^2(1-x)$$

Which is true if and only if x is large enough. To see that note that the inequality always hold when $x=1$ but it is never true when $x=0$ because

$$(s_0-b)^2(1+\delta)-\delta(r-b)^2 < (s_0-b)^2 \leftrightarrow (s_0-b)^2 < (r-b)^2$$

If $c \in [(r-b)^2, (1+\delta)(s_0-b)^2]$: Red always keep the standard and because of that Blue sets b at t = 1. Because that is also dominant when Blue is majority in the second period then Blue sets b and Red keeps b. To see that:

$$1+\delta-c>(1-(s_0-b)^2)(1+\delta)$$
$$\leftrightarrow c<(s_0-b)^2(1+\delta)$$

If $c>(1+\delta)(s_0-b)^2$: Both coalitions always keep s_0 because the cost is too high.

Summarizing:

If $s_0>\frac{r+b}{2}$ then $\underline{c}=(1+\delta)\Delta_b(s_0)-\delta\Delta$ and $\overline{c}=(1+\delta)\Delta_b(s_0)$.

If $s_0<\frac{r+b}{2}$ then $\underline{c}=\Delta_b(s_0)$ and $\overline{c}=(1+\delta)\Delta_b(s_0)$.

In addition, after we separate by scenario, $x^*(c)$ is equal to

In scenario 1:

$$x^*(c)=\frac{c-\delta\Delta_b(s_0)}{\delta\Delta_b(s_0)} \text{when } c\in[\Delta_b(s_0),(1+\delta)\Delta_b(s_0)]$$

In scenario 2:

$$x^*(c)=\begin{cases}\dfrac{c-\delta\Delta_b(s_0)}{\delta\Delta_b(s_0)} \text{ when } c\in[\Delta_b(s_0)\,,\Delta_r(s_0)]\\ \dfrac{c-\big((1+\delta)\Delta_b(s_0)-\delta\Delta\big)}{\delta\Delta} \text{ when } c\in[\Delta_r(s_0),(1+\delta)\Delta_b(s_0)]\end{cases}$$

In scenario 3:

$$x^*(c)=\begin{cases}\dfrac{c-\big((1+\delta)\Delta_b(s_0)-\delta\Delta\big)}{\delta(\Delta-\Delta_b(s_0)+c)} \text{ when } c\in[(1+\delta)\Delta_b(s_0)-\delta\Delta\,,\Delta_b(s_0)]\\ \dfrac{c-\big((1+\delta)\Delta_b(s_0)-\delta\Delta\big)}{\delta\Delta} \text{ when } c\in[\Delta_b(s_0),(1+\delta)\Delta_b(s_0)]\end{cases}$$

In scenario 4:

$$x^*(c)=\begin{cases}\dfrac{c-\big((1+\delta)\Delta_b(s_0)-\delta\Delta\big)}{\delta(\Delta-\Delta_b(s_0)+c)} \text{ when } c\in[(1+\delta)\Delta_b(s_0)-\delta\Delta\,,\Delta_b(s_0)]\\ \dfrac{c-\big((1+\delta)\Delta_b(s_0)-\delta\Delta\big)}{\delta\Delta} \text{ when } c\in[\Delta_b(s_0),\Delta]\\ 0 \text{ when } c\in[\Delta,(1+\delta)\Delta_b(s_0)]\end{cases}$$

In all these cases $\frac{\partial x^*(c)}{\partial c}\geq 0$. That proves the desired results. **End of the Proof.**

Proof of Proposition 1C: Blue's optimal strategy at t = 1 follows after we compare Blue's payoffs when it decides b, $r-\sqrt{c}$ or s_0. First we write general expressions for $U_B(b)$, $U_B(s_0)$ and $U_B\big(r-\sqrt{c}\big)$.

$$U_B(b)=\begin{cases}1-c+\delta\big(x+(1-x)(1-\Delta)\big) \text{ if } c<\Delta\\ 1+\delta-c \text{ if } c\geq\Delta\end{cases}$$

$$U_B(s_0)=\begin{cases}1-\Delta_b(s_0)+\delta((1-c)x+(1-x)(1-\Delta)) \text{ if } c<\Delta_b(s_0)\\ 1-\Delta_b(s_0)+\delta\big(x(1-\Delta_b(s_0))+(1-x)(1-\Delta)\big) \text{ if } c\in[\Delta_b(s_0),\Delta_r(s_0)]\\ (1-\Delta_b(s_0))(1+\delta) \quad \text{if } c>\Delta_r(s_0)\end{cases}$$

$$U_B(r-\sqrt{c})=\begin{cases}1-\Delta_b(r-\sqrt{c})-c+\delta\left((1-c)x+(1-x)\left(1-\Delta_b(r-\sqrt{c})\right)\right) \text{if } c<\dfrac{\Delta}{4}\\ \left(1-\Delta_b(r-\sqrt{c})\right)(1+\delta)-c \quad \text{if } c\geq\dfrac{\Delta}{4}\end{cases}$$

We first show that there exist $\underline{c}$ and $\overline{c}$ such that, for all values of x, Blue changes the standard to b when $c<\underline{c}$ and Blue keeps the standard as s_0 when $c>\overline{c}$. Later we discuss Blue´s strategy when $c\in[\underline{c},\overline{c}]$.

From Proposition 1G we know that Blue prefers $s_1=b$ to $s_1=s_0$ for all values of x when the cost is either smaller than $\Delta_b(s_0)$ (if $s_0>\frac{r+b}{2}$) or smaller than $(1+\delta)\Delta_b(s_0)-\delta\Delta$ (if $s_0<\frac{r+b}{2}$). In addition, when $c<\Delta/4$ Blue prefers $s_1=b$ to $s_1=r-\sqrt{c}$ when

$$U_B(b)>U_B(r-\sqrt{c})\leftrightarrow\Delta_b(r-\sqrt{c})+\delta cx>\delta(1-x)\left(\Delta-\Delta_b(r-\sqrt{c})\right)$$

This inequality holds for all values of x for a given c when $x=0$.[19] That is

$$\Delta_b(r-\sqrt{c})>\frac{\delta}{1+\delta}\ \Delta\leftrightarrow c<\Delta\left(1-\sqrt{\frac{\delta}{1+\delta}}\right)^2$$

We conclude that

$$\underline{c}=\begin{cases}\min\left\{\Delta\left(1-\sqrt{\dfrac{\delta}{1+\delta}}\right)^2,(1+\delta)\Delta_b(s_0)-\delta\Delta\right\} \text{ if } s_0\in\left[\dfrac{r+b}{2},r\right]\\ \min\left\{\Delta\left(1-\sqrt{\dfrac{\delta}{1+\delta}}\right)^2,\Delta_b(s_0)\right\} \text{ if } s_0\in\left[b,\dfrac{r+b}{2}\right]\end{cases}$$

On the other extreme of values of c, we know from Proposition 1G that when $c>(1+\delta)\Delta_b(s_0)$ Blue prefers to keep s_0 instead to set b for all values of x. In addition, after we compare $U_B(s_0)$ with $U_B(r-\sqrt{c})$ it follows that Blue always keeps s_0 when

$$U_B(s_0)>U_B(r-\sqrt{c})\leftrightarrow c>(1+\delta)\Delta_b(s_0)-(1+\delta)\Delta_b(r-\sqrt{c})$$

Which is evidently true when $c>(1+\delta)\Delta_b(s_0)$ in case that $(1+\delta)\Delta_b(s_0)>\Delta_r(s_0)$ and also evidently true when $c>\Delta_r(s_0)$, in case that $\Delta_r(s_0)>(1+\delta)\Delta_b(s_0)$. We conclude that

$$\overline{c}=\begin{cases}\Delta_r(s_0) \text{ if } s_0\in\left[b,\dfrac{1}{1+\sqrt{1+\delta}}r+\left(1-\dfrac{1}{1+\sqrt{1+\delta}}\right)b\right]\\ (1+\delta)\Delta_b(s_0) \text{ if } s_0\in\left[\dfrac{1}{1+\sqrt{1+\delta}}r+\left(1-\dfrac{1}{1+\sqrt{1+\delta}}\right)b,r\right]\end{cases}$$

[19] Clearly $\left(c-\Delta_b(r-\sqrt{c})+\Delta\right)$ is positive because $\Delta>\Delta_b(r-\sqrt{c})$.

At this point we are only missing the strategy followed by Blue in the first period when c takes intermediate values ($c \in [\underline{c}, \overline{c}]$). We have to distinguish scenarios in which s_0 is closer to b or closer to r. We impose that $s_0 \in \left[b, \frac{1}{1+\sqrt{1+\delta}} r + \left(1 - \frac{1}{1+\sqrt{1+\delta}}\right) b\right]$, the solutions for the rest of the cases are analogous.

We show that there exists a set of thresholds in the probability that Blue keeps its majority condition in the second period that substantially changes Blue's strategy in the first period. If x is larger than any of those thresholds then Blue moves the standard closer to its preferred ideology when compared to what Blue decides when x is lower than the threshold.

Because $s_0 \in \left[b, \frac{1}{1+\sqrt{1+\delta}} r + \left(1 - \frac{1}{1+\sqrt{1+\delta}}\right) b\right]$ then $\overline{c} = \Delta_r(s_0)$ but $\underline{c}$ is either equal to $= \Delta\left(1 - \sqrt{\frac{\delta}{1+\delta}}\right)^2$ or equal to $\Delta_b(s_0)$. Next we discuss these two possibilities separately.

If s_0 is close enough to b then $\underline{c} = \Delta_b(s_0)$. That implies that if Blue prefers b to s_0, it also prefers b to $r - \sqrt{c}$. Indeed, when $c \in [\Delta_b(s_0), \Delta_r(s_0)]$ these two preferences are

$$x > x^*(c) = \frac{c - \Delta_b(s_0)}{\delta \Delta_b(s_0)} \quad \text{(Blue prefers } b \text{ to } s_0)$$

$$x > x_2^*(c) = \begin{cases} 1 + \frac{1}{\delta}\left(1 - \frac{c(1+\delta) + \Delta}{2\sqrt{c\Delta}}\right) & \text{if } c \in \left[\Delta_b(s_0), \frac{\Delta}{4}\right] \\ 1 - \frac{1+\delta}{\delta}\left(1 - \sqrt{\frac{c}{\Delta}}\right)^2 & \text{if } c \in \left[\frac{\Delta}{4}, \Delta_r(s_0)\right] \end{cases} \quad \text{(Blue prefers } b \text{ to } r - \sqrt{c}\text{)}$$

In which $x^*(c) > x_2^*(c)$.[20]

Hence Blue prefers to set b than the other two options when $c < (1 + \delta x)\Delta_b(s_0)$. On the other side, we know that Blue prefers s_0 to b when $c \geq (1 + \delta x)\Delta_b(s_0)$. The question is whether that coalition prefers s_0 to $r - \sqrt{c}$ as well. This time the relevant comparison is

$$U_B(s_0) = 1 - \Delta_b(s_0) + \delta\big(x(1 - \Delta_b(s_0)) + (1 - x)(1 - \Delta)\big)$$

with

$$U_B(r - \sqrt{c}) = \begin{cases} 1 - \Delta_b(r - \sqrt{c}) - c + \delta\left((1 - c)x + (1 - x)\left(1 - \Delta_b(r - \sqrt{c})\right)\right) \text{if } c < \frac{\Delta}{4} \\ \left(1 - \Delta_b(r - \sqrt{c})\right)(1 + \delta) - c \quad \text{if } c \geq \frac{\Delta}{4} \end{cases}$$

Which defines the following concave function.

[20] $x_2^*(c)$ is a concave function in c with a slope smaller than $1/\delta\Delta_b(s_0)$ for all $c \in [\Delta_b(s_0), \Delta_r(s_0)]$. It is enough to prove that $\frac{\partial\left(1+\frac{1}{\delta}\left(1-\frac{c(1+\delta)+\Delta}{2\sqrt{c\Delta}}\right)\right)}{\partial c}_{c=\Delta\left(1-\sqrt{\frac{\delta}{1+\delta}}\right)^2} < \frac{1}{\delta\Delta_b(s_0)}$. But that is equivalent to $\Delta_b(s_0) < \Delta\left(1 - \sqrt{\frac{\delta}{1+\delta}}\right)^{\frac{3}{2}} /(1 - (1+\delta)\left(1 - \sqrt{\frac{\delta}{1+\delta}}\right)^2)$ which is true for all δ when $\Delta_b(s_0) < \Delta\left(1 - \sqrt{\frac{\delta}{1+\delta}}\right)^2$.

$$x > x_1^*(c) = \begin{cases} \dfrac{\Delta_b(s_0) + \delta\Delta - (1+\delta)\Delta_b\left(r - \sqrt{c}\right) - c}{\delta\left(\Delta - \Delta_b(s_0) + c - \Delta_b\left(r - \sqrt{c}\right)\right)} \text{ if } c < \dfrac{\Delta}{4} \\ \dfrac{\Delta_b(s_0) + \delta\Delta - (1+\delta)\Delta_b\left(r - \sqrt{c}\right) - c}{\delta\left(\Delta - \Delta_b(s_0)\right)} \text{ if } c \geq \dfrac{\Delta}{4} \end{cases} \text{ (Blue prefers } s_0 \text{ to } r - \sqrt{c}\text{)}$$

Knowing that a contradiction argument implies that $x_1^*(c) < x_2^*(c)$ then $\max\{x^*(c; s_0), x_2^*(c; s_0)\} = x^*(c; s_0)$ and $\min\{x_1^*(c; s_0), x_2^*(c; s_0)\} = x_1^*(c; s_0)$ which by itself implies that $x_2^*(c)$ is irrelevant to characterize Blue's strategy when s_0 is very close to b. Hence, Blue's strategy is summarized by Figure A.1.

<<Insert Figure A.1. about here>>

The previous mechanics can be applied to characterize the case in which s_0 is closer to $\frac{1}{1+\sqrt{1+\delta}} r + \left(1 - \frac{1}{1+\sqrt{1+\delta}}\right) b$. Then $\underline{c} = \Delta\left(1 - \sqrt{\frac{\delta}{1+\delta}}\right)^2 < \Delta_b(s_0)$ which implies that there exist $\hat{c} \in [\Delta_b(s_0), (1+\delta x)\Delta_b(s_0)]$ such that $x^*(\hat{c}) = x_1^*(\hat{c}) = x_2^*(\hat{c})$. The existence of $\hat{c}$ implies that this time

$$\max\{x^*(c; s_0), x_2^*(c; s_0)\} = \begin{cases} x_2^*(c; s_0) \text{ when } c < \hat{c} \\ x^*(c; s_0) \text{ when } c > \hat{c} \end{cases}$$

$$\min\{x_1^*(c; s_0), x_2^*(c; s_0)\} = x_1^*(c; s_0) \text{ for all values of } c.$$

Hence, Blue's strategy is summarized by Figure A.2

<<Insert Figure A.2 about here>>

The proof is completed after we note that the solution satisfies $s_0 < r - \sqrt{c}$ when $s_0 < \frac{r+b}{2}$ but satisfies $s_0 > r - \sqrt{c}$ when $s_0 > \frac{r+b}{2}$. **End of Proof.**

Proof of Proposition 3: We prove the next general formulation in which coalitions can be clustered to the left or to the right.

i. *When* $s_0 < 2b - r < b < r$ *then there exists* $\underline{c}$, $\overline{c}$ *and* $x^*(c)$ *when* $c \in [\underline{c}, \overline{c}]$ *with* $\frac{\partial x^*(c)}{\partial c} > 0$ *such that Blue sets*

$$s_1 = \begin{cases} b \text{ when } c < \underline{c} \\ \begin{cases} b \text{ if } x \geq x^*(c) \\ s_0 \text{ if } x < x^*(c) \end{cases} \text{when } c \in [\underline{c}, \overline{c}] \\ s_0 \text{ when } c > \overline{c} \end{cases}$$

ii. *When* $2b - r < s_0 < b < r$ *then there exists* $\underline{c}$, $\overline{c}, \overline{\overline{c}}$ *and* $x^*(c_1)$ *when* $c \in [\underline{c}, \overline{c}]$ *with* $\frac{\partial x^*(c)}{\partial c} > 0$ *and* $x^{**}(c_1)$ *when* $c \in [\overline{c}, \overline{\overline{c}}]$ *with* $\frac{\partial x^{**}(c)}{\partial c} < 0$ *such that Blue sets*

$$s_1 = \begin{cases} b \text{ when } c < \underline{c} \\ \begin{cases} b \text{ if } x \geq x^*(c) \\ s_0 \text{ if } x < x^*(c) \end{cases} \text{when } c \in [\underline{c}, \overline{c}] \\ \begin{cases} b \text{ if } x \leq x^{**}(c) \\ s_0 \text{ if } x > x^{**}(c) \end{cases} \text{when } c \in [\overline{c}, \overline{\overline{c}}] \\ s_0 \text{ when } c > \overline{\overline{c}} \end{cases}$$

iii. *When $b < r < s_0$ then Blue sets*

$$s_1 = \begin{cases} b \text{ when } c \leq (1+\delta)\Delta_b(s_0) \\ s_0 \text{ when } c > (1+\delta)\Delta_b(s_0) \end{cases}$$

We are interested in comparing the pay-offs obtained by Blue when that coalition sets $s_1 = b$ or ke eps $s_1 = s_0$ (coalitions are aligned). Those pay-offs are given by:

$$U_B(b) = \begin{cases} 1 - c + \delta\big(x + (1-x)(1-\Delta)\big) \text{ if } c < \Delta \\ 1 + \delta - c \text{ if } c \geq \Delta \end{cases}$$

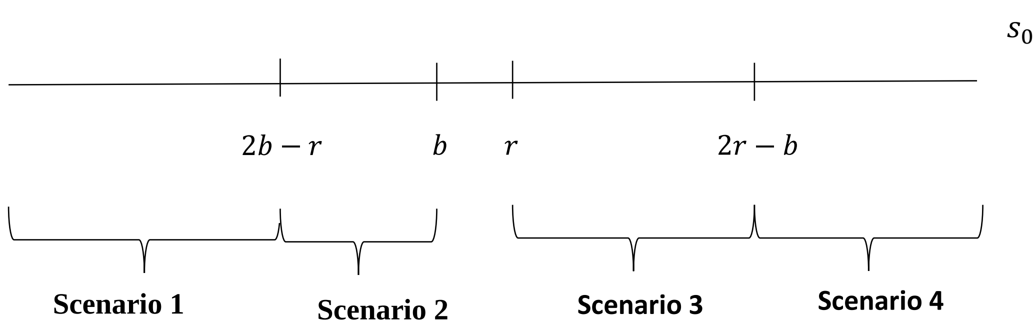
In addition, when $s_0 < b < r$ then

$$U_B(s_0) = \begin{cases} 1 - \Delta_b(s_0) + \delta((1-c)x + (1-x)(1-\Delta)) \text{ if } c < \Delta_b(s_0) \\ 1 - \Delta_b(s_0) + \delta\big(x(1-\Delta_b(s_0)) + (1-x)(1-\Delta)\big) \text{ if } c \in [\Delta_b(s_0), \Delta_r(s_0)] \\ (1-\Delta_b(s_0))(1+\delta) \quad \text{if } c > \Delta_r(s_0) \end{cases}$$

But when $b < r < s_0$ then

$$U_B(s_0) = \begin{cases} 1 - \Delta_b(s_0) + \delta\big((1-c)x + (1-x)(1-\Delta)\big) \text{ if } c < \Delta_r(s_0) \\ 1 - \Delta_b(s_0) + \delta\left((1-c)x + (1-x)\big(1-\Delta_b(s_0)\big)\right) \text{ if } c \in [\Delta_r(s_0), \Delta_b(s_0)] \\ \big(1-\Delta_b(s_0)\big)(1+\delta) \quad \text{if } c > \Delta_b(s_0) \end{cases}$$

To compare these expressions we discuss four scenarios that depend on the value of s_0.



Scenario 1: $s_0 < 2b - r < b < r \rightarrow \Delta < \Delta_b(s_0) < \Delta_r(s_0)$

If $c < \Delta$: b is a dominant strategy for Blue and r is a dominant strategy for Red which implies that Blue sets b at t = 1. To see that

$$1 - c + \delta\big(x + (1-x)(1-\Delta)\big) > 1 - \Delta_b(s_0) + \delta((1-c)x + (1-x)(1-\Delta))$$

$$\leftrightarrow c < \frac{\Delta_b(s_0)}{1 - \delta x}$$

Which evidently holds.

If $c \in [\Delta, \Delta_b(s_0)]$: Once more Blue sets b for all values of x. Indeed Blue sets b iff:

$$1 + \delta - c > 1 - \Delta_b(s_0) + \delta\big(x(1 - \Delta_b(s_0)) + (1 - x)(1 - \Delta)\big)$$

Which is always true because $x(1 - \Delta_b(s_0)) + (1 - x)(1 - \Delta) < 1$ and $c < \Delta_b(s_0)$.

If $c \in [\Delta_b(s_0), \Delta_r(s_0)]$: Blue sets b only if x is large enough otherwise it keeps s_0. Indeed Blue sets b iff:

$$1 + \delta - c > 1 - \Delta_b(s_0) + \delta\big(x(1 - \Delta_b(s_0)) + (1 - x)(1 - \Delta)\big)$$

$$\leftrightarrow c < (1 + \delta x)\Delta_b(s_0) + \delta(1 - x)\Delta$$

$$\leftrightarrow x > \frac{c - (\Delta_b(s_0) + \delta\Delta)}{\delta(\Delta_b(s_0) - \Delta)}$$

Which holds for all x when $c = \Delta_b(s_0)$ but it never holds for $c \in [(1 + \delta)\Delta_b(s_0), \Delta_r(s_0)]$ because the RHS becomes 1 when $c = (1 + \delta)\Delta_b(s_0)$.

If $c > \Delta_r(s_0)$: Both coalitions keep the standard because the cost of a change is too high. To see that

$$1 + \delta - c < \big(1 - \Delta_b(s_0)\big)(1 + \delta) \leftrightarrow c > (1 + \delta)\Delta_b(s_0)$$

Note that $(1 + \delta)\Delta_b(s_0) < \Delta_r(s_0) \leftrightarrow s_0 \in \left[\frac{\sqrt{1+\delta}}{\sqrt{1+\delta}-1} b - \frac{r}{\sqrt{1+\delta}-1}, 2b - r\right]$.

Scenario 2: $2b - r < s_0 < b < r \rightarrow \Delta_b(s_0) < \Delta < \Delta_r(s_0)$

If $c < \Delta_b(s_0)$: b is a dominant strategy for Blue and r is a dominant strategy for Red which implies that Blue sets b at t = 1. To see that

$$1 - c + \delta\big(x + (1 - x)(1 - \Delta)\big) > 1 - \Delta_b(s_0) + \delta((1 - c)x + (1 - x)(1 - \Delta))$$

$$\leftrightarrow c < \frac{\Delta_b(s_0)}{1 - \delta x}$$

Which evidently always hold.

If $c \in [\Delta_b(s_0), \Delta]$: Blue sets b only when x is large enough. To see that, Blue sets b iff:

$$1 - c + \delta\big(x + (1 - x)(1 - \Delta)\big) > 1 - \Delta_b(s_0) + \delta\big(x(1 - \Delta_b(s_0)) + (1 - x)(1 - \Delta)\big)$$

$$\leftrightarrow c < (1 + \delta x)\Delta_b(s_0) \leftrightarrow x > \frac{c - \Delta_b(s_0)}{\delta\Delta_b(s_0)}$$

Note that if $s_0 > \left(\frac{1+\sqrt{1+\delta}}{\sqrt{1+\delta}}\right) b - \frac{1}{\sqrt{1+\delta}} r \in [2b - r, b]$ then $(1+\delta)\Delta_b(s_0) < \Delta$ and then for $c \in [(1+\delta)\Delta_b(s_0), \Delta]$ Blue prefers b to s_0 for all values of x. But if $s_0 < \left(\frac{1+\sqrt{1+\delta}}{\sqrt{1+\delta}}\right) b - \frac{1}{\sqrt{1+\delta}} r$ then $(1+\delta)\Delta_b(s_0) > \Delta$ and Blue prefers b iff x is large enough.

If $c \in [\Delta, \Delta_r(s_0)]$: Blue sets b only if x is small enough otherwise it keeps s_0. Indeed Blue sets b iff:

$$1 + \delta - c > 1 - \Delta_b(s_0) + \delta\big(x(1 - \Delta_b(s_0)) + (1-x)(1-\Delta)\big)$$

$$\leftrightarrow c < (1 + \delta x)\Delta_b(s_0) + \delta(1-x)\Delta$$

$$\leftrightarrow x < \frac{(\Delta_b(s_0) + \delta\Delta) - c}{\delta(\Delta - \Delta_b(s_0))}$$

Which holds for all x when $c = (1+\delta)\Delta_b(s_0)$ because the RHS becomes 1 but it never holds for $c > \Delta_b(s_0) + \delta\Delta$ because the RHS becomes 0. Note that $(1+\delta)\Delta_b(s_0) < \Delta_r(s_0)$ because that is equivalent to $s_0 > \frac{\sqrt{1+\delta}}{\sqrt{1+\delta}-1} b - \frac{1}{\sqrt{1+\delta}-1} r$ which is true because

$$\frac{\sqrt{1+\delta}}{\sqrt{1+\delta}-1} b - \frac{1}{\sqrt{1+\delta}-1} r < 2b - r \leftrightarrow r > b$$

If $c > \Delta_r(s_0)$: Both coalitions keep the standard because the cost of a change is too high. To see that

$$1 + \delta - c < \big(1 - \Delta_b(s_0)\big)(1+\delta) \leftrightarrow c > (1+\delta)\Delta_b(s_0)$$

Which is always true given that $(1+\delta)\Delta_b(s_0) < \Delta_r(s_0)$.

Scenario 3: $b < r < s_0 < 2r - b \rightarrow \Delta_r(s_0) < \Delta < \Delta_b(s_0)$

If $c < \Delta_r(s_0)$: b is a dominant strategy for Blue and r is a dominant strategy for Red which implies that Blue sets b at t = 1. To see that

$$1 - c + \delta\big(x + (1-x)(1-\Delta)\big) > 1 - \Delta_b(s_0) + \delta((1-c)x + (1-x)(1-\Delta))$$

$$\leftrightarrow c < \frac{\Delta_b(s_0)}{1 - \delta x}$$

Which evidently always hold.

If $c \in [\Delta_r(s_0), \Delta]$: Blue always set b. To see that, Blue sets b iff:

$$1 - c + \delta\big(x + (1-x)(1-\Delta)\big) > 1 - \Delta_b(s_0) + \delta\big((1-c)x + (1-x)(1-\Delta_b(s_0))\big)$$

$$\leftrightarrow -c - \delta(1-x)\Delta > -\Delta_b(s_0) - \delta c x - \delta(1-x)\Delta_b(s_0)$$

$$\leftrightarrow c < \frac{\Delta_b(s_0) + (1-x)\delta(\Delta_b(s_0) - \Delta)}{1 - \delta x}$$

Which is always true because we are in a range in which $c < \Delta_b(s_0)$.

If $c \in [\Delta, \Delta_b(s_0)]$: Blue sets b for all values of x because:

$$1 + \delta - c > 1 - \Delta_b(s_0) + \delta\left((1-c)x + (1-x)\big(1 - \Delta_b(s_0)\big)\right)$$

$$\leftrightarrow -c \; > -\Delta_b(s_0) - \delta c x - \delta(1-x)\Delta_b(s_0)$$

$$\leftrightarrow c < \frac{\Delta_b(s_0) + (1-x)\delta\Delta_b(s_0)}{1 - \delta x}$$

Which is always true because we are in a range in which $c < \Delta_b(s_0)$.

If $c > \Delta_b(s_0)$: Both coalitions keep the standard if the cost is large enough otherwise they set b. To see that, Blue sets b iff

$$1 + \delta - c < \big(1 - \Delta_b(s_0)\big)(1+\delta) \leftrightarrow c > (1+\delta)\Delta_b(s_0)$$

Scenario 4: $b < r < 2r - b < s_0 \rightarrow \Delta < \Delta_r(s_0) \; < \Delta_b(s_0)$

If $c < \Delta$: b is a dominant strategy for Blue and r is a dominant strategy for Red which implies that Blue sets b at t = 1. To see that

$$1 - c + \delta\big(x + (1-x)(1-\Delta)\big) > 1 - \Delta_b(s_0) + \delta((1-c)x + (1-x)(1-\Delta))$$

$$\leftrightarrow c < \frac{\Delta_b(s_0)}{1 - \delta x}$$

Which evidently always hold.

If $c \in [\Delta, \Delta_r(s_0)]$: Blue always sets b. To see that, Blue sets b iff:

$$1 - c + \delta\big(x + (1-x)(1-\Delta)\big) > 1 - \Delta_b(s_0) + \delta\big((1-c)x + (1-x)(1-\Delta_b(s_0))\big)$$

$$\leftrightarrow -c - \delta(1-x)\Delta > -\Delta_b(s_0) - \delta c x - \delta(1-x)\Delta_b(s_0)$$

$$\leftrightarrow c \; < \frac{\Delta_b(s_0) + \delta(1-x)\Delta}{1 - \delta x}$$

Which is always true because we are in a range in which $c < \Delta_b(s_0)$.

If $c \in [\Delta_r(s_0), \Delta_b(s_0)]$: Blue sets b for all values of x because:

$$1 + \delta - c > 1 - \Delta_b(s_0) + \delta\left((1-c)x + (1-x)\big(1 - \Delta_b(s_0)\big)\right)$$

$$\leftrightarrow -c \; > -\Delta_b(s_0) - \delta c x - \delta(1-x)\Delta_b(s_0)$$

$$\leftrightarrow c < \frac{\Delta_b(s_0) + (1-x)\delta\Delta_b(s_0)}{1 - \delta x}$$

Which is always true because we are in a range in which $c < \Delta_b(s_0)$.

If $c > \Delta_b(s_0)$: Both coalitions keep the standard if the cost is large enough otherwise they set b. To see that, Blue sets b iff

$$1 + \delta - c < \big(1 - \Delta_b(s_0)\big)(1 + \delta) \leftrightarrow c > (1 + \delta)\Delta_b(s_0)$$

Summarizing:

If $s_0 < 2b - r$ then $\underline{c} = \Delta_b(s_0)$ and $\overline{c} = \Delta_r(s_0)$.

If $s_0 \in [2b - r, b]$ then $\underline{c} = \Delta_b(s_0)$ and $\overline{c} = \Delta$ and $\overline{\overline{c}} = \Delta_r(s_0)$

If $s_0 > r$ then $\underline{c} = \overline{c} = (1 + \delta)\Delta_b(s_0)$.

In addition, after we separate by scenario, $x^*(c)$ is equal to

In scenario 1:

$$x^*(c) = \frac{c - (\Delta_b(s_0) + \delta\Delta)}{\delta(\Delta_b(s_0) - \Delta)} \text{ when } c \in [\Delta_b(s_0), \Delta_r(s_0)]$$

In scenario 2:

$$x^*(c) = \frac{c - \Delta_b(s_0)}{\delta\Delta_b(s_0)} \text{ when } c \in [\Delta_b(s_0), \Delta]$$

And

$$x_1^*(c) = \frac{(\Delta_b(s_0) + \delta\Delta) - c}{\delta\big(\Delta - \Delta_b(s_0)\big)} \text{ when } c \in [\Delta, \Delta_r(s_0)]$$

It is direct that $\frac{\partial x^*(c)}{\partial c} > 0$ and $\frac{x_1^*(c)}{\partial c} < 0$. That proves the desired results. **End of the Proof.**

FIGURES

Figure 1. Timing of the Decisions (Discrete Model)

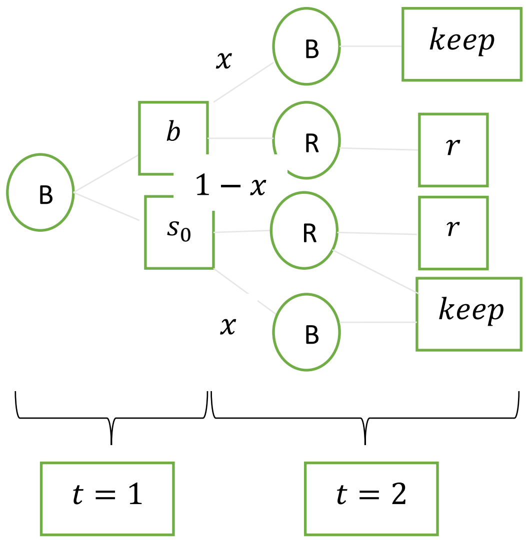


Figure 2. Majority decision at t = 1 when $s_0 \in \left[b, \frac{r}{1+\sqrt{1+\delta}} + \left(1 - \frac{1}{1+\sqrt{1+\delta}}\right) b\right]$.

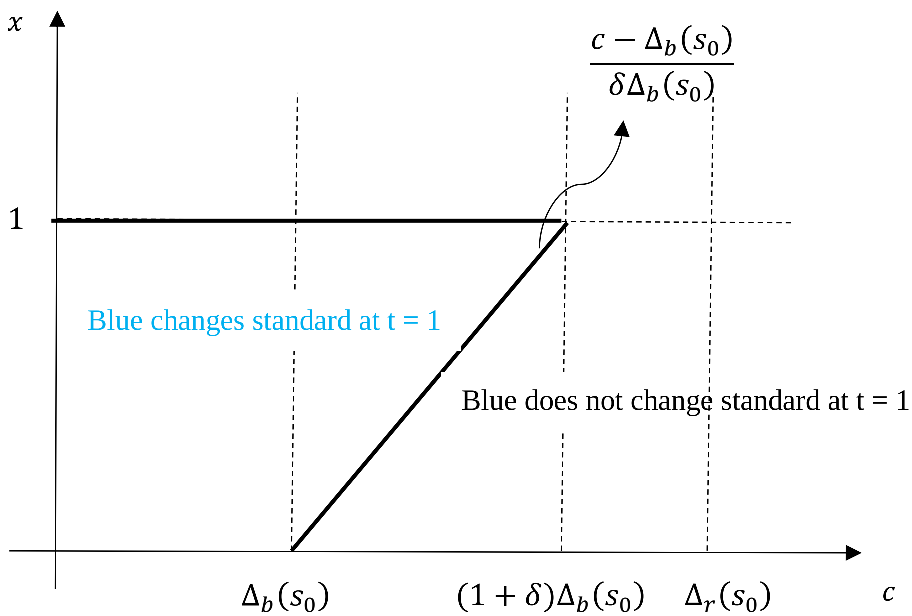


Figure 3. Majority decision at t = 1 when $s_0 \in \left[\frac{r}{1+\sqrt{1+\delta}} + \left(1 - \frac{1}{1+\sqrt{1+\delta}}\right) b, \frac{r+b}{2}\right]$

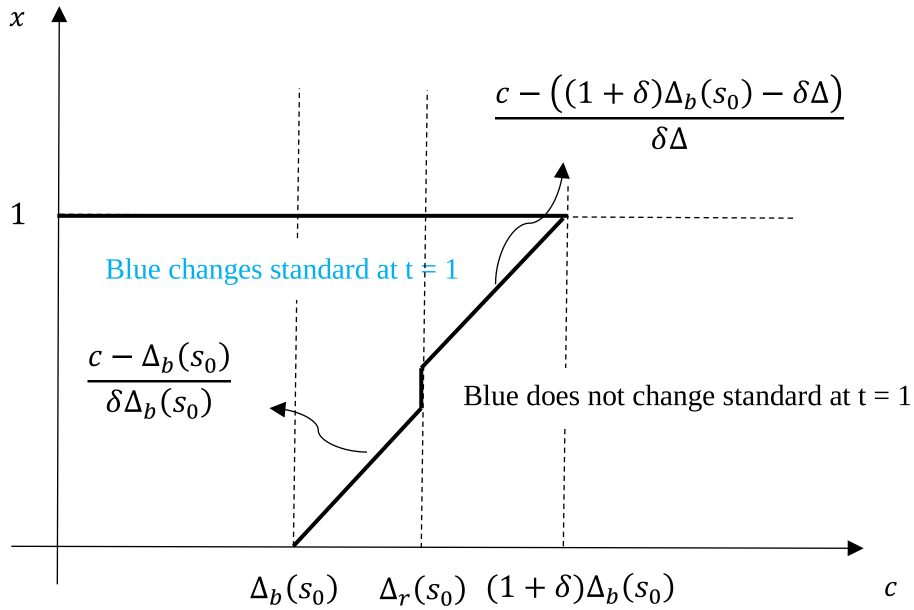


Figure 4. Areas of *maximum stability and maximum instability*

when $s_0 \in \left[b, \frac{r}{1+\sqrt{1+\delta}} + \left(1 - \frac{1}{1+\sqrt{1+\delta}}\right) b\right]$.

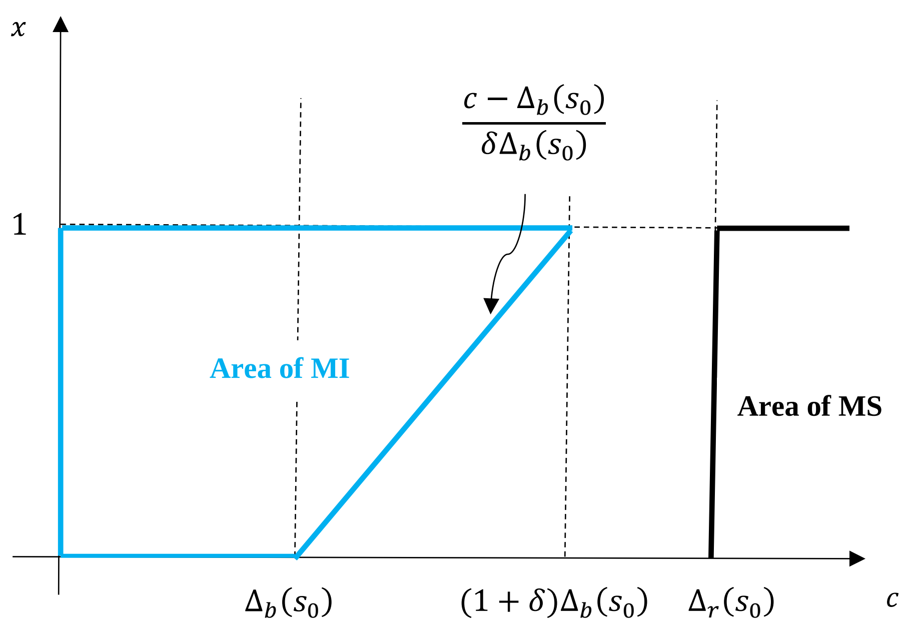


Figure 5. Areas of *maximum stability and maximum instability*

when $s_0 \in \left[\frac{r}{1+\sqrt{1+\delta}} + \left(1 - \frac{1}{1+\sqrt{1+\delta}}\right) b, \frac{r+b}{2}\right]$

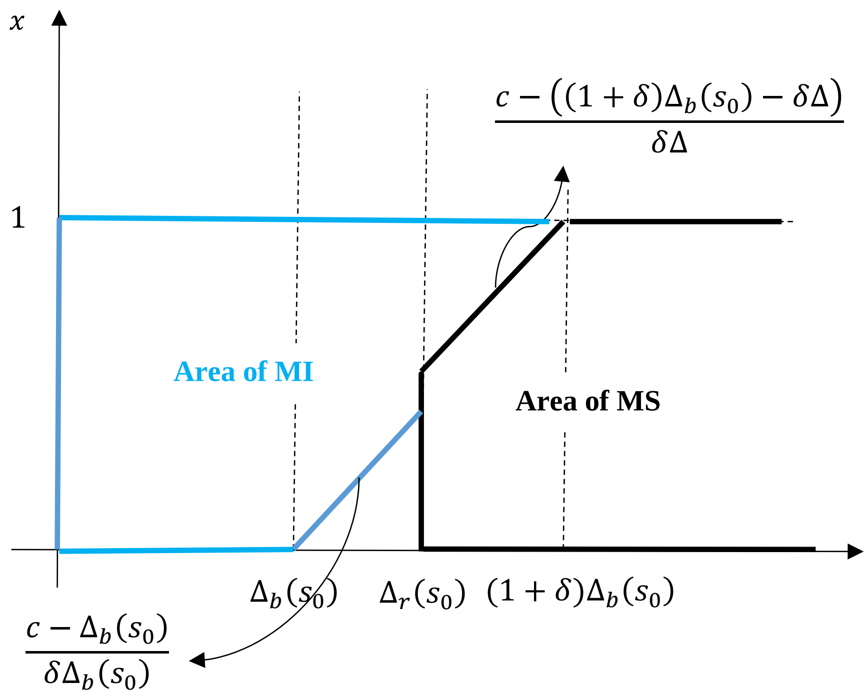


Figure 6. Area of MS decreases with the number of periods.

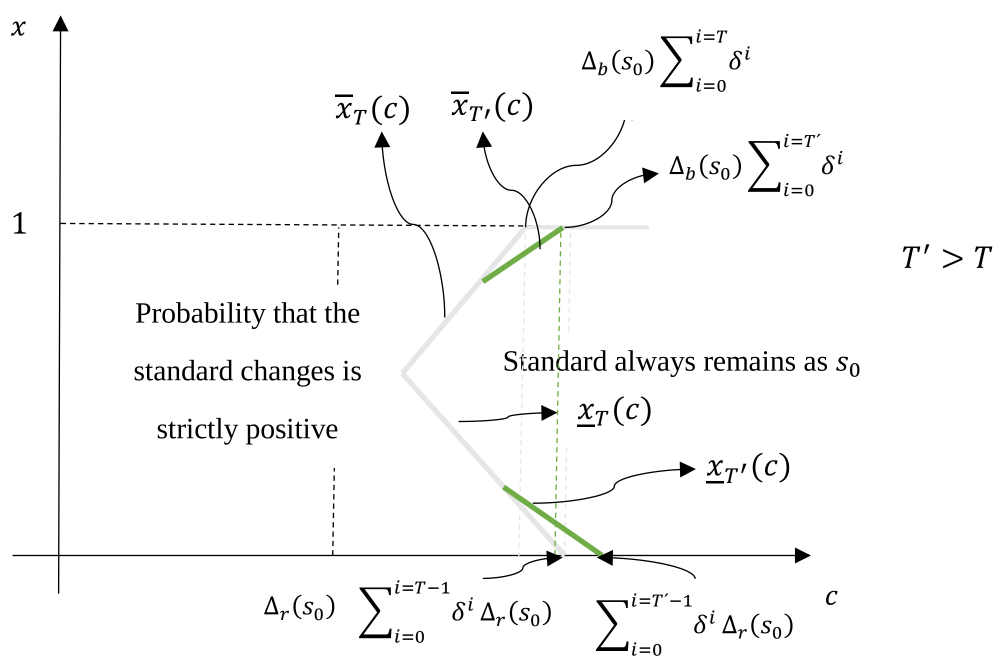


Figure 7. Timing of the Decisions (Continuous Model)

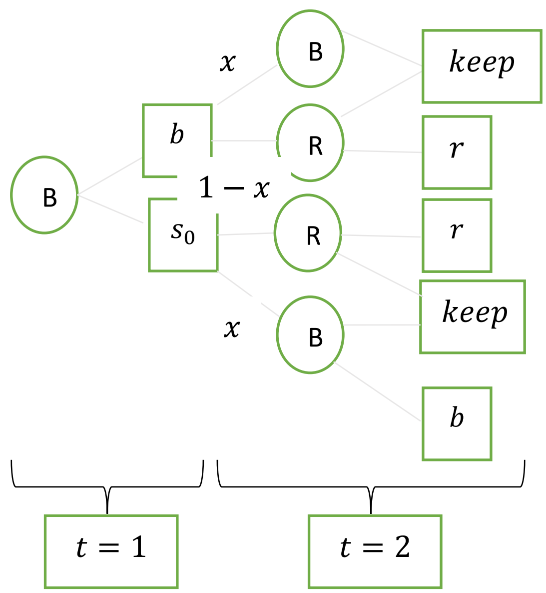


Figure 8. Blue sets b less frequently if r increases: $s_0 \in \left[\frac{r}{1+\sqrt{1+\delta}} + \left(1 - \frac{1}{1+\sqrt{1+\delta}}\right) b, \frac{r+b}{2}\right]$

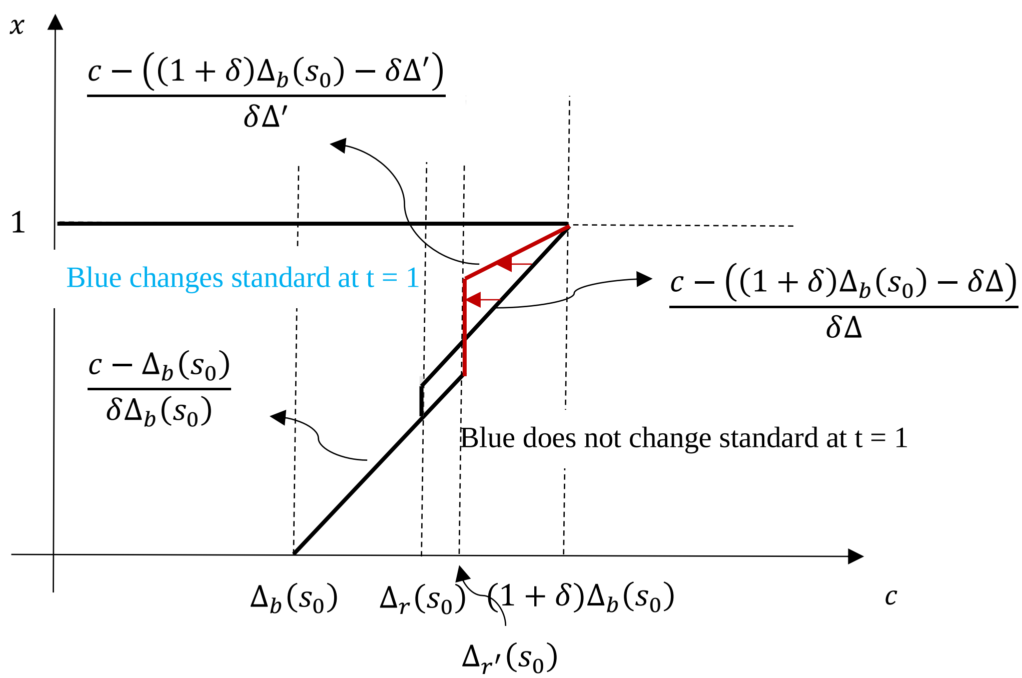


Figure 9. Blue sets b less frequently if r increases: $s_0 \in \left[\frac{r+b}{2}, \frac{r}{\sqrt{1+\delta}} + \left(1 - \frac{1}{\sqrt{1+\delta}}\right) b\right]$

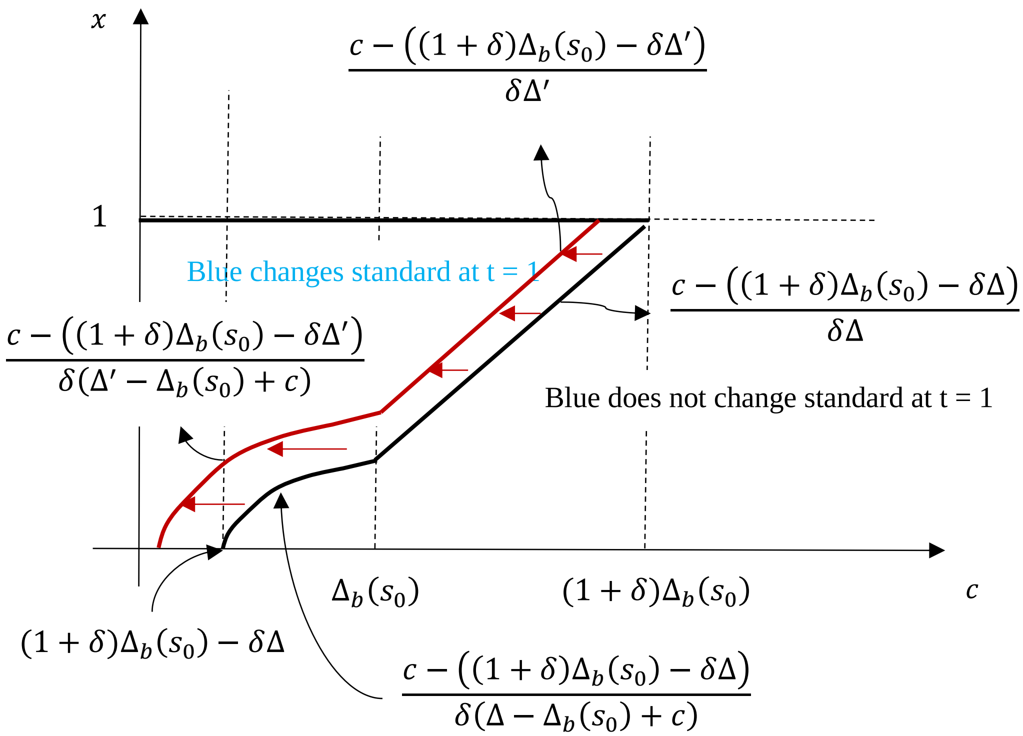


Figure A.1. Majority decision at t = 1 when $s_0 \in \left[b, (1-\sqrt{\frac{\delta}{1+\delta}})r + \sqrt{\frac{\delta}{1+\delta}}b\right]$

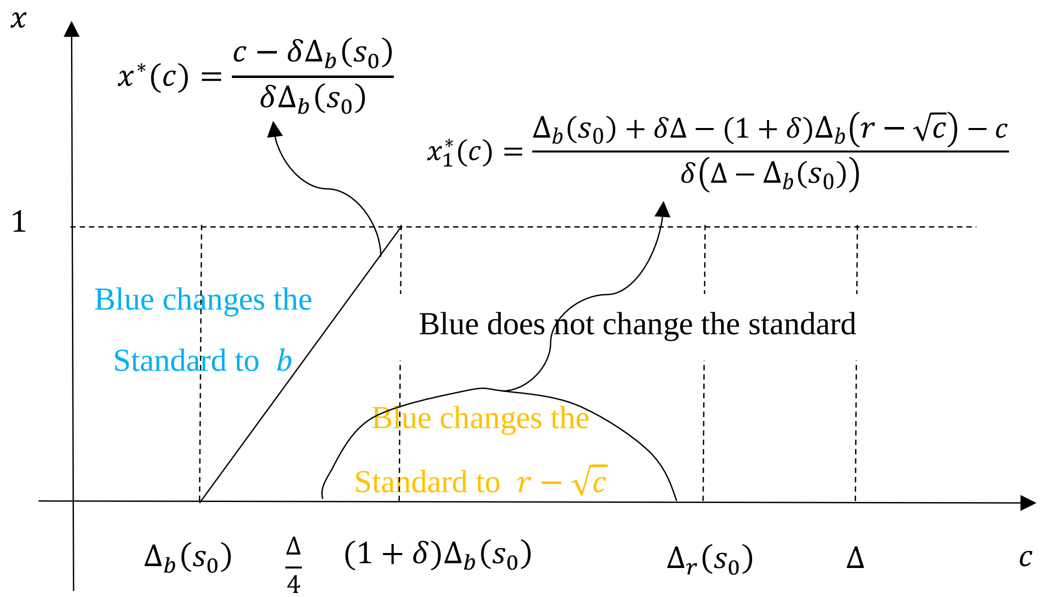


Figure A.2. Majority decision when $s_0 \in \left[\left(1-\sqrt{\frac{\delta}{1+\delta}}\right)r + \sqrt{\frac{\delta}{1+\delta}}b, \frac{1}{1+\sqrt{1+\delta}}r + \frac{\sqrt{1+\delta}}{1+\sqrt{1+\delta}}b\right]$