

# A Theory of Political Compromise\*

by

Avinash Dixit, Gene M. Grossman, and Faruk Gul  
Princeton University

May 1998

## Abstract

We study political compromise founded on tacit cooperation. Two political parties must share a fixed pie in each of an infinite sequence of periods. In each period, the party in power has ultimate authority to divide the pie. Power evolves according to a Markov process among a set of political states corresponding to different degrees of political “strength” for the two. The political strength of each party is a state variable, and the game is *dynamic*, rather than *repeated*. Allocations in an efficient, sub-game perfect equilibrium do not follow a Markov process. Rather, a party’s share reflects not only its current strength, but also how it got there in the recent history. We characterize the efficient division processes for majority rule and supermajority rule, and ask whether one regime allows greater compromise than the other.

**JEL Classifications:** D78, C73

**Keywords:** policy making, compromise, risk sharing, political parties, dynamic games

---

\*We thank Patrick Bolton, Elhanan Helpman, John Huber, Torsten Persson, Patrick Rey and seminar participants at the Copenhagen Business School, Georgetown, M.I.T., Princeton, GREMAQ (Toulouse) and the University of Michigan for their comments and suggestions, and the National Science Foundation for financial support.

# 1 Introduction

In any political system, authority to divide the political and economic spoils rests with the individual or individuals who are “in power”. This power may derive from electoral success, as in a democracy, or from military success, as in many autocracies. In any event, the prevailing rulers can grab everything that is up for grabs: they can decide all ideological issues to their own liking and distribute all economic surplus to themselves and their supporters. Of course, political fortunes may change over time; today’s opposition may become tomorrow’s rulers. Then the new rulers could use their power to their own advantage, fully undoing the decisions and allocations of the past.

Such behavior, and the violent swings of policy it suggests, are rarely observed in democratic polities. The individuals in power do indeed get more of the spoils, but they typically go some way to look after the interests of the opposition, giving it significant representation on legislative committees, and often striving to make policy by consultation and consensus rather than the exercise of raw power. In autocracies, extreme policy swings are more common, as rulers sometimes choose immoderate allocations and impose their ideological wills. But even here, concessions to out-of-power groups are not extraordinary, especially when the ruler has only a tenuous hold on the reigns of power.

A superficial answer would be that constitutions contain safeguards or checks and balances to prevent the tyranny of the majority. But similar compromises occur in polities with few or no constitutional constraints, for example in the U.K. Also, some legislative rules that grant power to minorities, for example cloture rules in the U.S. Senate, are procedures adopted by the legislature itself, not imposed by the constitution. In any case, the questions of how such safeguards operate, and how far they succeed in achieving their purpose of protecting minorities, should themselves be subjects of study.

Downs (1957) offered one possible explanation for policy stability, drawing on insights from Hotelling (1929). If individuals compete for power, their political plat-

forms may tend to converge. This tendency is most pronounced in democracies, where competition is regular and transparent. Candidates and parties who covet power *per se* will announce their intention to carry out the policies that carry the greatest public appeal. Even those with ideological leanings will shade their platforms in the direction of the median voter (Wittman, 1977, 1983).

But, as Alesina (1988) has noted, this explanation can be valid only if prospective rulers have the ability to commit their actions. Otherwise, the leaders will have incentive to resort to their own most-preferred policies once the political competition has been resolved. If those who will choose the leaders (through voting or otherwise) recognize this incentive, they will not give credence to any promises to pursue an alternative course.

The credibility of such promises must hinge on considerations such as the parties' desire to maintain their reputations with the voters over the long run. Alesina recognized that long-lived parties are also engaged in ongoing relationships with each other, and this by itself provides another explanation for stability. Such situations may also arise in some non-democracies, if the same individuals and groups vie for power in successive periods. In these settings, there can be mutual benefit from intertemporal smoothing. Each ruling party might be willing to compromise its actions, if it could trust its successors to follow suit. The question is whether such tacit cooperation can be sustained as an equilibrium of the political game.

Several preconditions for cooperative behavior are obvious. The ruling individuals must perceive an appreciable chance that their power will come to an end. They must be patient, so that future benefits can compensate immediate concessions. And they must foresee a possibility of regaining power, once it is lost. Without this last condition, they cannot reasonably expect their successors to return any favors, when the time comes for the *quid pro quo*.<sup>1</sup>

In this paper we attempt a detailed investigation of political compromise founded on tacit cooperation. We construct an infinite-horizon model in which there are two

---

<sup>1</sup>This last condition is what makes it less likely that dictators will engage in such compromise with their opponents.

political “parties.” In each period, a fixed sum is to be divided between the two, and in each period the power to make the division resides with one of the parties. An exogenous stochastic process governs the evolution of power. That is, we define a set of states – with elements corresponding to different degrees of political “strength” or popularity – in which each party has decision-making authority. Then we specify a Markov process that guides the transitions between states. We characterize the set of efficient allocation rules that might arise in equilibrium. An allocation rule dictates the division of the pie after every possible history of the political game. Equilibrium requires that a rule be self-enforcing; no party can ever have an incentive to cheat, given the repercussions that would ensue.

Our analysis offers a substantial generalization of Alesina (1988) in one important respect, while simplifying his model in another. In Alesina’s paper, there is a stable function linking the platforms of the two parties to their probabilities of winning any election. In equilibrium, when the parties’ platforms are constant, the variable indicating whether a particular party is in power is identically and independently distributed. In other words, there is no state variable in his analysis; a party’s chance of being in power in a given period bears no relation to whether it was so in the last. In our model, by contrast, the political environment evolves according to a stochastic process. Moreover, there can be an arbitrary number of states in this process, each with a different implication for what is likely to happen next. This gives rise to interesting policy dynamics, which are the focus of our analysis: we study how the compromise changes as parties’ fortunes ebb and flow.

What is crucial for such dynamics to be interesting is some persistence of political power. If the probability that a party wins the next election were independent of its vote share at the previous election, then all the states where it is in power could be collapsed into one, and the model would reduce to Alesina’s two-state case. In fact many empirical researchers have estimated Markov processes for vote shares in the US, and found considerable persistence; a recent study is Box-Steffensmeier and Smith (1996). They have also found mean-reverting tendencies so that at any time

the minority party sees a reasonable prospect of gaining power in the foreseeable future; see also Stokes (1966) and Bartels (1997).<sup>2</sup>

Although the underlying stochastic structure of our political process is Markov, we find that efficient allocation rules in general are not. The optimal compromise at any moment depends not only on which party is in power and with what degree of popularity, but also how it got there in the recent history. The compromise will be different when a party with a certain degree of strength was even stronger in the recent past compared to when a party with that same strength climbed there from a weaker position. What makes our problem difficult is the presence of the state variable. This makes the game *dynamic* rather than *repeated*, and renders the well-developed theory of tacit cooperation in repeated games (such as was applied by Alesina) inapplicable.

In order to concentrate on the dynamic aspects of the problem, we have chosen to ignore – at least in this initial exploration – the link between parties’ actions and the evolution of power. As we have indicated, the transition between states in our model is assumed exogenous. In reality, there may indeed be exogenous influences on the evolution of power. For example, voter’s tastes about ideological issues or their perceptions about parties’ competencies may fluctuate randomly in a democracy. Demographic turnover in the population, where children’s political attitudes combine their parent’s attitudes plus random shocks, provides perhaps the simplest visualization of our Markov process. And external events may impinge upon domestic politics in any political system. But policy actions do affect leaders’ survival probabilities, so we would hope to be able to endogenize political support in our future research. This model, with its exogenous process, will continue to constitute an important component of the more general one.

---

<sup>2</sup>John Huber, in a personal communication, related the probability of winning the next election to the vote share at the previous election for a sample of several parliamentary democracies with two-party systems, and found positive persistence. The level of statistical significance was only just under 15 percent, which is low by usual standards. Still, this means that with a probability of more than 85 percent, the relationship could not have arisen by chance, and we think this is good enough justification for constructing a theory of what happens in such circumstances.

The organization and central results of our paper are as follows. In Section 2, we lay out the basic model and introduce notation. Our main theorem, which characterizes an efficient equilibrium division process, is presented in Section 3. It associates with every state of the political system a particular allocation. This allocation is the one that would have left the ruling party indifferent between cooperating and not, had the state in question been the initial one. When the game begins with some particular initial state and a party in power, that party must be given at least the associated share that fulfills its incentive constraint. In subsequent periods, the shares are kept constant so long as this is feasible. Shares are altered only when keeping them constant would violate the incentive constraint for the party then in power, and only to the minimal extent necessary to satisfy the constraint. Thereafter, they are kept constant at the new level until another incentive constraint binds, when the minimal necessary change again is made. And so on.

In Section 4 we investigate the interesting special case that arises when the stochastic process takes the form which we call a “generalized random walk.” In a generalized random walk, the political system can evolve from any state only to one of two “neighboring” states, or it can remain where it was. The states can be interpreted as ordered degrees of popularity, and the restriction to imply that changes in popularity occur only gradually. We show that, in this case, the minimal incentive-compatible shares also can be ordered and illustrate the equilibrium dynamics in a simple diagram.

Section 5 contains an important extension. Although we will have shown by then that compromise can arise naturally as an equilibrium of the dynamic game, we recognize that “founding fathers” often write constitutional rules or develop institutional devices in an apparent attempt to foster even greater cooperation. Most of the clauses in the Bill of Rights of the U.S. constitution are intended to protect the rights of minorities. Rules that require legislative supermajorities for important changes in law or that require two branches of the government or two houses of parliament to concur before a bill can be passed are also aimed at preventing tyranny by the

majority. In our model, some of these rules can be captured by adding states in which no party has unilateral authority to set or change policy. In a supermajority regime, for example, if neither party's support exceeds the designated level, then both parties must agree to any change of the *status quo*. We are able to characterize the efficient compromises in such a regime, and compare them to the outcomes in a regime of simple majority rule. Interestingly, these devices intended to protect the welfare of minorities do not necessarily succeed in doing so; in fact, the division of the pie will often be more uneven (and hence less efficient) in a regime of supermajority rule than in one of simple majority rule.

In the Section 6 we discuss the relationship between our work and some others dealing with risk sharing (for consumers and workers) in the absence of commitment possibilities, and conclude with some suggestions for extensions and avenues for future research.

## 2 The Model

There are two political “parties”, party 1 and party 2. At each of an infinite sequence of dates  $t = 0, 1, 2, \dots, \infty$ , these parties will share a pie of size 1. Party  $i$  has a state-independent instantaneous utility function  $U^i(\cdot)$ , which is a function of the party's share of the pie in the period. We assume that  $U^i$  is continuous, increasing, and strictly concave. Without further loss of generality, we normalize instantaneous utility so that  $U^i(0) = 0$ , and  $U^i(1) = 1$ . The parties share a common discount factor  $\delta$ , and each seeks to maximize the present discounted value of its utilities.<sup>3</sup>

There are  $K$  possible states of the political system. Let  $S = \{1, 2, \dots, K\}$  denote

---

<sup>3</sup>Our analysis readily extends to the more general case where the size of the pie depends on the division, as can happen because of distorting taxes and transfers. Suppose  $y_2 = T(y_1)$ , where  $y_i$  is the amount received by party  $i$ , and  $T$  is a decreasing, concave function (the transformation frontier). We can define a new variable  $\tilde{y}_1 = 1 - T(y_1)$ , and a new utility function  $\tilde{U}_1(\tilde{y}_1) = U[T^{-1}(1 - \tilde{y}_1)]$ . It is easy to verify that  $\tilde{U}_1$  is increasing and strictly concave. The problem expressed in terms of  $\tilde{y}_1$  and  $y_2$  is mathematically identical to the one we solve.

the set of states, and let  $\{S^1, S^2\}$  be a partition of  $S$ . In period  $t$ , political “power” is described by the realization of a random variable  $X_t$ ,  $X_t \in S$ . If  $X_t \in S^1$ , then we say that party 1 is “in power.” This means that party 1 has the right to decide the division of the pie in period  $t$ . Similarly, if  $X_t \in S^2$ , then party 2 is in power, and this party decides how to split the pie in period  $t$ .

For ease of notation, we assume that  $X_0$  is degenerate, that is,  $\Pr[X_0 = x_0] = 1$  for some  $x_0 \in S$ . In other words, we index equilibrium outcomes by the realization of their initial state. Thereafter, an exogenous Markov process  $\phi$  guides the transitions of the political system. We let  $p_{kl}$  denote the transition probability  $\Pr[X_{t+1} = l \mid X_t = k]$ ; i.e.,  $p_{kl}$  is the probability that the system will move from state  $k$  in some period to state  $l$  in the next. As we remarked in the introduction, the assumption of an exogenous transition process is a strong one; it precludes any linkage between the actions taken by a party when in power and its prospects for remaining in power. However, a more complete analysis that would include such linkages would need to solve our problem along the way. And, we believe, the policy dynamics that result from exogenous influences on the evolution of power are interesting in their own right.

A history  $h_t = (x_0, x_1, \dots, x_t)$  refers to a possible realization of the political states  $X_0$  to  $X_t$  in the first  $t + 1$  periods. Let  $H_t$  denote the set of all  $t + 1$  period histories having positive probability; that is,  $H_t = \{h_t \mid \Pr[(X_0, X_1, \dots, X_t) = h_t] > 0\}$ .

In any period, the stage game is quite simple. The parties have full information about the current state and complete recall of the history until then. If the state is in  $S^1$ , then party 1 announces a division of the pie for that period. If the state is in  $S^2$ , it is party 2 that makes the announcement. A strategy for the in-power party, then, is a mapping from the history and the current state to a number in  $[0, 1]$ . The out-of-power party takes no action until it gains or regains power.

We study subgame perfect equilibria of this stochastic, dynamic game. Our focus on such equilibria reflects our belief that long-term political contracts cannot be written, especially when future issues and contingencies are difficult to foresee. Democracies have an underlying constitution, enforced by the judicial branch, that



designates the allocation of power, but places relatively few constraints on the uses of that power. In such circumstances, an agreement between political parties specifying what policy actions they should take would not be enforceable by any third party. Similarly, once a particular dictator has prevailed in the battle for power, there is no one to tell him what to do.

We are interested in the set of efficient equilibria. In dynamic games of complete information, just as in repeated games, the best cooperation is sustained by the implicit threat of the worst credible defection or the most severe credible punishment. These punishments are quite apparent in our model. Following any deviation from an equilibrium path, the deviant party can ensure itself of at least the expected utility associated with a random process that gives it the entire pie whenever it is subsequently in power, and zero otherwise. Since the party can achieve this unilaterally, no worse punishment would be credible. Therefore, the most severe punishment entails a permanent breakdown of cooperation: each party takes the entire pie when it has the opportunity to do so, and accepts zero when the other is in power. Given this anticipated punishment it is clear that, were a party to deviate from any equilibrium with tacit cooperation, it would immediately grab the whole pie.<sup>4</sup>

We will describe equilibrium outcomes as *division processes* which detail how much each party gets at every date  $t$  for all possible states at time  $t$  and all possible histories until then. More precisely, a division process  $\rho = \{Y_0, Y_1, Y_2 \dots\}$  is a stochastic process, where  $Y_t$  is a random variable representing the allocation to party 1 at time  $t$ , and each  $Y_t$  is measurable with respect to the  $\sigma$ -algebra generated by  $\{X_0, X_1, \dots, X_t\}$ . This means that, along an equilibrium path, once we know the realization of the states from  $X_0$  through  $X_t$ , we will know what allocation should emerge in period  $t$ . Since  $X_0$  is degenerate, so too will be the random variable  $Y_0$  describing the allocation in the initial state. We will denote the realizations of  $Y_t$  by  $y_t(h_t)$  and the initial allocation by  $y(x_0)$ .

---

<sup>4</sup>Subgame perfect equilibria protect against unilateral deviations. In the concluding section we consider joint deviations, and argue that efficient renegotiation-proof equilibria will have a similar form as our efficient subgame perfect ones.

Equilibrium outcomes satisfy a set of incentive compatibility constraints. These are constraints that rule out profitable deviations after every possible history. For any division process  $\rho$ , let us define the *payoffs* for the two parties:

$$V_t^1(\rho, h_t) = E \left[ \sum_{\tau \geq t} U^1(Y_\tau) \delta^{\tau-t} \mid (X_0, X_1, \dots, X_t) = h_t \right],$$

and

$$V_t^2(\rho, h_t) = E \left[ \sum_{\tau \geq t} U^2(1 - Y_\tau) \delta^{\tau-t} \mid (X_0, X_1, \dots, X_t) = h_t \right].$$

$V_t^i$  is the expected value for party  $i$ , discounted to time  $t$ , of pursuing the process  $\rho$  from  $t$  onward, if the history until  $t$  has been  $h_t$ . Incentive compatibility requires that the payoff to the party in power from continuing along the path  $\rho$  be at least as great as the payoff from grabbing more than the prescribed share and accepting the associated punishment. An efficient equilibrium must be supported by the most severe punishments available, which here consist of a permanent breakdown of cooperation. We use  $\bar{\rho} = \{\bar{Y}_0, \bar{Y}_1, \dots\}$ , where  $\bar{Y}_t = 1$  if  $X_t \in S^1$  and  $\bar{Y}_t = 0$  if  $X_t \in S^2$ , to denote the process associated with a complete lack of cooperation. Then a feasible division process is characterized by the incentive constraints  $V_t^i(\rho, h_t) \geq V_t^i(\bar{\rho}, h_t)$  for all  $t \geq 0$  and all  $h_t \in H_t$  such that  $x_t \in S^i$ .

Starting at time 0 in state  $x_0$ , one can trace out the whole Pareto frontier of payoffs, namely the set  $(v^1, v^2) \in R^2$  such that

$$v^2 = \max_{\rho \in \mathcal{F}} V_0^2(\rho, x_0) \quad \text{subject to} \quad V_0^1(\rho, x_0) \geq v^1,$$

where  $\mathcal{F}$  is the set of all feasible division processes; i.e., the processes for which  $V_t^i(\rho, h_t) \geq V_t^i(\bar{\rho}, h_t)$  for all  $t \geq 0$  and for all  $h_t \in H_t$  such that  $x_t \in S^i$ . Since  $\rho$  must be feasible,  $V_0^1(\bar{\rho}, x_0)$  is the smallest value  $v^1$  can take if  $x_0 \in S^1$ . Similarly for  $v^2$  if  $x_0 \in S^2$ . This leaves a degree of freedom in dividing payoffs between the parties; we will examine how this freedom can be exercised along efficient paths.

That completes the description of our model. Before we proceed to characterize the efficient equilibria, an observation is in order. We have chosen for expositional simplicity to endow our model with a specific stage game. In the stage game, only

the party in power acts, choosing an allocation for the current period. However, the results we shall present are more general than this. They apply to any game in which the party in power has the ultimate right to decide current policy. For example, it would not be possible to find equilibria that Pareto dominate the ones we shall describe if the stage game allowed for a series of negotiations between the parties, or for threats and promises by the out-of-power party before the ruling party acts. Ours are the most cooperative equilibria in any political game in which one party has final decision authority in every period and enforceable contracts are impossible.

### 3 Efficient Equilibria

In this section we describe the efficient political compromises. By an efficient process, we mean (as usual) a feasible process that maximizes the ex ante expected utility of one party subject to a given floor on the expected utility of the other. In other words, these are equilibria that are not Pareto dominated by any other equilibrium. We begin with a heuristic discussion and proceed to a formal statement of our main result.

The first point to be made is that no policy changes occur during an efficient process unless failure to change would mean that the party in power would be tempted to end the cooperation. Suppose to the contrary that there exists some history with realizations of the state variable in periods  $\tau$  and  $\tau + 1$  such that the parties' shares are not constant and the incentive-compatibility constraint does not bind on the party in power in period  $\tau + 1$ . Then the shares could be smoothed across these two realizations in such a way as to raise each party's expected utility as perceived from time  $\tau$ . This would increase the value of cooperation in all periods up to and including  $\tau$ , and would leave all incentive constraints until then satisfied. And since the incentive constraint does not bind at  $\tau + 1$  for the hypothesized path, a small change in the allocation to the party in power in  $\tau + 1$  would not cause a defection. The argument here is that the parties take advantage of every opportunity to smooth allocations, because their utilities are strictly concave.

Our second point follows from the first. We observe that in general an efficient division process cannot have a Markov structure. This can best be understood by means of an example. Suppose the states correspond to different percentages of popular support, and that popularity can swing by no more than two percentage points from one election to the next. Let the current state be one in which party 1 holds a 51% majority. Now compare two scenarios, one in which its position in the previous period also was a 51% majority and another in which it most recently enjoyed a 53% majority. In the former case, the majority party will be happy with whatever share of the pie it received in the previous period, since the continuation game looks the same now as then. This allocation to the majority party might not be too large, especially if its majority position is quite new and has a good chance of being reversed. In the second scenario, by contrast, the decline in the majority party's political strength causes the incentive constraint to be relaxed. Whatever share was sufficient to prevent defection when the party held a 53% majority will be more than enough to do so after its support has fallen. Since the incentive constraint will not bind in the current period, there can be no change in its allocation from the last. But the inherited share might be quite high, especially if the party recently held majorities even larger than 53%. In other words, a party's allocation typically will be higher in an efficient equilibrium when it achieves a given majority position from above than when it achieves it from below.

An efficient policy path has a reasonably simple form. To facilitate the description, let us associate with each state  $x_0$  in  $S$  a division  $y^*(x_0)$  with the following interpretation:  $y^*(x_0)$  is the first step in the (unique) efficient division process that maximizes the payoff to the party not in power first. This allocation must be just sufficient to stave off an immediate deviation by the party with the initial hold on power; otherwise, the share allocated to the party in power could be lowered without inducing a deviation, thereby increasing the expected utility of the other party.

We describe now how the policy evolves starting from some  $x_0$ . For concreteness, assume that party 1 holds power first. This party receives some initial allocation

$y(x_0)$  which is at least as large as  $y^*(x_0)$ . Thereafter, until the party loses power, its share never declines. If a state arises before party 1 ever loses power with the property that, had that state been the initial state, then party 1 would have required a larger share to induce its cooperation, that is, if  $y(x_0) < y^*(x_t)$ , then the party's share goes up to the larger amount, that is, to  $y^*(x_t)$ . There it again remains fixed, until either party 1 loses power or another state with a still larger initial requirement is realized.

Along any history in  $H_t$ , eventually a state  $x_t \in S^2$  will arise such that power shifts from party 1 to party 2. Had that state been the initial state, party 2 would have received at least  $1 - y^*(x_t)$  and party 1 at most  $y^*(x_t)$ . If  $1 - y^*(x_t)$  is higher than what party 2 received on the last date before it took power (namely,  $1 - Y_{t-1}$ ), then the share for party 2 must be increased (and that for party 1 reduced) to prevent a defection. But if  $1 - Y_{t-1} \geq 1 - y^*(x_t)$  or, equivalently,  $Y_{t-1} \leq y^*(x_t)$ , then the policy remains the same despite the change in government. Thereafter, during a phase in which party 2 remains in power, the policy evolves as described for party 1 previously; its share never falls (so that of party 1 never rises), and it only rises when a state occurs in which the division inherited from the previous period would not have been enough to prevent cheating, had the then prevailing state been the initial one.

The rule repeats along any history, which of course is nothing more than a sequence of phases in which one party or the other holds uninterrupted power separated by periods in which power changes hand.

We turn now to a more formal statement. For a given  $x_0 \in S^i$  and  $j \neq i$ , consider the problem

$$\max_{\rho \in \mathcal{F}} V_0^j(\rho, x_0) .$$

The solution is the feasible process that maximizes the payoff of the party not in power at date 0. In other words, this is the path that exercises the freedom of the Pareto frontier entirely in favor of the party not initially in power. In the appendix, we prove (see Lemma 1) that this solution exists and is unique. Denote it by  $\rho^* = \{y^*(x_0), Y_1^*, Y_2^*, \dots\}$ . The following theorem characterizes all efficient paths in terms

of the initial shares of  $\rho^*$ .

**Theorem 1** *The division process  $\rho = \{y(x_0), Y_1, Y_2, \dots\}$  is efficient if and only if*

$$(i) \ y(x_0) \begin{cases} \geq y^*(x_0) & \text{if } x_0 \in S^1 \\ \leq y^*(x_0) & \text{if } x_0 \in S^2 \end{cases}$$

and

$$(ii) \ Y_t = \begin{cases} \max[y^*(X_t), Y_{t-1}] & \text{if } X_t \in S^1 \\ \min[y^*(X_t), Y_{t-1}] & \text{if } X_t \in S^2, \end{cases}$$

for all  $t \geq 1$ .

The proof has three parts.<sup>5</sup> For concreteness, let  $X_t \in S^1$ . First we show that  $Y_t \geq y^*(X_t)$ . In Lemma 2 we establish that, if two efficient plans yield different payoffs to the party in power first, the one that gives it the greater present value of expected utility also awards it a higher initial share. That is, we prove that the initial share is like a “normal good”. Now recall that, by definition,  $y^*(x_t)$  is the first step in the efficient plan that gives the greatest payoff to party 2, and thus the lowest payoff to party 1, when the initial state is  $x_t$  (for  $x_t \in S^1$ ). Moreover, if a division process is efficient, it must also be efficient from date  $t$  onward following any history  $h_t \in H_t$ . Therefore,  $y(h_t) \geq y^*(x_t)$  for all  $h_t \in H_t$  and all  $x_t \in S^1$ .

Second, we show that for all  $t \geq 1$ ,  $Y_t \geq Y_{t-1}$ . Suppose to the contrary that there exists a history  $h_t$  with  $x_t \in S^1$  such that  $y(h_t) < y(h_{t-1})$ . Then it would be possible to raise  $y(h_t)$  and lower  $y(h_{t-1})$  to some common intermediate value in such a way that expected utility at  $t - 1$  rises for both parties while that for party 1 also rises at  $t$ . But then the payoffs of both parties could be increased without violating incentive compatibility at any date. This contradicts the assumed efficiency of  $\rho$ .

Finally, we show that for  $t \geq 1$ ,  $Y_t \leq \max[Y_{t-1}, y^*(X_t)]$ . Again suppose to the contrary that there exists a history  $h_t$  such that  $y(h_t) > \max[y(h_{t-1}), y^*(x_t)]$ . Then it would be possible to reduce  $y(h_t)$  and increase  $y(h_{t-1})$  at least slightly in such a way as to raise payoffs  $V_{t-1}^i$  and  $V_{t-1}^j$  without violating feasibility. This follows from

---

<sup>5</sup>All proofs are located in the appendix.

the desirability of intertemporal smoothing coupled with the slack in the incentive-compatibility constraint that is implied by  $y(h_t) > y^*(x_t)$ .

Taken together these three statements imply  $Y_t = \max[Y_{t-1}, y^*(X_t)]$ . Analogous arguments can be used to establish that  $Y_t = \min[Y_{t-1}, y^*(X_t)]$  for  $X_t \in S^2$ .

Theorem 1 has several notable implications. The first is that all efficient policy paths look the same after the first time a binding incentive-compatibility constraint is reached. Consider a pair of efficient equilibria that differ in the payoff levels of the two parties. To the extent that one process awards greater expected utility to the party in power first (say party 1), it does so using higher front-loaded shares. The two paths will start from different initial allocations, but as soon as the paths reach a state in which, in both cases, the party in power is just indifferent between defecting from cooperation and not, the efficient processes coincide thereafter. The result follows from the fact that there is a unique way to award a party a specified payoff most efficiently. Since, at a binding incentive constraint, the party in power must get the expected utility associated with the deviation path  $\bar{\rho}$ , the processes must be the same as soon as this happens.

The second implication is that the character of efficient equilibria does not depend on the specifics of the Markov process. Notice that we have made no assumptions about the transition probabilities, nor do they appear in Theorem 1. We can say more about the policy dynamics, however, if we are willing to impose some structure on the process generating  $X_t$ . In particular, we can determine a complete ordering of the  $y^*(k)$  if political power follows what we have termed a generalized random walk. This ordering allows a concise description of the equilibrium policy dynamics, which we provide in the next section.

## 4 Political Power as a Generalized Random Walk

Thus far we did not impose any structure on the states  $S$  or on the transitions between them. These considerations affect the initial share functions  $y^*(\cdot)$ . Therefore it is of interest to find some sufficient conditions on the states and transitions that allow us

to order the  $y^*(x)$ 's.

Consider two states  $k, \ell \in S^1$ . Compare the consequences of starting the game in either of these two states, of course with party 1 initially in power. If this party will surely hold power longer starting in  $\ell$  than in  $k$ , its incentive constraint will be more severe in  $\ell$  than in  $k$ . Then, presumably, it should be given more in  $\ell$  than in  $k$  to stave off defection, that is,  $y^*(\ell) > y^*(k)$ . We now provide a sufficient condition for this to happen.

To this end, let  $k$  be some state in  $S^1$ . Define the set  $P(k)$  as the set of states  $\ell \in S^1$  such that every positive-probability path from  $\ell$  to  $S^2$  goes through  $k$ . Call  $P(k)$  the set of “precursors” of  $k$ . From any state  $\ell$  in  $P(k)$ , party 1 can fall from power only if it first reaches the state  $k$ . This means that loss of power appears unambiguously more distant as seen from state  $\ell$  as compared to state  $k$ . Thus we have

**Proposition 1** *For all  $\ell \in P(k)$ ,  $y^*(\ell) \geq y^*(k)$ .*

This proposition defines a binary relation  $P$  over states, but it is only a partial order relation. Next we find a sufficient condition for the relation to be a complete order. For this we define a *generalized random walk*. We order the states so that  $k \in S^1$  for  $k = 1, \dots, m$  and  $k \in S^2$  for  $k = m + 1, \dots, K$ . The process generating  $X_t$  is a generalized random walk if, for some such ordering, (i)  $p_{ij} = 0$  for all  $i \in \{2, \dots, K - 1\}$  and  $j \notin \{i - 1, i, i + 1\}$ , (ii)  $p_{ij} = 0$  for  $i = 1$  and  $j \notin \{1, 2\}$ , and (iii)  $p_{ij} = 0$  for  $i = K$  and  $j \notin \{K - 1, K\}$ . In a generalized random walk power evolves only gradually; the political system moves to a neighboring state, if it moves at all.<sup>6</sup>

If  $X_t$  follows a generalized random walk, the  $y^*(k)$  can be given a complete ordering. From each state  $k \in S^1$ , the system must pass through  $k + 1$ ,  $k + 2$  etc.

---

<sup>6</sup>Notice that we do not impose any symmetry assumptions here; e.g., leftward movements from any state need not be as likely as rightward movements. We also allow a positive probability of staying in the original state. That is why we call this a “generalized” random walk. The concept is similar to the “simple random walk with variable probabilities” in Cox and Miller (1965, p. 89), except that we have reflecting barriers at both ends.



before party 1 can fall from power. Meanwhile, state  $k$  stands between all states with indexes less than  $k$  and the party's loss of power. Thus,  $P(k) = \{1, \dots, k - 1\}$  for all  $k \leq m$ ; the ordering of states in  $S^1$  reflects decreasing political strength for party 1. Similarly, for states in  $S^2$ , the strength of party 1 falls as the index rises, because each higher-numbered state means more intermediate ones to pass through before the party can regain control. Applying Proposition 1, we have  $y^*(1) \geq y^*(2) \geq \dots \geq y^*(K)$ .

Using this ordering, we can illustrate some examples of efficient policy dynamics. Figures 1 and 2 show possible levels of  $y^*(k)$  for the different values of  $k$ . We may think of the different states  $k$  as representing, for example, different shares of the popular vote, with  $k \leq m$  corresponding to states in which the vote share of party 1 exceeds fifty percent and states with  $k \geq m + 1$  to those in which its vote share falls short of fifty percent.

In Figure 1,  $y^*(m) > y^*(m + 1)$ . This implies that each party would defect immediately even in its weakest ruling state, were it not to receive something more than what would be left over if the other received the least it would accept in *its* weakest ruling state. The arrows show a possible policy trajectory. Starting in state  $a$ , with party 1 in power, suppose the party's majority widens for a while. Then the division will tilt increasingly in its favor as long as its strength continues to grow, until its vote share peaks at some level, say in state  $b$ . Then, as the parties' popularities slowly reverse, the policy remains fixed, until power actually changes hands. A discrete policy shift coincides with the change in majority party, followed by further more-gradual gains for party 2. In short, policy evolves like the turnings of a ratchet; it shifts in favor of the party in power whenever that party attains a new largest majority since its last time in the minority, and remains unchanged as a party's support recedes, or when its gains serve only to restore recent losses.

We can offer some evidence of such outcomes. In Mexico, the Institutional Revolutionary Party (PRI) had a lock on power for many decades, with the result that all patronage jobs went to its activists and public projects to areas of its support. Recently its strength has eroded, but it has not shared the spoils with the opposition

parties to any significant extent. Next, we point out the phenomenon of the “first hundred days,” where a party that comes into power after a long period out of power begins by enacting major policy changes in its first few weeks and months. This conforms well with the story of Figure 1.

Of course, these isolated instances do not constitute a test of the theory. That would require a large systematic empirical project that is clearly beyond the scope of this paper. However, our findings suggest how such a test should proceed: the relevant independent variable for explaining policy outcomes ought not be the majority party’s current electoral strength, but the maximum political support it has enjoyed in this its most recent stint in power.

Figure 2 depicts a different possible dynamic. Here,  $y^*(m) < y^*(m + 1)$ ; that is, a deviation is not tempting for a party with a small majority even if that party were to receive what remains when the other satisfies its minimum requirement in its weakest ruling state. Now, policy jumps might or might not coincide with switches in power from one party to the other, depending upon the history prior to the switch. If party 1 had achieved a large majority (such as  $a$ ) before it fell from grace, then a jump from  $y^*(a)$  to  $y^*(m + 1)$  would occur at the moment party 2 retook the helm. But if party 1 had achieved only a modest majority (such as at  $b$ ) before party 2 experienced its resurgence, then the division of spoils would remain stable right through a change in government, until such a time as one party’s majority grew sufficiently to command a policy change.

At the extreme, if the parties are sufficiently patient, it is possible that the sum of what each must receive in its most powerful state is less than one. That is, if  $\delta$  is sufficiently close to 1, we can have  $y^*(1) < y^*(K)$ . Then the policy of keeping  $Y_t$  forever constant at any number between  $y^*(1)$  and  $y^*(K)$  satisfies all incentive constraints and provides first-best cooperation.

## 5 Supermajority Rule

Political institutions sometimes limit a majority party's authority to set new policies. The U.S. Constitution, for example, contains several "checks and balances" on the exercise of power. New bills need the President's approval, or else sufficient legislative support to override his veto. International treaties require the votes of two thirds of the senators to be ratified. And so on.

Many of the limitations on power, such as those cited in the last paragraph, take a similar form. Typically, new policies can be implemented unilaterally by a party if its political strength is sufficiently great. Otherwise, change requires the acquiescence of the opposition. We will refer to such institutions broadly as regimes of *supermajority rule*. In this section, we extend our model to allow for such rules. We characterize efficient equilibria and compare the policy outcomes to those that arise under simple majority rule.

### Characterization

Recall that for states of the political system in  $S^1$ , we have assumed that party 1 can unilaterally set the current policy. These states we now associate with situations of strong support for party 1. In such circumstances, party 1 can choose any division of the pie in period  $t$ , as before. Similarly, for states in  $S^2$ , party 2 enjoys sufficient strength to satisfy any supermajority requirements. In these states, party 2 sets the allocation  $Y_t$  unilaterally. What is new in our modified model is the addition of a third set of states, labelled  $S^0$ . In these states, neither party has sufficient strength to effect a change in policy. Then the *status quo* continues unless both parties agree to a change.

Both the political procedure and the mathematical formalism of this require some clarification. Procedurally, the division process  $\rho$  can be called a "long term budgetary plan," and the division  $Y_t$  in a particular period can be called a "budget." We think of the long term plan as only an implicit cooperative agreement between the parties.

In each period, the actual shares  $(y_1, y_2)$  for the period must be chosen by enacting a budget. In states when some party has a supermajority, it can enact a budget on its own. In states when neither party has a supermajority, unanimity is required to pass a budget, and if no budget is passed, then the previous period's division remains in effect:  $Y_t = Y_{t-1}$ . In other words, we assume that a continuation resolution is automatic.

More formally, the stage game in period  $t \geq 1$  involves announcements by each party of a proposed allocation for the period. If  $X_t \in S^i$  for  $i = 1$  or  $2$ , then the proposal by party  $i$  is implemented. If  $X_t \in S^0$  and the proposals coincide, the common proposal is put into effect. Otherwise,  $Y_t = Y_{t-1}$ . We restrict attention to processes that have an initial realization  $x_0$  in  $S^1$  or  $S^2$ ; otherwise we would have to make some arbitrary assumption about the initial *status quo*.

As with the simple majority regime, we have stipulated a particular stage game, but the same dynamic equilibrium can be supported by a different stage game. What is crucial is that in the  $S^0$  states, both parties possess the power to veto any *change* of the division from the previous period.

We can now define the defection process  $\bar{\rho}^s = \{\bar{Y}_0^s, \bar{Y}_1^s, \dots\}$  for the supermajority regime. At any time  $t$ , if party 1 has the current supermajority or had the last supermajority, then set  $\bar{Y}_t^s = 1$ . Whereas if party 2 has the current supermajority or had the last supermajority, then set  $\bar{Y}_t^s = 0$ . This process differs from the corresponding process for a regular majority regime in as much as the allocation changes only when a new party captures supermajority support.

Next we specify the incentive-compatibility constraints for the candidate efficient division process  $\rho$ . For  $i = 1, 2$ , we require  $V_t^i(\rho, h_t) \geq V_t^i(\bar{\rho}^s, h_t)$  for all histories  $h_t \in H_t$  such that  $x_t \in S^i$ . This means that a party in a supermajority state must have no unilateral incentive to deviate, considering the punishment that would (eventually) ensue. For histories  $h_t \in H_t$  such that  $x_t \in S^0$  we have two cases to consider. If the process  $\rho$  has  $y(h_t) \neq y(h_{t-1})$  then we must have  $V_t^1(\rho, h_t) \geq V_t^1(\bar{\rho}^s, h_t)$  and  $V_t^2(\rho, h_t) \geq V_t^2(\bar{\rho}^s, h_t)$ ; *both* parties must accede to any change in the allocation,

and so neither can have any incentive to deviate. Alternatively, if the process  $\rho$  has  $y(h_t) = y(h_{t-1})$ , then no incentive constraint applies. Since constancy of shares is the default action, it does not require ratification by either party.

Now we are ready to study efficient equilibrium policy processes. Once again, it helps to define a set of allocations which are the initial steps in the efficient processes starting from the various states in  $S^1$  or  $S^2$  that maximize the payoff of the party out of power first. These are the analogs to  $y^*(x_0)$ , and we denote them by  $y^{*s}(x_0)$ . Formally, we define  $y^{*s}(x_0)$  as the first element in  $\rho^{*s} \equiv \{y^{*s}(x_0), Y_1^{*s}, Y_2^{*s}, \dots\}$ , where  $\rho^{*s}$  solves

$$\max_{\rho \in \mathcal{F}^s} V_0^j(\rho, x_0)$$

for  $x_0 \in S^i$  and  $j \neq i$ . Here  $\mathcal{F}^s$  is the set of feasible division processes; i.e., those that satisfy all of the incentive-compatibility constraints of the supermajority regime.

Theorem 2 characterizes the efficient division processes under supermajority rule.

**Theorem 2** *The division process  $\rho = \{y(x_0), Y_1, Y_2, \dots\}$  is efficient under supermajority rule if and only if*

$$(i) \ y(x_0) \begin{cases} \geq y^{*s}(x_0) & \text{if } x_0 \in S^1 \\ \leq y^{*s}(x_0) & \text{if } x_0 \in S^2 \end{cases}$$

and

$$(ii) \ Y_t = \begin{cases} \max[y^{*s}(X_t), Y_{t-1}] & \text{if } X_t \in S^1 \\ \min[y^{*s}(X_t), Y_{t-1}] & \text{if } X_t \in S^2 \\ Y_{t-1} & \text{if } X_t \in S^0 \end{cases}$$

for all  $t \geq 1$ .

An efficient process begins by giving a party with an initial supermajority enough to deter its immediate defection at the start of the game. Thereafter, policy changes occur only when some party enjoys a supermajority state. Even then, the *status quo* persists unless the previous split of the pie would not have been enough to deter the strong party from grabbing the pie, had the current state been the starting point.

The form of an efficient process under supermajority rule resembles that under simple majority rule. The new feature is the addition of states in which neither party

has authority to change policy unilaterally. Our theorem indicates that no policy changes take place in such states, despite the fact that changes are possible if the two cooperating parties agree to them.<sup>7</sup>

Policy stability in states in which neither party has a strong hold on power might appear to the observer as political “gridlock”. But here the stability is efficient. Suppose there were a history  $h_t$  such that  $x_t \in S^0$  and  $y(h_t) \neq y(h_{t-1})$ . Then, since  $V^1$  and  $V^2$  are strictly concave, the allocations in periods  $t$  and  $t - 1$  could each be replaced by a common intermediate value in such a way that the payoffs of both parties rise at dates on or before  $t - 1$ . No feasibility constraint at  $t - 1$  or before would be violated. Moreover, the new process would be feasible after history  $h_t$ , because holding constant the shares requires no ratification when neither party enjoys a supermajority. It follows that the proposed process with  $y_t(h_t) \neq y_{t-1}(h_{t-1})$  is not efficient.

Although the forms of the processes for the majority and supermajority regimes are similar, the outcomes need not be. The degree of compromise that can be accommodated under either regime depends upon the temptation that a powerful party has to cheat. Here, there are offsetting considerations at work. Compare two regimes: a simple majority regime when the sets of states  $S^1$  and  $S^2$  together cover all states, and a supermajority regime where some of the states in  $S^1$  and  $S^2$  under this simple majority regime are reassigned to  $S^0$ . Now consider a state in which party 1 has power in either regime; i.e., its political strength is sufficient for unilateral action even under supermajority rule. The powerful party will have a greater temptation to cheat in the supermajority regime, because then a deviation could not be punished by party 2 until that party had attained a supermajority of its own. In contrast, party 1 typically would begin to pay for a deviation sooner under a simple majority rule, as party 2 could retaliate as soon as it had gained the narrowest of political advantages.

The force operating in the other direction arises for states belonging to  $S^0$  under

---

<sup>7</sup>In other words, because smoothing over time is desirable, an efficient process proposes no change in the  $S^0$  states, so along the equilibrium path the power of either party to veto a change is never used. Of course it is important to provide such a power, to rule out off-equilibrium deviations.

supermajority rule. Consider one such state in which party 1 has a majority but not a supermajority. Under simple majority rule, party 1 could grab the pie in this state, and so an incentive constraint might bind. In a supermajority regime, the rules prevent party 1 from grabbing the pie without its rival's acquiescence. In fact, neither party could gain by refusing to ratify the budget implied by the efficient budget plan, because the plan preserves the *status quo*, much as the defection process would. The incentive constraints never bind for states in  $S^0$  under supermajority rule. Loosely speaking, then, the supermajority regime imposes fewer (binding) constraints, but those it imposes are more severe.

We will return to the comparison between regimes in a moment. First, let us depict an equilibrium process under supermajority rule when political power follows a generalized random walk. Figure 3 is similar Figures 1 and 2. Along the horizontal axis is an ordering of the states, with the strength of party 1 falling monotonically as we move to the right. For  $i \in \{1, \dots, r\}$ , party 1 has a supermajority, while for  $i \in \{R, \dots, K\}$  party 2 has a supermajority. The figure shows  $y^{*s}(x_0)$  for all states in  $S^1$  or  $S^2$ . The figure also shows some intermediate states which are neither in  $S^1$  nor in  $S^2$ . These are states in which neither party enjoys supermajority support. For  $i \in \{r + 1, \dots, m\}$ , party 1 holds a modest majority, but not a supermajority. Similarly, for  $i \in \{m + 1, \dots, R - 1\}$  party 2 is in this position. As before, the system can evolve from any state only to an adjacent one.

Suppose the system begins with party 1 enjoying a supermajority, such as at point  $a$ . Party 1 must receive an initial allocation at least as large as  $y^{*s}(a)$ . If it gets exactly this amount, and if the party's political strength begins to grow, then its share will rise initially, as indicated by the arrow. Now suppose the strength of party 1 peaks at some level, say at  $b$ . Then the division will remain constant through any ebbs and flows of its political strength (that do not bring it beyond its previous peak strength), at least until party 2 achieves its smallest supermajority. Then party 2 must receive at least  $1 - y^{*s}(R)$ . If the inherited policy gives it less than this amount, the policy must change discretely, as shown in the figure. What distinguishes Figure

3 from Figures 1 and 2 is the existence of the intermediate states (those in  $S^0$ ) at which policy never changes.

## Regime Comparison

We have characterized the efficient compromise processes under each of two regimes, simple majority and supermajority, assuming a very general structure of the model, with any number of states and arbitrary transition probabilities, etc. We also pointed out two opposing forces that affect the relative degrees of compromise achievable under the two. On the one hand, when one party currently has a sufficiently large majority, it can deviate to grab the entire pie and expect to keep it for longer under supermajority rule than under simple majority rule. On the other hand, when a party has a small majority, it can choose to grab the whole pie under simple majority rule but not under supermajority rule. To go beyond the mere recognition of these opposing forces and determine whether one or the other is unambiguously stronger, the general model proves too complex. Therefore we turn to a simpler case, where the issues can be meaningfully posed and answered.

To distinguish between states where each of the parties has a simple majority and a supermajority, we need a minimum of four states, and that is what we assume. We label the states 1 through 4. In state 1, party 1 has complete power under either regime, as does party 2 in state 4. In state 2, party 1 enjoys a modest but not overwhelming majority. In this state, the party can change policy unilaterally under simple majority rule, but requires its rival's acquiescence under supermajority rule. The same is true for party 2 in state 3. Thus,  $S^1 = \{1\}$ ,  $S^2 = \{4\}$ , and  $S^0 = \{2, 3\}$ . The political system evolves among the four states according to a generalized random walk.

For further simplicity, we assume symmetry between the parties. Thus we give them the same utility functions,  $U^1(\cdot) = U^2(\cdot) = U(\cdot)$ , and make the Markov process symmetric, that is, set  $p_{11} = p_{44}$ ,  $p_{12} = p_{43}$ ,  $p_{21} = p_{34}$ , and so on.

We need a criterion or metric to compare the two regimes. For this purpose, we



define an aggregator of the parties' ex ante payoffs, denoted  $W(EV_0^1, EV_0^2)$ , where the expectation is taken over the possible values of  $x_0$ .<sup>8</sup> We take  $W(\cdot)$  to be symmetric and quasi-concave. One interpretation of  $W$  is that it represents the parties' common objective function when they establish "the rules of the game" behind a veil of ignorance. At this stage, they might be ignorant about which of the two will capture power first.

Finally, we compare the two regimes after imposing a common distribution for the initial state. As before, we restrict the starting state to be one of those in which one party or the other enjoys a supermajority. This allows us to avoid specifying what would happen at date 0 if neither party had a supermajority and one failed to approve the initial allocation. To preserve symmetry, we give states 1 and 4 an equal probability of happening first. Now, under both regimes, the initial allocation to the party in power can be chosen optimally, so the comparison is driven by the logic of compromise and not by an arbitrary choice of an initial *status quo*. Let  $\hat{W}$  be the maximum of  $W$  that can be achieved in a regime of simple majority rule. Similarly, let  $\hat{W}^s$  be the maximum value that can be attained under supermajority rule. We find that simple majority rule allows greater political compromise than supermajority rule in the following sense.

**Proposition 2** *In a symmetric polity with four political states,  $\hat{W} \geq \hat{W}^s$ .*

Our method of proof is to identify a pair of processes  $\tilde{\rho}(1)$  and  $\tilde{\rho}(4)$  that are feasible under majority rule when  $x_0 = 1$  and  $x_0 = 4$ , respectively, which together provide the same ex ante welfare as  $\hat{\rho}^s(1)$  and  $\hat{\rho}^s(4)$ , the optimal plans under supermajority rule. Since  $\tilde{\rho}(1)$  and  $\tilde{\rho}(4)$  are not necessarily optimal in a simple majority regime, the maximum attainable  $W$  under majority rule typically exceeds  $\hat{W}^s$ .

We can focus on the case where  $x_0 = 1$ , because the arguments for  $x_0 = 4$  are analogous. We define  $y^{*s} \equiv y^{*s}(1)$  for the supermajority regime and consider first

---

<sup>8</sup>Until now, we have assumed that the initial state has a degenerate distribution and have indexed equilibria by the realization of  $X_0$ . For the purpose of the regime comparison, it makes sense to take an ex ante view, whereby the initial state is random.

the properties of  $\hat{\rho}^s(1) \equiv \{\hat{y}_0^s(1), \hat{Y}_1^s, \dots\}$ . Feasibility requires that  $\hat{y}_0^s(1) \geq y^{*s}$ . But the maximization of  $W$  requires  $y_0(1)$  as close to one half as possible.<sup>9</sup> Therefore,  $\hat{y}_0^s(1) = \max[y^{*s}, \frac{1}{2}]$ . The process that maximizes  $W$  under supermajority rule must be efficient. Then Theorem 2 implies that, for  $t \geq 1$ , the allocation changes only when a new party captures a supermajority. Moreover, the symmetry implies that the party that enjoyed the last supermajority always receives the share  $y^{*s}$ . We construct an alternative plan  $\tilde{\rho}(1)$  by setting  $\tilde{y}_0(1) = y^{*s}$  and assuming that for  $t \geq 1$ , the majority party receives  $y^{*s}$ , no matter what the size of the majority. This process is similar to  $\hat{\rho}^s(1)$ , except that it switches a parties' share from  $y^{*s}$  to  $1 - y^{*s}$  as soon as its rival achieves a simple majority, rather than waiting until the rival gains a supermajority. Our proof establishes that  $\tilde{\rho}(1)$  and the analogously defined  $\tilde{\rho}(4)$  are feasible under majority rule (when the initial states are 1 and 4, respectively), and that together they yield the same ex ante aggregate welfare as  $\hat{\rho}^s(1)$  and  $\hat{\rho}^s(4)$ .

It is instructive to consider why the process  $\tilde{\rho}(1)$  is feasible under simple majority rule, but not under supermajority rule. Compared to  $\hat{\rho}^s(1)$ , the plan  $\tilde{\rho}(1)$  accelerates the time when a party with a large majority must pay back its rival for its favorable current allocation. This reduces expected utility for the party holding a supermajority. Since party 1 is just indifferent at date 0 between cooperating and not when the state is 1 and the process is  $\hat{\rho}^s(1)$  under supermajority rule,  $\tilde{\rho}(1)$  is not feasible in this regime.

However, a deviation is less valuable to a party holding a supermajority when the regime is one of simple majority rule than when it is one of supermajority rule. Under majority rule, the deviant gets 1 only so long as it holds on to some sort of majority. As soon as the rival achieves even a modest electoral victory, the rival takes the pie in a punishment phase. Whereas under supermajority rule, the deviant can continue to take the pie even after its rival has achieved a small majority, by blocking any change to the *status quo*. Only when the rival attains a supermajority does the punishment

---

<sup>9</sup>This follows from the symmetry and quasi-concavity of  $W$ , the fact that for any efficient  $\rho = \{y_0(1), Y_1, Y_2, \dots\}$ ,  $y_0(1) \geq \frac{1}{2}$  if and only if  $V_0^1(\rho, 1) \geq V_0^2(\rho, 1)$ , and the fact that  $V_0^2(\rho, 1)$  falls with a rise in  $y_0(1)$ .

set in. Moreover, the regime switch (from supermajority to simple majority) effects a greater reduction in the value of cheating than the switch from  $\hat{\rho}^s(1)$  to  $\tilde{\rho}(1)$  does to the value of cooperation. Thus, a party with a current supermajority elects to cooperate when the budget plan is  $\tilde{\rho}(1)$  and the regime is one of majority rule.<sup>10</sup>

Next we explain why  $\tilde{\rho}(1)$  and  $\tilde{\rho}(4)$  yield the same ex ante aggregate welfare as  $\hat{\rho}^s(1)$  and  $\hat{\rho}^s(4)$ . The ex ante expected payoff for party 1 is  $EV^1 = \frac{1}{2}V_0^1(\tilde{\rho}(1), 1) + \frac{1}{2}V_0^1(\tilde{\rho}(4), 4)$ ; i.e., the payoff starting from states 1 and 4, weighted by the probability of each initial state. That of party 2 is  $\frac{1}{2}V_0^2(\tilde{\rho}(1), 1) + \frac{1}{2}V_0^2(\tilde{\rho}(4), 4)$ . But the symmetry implies  $V_0^1(\tilde{\rho}(4), 4) = V_0^2(\tilde{\rho}(1), 1)$  and  $V_0^2(\tilde{\rho}(4), 4) = V_0^1(\tilde{\rho}(1), 1)$ . Similar equalities hold for  $\hat{\rho}^s(1)$  and  $\hat{\rho}^s(4)$ . So the comparison of ex ante welfare under the alternative plans hinges on a comparison of the joint utility of the two parties that is achieved by each plan starting from state 1. Note that the sum of the parties' expected utilities equals the expectation of the sum of their utilities.<sup>11</sup> Moreover, each of the plans  $\tilde{\rho}(1)$  and  $\hat{\rho}^s(1)$  gives the same aggregate utility  $U(y^{*s}) + U(1 - y^{*s})$  in every period. It follows that the plans provide the same ex ante aggregate welfare  $W$ .

Our finding that  $\hat{W} \geq \hat{W}^s$  calls into question an often-heard rationale for supermajority rules. These rules frequently are offered as a way to prevent tyranny of the majority. But our example suggests that they may have unintended consequences. In our example, a supermajority rule impedes political compromise, by strengthening the hand of any party that manages to achieve a supermajority. With less compro-

---

<sup>10</sup>For the feasibility of  $\tilde{\rho}$  under simple majority rule, we also need to confirm that no incentive constraints are violated in political states 2 and 3. This is done in the appendix.

<sup>11</sup>That is,

$$\begin{aligned}
V_0^1(\tilde{\rho}(1), 1) + V_0^2(\tilde{\rho}(1), 1) &= \sum_{t=0}^{\infty} \delta^t \sum_{h_t \in H_t} \Pr(h_t | x_0 = 1) \cdot U^1[\tilde{y}(h_t)] + \\
&\quad \sum_{t=0}^{\infty} \delta^t \sum_{h_t \in H_t} \Pr(h_t | x_0 = 1) \cdot U^2[1 - \tilde{y}(h_t)] \\
&= \sum_{t=0}^{\infty} \delta^t \sum_{h_t \in H_t} \Pr(h_t | x_0 = 1) \cdot \{U[\tilde{y}(h_t)] + U[1 - \tilde{y}(h_t)]\}
\end{aligned}$$

mise possible, the policy in every period caters more to the party that last had a large majority and less to the other party than would be the case under simple majority rule. In this sense, ordinary majority rule does more to protect the weak than supermajority rule, and symmetric parties would be unanimous in their preference for it if the question were put to them before each knew its political strength.<sup>12</sup>

The astute reader may be wondering at this point why we compared majority rule with supermajority rule and did not entertain a simple arrangement that could guarantee an efficient outcome in our model. Why not write into the constitution that the parties' shares are fixed and immutable, no matter what the prevailing strengths of the two parties? This "takes the issue out of politics," and may seem an attractive solution to the commitment problem. But in reality there are strong arguments against such a rigid constitution. Circumstances of the parties and the economy may change in a way that makes it optimal to change the division. To the extent that such changes are foreseen, at least as the states of a known stochastic process, a more complex efficient division rule can be laid down once and for all. But if even the possible states and the stochastic laws can undergo unforeseeable changes, then a more flexible constitution open to revision is needed. Even if such flexibility is not present, extreme circumstances create an irresistible pressure for change. Therefore we reject the apolitical solution.

## 6 Relation to the Literature

Our theoretical structure bears similarity to that of some other models of self-enforcing risk-sharing arrangements, most notably Thomas and Worrall (1988) and Kocherlakota (1996).

Thomas and Worrall consider a tacit wage contract offered by a risk-neutral firm to a risk-averse worker. The wage in the outside or spot market fluctuates stochastically.

---

<sup>12</sup>We have made the comparison between regimes holding the transition probabilities fixed. If the evolution of political strength were endogenous, it is possible that the transition processes would differ under alternative rules for delegating power.

When that wage is sufficiently high, the worker can walk away forever from the relationship with this firm. When it is sufficiently low, the firm can replace the worker with a new one. They characterize efficient contracts subject to the worker's and firm's incentive constraints, that is, each must receive at all times at least the expected present value associated with a permanent move to the spot market.

Kocherlakota considers an ongoing interaction between two individuals, each of whom receives an income stream that fluctuates stochastically. Each can consume his own endowment (autarky), but being risk-averse, both stand to gain by sharing. Perfect, externally enforced, Arrow-Debreu insurance would require transfers at all dates and in all states so as to equalize the ratios of the two consumers' marginal utilities across dates and states. However, without external enforcement of insurance contracts, such a first-best need not be compatible with the two individuals' incentives. When one of them gets a large positive income shock, he may do better by consuming it all and refusing to make the stipulated transfer to the other, even if this causes the tacit agreement to collapse and autarky to prevail for ever after.

Kocherlakota characterizes incentive compatible allocations, such that at all dates and in all states neither party gains by renegeing in this way. He finds two important results: (1) The requirement of incentive compatibility gives rise to persistence in consumption. If a shock cannot be smoothed across states because of incentive constraints, some mutual gain can be achieved by smoothing it over time. That is, if one individual has high income today and must be given more utility right now to prevent his renegeing, this is better done by giving a little more consumption today and the implicit promise of a little more tomorrow when he might not have such a high income, than by giving a lot of consumption today and decreasing it tomorrow if his income falls. (2) The allocation rule has limited memory. If person 1's incentive constraint binds today, then the history of past periods when this person's incentive constraint was slack is irrelevant. Thomas and Worrall find a wage-persistence property similar to (1).

In our model, the individuals are replaced by political parties, and we found

similar results about their tacit arrangements for efficient and incentive-compatible sharing of political and economic spoils. However, the political setting differs in some essential ways from that of the macrodynamics of consumption and wages, and yields some new results.

Most importantly, Thomas and Worrall (1988) and Kocherlakota (1996) assume that the stochastic shocks to income are independently and identically distributed over periods. We allowed persistence of political power: the shocks follow a Markov chain. Thomas and Worrall assert that allowing spot wages to follow a Markov process would not change their results significantly, but in fact it introduces an important new consideration. In their paper, a worker has more to gain by switching to the spot market when the wage there is high than when it is low, but solely because of the extra value of this period's earnings; the expectations for future periods are the same in the two cases. Similarly, Kocherlakota's individuals have greater incentive to renege when the current income is high solely because of the extra utility of this period's consumption.

In our model, we ordered the states by a measure of the strength of a party's political power. We supposed that a party might have a more solid hold on power in state 1 than in state 2, in the sense that the next date at which it will lose power is stochastically more distant when starting in state 1 than starting in state 2. If the party defects to grab everything in state 1, therefore, it expects to keep getting everything for longer than it would starting in state 2. Thus we identified a much stronger dynamic source of gain, extending for many periods, that makes the incentive constraint of the party in power tighter in state 1 than in state 2. To maintain incentive compatibility of the compromise allocation, the party in power must be given more in state 1 than in state 2. We obtained a clear and intuitive sufficient condition (Proposition 1) for ordering the share of the party in power.

Next, we extended the analysis to consider the inherently political question of designing the rules of the sharing game, that is, a constitution. If the constitution stipulates that the *status quo* can only be overturned by a supermajority, then there

may be times when neither party can grab the whole pie. A party that reneges at such times can only block anticipated changes in the allocation of the spoils. The existence of these states enlarges the zone of persistence of consumption, and the allocations when one party has a supermajority are in turn affected by the expectation of such persistence. The party with the current supermajority knows that the allocation it creates is unlikely to be overturned for a long time. Therefore it has greater temptation to renege, and must be given a larger share to preserve its incentive to continue cooperation, than would be the case under a simple majority constitution. Thus the supermajority rule can lead to less sharing and therefore worse allocations.

Finally, we used different techniques to prove our characterization theorems. Our proofs, which use only the concavity of the utility functions and do not require differentiability, are at once simpler and more general.

## 7 Future Research

In conclusion, we suggest some extensions and directions for future research, building on our model of self-enforcing compromise.

The first extension concerns our equilibrium concept. Our efficient subgame perfect equilibria are sustained by the drastic punishment of a total collapse of cooperation following any deviation. Once such a punishment path ensues, neither party can by itself do anything to restore cooperation. But the two may get together and agree to a better alternative. The expectation of such a development will in turn reduce the severity of the punishment that a deviator reckons with. In other words, our subgame perfect equilibria are not in general renegotiation-proof.

There is no broadly accepted theory of renegotiation in repeated and dynamic games; see Fudenberg and Tirole (1991, pp. 174–182) and Myerson (1991, pp. 408–412) for discussions of alternative concepts. But all approaches have one thing in common – they entail a more severe incentive compatibility constraint than does the requirement of subgame perfectness. This means that our definition of a feasible division process  $\rho$ , namely one satisfying  $V_t^i(\rho, h_t) \geq V_t^i(\bar{\rho}, h_t)$  for all  $t \geq 0$  and

all  $h_t \in H_t$  such that  $x_t \in S^i$ , must be amended to substitute a larger right hand side in the inequality. However, the left hand side is unchanged, and therefore all the arguments about the gains from intertemporal smoothing are also unchanged. Therefore efficient renegotiation-proof equilibria will have the same general form as our subgame perfect ones. They will be characterized by a set of numbers  $y^r(x_0)$  corresponding to possible initial states  $x_0$ . These will be the smallest shares that must be awarded to the party initially in power. Thereafter the share of a party in power will never decrease, increasing only if its incentive constraint becomes tighter, and so on. Of course the actual numbers  $y^r(x_0)$  for the renegotiation-proof case will be different from the  $y^*(x_0)$  of the subgame perfect case. In the former, with a tighter incentive constraint, the party in power must be given a greater share; therefore less cooperation can be sustained.

A second extension would broaden the set of political institutions. These institutions in reality have a more complex structure than our assumption that the majority (or the supermajority as the case may be) makes the decision. The legislature may be bicameral; power may be divided between the legislature and the executive with veto and override provision; power may be divided between a central government and regional governments; and so on. Future work may consider these variants and their effects on political compromise. Many of these institutions increase the possibility of divided government. As we have seen in the case of supermajority regimes, this may or may not be conducive to better outcomes.

Another extension would expand the number of political parties beyond the two considered here. If there were several such parties, there might be many states in which none has the political strength to set policy unilaterally. Then coalitions would be needed to constitute a legislative majority. Members of a ruling coalition could share the spoils among themselves, but efficiency would require compromise with non-members for intertemporal smoothing. Since coalitions might dissolve and change over time, each party would have to take account of this in evaluating the ongoing tacit cooperation. Presumably, rules governing the process of coalition formation would



be important for determining the ability to compromise. But the most important extension of our model would endogenize the political process by which parties gain or lose power. This would make the transition probabilities of our Markov process depend on voters' retrospective or prospective judgments about the performance of the parties. In a retrospective approach, aggregate output might follow a Markov process, and the party in power be more likely to retain or even increase its power when the aggregate output is high. In a prospective approach, the policy competence of the party in power might affect the aggregate output process, and voters use the output realization to draw inferences about competence.

Even without fluctuations in aggregate outcomes, the distribution of a fixed total can have a feedback effect on voting. What we have modeled as the parties' payoffs ultimately go back to their supporters. A party that gets a larger share can spread it further and perhaps attract more voters at the margin. Alternatively, voters might simply punish at the polls a party that shows itself to be excessively greedy. There is probably an element of realism in such voting behavior even though it typically does not benefit voters directly and immediately in material terms. Such a social norm can have the value of being conducive to greater compromise and therefore better outcomes for all in the long run.

## References

- [1] Alesina, Alberto (1988), "Credibility and Policy Convergence in a Two-Party System with Rational Voters," *American Economic Review*, 78, 796-805.
- [2] Bartels, Larry M. (1997), "Electoral Continuity and Change, 1868-1996," paper presented at the Donald Stokes Memorial Panel, American Political Science Association, Washington, DC, August 1997.
- [3] Box-Steffensmeier, Janet M. and Smith, Renée M. (1996), "The Dynamics of Aggregate Partisanship," *American Political Science Review*, 90, 567-580.
- [4] Cox, D.R. and Miller, H.D. (1965), *The Theory of Stochastic Processes*, London: Chapman and Hall.
- [5] Downs, Anthony (1957), *An Economic Theory of Democracy*, New York: Harper Row.
- [6] Fudenberg, Drew and Tirole, Jean (1991), *Game Theory*, Cambridge, MA: MIT Press.
- [7] Hotelling, Harold (1929), "Stability in Competition," *Economic Journal*, 39, 41-57.
- [8] Kocherlakota, Narayana R. (1996), "Implications of Efficient Risk Sharing without Commitment," *Review of Economic Studies*, 63, 595-609.
- [9] Myerson, Roger B. (1991), *Game Theory: Analysis of Conflict*, Cambridge, MA: Harvard University Press.
- [10] Stokes, Donald E. (1966), "On the Existence of Forces Restoring Party Competition," in *Elections and the Political Order*, Angus Campbell, Philip E. Converse, Warren E. Miller, and Donald E. Stokes eds., New York: Wiley, 180-193.
- [11] Thomas, Jonathan, and Worrall, Tim (1988), "Self-Enforcing Wage Contracts," *Review of Economic Studies*, 55, 541-554.

- [12] Wittman, Donald (1977), "Candidates with Policy Preferences: A Dynamic Model," *Journal of Economic Theory*, 15, 86-103.
- [13] Wittman, Donald (1983), "Candidate Motivation: A Synthesis of Alternatives," *American Political Science Review*, 77, 142-157.

## Appendix

### Proof of Theorem 1:

First we prove that every efficient division process has the specified form (the “only if” part). In fact we consider a more general problem of maximizing the expected present value of the party not in power at date 0, subject to giving the one in power a payoff at least equal to an arbitrarily specified value  $\alpha$  (which need not equal the payoff under the punishment division process  $\bar{\rho}$ ).

Since the values of period utilities  $U^i$  are bounded in  $[0,1]$ , and the payoff functions  $V_t^i$  are defined as the expected present values of utilities, the values of  $V_t^i$  are also bounded, and lie in the range  $[0, 1/(1 - \delta)]$ . Therefore, for any  $\alpha \in [0, 1/(1 - \delta)]$ , any initial state where party  $i$  is in power ( $x_0 \in S^i$ ), and  $j \neq i$ , we can define

$$\mathcal{L}(\alpha) = \max_{\rho \in \mathcal{F}} V_0^j(\rho) \quad \text{subject to} \quad V_0^i(\rho) \geq \alpha.$$

A division process that solves this problem for some  $\alpha$  will be labelled *efficient*. We first examine existence and uniqueness.

**Lemma 1:** For any  $x_0 \in S^i$  and  $\alpha \in [0, 1]$ , either  $\mathcal{L}(\alpha)$  has a unique solution, or  $V_0^i(\rho) < \alpha$  for all  $\rho \in \mathcal{F}$ .

**Proof of Lemma 1:** Obviously, if  $V_0^i(\rho) < \alpha$  for all  $\rho \in \mathcal{F}$ , then the maximization problem has no solution. So consider the opposite case, and define

$$M(\alpha) = \sup_{\rho \in \mathcal{F}} V_0^j(\rho) \quad \text{subject to} \quad V_0^i(\rho) \geq \alpha.$$

Specifying a division process is equivalent to specifying a sequence of real numbers, one after each  $t$ -period history of the Markov process  $\phi$ . Therefore each  $\rho$  can be identified with an element of  $[0, 1]^\infty$ . Endow this space with the product topology. Then the payoff functions  $V_0^1, V_0^2$  are continuous, and the feasible set  $\mathcal{F} = \{\rho \mid V_t^i(\rho, h_t) \geq V_t^i(\bar{\rho}, h_t) \text{ for all } t \geq 0 \text{ and for all } h_t = (x_0, x_1, \dots, x_t) \text{ such that } x_t \in S^i\}$ , is compact. Consequently, there exists a subsequence  $\rho^{(n)} \in \mathcal{F}$  such that  $\rho^{(n)} \rightarrow \rho \in \mathcal{F}$  and  $V^j(\rho, x_0) = M(\alpha)$ . This is then a solution to  $\mathcal{L}(\alpha)$ . Uniqueness follows immediately from the strict concavity of  $V^i$  and  $V^j$  and the convexity of  $\mathcal{F}$ . ■

We denote the unique solution of  $\mathcal{L}(\alpha)$  by  $\rho^\alpha = \{y^\alpha(x_0), Y_1^\alpha, Y_2^\alpha, \dots\}$ .

In particular, if we set  $\alpha = V_0^i(\bar{\rho}, x_0)$ , the constraint on the maximization problem is the incentive compatibility constraint of the party initially in power. This is automatically satisfied by the trigger division process, so  $\mathcal{L}(V_0^i(\bar{\rho}, x_0))$  has a unique solution. It is what we denoted by  $\rho^* = \{y^*(x_0), Y_1^*, Y_2^*, \dots\}$ .

Now we proceed to the necessity part of the theorem. We first state and prove two lemmata:

**Lemma 2:** Let  $\alpha_1, \alpha_2$  be any two numbers in  $[0, 1]$ . For  $m = 1, 2$ , denote the unique solutions to  $\mathcal{L}(\alpha_m)$  by  $\rho^m = \{y^m(x_0), Y_1^m, Y_2^m, \dots\}$ . Then  $\alpha_1 > \alpha_2$  implies

$$y^1(x_0) \geq y^2(x_0) \text{ if } i = 1 \quad \text{and} \quad y^1(x_0) \leq y^2(x_0) \text{ if } i = 2.$$

This says that in any given initial state, if you tighten the constraint on the payoff of the player in power, you have to start him off with a larger share.

**Proof of Lemma 2:** We consider only the case  $i = 1$  and  $j = 2$ ; the proof the other way round is symmetric. Define

$$W^2(\beta) = \max_{\rho \in \mathcal{F}} \mathbb{E}_{X_1} [V_1^2(\rho, (x_0, X_1))] \quad \text{subject to} \quad \mathbb{E}_{X_1} [V_1^1(\rho, (x_0, X_1))] \geq \beta.$$

The argument used in the proof of Lemma 1 establishes that if the feasible set of this maximization problem is non-empty, then it has a unique solution; that is,  $W^2(\beta)$  is attained for some  $\rho$ . Then we can re-write the maximization problem  $\mathcal{L}(\alpha)$  as

$$\max_z \left\{ U^2(1 - z) + \delta W^2 \left( \frac{\alpha - U^1(z)}{\delta} \right) \right\}.$$

Since  $\rho^1$  is the unique solution to  $\mathcal{L}(\alpha_1)$ , we have

$$U^2(1 - y^1(x_0)) + \delta W^2 \left( \frac{\alpha_1 - U^1(y^1(x_0))}{\delta} \right) > U^2(1 - y^2(x_0)) + \delta W^2 \left( \frac{\alpha_1 - U^1(y^2(x_0))}{\delta} \right).$$

Rearranging yields

$$\delta \left[ W^2 \left( \frac{\alpha_1 - U^1(y^1(x_0))}{\delta} \right) - W^2 \left( \frac{\alpha_1 - U^1(y^2(x_0))}{\delta} \right) \right] > U^2(1 - y^2(x_0)) - U^2(1 - y^1(x_0)).$$

Similarly, since  $\rho^2$  is the unique solution of  $\mathcal{L}(\alpha^2)$ , we have

$$\delta \left[ W^2 \left( \frac{\alpha_2 - U^1(y^2(x_0))}{\delta} \right) - W^2 \left( \frac{\alpha_2 - U^1(y^1(x_0))}{\delta} \right) \right] > U^2(1 - y^1(x_0)) - U^2(1 - y^2(x_0)).$$

Combining these two, we get

$$\begin{aligned} & W^2 \left( \frac{\alpha_1 - U^1(y^1(x_0))}{\delta} \right) - W^2 \left( \frac{\alpha_2 - U^1(y^2(x_0))}{\delta} \right) \\ > & W^2 \left( \frac{\alpha_2 - U^1(y^1(x_0))}{\delta} \right) - W^2 \left( \frac{\alpha_2 - U^1(y^2(x_0))}{\delta} \right) \end{aligned} \quad (\text{A1})$$

The utility functions  $U^i$  are concave; this ensures that the set  $\mathcal{F}$  of feasible division processes is convex, and the function  $W^2$  is concave. Therefore

$$W^2(a + \Delta) - W^2(a) > W^2(b + \Delta) - W^2(b) \text{ for } a > b \text{ implies } \Delta < 0.$$

Using this with suitable interpretations of  $a$ ,  $b$  and  $\Delta$  in (1) yields

$$\frac{1}{\delta} [ U^1(y^2(x_0)) - U^1(y^1(x_0)) ] < 0.$$

Since  $U^1$  is increasing, this implies  $y^1(x_0) > y^2(x_0)$ . ■

**Lemma 3:** Let  $h_t = \{x_0, x_1, \dots, x_t\} \in H_t$ . Let  $\rho = \{y_0, Y_1, Y_2, \dots\}$  be a division process. If  $y(h_t) \neq y(h_{t-1})$ , then there exists another division process  $\hat{\rho}$  such that  $V_0^j(\hat{\rho}, x_0) > V_0^j(\rho, x_0)$  for  $j = 1, 2$ .

This simply says that intertemporal smoothing is beneficial. In specific contexts in the proof of the theorem, its feasibility still needs to be checked.

**Proof of Lemma 3:** Let  $p = p_{lm}$  for  $l = x_{t-1}$  and  $m = x_t$ . Define

$$\gamma \equiv [\delta p y(h_t) + y(h_{t-1})]/(\delta p + 1),$$

and define  $\hat{\rho}$  by  $\hat{\rho}(h_{t-1}) = \hat{\rho}(h_t) = \gamma$  and  $\hat{\rho} = \rho$  for all other dates and/or histories. The result follows from the strict concavity of the  $V^i$ . ■

**Proof of Theorem:** We will consider the case  $X_t \in S^1$ ; the proof for the other case is symmetric. To prove

$$Y_t = \max[y^*(X_t), Y_{t-1}],$$

we need to prove three things.

Part I: For all  $t \geq 0$ ,  $Y_t \geq y^*(X_t)$

Suppose for any  $t$  and a history  $h_t \in H_t$  we have  $y(h_t) < y^*(x_t)$ . Let  $\alpha$  be the value of  $V_t^1(\rho)$  after history  $h_t$ , and  $\alpha^*$  the corresponding value of  $V_t^1(\bar{\rho})$ .

The continuation of  $\rho$  from  $(t, h_t)$  must be efficient; therefore it solves  $\mathcal{L}(\alpha)$ . And, since  $\phi$  is Markov,  $\alpha^* = V_0^1(\bar{\rho})$  if the starting state is  $x_t$ . Our assumption that  $y(h_t) < y^*(x_t)$ , using Lemma 2, implies  $\alpha \leq \alpha^*$ , and uniqueness of efficient processes (Lemma 1) strengthens this to  $\alpha < \alpha^*$ . But then  $V_t(\rho) < V_t(\bar{\rho})$ , so  $\rho \notin \mathcal{F}$ , contradicting feasibility.

Part II: For all  $t \geq 1$ ,  $Y_t \geq Y_{t-1}$

If for some history  $h_t$  we have  $y(h_t) < y(h_{t-1})$ , then define  $\gamma$  as in Lemma 3. By concavity of  $U^1$ ,

$$U^1(h_{t-1}) + \delta p U^1(h_t) < U^1(\gamma) + \delta p U^1(\gamma),$$

and the values on all other realizations are equal. Therefore the  $\hat{\rho}$  constructed in Lemma 3 is feasible, and there it was shown to give higher expected utilities for both players. This contradicts the efficiency of  $\rho$ .

Part III: For all  $t \geq 1$ ,  $Y_t \leq \max[y^*(X_t), Y_{t-1}]$

Suppose for some  $t$ , and a history  $h_t \in H_t$  we have  $y(h_t) > y^*(x_t)$ . Defining  $\alpha$  and  $\alpha^*$  as in Part I, above, here we have  $\alpha > \alpha^*$ .

Suppose we simultaneously have  $y(h_t) > y(h_{t-1})$ , and define  $\hat{\rho}$  as in Part II above. For some  $\lambda \in [0, 1]$ , define  $\rho' = \lambda \rho + (1 - \lambda) \hat{\rho}$ . Then  $V_{t-1}^1(\rho', h_{t-1}) > V_{t-1}^1(\rho, h_{t-1})$  because of the benefit from smoothing. Starting at  $(t, h_t)$ , we have  $V_t^1(\rho', h_t) < V_t^1(\rho, h_t)$ , but because  $U^1$  is continuous the difference can be made arbitrarily small by choosing  $\lambda$  sufficiently close to 1. Also, under the provisional assumptions of

this part,  $V_t^1(\rho, h_t) > \alpha^* = V_t^1(\bar{\rho}, h_t)$ . Therefore  $V_t^1(\rho', h_t) \geq V_t^1(\bar{\rho}, h_t)$ . For all other dates and histories,  $\rho'$  and  $\rho$  give equal values. Therefore  $\rho'$  is feasible.

Then by the strict concavity of  $U^1$  and  $U^2$ , we get  $V_0^j(\rho', x_0) > V_0^j(\rho, x_0)$  for  $j = 1, 2$ . This contradicts the efficiency of  $\rho$ .

This completes the verification of all cases, and completes the proof of the “only if” part.

To show the “if” part, namely that every division process of this form is efficient, consider any process  $\rho = \{y_0, Y_1, Y_2, \dots\}$  satisfying the conditions (i) and (ii) in the statement of the theorem. For ease of notation suppose  $x_0 \in S^1$ ; the argument for  $x_0 \in S^2$  is analogous. Let  $\alpha = V^1(\rho, x_0)$  and suppose  $\hat{\rho} = \{\hat{y}_0, \hat{Y}_1, \hat{Y}_2, \dots\}$  solves the problem  $\mathcal{L}(\alpha)$ . By the “only if” part of the theorem proved above,  $\hat{\rho}$  satisfies the conditions (i) and (ii).

Both  $\rho$  and  $\hat{\rho}$  satisfy the same stochastic difference equation (ii), and its dynamics are such that along any path, the process with the higher initial share of party 1,  $y_0$  or  $\hat{y}_0$ , will also have a uniformly higher share,  $y_t$  or  $\hat{y}_t$ , up to some time  $t$  and the two will coincide thereafter. Therefore

$$V^1(\hat{\rho}, x_0) \geq V^1(\rho, x_0) \text{ if and only if } \hat{y}_0 \geq y_0$$

and

$$V^2(\hat{\rho}, x_0) \geq V^2(\rho, x_0) \text{ if and only if } \hat{y}_0 \leq y_0.$$

In the present context, the constraint in  $\mathcal{L}(\alpha)$  gives us  $V^1(\hat{\rho}, x_0) \geq V^1(\rho, x_0)$ , while the fact that  $\hat{\rho}$  is efficient gives us  $V^2(\hat{\rho}, x_0) \geq V^2(\rho, x_0)$ . The two together imply  $\hat{y}_0 = y_0$ . Then  $\hat{\rho} = \rho$ , establishing the efficiency of  $\rho$ . ■

## Proof of Proposition 1:

Let  $P(k) = \{\ell^1, \ell^2, \dots, \ell^m\}$ , where  $m$  is finite (because  $S^1$  is). By reordering the state labels if necessary, let  $\ell^1, \ell^2, \dots, \ell^j$  be the ones with the (equal) lowest value of  $y^*$  in  $P(k)$ , that is,

$$\{\ell^1, \ell^2, \dots, \ell^j\} = \arg \min\{y^*(\ell^i) \mid \ell^i \in P(k)\}.$$

Of course the minimizer may be unique, in which case  $j = 1$ ; this case is covered by the more general argument below.

We prove that  $y^*(\ell^1) \geq y^*(k)$ ; this immediately implies the desired result that  $y^*(\ell) \geq y^*(k)$  for all  $\ell \in P(k)$ . The proof is by contradiction; therefore we begin by supposing that  $y^*(\ell^1) < y^*(k)$ .

Consider the stochastic process that begins at  $\ell^1$ , and start the division process with the initial  $y$  equal to  $y^*(\ell^1)$ . The next step can only be to a state in  $\{k\} \cup P(k)$ , because if any other step could occur with positive probability, then there would be a positive probability route to  $S^2$  that avoids  $k$ . The possible steps can be divided into three kinds: (i) Those to some  $\ell^i$  for  $i = 1 \dots j$ . In these states,

$y^*(\ell^i) = y^*(\ell^1)$  by definition. Then at the next date, by Theorem 1 the initial  $y$  equals  $y^*(\ell^i)$  so the incentive constraint at  $\ell^i$  is binding. (ii) Those to some  $\ell^i$  for  $i = j + 1, \dots, m$ , where  $y^*(\ell^i) > y^*(\ell^1)$  by definition. Therefore at the next date, by Theorem 1 the initial  $y$  equals  $y^*(\ell^i)$  and again the incentive constraint at  $\ell^i$  is binding. (iii) The step to  $k$ , where  $y^*(k) > y^*(\ell^1)$  by our contradiction assumption. So yet again by Theorem 1 the starting  $y$  at the next date equals  $y^*(k)$  and the incentive constraint is binding.

Now consider the recursion relation for the efficient division process:

$$V(\ell^1) = U^1[y^*(\ell^1)] + \delta \left[ \sum_{i=1}^j p_{1i} V(\ell^i) + \sum_{i=j+1}^m p_{1i} V(\ell^i) + p_{1k} V(k) \right].$$

Along the punishment path that would ensue if party 1 were to deviate at date 0,

$$\bar{V}(\ell^1) = U^1(1) + \delta \left[ \sum_{i=1}^j p_{1i} \bar{V}(\ell^i) + \sum_{i=j+1}^m p_{1i} \bar{V}(\ell^i) + p_{1k} \bar{V}(k) \right].$$

Because the incentive constraints bind at date 1 for all the relevant states,  $V(\ell^i) = \bar{V}(\ell^i)$  for  $i = 1, 2, \dots, m$ , and  $V(k) = \bar{V}(k)$ . Now comparing the two recursion equations above immediately gives  $y^*(\ell^1) = 1$ . But this contradicts our assumption that  $y^*(\ell^1) < y^*(k)$ . ■

## Proof of Theorem 2:

Let  $\mathcal{F}^s$  be the set of feasible division processes, that is,  $\rho = (y_0, Y_1, Y_2, \dots)$  such that, for all  $t$ , and for all  $h_t = (x_0, x_1, \dots, x_t) \in H_t$ , (a) if  $x_t \in S^i$  we have  $V_t^i(\rho, h_t) \geq V_t^i(\bar{\rho}^s, h_t)$  and (b) if  $x_t \in S^0$  and  $y(h_t) \neq y(h_{t-1})$ , then we have both  $V_t^1(\rho, h_t) \geq V_t^1(\bar{\rho}^s, h_t)$  and  $V_t^2(\rho, h_t) \geq V_t^2(\bar{\rho}^s, h_t)$ .

Let  $\mathcal{F}^c$  be the set of division processes that make no change in the shares when the current state is not a supermajority state, that is,  $\rho = (y_0, Y_1, Y_2, \dots)$  such that for all  $t \geq 1$  and  $x_t \in S^0$ , we have  $Y_t = Y_{t-1}$ . Let  $\mathcal{F}^{sc} = \mathcal{F}^s \cap \mathcal{F}^c$ .

For  $\alpha \in [0, 1/(1 - \delta)]$ ,  $i, j \in \{1, 2\}$ ,  $j \neq i$ , and  $x_0 \in S^i$ , define

$$\mathcal{L}^s(\alpha) = \max_{\rho \in \mathcal{F}^s} V_0^j(\rho, x_0) \quad \text{subject to} \quad V_0^i(\rho, x_0) \geq \alpha.$$

First we prove the following:

**Lemma 1<sup>s</sup>:** For any  $x_0 \in S^i$  ( $i = 1, 2$ ) and  $\alpha \in [0, 1/(1 - \delta)]$ , either  $\mathcal{L}^s(\alpha)$  has a unique solution, or  $V_0^i(\rho, x_0) < \alpha$  for all  $\rho \in \mathcal{F}^s$ .

**Proof of Lemma 1<sup>s</sup>:** Obviously if  $V_0^i(\rho, x_0) < \alpha$  for all  $\rho \in \mathcal{F}^s$  then the maximization problem has no solution, so consider the opposite case and define

$$M(\alpha) = \sup_{\rho \in \mathcal{F}^s} V_0^j(\rho, x_0) \quad \text{subject to} \quad V_0^i(\rho, x_0) \geq \alpha.$$



First we show that  $\mathcal{F}^s$  is compact. Note that the set of all division processes is compact since  $y(h_t) \in [0, 1]$  for all  $h_t \in H_t$ . Therefore it is enough to show that the subset  $\mathcal{F}^s$  is closed. Assume that a sequence of processes  $\rho^{(n)} = \{y_0^{(n)}, Y_1^{(n)}, Y_2^{(n)}, \dots\} \in \mathcal{F}^s$  converges to some  $\rho = \{y_0, Y_1, Y_2, \dots\}$ . We will show that  $\rho \in \mathcal{F}^s$ . To do so we need to show that  $\rho$  satisfies the weak inequalities stated above in the definition of  $\mathcal{F}^s$ . This follows immediately from the continuity of  $V_t^i$  and the corresponding inequalities for  $\rho^{(n)}$ . Note that we may have  $y^{(n)}(h_t) \neq y^{(n)}(h_{t-1})$  for all  $n$  but  $y(h_t) = y(h_{t-1})$  in the limit. But in this situation  $\rho$  does not have to meet any incentive constraints and therefore belongs to  $\mathcal{F}^s$  anyway.

However,  $\mathcal{F}^s$  is not in general convex because of the form of the incentive constraints in states  $x \in S^0$ , and we need convexity in the argument below. We prove that any solution to  $\mathcal{L}^s(\alpha)$  maintains  $y(h_t) = y(h_{t-1})$  in these states. That is to say, we show that if  $\rho$  solves  $\mathcal{L}^s(\alpha)$ , then  $\rho \in \mathcal{F}^{sc}$ . It is easy to verify that  $\mathcal{F}^{sc}$  is convex, which suffices for our argument.

Proving that  $\rho \in \mathcal{F}^{sc}$  is straightforward: if  $y(h_t) \neq y(h_{t-1})$  for some  $h_t = (x_0, x_1, \dots, x_t) \in H_t$  such that  $x_t \in S^0$ , then replacing both  $y(h_{t-1})$  and  $y(h_t)$  by

$$y = \frac{1}{1 + \delta p} y(h_{t-1}) + \frac{\delta p}{1 + \delta p} y(h_t)$$

where  $p = \Pr[X_t = x_t | X_{t-1} = x_{t-1}]$  strictly improves both payoffs at time  $t - 1$  given history  $h_{t-1}$  and removes incentive constraints associated with history  $h_t$  at time  $t$ . Thus the new process is also feasible and has increased payoffs, which implies that the original  $\rho$  could not have been optimal. Thus  $\rho$  solves

$$\max_{\rho \in \mathcal{F}^{sc}} V_0^j(\rho, x_0) \quad \text{subject to} \quad V_0^i(\rho, x_0) \geq \alpha.$$

Then the uniqueness of the solution to  $\mathcal{L}^s(\alpha)$  follows from the strict concavity of the  $U^i$  and the convexity of  $\mathcal{F}^{sc}$ .

The remainder of the proof of Theorem 2 follows from the corresponding arguments used in the proof of Theorem 1; we need only replace  $\mathcal{F}$  with  $\mathcal{F}^{sc}$ .  $\square$

## Proof of Proposition 2:

Let  $\hat{\rho}^s(1)$  be the solution to the problem of maximizing  $V_0^2(\rho, x_0 = 1)$  subject to  $\rho \in \mathcal{F}^s$ . The existence and uniqueness of this solution was established by Theorem 2. By the same theorem,  $\hat{\rho}^s(1)$  has the following form:

$$\hat{Y}_t^s = \begin{cases} y^* & \text{if } X_t = 1 \\ \hat{Y}_{t-1}^s & \text{if } X_t \in \{2, 3\} \\ 1 - y^* & \text{if } X_t = 4. \end{cases}$$

Define  $\tilde{\rho}(1)$  as follows:

$$\tilde{Y}_t^s = \begin{cases} y^* & \text{if } X_t \in \{1, 2\} \\ 1 - y^* & \text{if } X_t \in \{3, 4\}. \end{cases}$$

As noted in Section 5,

$$V_0^1(\hat{\rho}, 1) + V^2(\hat{\rho}, 1) = [U(y^*) + U(1 - y^*)]/(1 - \delta) = V_0^1(\tilde{\rho}, 1) + V_0^2(\tilde{\rho}, 1).$$

Thus we will complete the proof that  $\hat{W} \geq \hat{W}^s$  by showing that  $\tilde{\rho} \in \mathcal{F}^s$ .

First we will establish that  $V_0^1(\tilde{\rho}, 1) \geq V_0^1(\bar{\rho}^s, 1)$  and hence party 1 has no incentive to deviate at time 0. This will also establish that party 1 (and by symmetry party 2) has no incentive to deviate after any history in which it has the current supermajority.

Since  $\hat{\rho} \in \mathcal{F}^s$ , we have

$$\begin{aligned} V_0^1(\tilde{\rho}, 1) - V_0^1(\bar{\rho}, 1) &\geq V_0^1(\tilde{\rho}, 1) - V_0^1(\bar{\rho}, 1) - V_0^1(\hat{\rho}, 1) + V_0^1(\bar{\rho}^s, 1) \\ &= \sum_{t=0}^{\infty} \mathbb{E}[U(\tilde{Y}_t) - U(\bar{Y}_t) - U(\hat{Y}_t^s) + U(\bar{Y}_t^s)] \delta^t \end{aligned}$$

Let  $Z_t = U(\tilde{Y}_t) - U(\bar{Y}_t) - U(\hat{Y}_t^s) + U(\bar{Y}_t^s)$  and

$$R_1 = \sum_{t=0}^{\infty} \mathbb{E}[Z_t] \delta^t.$$

Observe that if either party has a supermajority, or the party that has a weak majority is also the last party to have a supermajority, then  $Z_t = 0$ . On the other hand, if party 2 has a weak majority and party 1 was the last to have a supermajority, then

$$Z_t = z \equiv U(1 - y^*) - U(y^*) + 1.$$

Similarly, if party 1 has a weak majority and party 2 was the last to have a supermajority, then  $Z_t = -z$ .

Let

$$R_4 = \mathbb{E} \left[ \sum_{l=t}^{\infty} z_l \delta^{l-t} \mid X_t = 4 \right]$$

By symmetry,  $R_4 = -R_1$ . Define two stochastic processes  $\tau, \tau'$  as follows:  $\tau_t = 1$  if and only if  $X_l < 4$  for all  $l \leq t$ , otherwise  $\tau_t = 0$ ; while  $\tau'_t = 1$  if and only if  $t$  is the first time state 4 is reached, otherwise  $\tau'_t = 0$ . Then we can rewrite  $R_1$  as follows:

$$\begin{aligned} R_1 &= \sum_{t=0}^{\infty} \mathbb{E}[Z_t \tau_t] \delta^t + R_4 \sum_{t=0}^{\infty} \mathbb{E}[\tau'_t] \delta^t \\ &= \sum_{t=0}^{\infty} \mathbb{E}[Z_t \tau_t] \delta^t - R_1 \sum_{t=0}^{\infty} \mathbb{E}[\tau'_t] \delta^t \end{aligned}$$

Hence

$$R_1 = \frac{\sum_{t=0}^{\infty} \mathbb{E}[Z_t \tau_t] \delta^t}{1 + \sum_{t=0}^{\infty} \mathbb{E}[\tau'_t] \delta^t}$$

Observe that  $\tau'_t \geq 0$  for all  $t \geq 0$ , and either  $Z_t \geq 0$  or  $\tau_t = 0$ . Hence  $R_1 \geq 0$  and therefore  $V_0^1(\tilde{\rho}, 1) \geq V_0^1(\bar{\rho}, 1)$  as desired.

To conclude the proof, we show that party 1 has no incentive to deviate in state 2 and hence by symmetry party 2 has no incentive to deviate in state 3. Let

$$\tilde{Z}_t = \begin{cases} U(y^*) - 1 & \text{if } X_t \in \{1, 2\} \\ U(1 - y^*) & \text{if } X_t \in \{3, 4\} \end{cases}$$

Then  $V_0^1(\tilde{\rho}, 1) - V_0^1(\bar{\rho}, 1) =$

$$Z_0 + \delta p_{11} [V_0^1(\tilde{\rho}, 1) - V_0^1(\bar{\rho}, 1)] + \delta p_{12} \mathbf{E} \left[ \sum_{t=1}^{\infty} Z_t \delta^t \mid X_t = 2 \right]$$

or

$$\delta p_{12} \mathbf{E} \left[ \sum_{t=1}^{\infty} Z_t \delta^t \mid X_t = 2 \right] = -Z_0 + (1 - \delta p_{11}) [V_0^1(\tilde{\rho}, 1) - V_0^1(\bar{\rho}, 1)]$$

Since  $Z_0 \leq 0$  and by the argument above  $V_0^1(\tilde{\rho}, 1) - V_0^1(\bar{\rho}, 1) \geq 0$ , it follows that

$$\mathbf{E} \left[ \sum_{t=1}^{\infty} Z_t \delta^t \mid X_t = 2 \right] \geq 0 ,$$

which is precisely the condition needed to deter party 2 from deviating in state 2. ■

Figure 1

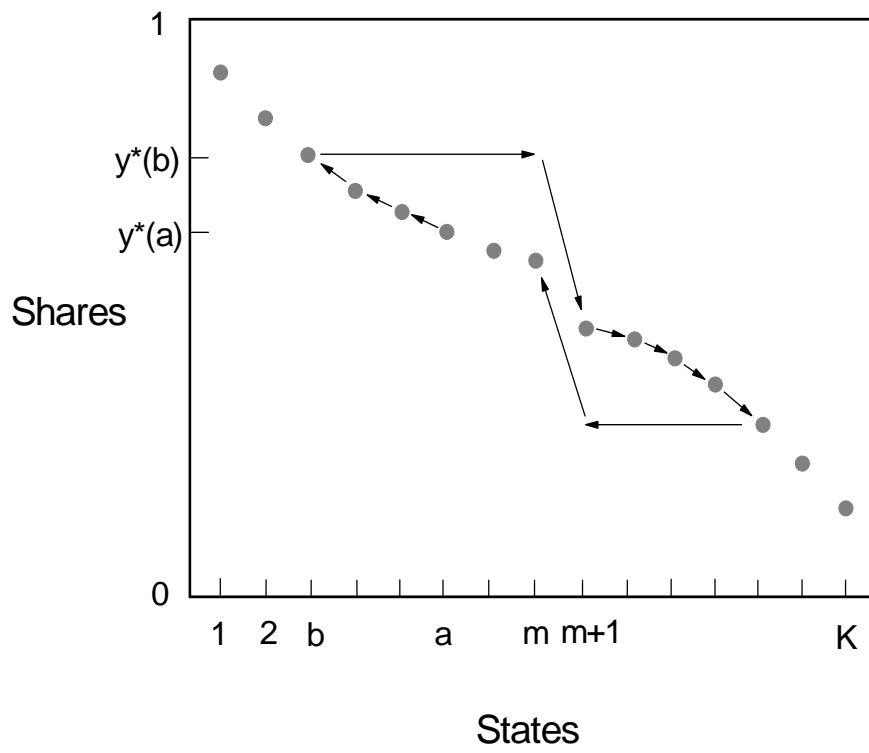


Figure 1:

Figure 2

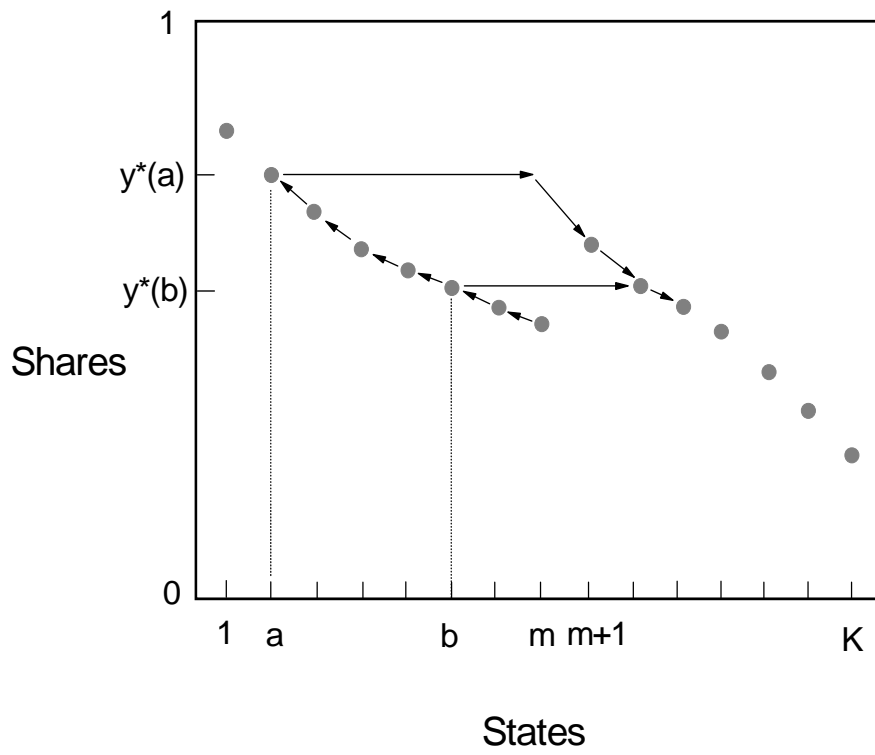


Figure 2:

Figure 3

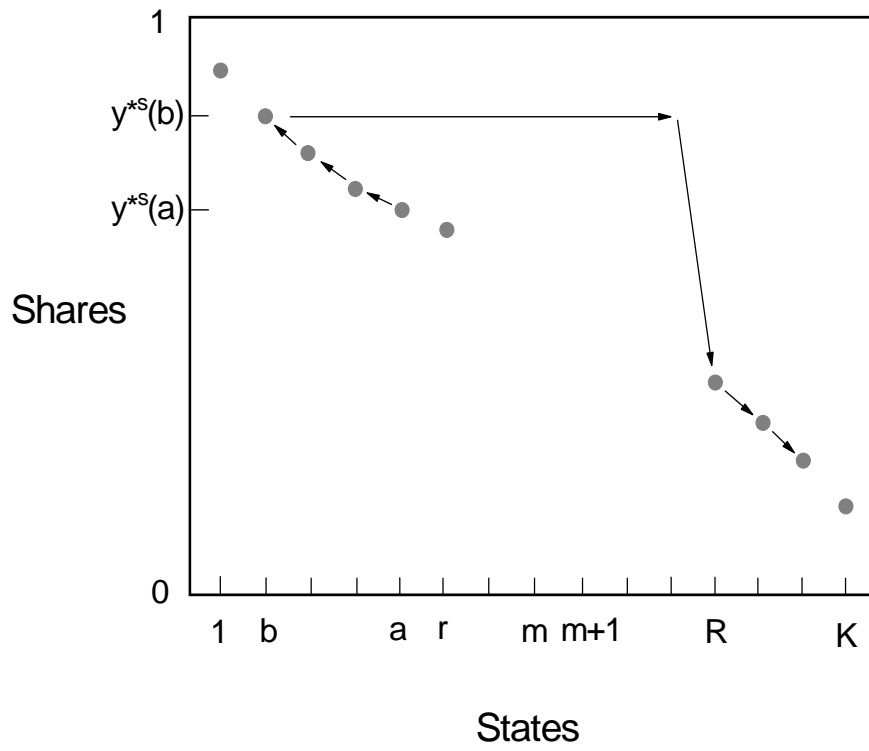


Figure 3: