Choosing How to Choose: Self-Stable Majority Rules

Salvador Barbera and Matthew O. Jackson *

August 2000
revised: July 10, 2002

Abstract

We consider the endogenous choice of a voting rule, characterized by the majority size needed to elect change over the status quo, by a society who will use the rule to make future decisions. Under simple assumptions on the uncertainty concerning the future alternatives that will be voted upon, voters' have induced preferences over voting rules that are single-peaked and intermediate. We explore the existence of self-stable voting rules, i.e., voting rules such that there is no alternative rule that would beat the given voting rule if the given voting rule is used to choose between the rules. There are situations where self-stable voting rules do not exist. We explore conditions that guarantee existence, as well as issues relating to efficiency and constitutional design.

*Barbera is at CODE, Departament d'Economia i d'Historia Economica, Universitat Autonoma de Barcelona, 08193 Bellaterra, Spain, (Salvador.Barbera@uab.es) and Jackson is at the division of Humanities and Social Sciences, 228-77, California Institute of Technology, Pasadena, California 91125, USA, (jacksonm@hss.caltech.edu). We thank Randy Calvert, Dan Levin, and Jean Mercier Ythier for very helpful discussions of the paper. We also thank Danilo Coelho, Gabrielle Demange, Anke Gerber, Annik Laruelle, Eric Maskin, Charlie Plott, Cheng-Zhong Qin, Jim Snyder, Guofu Tan, Federico Valenciano, and Peyton Young for helpful conversations and suggestions, and Takehiko Yamato for detailed comments on an earlier draft. Financial support under NSF grant SES-9986190 and under DCGYT Direcció General de Recerca projects PD-98-0870 and SGR-980062 is gratefully acknowledged. This project was initiated during a visit of Barbera to Caltech.

1 Introduction

Different societies use different voting rules to make collective decisions. How do they come to choose the voting rule that will guide their subsequent decision making? Should the choice of a choice rule be treated as a special type of decision, or is it just one more instance of the many issues that a society has to face? If the choice of voting rule deserves a special treatment, what are the grounds for it, and what type of considerations are to be added to those already incorporated into the rules for day-to-day decisions? These are the questions we address in this paper. Let us expand on their relevance, outline our approach and some of our results.

The voting rule that a society uses to make collective decisions has consequences. For instance, a society that uses simple majority might reach a different decision than if it required two thirds of the voters to approve change over status quo for a proposed change to be accepted. The extensive social choice literature has analyzed the properties of many voting rules, and it is well understood that the rules to be used make a difference as soon as society members have conflicting interests.

Taking for granted that voting rules matter, the question arises as to who chooses the rules of choice, and on what grounds. We consider here that the members of society vote over voting rules. This is the case in many situations where democratic institutions are in place: they include procedures for constitutional changes.[1] There are many important questions that come to mind and we focus on several of them:

- When do self-stable voting rules exist - that is, are there rules such that if such a rule is used, then no other rule would defeat it?
- What sorts of constitutions (specifying one voting rule for standard decisions and a different voting rule for voting over constitutions) are self-stable?
- When is simple majority rule (which is attractive for its efficiency) self-stable and when can it be included in a self-stable constitution?

To give more perspective to these questions, let us describe the model in which we answer them. To explore voting over voting rules, we need to study a context in which preferences over voting rules and votes over voting rules can be naturally considered. We do this by taking the following perspective. Two alternatives are to be voted upon in in the future, and

[1]Even when citizens are not actually allowed to change the rules, understanding whether or not they would like to change them, and in what direction, is a relevant standard of evaluation.

perhaps a number of such choices will be faced. In each instance, one of these alternatives will be the status quo and the other will be some change. A voting rule is simply characterized by the minimal number of votes required in favor of change in order to select change. Each voter is characterized by a probability that they will support the status quo. To fix ideas, think of a legislature just beginning a term. Legislators know that they will face votes on bills which would introduce changes to current systems. Some legislators might be very satisfied with the current system and thus expect to favor a relatively low percentage of the bills that will be voted upon. These legislators would have a relatively high probability of supporting the status quo and a low probability of supporting change. They would end up preferring voting rules that make change difficult.[2] Other legislators might be very dissatisfied with the current system and thus expect to favor a high percentage of the bills that will be faced. These legislators would have a relatively low probability of supporting the status quo and a high probability of supporting change. They would end up preferring voting rules that make change easy.[3] Voters' beliefs over their relative likelihood of supporting change versus the status quo induces preferences over voting rules. We examine these indirect preferences over voting rules in some detail, and it turns out that this very simple model of uncertainty ends up producing a remarkable structure to preferences over voting rules, as discussed in the sections that follow.

While the discussion of the uncertain preferences above is a bit abstract, it is clear that there are many decisions which fit exactly into this framework. For instance, examples include voting over whether to admit new members to a club or organization (or to grant tenure), or a vote by a parliament or legislature on whether or not to go to war, a vote by shareholders of a company on whether to approve a plan proposed by their board, etc.. Nevertheless, an important consideration that this model abstracts away is agenda formation. As agendas can be endogenous, agenda formation may interact with the voting rule in interesting ways. In our analysis here, we ignore where the agenda comes from with the understanding that analyzing voting over voting rules with exogenous agendas is a useful

[2]The term "relatively" is used in this discussion for a reason. As we will see below, the preferences over voting rules will depend on the relative comparisons of various voters' probabilities of supporting change, and cannot simply be determined by the absolute probabilities. For instance, even if a voter has a low probability of supporting change; if that probability is much higher than any other voter's probability, then that voter will be the most likely to be among the supporters of a change whenever change is supported. This would mean that the voter should favor voting rules that make change easy, even though their absolute probability of supporting change is low.

[3]For much more discussion of and motivation for this model of a choice over two alternatives that are currently uncertain in voters' minds, see the introductory chapter in Niemi and Weisberg (1972), as well as the chapters by Badger (1972), Curtis (1972), and Schofield (1972).

first step, there are important situations where agendas are naturally viewed as exogenous to the voters, and expanding the model to account for endogenous agendas provides for interesting further research.

Knowing how the members in our stylized society have induced preferences over different voting rules, we can ask which rules are likely to be chosen, or more specifically, which rules are likely to persist. Given some voting rule in place, by imposition or by a previous act of choice, would this rule prevail against others, when the rule itself is used to determine whether to keep it as the status quo, or change it for some of its potential contenders? Since we concentrate on the choice among majoritarian rules, our problem is that of choosing the majority size required to change the status quo. Then, our previous question can be rephrased as follows. Given a majority size s, will it be the case that no alternative majority size s' is preferred over s by s or more voters? If s is not so defeated by any other rule, we say that the rule s is *self-stable*.

We argue that self-stability is a desirable property of voting rules. If a society uses a rule that is not self-stable, then there will be a large enough group of voters who all prefer some alternative voting rule to be successful in changing the rule. Hence, societies which value the type of stability embodied in the continued use of the same rules of the game, should settle for voting rules which are self-stable, when available. Also, one may argue that self-stable voting rules are the ones that should emerge over time, as rules that are not self-stable will be transient while voting rules that are self-stable will be absorbing states.

Since self-stability is a property that depends on the preferences of voters regarding alternative voting methods, different rules will be self-stable depending on the type of society. We thus provide a careful study of the variables that determine which rules are self-stable. As we shall see, for some societies no voting rule is self-stable. Moreover, our analysis makes it apparent that this lack of self-stability is a robust phenomenon, not an occasional anomaly. However, as we show, the non-existence problem occurs in less than $1/2$ of all potential societies, in a well-defined sense.

If some societies do not have the possibility of choosing any voting rule that is self-stable, does this mean that they are doomed to an endless and hopeless drift from an unstable rule to another unstable one? Certainly not. If this was the end of our story, we might well have discovered an imaginary difficulty, since we do not often observe a continual change of rules within societies which resort to voting as a means to solve conflict. But notice that, in actual constitutions, special rules are often set in order to change the rules of choice. Our discovery that societies may face situations where no single rule would be self-stable suggests an explanation for this fact. If the choice of only one rule for decisions of any type may be a source of instability, then it is natural to set at least two rules: one to be used when making standard choices, and a different one to use when it comes to change the preceding rule.[4]

Consider, then, a stylized constitution, given by two rules (s, S). The first rule, s, is used for all decisions except decisions over voting rules. The second rule, S, is used to make any decisions over the constitution, and in particular is used to make any change of s. For example, in many real situations s is simple majority (one half of the voters) and S is two thirds, so a super-majority is needed to change the rules. We say that a constitution (s, S) is self-stable if no voting rule s' would gather the support of S or more voters when compared to s. Hence, a constitution is self-stable if its rule s to make standard decisions cannot be challenged effectively under the prescribed procedure for constitutional changes. There are always self-stable constitutions for any society, as we shall show.

We explore also explore which super-majorities S are such that the (simple majority, S) constitution is self-stable. This is an attractive type of constitution because of the efficiency of simple majority rule.

Our research takes the view that 'choosing how to choose' is an issue that calls for the treatment of institutions as endogenous variables, and not as exogenously given data. Thus, we see it as part of a broad and ambitious research program of not only understanding normative or positive properties of institutions and mechanisms, but also how they come to take certain forms when individuals in the society have personal stakes in the design of the institution and can affect it. To some extent this presents a 'chicken and egg' dilemma, as the existing institutional environment to a large extent determines what institutional changes can take place, and are also the result of previous institutional change. Economics has a tradition of dealing with problems of this kind by resorting to appropriate fixed-point notions, and self-stability can viewed in this light.

The Related Literature

Since we concentrate on the forces that drive the choice among different majority sizes, we are naturally led to identify those instances where simple majority would arise as the choice of society. There are many angles from which one can provide theoretical reasons to support and explain the widespread use of simple majority, and its appeal as a reference point even when it is not actually used (e.g. see Schofield (1972)). One is axiomatic: simple

[4]We do not claim that the lack of self-stability is the only relevant reason to establish special rules for constitutional change. Other classical explanations may concur. For example, if existing constitutions are viewed as the result of some accumulated wisdom, requiring a large majority to change it may also be interpreted as a precaution against a short sighted rejection of past experience. What we claim is that self-stability is a relevant consideration as well, and one that has not been put forth before.

majority is the unique rule to satisfy some combinations of attractive desiderata, as shown in May's Theorem (1952). Another classical justification is referred to as the Condorcet jury theorem: the simple majority winner is the maximum likelihood estimator of the correct decision, in societies where all voters have the common aim of choosing the alternative that is correct, but differ in their perceptions due to noisy information (e.g., see Condorcet (1785), or Young and Levenglick (1978)). Our approach differs from the axiomatic, because we treat the decision rules as choice variables, and from the jury theorem approach because our voters may have conflicting objectives. In fact, the model we work with was first proposed in the early seventies in a series of brilliant papers (most of which are collected in a volume edited by Niemi and Weisberg (1972)), inspired by a seminal work of Rae (1969), whose purpose was to justify the use of simple majority from a novel viewpoint. Our analysis reinforces the conclusions of this early literature, by pointing at new arguments and situations under which simple majority would be the only self-stable rule. Moreover, the special nature of simple majority is also reinforced by the fact that it is the only rule for day-to day decisions which constitutes, when coupled with unanimity, a self-stable constitution valid for all societies. This and other side results make it clear that simple majority is special in many ways. Yet, our analysis also clarifies that alternative majority sizes may have their own advantages, and in particular may be self-stable for societies where simple majority would not satisfy this desirable property.

Constitutional design and properties of voting rules are topics that have been extensively studied in political science and social choice theory. Such studies go back to the classics, such as Rousseau (1762) who explicitly discussed how the size of a majority required in a voting rule should be related to the importance of the question at hand.[5] Buchanan and Tullock (1962) were the first to raise the issue of choosing how to choose, and they raise it explicitly in the context of constitutional design. To quote from their work (page 6):

"When we recognize that "constitutional" decisions themselves, which are necessarily collective, may also be reached under any of several decision-making rules, the same issue is confronted all over again. Moreover, in postulating a decision making rule for constitutional choices, we face the same problem when we ask: How is the rule itself chosen?"

While Buchanan and Tullock raise the issue of choosing how to choose, they end up stepping around it and instead focussing on the role of full consent in decision making, including decisions regarding constitutional choice. This holds with their central thesis that

[5]Some recent references from the large literature that relates to issues regarding majority size includes Caplin and Nalebuff (1988), Austen-Smith and Banks (1997), Feddersen and Pesendorfer (1998), and Dasgupta and Maskin (1998).

full consent is the only manner in which proper improvements can be reached due to the external costs that will be forced on individuals through other decision making rules such as simple majority rule. In contrast, the approach that we take here of identifying self-stable voting rules and constitutions, addresses this problem of choosing voting rules and constitutions head-on. Our central purpose is to analyze such questions of choosing how to choose and to provide a framework where a fuller understanding of such choices can be developed.

While such exploration of voting over voting rules is relatively new, there is a predecessor to our paper. The first work that is concerned with carefully modeling how societies choose how they will choose is Koray (2000). Koray outlines a method for viewing social choice functions themselves as alternatives, so that one can ask whether a social choice function always selects itself. He shows that given enough richness of preferences the only self-selective social choice functions are dictatorial. While our self-stability notion is similar in spirit to self-selection, there are differences in the concepts and settings so that such impossibility results are not an issue in our analysis. First, our concept of self-stability only requires that a voting rule should not be beaten by another rule when the given rule is used, which is different from saying that a rule must select itself. Another way to say this is that in our setting there is a special standing to the status-quo alternative, which can provide an asymmetry not present in the more abstract social choice setting analyzed by Koray. Second, the underlying setting here considers votes over two (possibly uncertain) alternatives at a time, rather than making selections from three or more (known) alternatives. These differences reflect very different idealized applications, and end up providing dramatic differences in the model and the results, so that the only real tie between our study and Koray's is in the common interest of endogenizing the way in which societies make choices.

There are previous studies of votes to be taken over pairs of alternatives that are still unknown, but over which voters have beliefs about which they will prefer. In particular, Rae (1969) conjectured that in such settings simple majority rule was the one that would maximize total societal welfare, and argued this through some examples. The properties of simple majority rule were more extensively modeled and verified by Badger (1972) and Curtis (1972). In doing this, Badger and Curtis studied properties of preferences over voting rules that are useful in our study here and are noted below as Theorem 1. While the Badger and Curtis papers provide some useful tools and a nice point of departure for our analysis, they are not concerned with self-stability of voting rules.

2 Definitions

Voters and Alternatives

$N = \{1, \ldots, n\}$ is a set of *voters*.

The voters will face votes over pairs *alternatives*. We denote the terms of these pairwise choices as a and b. Alternative a is interpreted as the *status-quo*. Alternative b is interpreted as a *change*.

Voting Rules

Each voter casts a vote in $\{a, b\}$.

A *voting rule* is characterized by a number $s \in \{1, \ldots, n\}$.[6] If at least s voters say "b" then b is elected, and a is elected otherwise.

Some examples of voting rules are as follows.

If $s = 1$, then b is elected whenever there is at least one voter for change, and so a is elected only when it is unanimously supported.

If $s = n$, then b is elected if there is unanimous support for change, and a is elected as soon as at least one a voter supports it.

If n is odd and $s = \frac{n+1}{2}$ or n is even and $s = \frac{n}{2} + 1$, then the voting rule is the standard or simple *majority rule*.[7]

As simple majority rule is referred to at several points in what follows we denote it by s^{maj}. Thus, $s^{\text{maj}} = \frac{n+1}{2}$ if n is odd and $s^{\text{maj}} = \frac{n}{2} + 1$ if n is even.

Note that our definition of a voting rule presumes anonymity. We discuss this property in the concluding remarks.

Voter Preferences

At the time a voting rule s is chosen, voters have expectations over the future issues that will be voted on, but do not know the exact realization. We consider a stark model where voters are simply characterized by a parameter $p^i \in (0, 1)$. This represents the probability

[6] Allowing for $s = 0$ or $s = n + 1$ results in degenerate voting rules that always choose b or always choose a, respectively. We focus on rules where there is a real choice to be made.

[7] When n is even, there are two possible choices: $\frac{n}{2}$ and $\frac{n}{2} + 1$ depending on which alternative wins in the case of a tie. For simplicity, we break ties in favor of the status quo in this case. None of the analysis that follows is dependent on tie-breaking conventions.

that they will prefer b to a at the time of the vote.[8] Such an uncertain setting was first considered in Badger (1972) and Curtis (1972), and we will make use of some of their results in what follows.

The realizations of voters' support for the alternatives are independent. For instance, the probability that voters 1 and 2 support b while voter 3 supports a is $p_1 p_2(1 - p_3)$.

A voter gets utility 1 if his preferred alternative is chosen in the vote, and utility 0 otherwise.[9]

The *society* of voters is represented by a set of voters N and a vector $p = (p_1, \ldots, p_n)$.

In what follows, we treat the society (N, p) as given and so will often suppress the fact that preferences will depend on these parameters, except where we want to specifically point out this dependence.

As we shall see, the simple identification of a voter with the probability that he or she will support change rather than the status quo is sufficient to induce significant structure on voters' preferences over voting rules. The probability attached to each voter may be justified in several ways. For example, suppose that each voter has well defined preferences over alternative decisions. If the voter knows the existing social situation at a point in time and also has an idea of the kind of changes that may be proposed, then the voter may have an estimate of how often he or she will be facing a desired change, and how often it would rather see the status quo being maintained.[10]

Induced Preferences over Voting Rules

The simple structure described above induces preferences over voting rules for every voter. That is, given that voters have beliefs about the likelihood of different patterns of support for a and b, a voter can calculate his or her expected utility under each voting rule s.

---

[8] Extensions to the case where p_i can be 0 or 1 are straightforward. These cases complicate some of the calculations and proofs when we divide by p_i or $1 - p_i$, but are still easily directly handled as special cases. To keep an uncluttered exposition, we leave the cases where $p_i = 0$ or 1 to the interested reader.

[9] This presumes that a voter cares as much for getting change when preferring change over the status quo, as the voter cares for preserving the status-quo when preferring the status quo over change. We discuss the role of this assumption in detail in the concluding remarks.

[10] Governments, of course, change and the status quo will evolve over time and hence so will voters' preferences. In this stark model, the probability to favor change that characterizes our stylized voters summarizes a lot of information about preferences, expectations about the likely proposals they will face, expectations about the probability that governments (committees, legislatures, etc.) change ,and expectations regarding how the status quo may evolve. This summary information will also reflect the time horizon for which all these expectations are taken to be relevant. We understand that richer conclusions may be reached by more explicit models, but we emphasize here that the simplicity of this summary representation provides substantial insight into the issue of stability of voting rules.

Let $U_i(s)$ be the expected utility of voter i if voting rule s is used. This is expressed as follows. For any $k \in \{0, \ldots, n-1\}$, let $P_i(k)$ denote the probability that exactly k of the individuals in $N \setminus \{i\}$ support the change. We can write

$$P_i(k) = \sum_{C \subset N \setminus \{i\} : |C| = k} \times_{j \in C} p_j \times_{\ell \notin C} (1 - p_\ell). \tag{1}$$

and

$$U_i(s) \;=\; p_i \sum_{k=s-1}^{n-1} P_i(k) \;+\; (1 - p_i) \sum_{k=0}^{s-1} P_i(k). \tag{2}$$

Self-Stable Voting Rules

Given that voters' have induced preferences over voting rules, we also study how voters would vote when voting over voting rules. Our particular interest is in voting rules that are robust in the sense that if they are used, then no other voting rule could be put forth as a challenger and defeat the given voting rule viewed as the status quo.

A voting rule s is *self-stable (for society p)* if $\#\{i \mid U_i(s') > U_i(s)\} < s$ for every $s' \neq s$.

The property of self-stability ensures that a given voting rule would be robust to change if used for making decisions.[11]

3 Properties of Preferences over Voting Rules

Before examining the issue of the existence of self-stable voting rules, we establish some properties of voters' preferences over voting rules that will be instrumental in understanding self-stability.

Single Peaked Preferences

The usual definition of single-peaked preferences requires that all alternatives can be ranked from left to right, that one alternative $\widehat{s}$ is best, and that the alternatives that one encounters by moving leftward (or rightward) away from $\widehat{s}$ are considered worse and worse. Our definition here will be slightly weaker, as it allows a voter to have two peaks.[12] In

[11] As with many equilibrium concepts, we do not model how one might reach a self-stable voting rule, nor do we model the choice among them if there are several. What we can say is that a self-stable rule would stay in place if reached, while other rules would tend not to.

[12] These could be referred to as single-plateaued preferences following the literature. However, given that such indifference can only occur between two points and happens non-generically (in p) we stick with the term single-peaked.

particular, it is possible that $U_i(\widehat{s}) = U_i(\widehat{s} - 1)$. For instance, in a society where n is even and each $p_i = p$ for all i, all individuals will be indifferent between $n/2$ and $n/2 + 1$.

U_i is *single-peaked* if there exists $\widehat{s} \in \{1, \ldots, n\}$ with $U_i(\widehat{s}) \geq U_i(s)$ for all $s \in \{1, \ldots, n\}$ such that $U_i(s) > U_i(s-1)$ for any $\widehat{s} > s > 1$ and $U_i(s-1) > U_i(s)$ for any $n \geq s > \widehat{s}$.

Let $\widehat{s}_i$ denote the *peak* of voter i.

In the case where a voter has twin-peaks, the definition above selects the higher of the two peaks as $\widehat{s}_i$. This is simply a convention and does not matter in any of the results that follow.

The following theorem is proven in Badger (1972). We include a proof in the appendix, for completeness.

THEOREM **1** [Badger (1972)] *For any society, every voter's preferences over voting rules are single-peaked.*

Intermediate Preferences and Single Crossing

While Theorem 1 tells us that each voter's preferences over voting rules have the nice property of single-peakedness, the following theorem tells us about how different voters' preferences are related to each other. There are two properties that are useful in noting.

A society of voters has preferences satisfying the *single crossing property* if for any i and j with $p_j \geq p_i$,

$$U_i(s) - U_i(s') \geq U_j(s) - U_j(s')$$

for all $s \geq s'$.

As we shall see, the single crossing property is satisfied in this model. The single crossing property allows us to order preferences over voting rules in terms of the p_i's; but more importantly also implies that the preferences are intermediate.

A society of voters has *intermediate preferences* if for any i, j, k with $p_j \geq p_k \geq p_i$:

- $U_i(s) \geq U_i(s')$ and $U_j(s) \geq U_j(s')$ imply that $U_k(s) \geq U_k(s')$, and
- $U_i(s) > U_i(s')$ and $U_j(s) > U_j(s')$ imply that $U_k(s) > U_k(s')$.

Intermediate preferences are usually defined by requiring that there exists some ordering over individuals so that when two individuals have the same ranking over two alternatives, then individuals between them in the ordering have that same ranking (e.g., see Grandmont

(1978)). Here the natural ordering over individuals is in terms of their p_i's, the distinguishing characteristic of voters, and so we take the shortcut of defining intermediate preferences directly in terms of that ordering. Hence, a society will have intermediate preferences over voting rules if whenever two voters with p_i and p_j agree on how to rank two rules s and s', then all voters with probabilities p_k between p_i and p_j will also agree on the way to rank these two rules. The simple model we are considering has the following strong feature.

THEOREM **2** *Every society has preferences over voting rules that satisfy the single crossing property and are intermediate.*

The proof of Theorem 2 appears in the appendix. The intuition for why the voters' peaks over voting rules follow an inverse order to the voters p_i's (Corollary 1 below), is fairly straightforward, as voters with higher p_i's are more likely to favor change and thus will be in favor of a lower quota than voters who are less likely to favor change. While ordering the peaks is intuitive and useful, we emphasize that Theorem 2 has much stronger implications, as it relates preferences over arbitrary values of s and s', including those falling on opposite sides of a set of voters' peaks. This additional structure will also be useful in what follows. The proof of these aspects of preferences builds inductively from preferences over adjacent voting rules, and involves direct comparison of the expressions of differences in expected utilities for different voters. Details are in the appendix.

As just mentioned above, Theorem 2 has the following useful corollary (see the proof of Theorem 2).

COROLLARY **1** *For any society, $\widehat{s}_i \geq \widehat{s}_j$ whenever $p_j \geq p_i$.*

There are some other facts about the location of the voters' peaks that are worth emphasizing. The *relative* ordering of p_i's is not only important in determining the relative ordering over the $\widehat{s}_i$'s, but it is also critical in determining the actual values of the $\widehat{s}_i$'s. This is seen in the following theorem, which states that *regardless* of p, there is always some voter who has a peak at least as high as s^{maj} and some other voter who has a peak no higher than s^{maj}.

THEOREM **3** *For any society there exist i and j such that $\widehat{s}_i \geq s^{\mathrm{maj}} \geq \widehat{s}_j$.*

The proof of Theorem 3 is based on the following reasoning. The unique maximizer of $\sum_i U_i(s)$ is s^{maj}, since s^{maj} chooses the alternative that will result in the largest group of voters who get utility 1 for each realization of preferences over a and b (see the concluding remarks). Thus, if some voter's expected utility is increased by moving to an s that is higher than s^{maj}, then some other voter's expected utility must fall as the result of such a move. The same is true in reverse. So there is at least one voter with a peak at least as high as s^{maj} and at least one voter with a peak no higher than s^{maj}. The complete proof, taking into account the possibility of twin-peaks appears in the appendix.

Note that by combining Corollary 1 with Theorem 3, we know that the voter who has the highest p_i must have a $\widehat{s}_i$ which is no higher than s^{maj} and the voter who has the lowest p_i must have a $\widehat{s}_i$ that is at least as high as s^{maj}, *and this is true regardless of* p.

4 Self-Stable Voting Rules

With some understanding of voters' preferences over voting rules under our belts, we now examine the issue of existence of self-stable voting rules in some detail.

We begin by considering the special case where all voters have the same p_i. This is of some interest where this common p is an indicator of the average propensity to favor change of a society's representative voter. It is also worth considering as an exercise, since the reasoning required for this simple case extends to the analysis of more heterogeneous societies. Moreover, the conclusion we reach may seem counterintuitive at first, although it is easy to reach after some reflection.

Theorem 4 *If $p_i = p_j$ for all i and j, then s^{maj} is the unique self-stable rule if n is odd, while both s^{maj} and $s^{\mathrm{maj}} - 1$ are self-stable if n is even.*

Thus, simple majority is the unique self-stable voting rule whenever all voters have the same probability of choosing change, irrespective of what this probability might be. One might have guessed that societies where all voters are very likely to want changes would prefer low values of s, that is low barriers to change, and that homogeneously conservative societies would favor high values of s. But this is not the case. Actually, in homogeneous societies, all voters have their peak at $\widehat{s}_i = s^{\mathrm{maj}}$, and thus simple majority is the consensus choice of rule. This is a simple corollary of Theorem 3. What actually matters is not the absolute values of the p's but their values relative to those of other voters. For instance, consider a society where $p_i = .01$ for each i and so voters are very conservative and very likely to support the status-quo. In this case, shouldn't it be that voters all prefer a high quota s as they each know they are likely to support the status quo? The answer is no and the reasoning lies in the answer to the following question. Which alternative would a voter prefer society to choose in a generic realization where k voters end up supporting a and $n-k$ voters end up supporting b? That is, the voter can think of the different scenarios possible for numbers of voters supporting a and b, and then ask which side he is most likely to fall on in each scenario. Given the symmetry in p_i's, conditional on this realization of preferences it is most likely that the voter is in the larger of the two groups. So, the voter would like society to choose a in scenarios where $k > n-k$ and society to choose b in scenarios where $k < n-k$, and is indifferent if $k = n-k$. Thus, the voter would like society to choose in favor of the majority as that is where the voter is most likely to be in any realization. Once one understands the above reasoning, then Theorem 3 and the importance of relative comparisons becomes clear.

Dichotomous Societies

Theorem 4 showed that there always exists a self-stable voting rule in a homogeneous society. Next, let us examine societies with some heterogeneity, starting with the next simplest case where society splits into two groups.

A society (N, p) is *dichotomous* if there exists $N^1 \neq \emptyset$, $p^1 \in (0,1)$, $N^2 \neq \emptyset$, and $p^2 \in (0,1)$ such that $N = N^1 \cup N^2$, $p_i = p^1$ for all $i \in N^1$, $p_i = p^2$ for all $i \in N^1$.

A dichotomous society is thus one which can be divided into two groups such that members of the same group have the same p_i's. For such a society, let n^1 and n^2 denote the respective cardinalities of N^1 and N^2.

Example 1.

$N^1 = \{1, \ldots, 4\}$ and $N^2 = \{5, \ldots, 10\}$ with $p^1 = .01$ and $p^2 = .99$.

In this society, $\widehat{s}^1 = 8$ and $\widehat{s}^2 = 4$.

Let us examine why $\widehat{s}^2 = 4$ (the intuition for $\widehat{s}^1 = 8$ is similar). This can be verified by direct calculations, but also can be seen in an intuitive manner. Let us consider a voter in N^2. Consider a scenario where exactly three voters end up supporting change. Given the extreme values of $p^1 = .01$ and $p^2 = .99$, if there are three voters who end up supporting change, it is very likely that all of those voters are from N^2. Given that there are six voters in N^2 this leads to a probability of nearly 1/2 that a voter in N^2 would assign to supporting change conditional on three voters supporting change. Although this probability is nearly 1/2, it is still less than 1/2 due to the small probability that some of the voters in N^1 will be among those supporting change. So, a voter in N^2 will prefer that society choose the status quo conditional on three voters supporting change. If we consider a scenario where exactly four voters end up supporting change, then the conditional probability that a voter in N^2 would assign to being one of the supporters of change is nearly 2/3. Since it is above 1/2, a voter in N^2 will prefer that society choose change conditional on four voters supporting change. Given these two observations it follows that $\widehat{s}_2 = 4$. Similar reasoning leads to $\widehat{s}_1 = 8$.

Generally we can think of a voter considering each possible scenario of numbers of supporters for each of the alternatives. For each scenario the voter determines which group they are more likely to fall in. The voter's most preferred voting rule ($\widehat{s}_i$) corresponds to the scenario with the smallest sized group supporting change for which the voter finds it more likely that he or she will support change. We can see that if the voting rule is raised or lowered from 4, then there will be some scenarios where the choice will be made in favor of the group that the voter finds it less likely that he or she will fall in. We can also see why it is rare for a voter to have twin peaks - as that can only happen in a case where the voter assigns probability of exactly 1/2 to each of the two groups in some scenario.

In this example, $\{7, 8\}$ is the set of self-stable voting rules. It is easy to see that 8 is self-stable as only group N^2 would like to change voting rules if 8 is used, but then they only have 6 members and so are too small to make the change under a rule of 8. The same is true of quota 7, and although in that case group N^1 would like to raise the quota from 7 to 8 it is too small to do so. To see that no other rule is stable, note that 4 is unanimously preferred to any smaller rule, and 8 is unanimously preferred to any larger rule. So the only other candidates for self-stability are the quotas 4, 5, and 6. However, 5 and 6 are not stable because N^2 prefers 4 and has enough voters to move the quota to 4. 4 is not stable since group N^1 would have enough voters to increase the quota.

While in the above example, there were two self-stable voting rules, one could imagine that if p^1 were larger then a situation would arise where $6 \geq \widehat{s}^1$ and $4 \geq \widehat{s}^2$. If that were possible, then there would not exist a self-stable voting rule. Then the 6 people in group 2 would want to move the rule from $\widehat{s}^1$ down to $\widehat{s}^2$ and the 4 people in group 1 would want to move the rule up from $\widehat{s}^2$ to $\widehat{s}^1$. As it turns out, however, such examples do not exist and there always exists a self-stable voting rule in a dichotomous society. This is stated in the following theorem.

THEOREM **5** *A dichotomous society has at least one self-stable voting rule.*

The proof of Theorem 5 involves explicit examination of voters' conditional probabilities that they will support alternative b if k voters support b (the types of conditional probabilities discussed in Example 1). The proof is complex and provided in the appendix. Very roughly, it works by relating the conditional beliefs of the two groups to each other. The main case that has to be ruled out to establish existence is where $n^2 \geq \widehat{s}^1$ and $n^1 \geq \widehat{s}^2$, when $\widehat{s}^1 \neq \widehat{s}^2$.[13] If the beliefs of N^1 are such that $n^2 \geq \widehat{s}^1$, this means that the voters in N^1 have *relatively* high beliefs that they will be among the supporters of b. This implies that the voters in N^2 have *relatively* low beliefs that they will be among the supporters of b, and so $\widehat{s}^2$ will be high enough to be larger than n^1. The challenge in the proof is to show that these relative statements translate into absolute statements about the relationship between $\widehat{s}^1$ and $\widehat{s}^2$ and their comparison to n^1 and n^2.

Self-Stability in More Heterogeneous Societies

While there always exists a self-stable voting rule in any dichotomous society, the same is not true more generally. The reasoning behind the proof of Theorem 5 cannot be extended, as tying down the beliefs of one group of voters no longer determines what the conditional beliefs of the remainder must be, given the possible heterogeneity of the remainder. In fact, the existence of self-stable voting rules is no longer guaranteed if society has at least three levels taken on by the p_i's. This is illustrated in the following counter-example to the general existence of self-stable voting rules.

Example 2.

$N = \{1, \ldots, 5\}$. $p_1 = p_2 = p_3 = 1/2$, $p_4 = 3/8$, and $p_5 = 3/16$.

Direct calculations lead to $\widehat{s}_1 = \widehat{s}_2 = \widehat{s}_3 = 2$, $\widehat{s}_4 = 3$ and $\widehat{s}_5 = 4$.

Let us verify that there is no self-stable voting procedure. All voters want to raise the quota from 1 and lower it from 5. That leaves the quotas of 2,3, and 4 to be checked as the only possibilities for self-stable voting rules. Voters 1 to 3 would vote to lower it from 3 to 2, voters 1 to 4 would vote to lower it from 4 to 3, and voters 3 and 4 would vote to raise it from 2 to 3. Thus, no voting rule is self-stable.

The possibility that a society may not have a self-stable voting rule is striking, to say the least. In order to understand its implications, it is worth discussing more extensively when

[13] If, for instance, $\widehat{s}^1 > n^2$ then $\widehat{s}^1$ would be self-stable. So it would have to be that both $n^2 \geq \widehat{s}^1$ and $n^1 \geq \widehat{s}^2$ for there not to exist a self-stable rule. Without loss of generality let $\widehat{s}^2 > \widehat{s}^1$, as the case where $\widehat{s}^2 = \widehat{s}^1$ would lead to unanimity and thus self- stability. So, to see that if a case existed where $n^2 \geq \widehat{s}^1$ and $n^1 \geq \widehat{s}^2$ when $\widehat{s}^1 \neq \widehat{s}^2$, then there would not exist a self-stable voting rule, note that there would be unanimous support for change of any s that lies outside of the range between (and including) $\widehat{s}^1$ and $\widehat{s}^2$. Also N^1 would want to change away (and could change) from and s such that $\widehat{s}^2 \geq s > \widehat{s}^1$. Finally, N^2 would want to change from $\widehat{s}^1$.

this phenomenon can occur, and what practical implications it may have.

We begin by identifying some conditions on the distribution of p's that guarantee the existence of self-stable rules.

Say that a society is *symmetric* if when voters are labeled such that $p_i \geq p_j$ when $i > j$, it follows that $p_i = 1 - p_{n-i}$.

PROPOSITION **1** *If a society is symmetric, then* s^{maj} *is a self-stable voting rule.*

This fact is an easy corollary of Theorem 6, below.

Let $\widehat{s}_{\text{med}}$ denote the median of $(\widehat{s}_1, \ldots, \widehat{s}_n)$, i.e., the median of the peaks of the voters. Note the following simple condition for existence.

Remark: If $\widehat{s}_{\text{med}} \geq s^{\text{maj}}$, then $\widehat{s}_{\text{med}}$ is self-stable.

When $\widehat{s}_{\text{med}} \geq s^{\text{maj}}$, then at most half of the population would like to lower the rule below the median, and at most half would like to increase it above the median, and since $\widehat{s}_{\text{med}} \geq s^{\text{maj}} \geq \frac{n+1}{2}$, it follows that $\widehat{s}_{\text{med}}$ be self-stable.

The above remark points at some asymmetry in the role of majority sizes that are above s^{maj} and those that are below. This asymmetry is further emphasized by the following result, which also establishes that, although instances of nonexistence of self-stable rules may be robust, there are more societies where self-stable rules exist than for where they fail to exist.

THEOREM **6** *If there does not exist a self-stable voting rule for a society* (N, p)*, then there exists a self-stable voting rule for the society* $(N, \overline{p})$*, where* $\overline{p}$ *is defined by* $\overline{p}_i = 1 - p_i$ *for each* i*. Moreover,* $\widehat{s}_{\text{med}}$ *is self-stable for society* $(N, \overline{p})$*.*

The proof of Theorem 6 is relatively simple and so we just provide a sketch. It follows from the observation that the setting we are examining is symmetric in the following way: if in society (N, p) voter i would like society to choose b conditional only on knowing that s voters out of society favor b, then in society $(N, \overline{p})$ voter i would like society to choose a conditional only on knowing that s voters out of society favor a. This implies that if $\widehat{s}_i$ is i's peak under society (N, p), then $n - \widehat{s}_i + 1$ is i's peak under society $(N, \overline{p})$. To establish the theorem, note that non-existence of a self-stable voting rule implies that $\widehat{s}_{\text{med}}$ is no larger than $\frac{n}{2}$, as otherwise it would be self-stable. The reasoning above then implies that $\widehat{s}_{\text{med}}$ for society $(N, \overline{p})$ is larger than $\frac{n}{2}$, and so is stable.

Despite its simple proof, Theorem 6 has some powerful implications. It implies that non-existence is a problem for less than "half" of the potential societies in terms of the p's.

As already remarked, Proposition 1, which asserts the existence of self-stable voting rules for symmetric societies, is an easy corollary of Theorem 6.

Before proceeding further, let us make some additional remarks about self-stable voting rules.

When self-stable voting rules exist, there may be a number of them. Moreover, the set of self-stable voting rules need not be an interval, nor need it include s^{maj}. These points are illustrated in the following example.

Example 3.

The society (N, p) is dichotomous.

$N^1 = \{1, \dots, 5\}$ and $N^2 = \{6, \dots, 16\}$ with $p^1 = .01$ and $p^2 = .99$.

Here $\widehat{s}^1 = 14$ and $\widehat{s}^2 = 6$.

It follows that $\{6, 12, 13, 14\}$ is the set of self-stable voting rules.

It is clear that the set of self-stable voting rules will consist of a set of intervals, each of which includes at least one $\widehat{s}_i$. This puts an upper bound on the number of disjoint intervals that can be included, at the number of distinct p_i's that are present in the society.

As we have seen so far, self-stable rules will exist for many, but not all, societies. Does this mean we should take the possibility that a society might not find a self-stable rule as a serious threat to the stability of actual constitutions or is it just a curiosity? Next, we will show that this possibility of instability that we have unearthed may help explain the reason why many actual societies resort to special rules when it comes to changing the voting rules. This is discussed in the next section where we turn to the definition and the characterization of self-stable constitutions.

Sub-Majority Voting Rules

Before turning to the question of constitutional design, let us comment on some problems related to the choice of majority sizes smaller than s^{maj}.

Rules with $s < s^{\mathrm{maj}}$ can be problematic in the following sense. Consider a situation where a and b are each supported by half of the population. A vote under s will result in b becoming the new status quo. But then, with b as the new status quo, the other half of the voters would support (and could effect) change back to a if it is proposed for a vote against b. Thus, there is the potential to continuously cycle back and forth between a and b as the status quo. This, of course, is only a potential problem of sub-majority rules.

Note that there are two caveats to the above noted difficulty with sub-majority rules. First, for some alternatives it may not be possible to make reversals. For instance, if a is a current membership of a society and b is a question to include a new member, it may not be permitted to later vote to revoke membership. There are many such examples of decisions which cannot be reversed, such as a vote to tenure a faculty member, or a vote to declare war, etc.. Second, the difficulty requires that one reasonably expect that the reversed proposal be made, and so the agenda control becomes important. It may be that the agenda is controlled in manners so that once b has been voted for, a is never again pitted as an alternative. We have abstracted away from the agenda in our model, and a more complete analysis of the potential instability of sub-majority rules demands a careful modeling of the agenda.

In spite of these caveats, this possible chaotic behavior associated with sub-majority rules is not to be taken lightly, and so we discuss its importance with regards to our analysis from two different perspectives.

First, we consider the possibility that a society somehow precludes itself from ever selecting a sub-majority rule. If this is the case, then the existence of self-stable voting rules is ensured. To see this, consider such a society. The preferences of voters over the restricted set of s's ($s \geq s^{\text{maj}}$) are still single peaked. Voters whose unrestricted peaks were at least s^{maj} have the same peak on the restricted set, while voters whose peaks were below s^{maj} now have s^{maj} as a peak. The median of the restricted peaks will be self-stable over the restricted set of voting rules. This leads to the following theorem.

Theorem 7 *For any society where only $s \geq s^{\text{maj}}$ are admissible voting rules, $\widehat{s}_{\text{med}}$ (defined relative to restricted preferences) is a self-stable voting rule.*

The proof follows the argument of the remark we made earlier in this section. Hence, for societies who find that s's below s^{maj} should be excluded from the analysis on a priori grounds, we can still offer a theory for the choice of rules, and add a new property, that of self-stability, to the already well justified choice of the median based-rules.

Despite the reassuring nature of the above theorem, we should not close our eyes to the possibility that societies might consider sub-majority rules, in spite of their potential for chaotic decision making behavior. From a positive viewpoint, there is evidence of chaotic behavior, cycles and swings in many collective-decision making processes. We shall not insist on that, but a theory based on the idea that voters actually vote on how to vote should not exclude the possibility that some voters would support rather unstable arrangements. It is also rare that the possibility of a sub-majority voting rule is prohibited by fiat. More importantly, a democratic constitution is open to amendment. If under current rules, a sufficient majority prefers a sub-majority voting rule there may be nothing to prevent them from making that change. This is precisely the difficulty that arises with the lack of existence of self-stable rules: some super-majority prefers a sub-majority rule, and they have sufficient numbers to enact that change to the current system.

Finally, notice that in example 3, a rule with s below s^{maj} (s=6) emerges as self-stable, along with others involving values above majority.[14] Excluding these low value rules a priori will deprive us of knowing all possible stable arrangements, when they exist.

5 Self-Stable Constitutions

We now explore the consequences of admitting constitutions that allow for different voting rules to be used for making different types of decisions. This will solve the existence problems noted in the previous section, and provides an explanation for such structure in observed constitutions. A constitution can specify one voting rule to be used on all issues except for the change of this "standard" rule, where a different rule may be used. The analysis above suggests that there are situations where this would be natural, and provides an understanding of why this should be the case.

A *constitution* is a pair of voting rules (s, S), where s is to be used in votes over the issues a, b and S is to be used in any votes regarding changes from s to any other rule s'. [15][16]

---

[14] Moreover, there exist examples where the only self-stable rules are sub-majority rules. For example, consider a dichotomous society with $N^1 = \{1, 2\}$ and $N^2 = \{3, \ldots, 7\}$ with $p^1 = .3$ and $p^2 = .5$. Straightforward calculations lead to $\widehat{s}^1 = 5$ and $\widehat{s}^2 = 3$. There $s = 3$ is the only self-stable rule.

[15] There are many examples where such pairs of rules are in place. For instance, the U.S. senate uses simple majority rule, and a 67/100 rule to change the senate rules. In fact, under the filibusters that are possible in the senate, one needs 60/100 votes to call a vote and so the effective voting rule might be thought of as a (60,67) rule rather than a (50,67) rule. An interesting (unstable) example arose recently in California. Under the law until 2000, school bond and tax issues required a 2/3 majority of the participating voters to pass. So we might think of these votes as having the rule s be 2/3 of the voters. However, propositions (initiatives that may be placed on the ballot through a variety of means) in California may be passed with a 1/2 majority. In particular, one can place a proposition on the ballot which changes the vote required on such issues. Thus, one can amend the rule by a 1/2 vote. Thus we can think of S as being 1/2 of the voters. In fact in the 2000 election, Proposition 39 suggested changing the voting rule on school bond and tax issues from 2/3 to 55%. Interestingly, Proposition 39 passed with 53.4% (as reported by the Secretary of State of California) of the vote. Having a 2/3 majority voting rule that can be amended by a 1/2 vote is inherently unstable.

[16] As pointed out to us by Randy Calvert, one could also think of a more general nesting of rules, where one thinks of a voting rule S' to amend (s, S), and so on; and it might be interesting to consider when these may be truncated (as effectively the case of a pair means that the same S is used for all higher orders).

A constitution $(s, S))$ is *self-stable* if $\#\{i \mid U_i(s') > U_i(s)\} < S$ for any s'..

Self-stability of a constitution requires that the preferences of voters be such that there does not exist a voting rule s' that would defeat the constitution's prescribed voting rule s to be used for choices over issues, when these two voting rules are compared under the constitution's voting rule S, to be used for choices over rules. So, a self-stable constitution is one that would not be changed once in place.

THEOREM **8** *For any society, the constitutions (s^{maj}, n) and $(\widehat{s}_{\mathrm{med}}, S)$ for any $S \geq s^{\mathrm{maj}}$ are self-stable.*

Theorem 8 follows is a straightforward consequence of our theorems on intermediate preferences (Theorem 2) and on relative positioning of voter's peaks (Theorem 3), and so we simply offer a description of the proof as follows. The self-stability of (s^{maj}, n) follows from the observation that by Theorems 3 and 2 there is always at least one voter who will wish to keep the voting rule over issues no higher than s^{maj} and at least one who will wish to keep the voting rule no lower than s^{maj}. Thus, there is no unanimous consent to raise or lower the voting rule from s^{maj}. The self-stability of $(\widehat{s}_{\mathrm{med}}, S)$ with $S \geq s^{\mathrm{maj}}$ follows from Theorem 2 and the definition of $\widehat{s}_{\mathrm{med}}$, as by intermediate preferences fewer than $n/2$ voters will prefer to raise the voting rule from $\widehat{s}_{\mathrm{med}}$, and similarly fewer than $n/2$ voters will prefer to lower the voting rule from $\widehat{s}_{\mathrm{med}}$.

Theorem 8 is essentially tight in the sense that for any (s, S) that does not coincide with either (s^{maj}, n) or $(\widehat{s}_{\mathrm{med}}, S)$ with $S \geq s^{\mathrm{maj}}$, there is some situations in which (s, S) is not self-stable (with a single exception that $(s^{\mathrm{maj}}, n-1)$ is always self-stable whenever n is odd).[17] Let us be more explicit. First, consider (s, S) with some $s \neq s^{\mathrm{maj}}$. If p is such that $p_i = p_j$ for all i and j then $\widehat{s}_i = s^{\mathrm{maj}}$ for all i and so any (s, S) for which $s \neq s^{\mathrm{maj}} = \widehat{s}_{\mathrm{med}}$ will be unstable regardless of S. So we need only consider (s, S) where $s = s^{\mathrm{maj}}$ or $s = \widehat{s}_{\mathrm{med}}$. We can see the problem with $(\widehat{s}_{\mathrm{med}}, S)$ where $S < s^{\mathrm{maj}}$ from Example 2, as it is possible to have societies where a near majority prefers to move the voting rule away from $\widehat{s}_{\mathrm{med}}$.[18] Finally, when considering (s^{maj}, S) with $S < n$ ($S < n-1$ if n is odd), consider a society where voter 1 has p_1 near 0 (and the same for voter 2 in the case of n being odd), and all other voters have the same p_i near 1. For high enough p_i, voters i will have probability greater than 1/2

[17]Note that any Pareto optimal s is stable when put together with n. The claim here is that s^{maj} is the only s that is Pareto optimal for all societies.

[18]More generally, consider a society with a single voter who has the median preferences and other voters who have extreme p_i's near 0 and 1, who will prefer to lower or raise the voting rule. In particular, the voters with p_i's near enough to 1 will prefer an $s < \widehat{s}_{\mathrm{med}}$ over $\widehat{s}_{\mathrm{med}}$ and there will be at least $s^{\mathrm{maj}} - 1$ such voters.

of supporting change when there are $n/2$ supporters of change if n is even and when there are $(n-1)/2$ supporters of change when n is odd. This leads to peaks of $s^{\text{maj}} - 1$ for the voters with p_i near 1, and so they will vote to decrease the voting rule if it is set at s^{maj}.

In summary. For each society, there will always be at least two self-stable constitutions (three when n is odd).

Although we have treated the constitutions (s^{maj}, n) and $(\widehat{s}_{\text{med}}, S)$ on equal footing in the statement of Theorem 8, notice the following essential difference. The constitution $(\widehat{s}_{\text{med}}, S)$ varies across societies, since $\widehat{s}_{\text{med}}$ depends on the distribution of p_i's. On the other hand, (s^{maj}, n) is the same across all societies of the same size. Hence, (s^{maj}, n) is a stable constitution regardless of the society, while a constitution of the form $(\widehat{s}_{\text{med}}, S)$ is by definition tailored to a specific society.

Simple Majority Constitutions

The self-stability of constitutions involving simple majority rule is of particular interest because of the prominence of simple majority in actual constitutions and its special properties including overall efficiency (see the concluding remarks for additional discussion). We have just seen that the particular constitution (s^{maj}, n) is, in addition, very robust, since it self-stable for any society. We now explore the conditions on the distribution of p_i's that are sufficient for other constitutions (s^{maj}, S) to be self-stable for values of $S < n$.

Let $z_i = \frac{p_i}{1-p_i}$. Thus, z_i represents the ratio of the probability that i supports change compared to the probability that i supports the status quo. Any positive number is a potential z_i.

THEOREM **9** *For any society with even n the constitution (s^{maj}, S) is self-stable if*

$$S > |\{i : \sum_{C \subset N, |C| = \frac{n}{2}, i \in C} (\times_{j \in C} z_j) \geq \sum_{C \subset N, |C| = \frac{n}{2}, i \notin C} (\times_{j \in C} z_j)\}| > n - S. \tag{3}$$

Note that (3) can be rewritten as

$$S > |\{i : z_i \geq \sum_{k \neq i} \lambda_k^i z_k\}| > n - S, \tag{4}$$

where

$$\lambda_k^i = \frac{2}{n} \frac{\sum_{|C| = \frac{n}{2} - 1; i, k \notin C} (\times_{j \in C} z_j)}{\sum_{|C| = \frac{n}{2} - 1; i \notin C,} (\times_{j \in C} z_j)}.$$

Here, the λ_k^i are weights such that $\sum_{k \neq i} \lambda_k^i = 1$, and so $\sum_{k \neq i} \lambda_k^i z_k$ is a weighted average of z_k's over k's other than i. Thus, (4) says roughly that the number of voters with above average z_i's is not too high and not too low. It can be shown that this is also equivalent to having the number of voters with below average z_i's not be too high or too low.

Condition (3) is almost a necessary condition as well, except for the possibility that one particular voter (the $n - S$-th voter when ordered in terms of decreasing p_i's) has peak exactly at s^{maj} which allows for a slightly weaker condition.

To see the implications of Theorem 9, let us consider the constitution where $s' = 2n/3$. That constitution is stable, provided there are at least 1/3 of the voters who do not wish to raise the voting rule from s^{maj} and at least 1/3 of the voters who do not wish to lower it from s^{maj}. The proof of the theorem involves showing that these are equivalent to the inequalities relating the z_i's. The requirements of the theorem are then that at least 1/3 and no more than 2/3 of the voters have a z_i that is bigger than the weighted average of the other voters' z_i's. This is in effect a limitation on the skewness of the distribution of the z_i's (or, in effect, the p_i's). If the distribution of z_i's is not too skewed, then $(s^{\text{maj}}, 2n/3)$ will be self-stable.

6 Concluding Remarks

There are a number of dimensions along which our model has been simplified, such as taking the agenda to be a binary one, taking the agenda to be exogenous, examining only anonymous voting rules, considering non-repeated environments, and considering a fixed population of voters. Relaxing some of these restrictions provides a rich agenda for further research. Some of these dimensions we have little feeling for at this point, while others we have given some preliminary thought to as we briefly discuss below.

Anonymity

Anonymity has been presumed in our analysis through the definition of a voting rule. Allowing for non-anonymous rules can help with self-stability. For instance, dictatorial rules are clearly self-stable. There are contexts where non-anonymous rules are quite natural, as when certain voters have seniority or other priority in a society. Thus, it would be interesting to consider various non-anonymous rules (other than dictatorships) to understand their self-stability properties.[19] This issue alone presents a fairly rich agenda for further study of endogenous voting rules.[20]

[19]The class of voting rules can be modeled by the class of simple games, which list the coalitions of agents can enforce change over the status quo. The extension of self-stability is then to consider the voting rules for which there is no alternative preferred by all members of some winning coalition.

[20]A new paper by Sosnowska (2002) extends the model here to weighted voting rules.

Large Numbers

We have deliberately worked with finite societies for two reasons. First, there are many applications where the society in question is small and not well approximated by an infinite society. Second, if one worked with a continuum society (or some other infinite model), then, without making additional assumptions about the distribution of the underlying uncertainty, a (suitable) law of large numbers would eliminate the uncertainty over the proportion of society supporting change over the status quo. This uncertainty is the critical aspect that makes for non-degenerate and interesting voters' preferences over voting rules.

While we chose to work with a finite model, it is still interesting to ask questions about large societies.[21] Example 2 extends when the society is replicated a large number of times, and so general existence of self-stable voting rules will not come simply from considering a large society. However, there may be some interesting conditions that are sufficient for self-stability one can obtain from looking at large societies.[22]

Preferences

Assuming that a voter gets a utility of 1 when his preferred alternative is selected and 0 otherwise involves more than a normalization. Instead, it could be that when voter i supports a then i gets utility 1 if a is selected and 0 if b is selected, while when voter i supports b then i gets utility x_i when b is selected and 0 if a is selected.

This more general setting leads to changes in the analysis in the following ways.

First, Theorem 1 on single-peaked preferences goes through unaltered and it is easily checked that the proof works with only slight modification.

Second, the extension of Theorem 2 on intermediate preferences is a more complicated matter. There are now two characteristics that distinguish voters and so finding an ordering on voters for which their preferences are intermediate is more delicate. In the case where $x_i \geq x_j$ whenever $p_i \geq p_j$ (so that voters who are more likely to support alternative b care relatively more about alternative b), preferences are still intermediate. Again, for this situation the proof goes through with very little modification. This would seem to be a natural condition. However, if there is no such relationship between the x_i's and the p_i's, then preferences may fail to be intermediate, and it is easy to construct counter-examples.

[21] See Schofield (1971) for some calculations concerning voters' preferences in large heterogeneous societies.

[22] One possibility is to think about conditions on the distributions of p_i's, in an analogous way that conditions identified by Caplin and Nalebuff (1988) on distributions of preferences suffice for an alternative with nice properties in their setting.

The existence of self-stable voting rules with dichotomous preferences, Theorem 5, depends on the property that a voter cares (in expectation, at least) similarly for having a win when the voter supports a and having b win when the voter supports b. Without that assumption, examples can be constructed where there does not exist a self-stable voting rule. However, Theorem 6 extends under an ordering that preserves the intermediate preferences.

Another aspect of preferences that might be due for further consideration is the assumption of the independence of the probabilities that the voters support change. This assumption played a role in our proof of single-peakedness. Most importantly, this ensures that likelihood that a voter has that they support change conditional on k voters supporting change is be monotone in k. With certain forms of correlation, this conditional probability may no longer be monotone. While arbitrary forms of correlation could be difficult to accommodate, there are natural ones which still allow for such monotonicity and would thus still be tractable.

Dynamics and the Time Perspective

In our analysis we have considered the choice of a voting rule at a time where preferences over the alternatives to be voted on are uncertain, but voters have probability assessments over their likelihoods of supporting change over the status quo.

What if we consider different time perspectives?

First, if we move back in time to an ex-ante point of view (before voters know the p_i's), then simple majority rule becomes attractive as it maximizes total utility, and hence ex-ante utility if voters have similar beliefs about what their p_i's are likely to be. However, unless s^{maj} turns out to be self-stable or else voters can bind themselves not to make future changes, one has to address the issue of changes that can be made once the p_i's are known. So, we still need to consider the interim perspective, unless there is a way to make future changes in rules impossible.

We can also consider moving forward in time to an ex-post point of view, where voters know which alternatives they support. If a voting rule is already in place from the interim perspective, there will be no changes made. To see this, note that the voters in favor of changing the voting rule will be the voters who are not getting the choice that they want. If they are not getting the choice that they want under the current voting rule s, then they must be a group of less than s voters and so could not affect a change.

While the ex-post perspective has no implications for self-stable voting rules, it can change the outcome for some self-stable constitutions. In particular, it can destabilize constitutions of the form (s, s') where $s > s'$. In such a case, if there turn out to be s'' voters in favor of change, where $s > s'' \geq s'$, then the s'' voters will vote to change the voting rule to s''. This problem arises only for constitutions where $s > s'$, and so if a constitution has $s' \geq s$, then there will be no incentive for voters to change it ex-post.

An issue that is of interest for future investigation arises in a dynamic situation where voters face a sequence of choices. Consider a situation where voters already know their preferences on a current choice, but only have limited knowledge over future choices. Here there is a possibility of a voter's preferences over voting rules going against their immediate (myopic) wishes, due to the longer horizon. This issue is pointed out in the following example.

Example 4.

$N = \{1, \ldots, 8\}$. $p_1 = p_2 = p_3 = .01$, $p_4 = p_5 = .49$, and $p_6 = p_7 = p_8 = .99$.

In this case, direct calculations lead to $\widehat{s}_1 = \widehat{s}_2 = \widehat{s}_3 = 2$, $\widehat{s}_4 = \widehat{s}_5 = 5$ and $\widehat{s}_6 = \widehat{s}_7 = \widehat{s}_8 = 7$.

The self-stable voting rules are $\{5, 6, 7\}$.

Now let us adopt a longer time perspective. Suppose we are using the voting rule $s = 5$.

Consider a realization where society evenly splits between a and b, and in particular voters 1 to 3 and 6 support a and 4, 5, 7 and 8 support b. Under the voting rule $s = 5$ the status quo a will be chosen.

Consider the incentives of the voters to change the voting rule. If the rule is changed from 5 to 4 then b would be chosen instead of a.

Voters 4 and 5 will gain 1 unit of utility today, but lose $U_4(5) - U_4(4)$ in each of the future periods. This expression is close to 0, and so for all except very high discount factors, the time discounted future loss in expected utility from changing from a voting rule of 5 to 4 is less than 1. So, for discount factors that are not too high, voters 4 and 5 will support a change to a voting rule of 4. Voters 7 and 8 would gain today and in the future such a change. Voter 6 would lose 1 unit of utility today from such a change, but would gain $U_6(5) - U_6(4)$ discounted over the future horizon. While $U_4(5) - U_4(4)$ is close to 0, $U_6(5) - U_6(4)$ is not. Thus for some range of discount factors, voters 4 to 8 would all support a change from 5 to 4.

Simple majority as a special rule: Efficiency and Immunity to Vote Trading

We close by pointing out two related important features, arising in the model that we used here, that give a special status to s^{maj}.

A first property was the focus of Rae (1969) (and later proven by Curtis (1972) and Badger (1972)). It is that s^{maj} maximizes total societal welfare, and it is essentially the only rule to do so. More precisely, for any society (N, p), s^{maj} uniquely maximizes $\sum_i U_i(s)$, unless

n is even in which case $\frac{n}{2}$ also maximizes $\sum_i U_i(s)$. This property can also be seen as an ex ante efficiency property (from the perspective before voters know their p_i's, where they each have the same beliefs). There is a fairly direct proof of this property.[23] Majority rule maximizes total societal welfare at each ex post realization of preferences over a and b since it maximizes the number of 1's realized as utility compared to 0's that are realized. Any other rule (except $\frac{n}{2}$ when n is even) realizes a lower total utility at some ex post realization of preferences as it will select one of the alternatives when a minority supports it (favoring a if $s > s^{\text{maj}}$ or b if $s < s^{\text{maj}}$).

A second property that is related to the utilitarian property of simple majority rule, is that it is the only voting rule that is immune to the trading of votes by the grand coalition. This point is made more precisely below. The intuition behind it is straightforward but is still worth making, as it suggests that the utilitarian property of simple majority rule not only has welfare implications, but also implications in terms of incentives.

Describe a realization of preferences by $C \subset N$, where C is the set of voters who support the change over the status quo.

Consider the possibility that voters trade votes as follows. For each C that could be realized, a randomization device says "trade" with some probability p_C and "no trade" with probability $1 - p_C$. "Trade" is then interpreted as having the voters in the minority group reverse their votes. Under "no trade", all voters vote in their own interest.

In the above situation vote trading is described by a randomization device, where in some (random) circumstances the minority of voters are called on to reverse their votes. They do this with the expectation that there are other situations where they will be in the majority and benefit from having the minority reverse their votes. If the vote is only a one time happening, then one would need a commitment to abide by the randomization device in order to have such a scheme work. If there is a sequence of choices to be made (and not too much discounting by voters), then the p_C's can be thought of as frequencies and votes can indeed be traded in a quid-pro-quo or self- sustaining manner.

So a vote trading scheme is a list $\{p_C\}_{C \subset N}$ of the probabilities that votes are traded for the various possible realized groups supporting change.

Note that when "trade" is in effect, the votes are cast unanimously and in favor of what the majority of voters prefers. So effectively, a vote trading scheme transforms a voting rule to choose as s^{maj} would in a limited (random) set of circumstances. For every voting rule except s^{maj} this reverses the outcome in some scenarios.

[23] Much more complex proofs appear in Curtis (1972) and Badger (1972).

A voting trading scheme $\{p_C\}_{C\subset N}$ *upsets* a voting rule s, if every voter's expected utility is higher under the vote trading scheme given the rule s than if they voted in their own interest under s.

A voting rule s is *immune to vote trading* if it is not upset by any vote trading scheme.

It is now easy to verify that simple majority rule is the only voting rule that is immune to vote trading, unless n is even in which case $n/2$ is also immune to vote trading. (This is proven in the appendix.)

While simple majority rule is uniquely immune to vote trading, we should point out that the definition requires that all voters be made better off. We can also consider vote trading by sub-coalitions. It turns out that no voting rule is immune to vote trading by sub-coalitions (provided that $n \geq 4$). This is seen as follows: pick an arbitrary sub-coalition of three voters, say $\{1, 2, 3\}$. Operate a vote trading scheme as follows. Consider all situations where $\{1, 2\}$ are on the losing side of the vote and 3 is on the winning side and is pivotal - i.e., 3 is able to change the vote by reversing her vote. With $n \geq 4$ there is at least one such a scenario for every voting rule s. Have a probability p_3 that 3 reverses her vote conditional on being in such a situation. Choose corresponding p_1 and p_2's. These can be chosen so that the probability that each of the voters 1,2, and 3 unconditionally expects to be on the benefiting side of the vote more often than in the position of being called on to reverse their vote. Thus, each of the voters in $\{1, 2, 3\}$ would strictly benefit from following this scheme.

As we have seen simple majority is not perfect. It is not always self-stable and it is subject to some forms of vote trading. But it is certainly a salient rule, for many reasons, including those mentioned above, and can always be made part of a self-stable constitution. In fact, the lack of self-stability of majority rule (as well as others) provides insight into the prevalence of constitutions with separate voting rules for standard decisions versus amendments of the constitution.

References

Austen-Smith, D. and J. Banks (1996) "Information Aggregation, Rationality and the Condorcet Jury Theorem," *American Political Science Review*, vol. 90, pp. 34–45.

Badger, W.W. (1972) "Political Individualism, Positional Preferences, and Optimal Decision-Rules," in *Probability Models of Collective Decision Making*, edited by R.G. Niemi and H.F. Weisberg, Merrill Publishing: Columbus Ohio.

Buchanan, J.M. and G. Tullock (1962) *The Calculus of Consent: Logical Foundations of Constitutional Democracy* University of Michigan Press: Ann Arbor.

Caplin, A. and B. Nalebuff (1988) "On 64%-Majority Rule." *Econometrica*, vol. 56, pp 787–814.

Condorcet, Marquis de (1785) *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*, Paris.

Curtis, R.B. (1972) "Decision Rules and Collective Values in Constitutional Choice," in *Probability Models of Collective Decision Making*, edited by R.G. Niemi and H.F. Weisberg, Merrill Publishing: Columbus Ohio.

Dasgupta, P. and E. Maskin (1998) "Robustness of Majority Rule," mimeo.

Feddersen, T.J. and W. Pesendorfer (1998) "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts," *American Political Science Review*, vol. 92, pp. 23–36.

Grandmont, J.M. (1978) "Intermediate Preferences and the Majority Rule," *Econometrica*, vol. 46, pp. 317–330.

Koray, S. (2000) "Self-Selective Social Choice Functions Verify Arrow and Gibbard-Satterthwaite Theorems," *Econometrica*, vol. 68, pp. 981–996.

Niemi and Weisberg (1972) "Substantive Applications of Collective Decision-Making," in *Probability Models of Collective Decision Making*, edited by R.G. Niemi and H.F. Weisberg, Merrill Publishing: Columbus Ohio.

Rae, D. (1969) "Decision Rules and Individual Values in Constitutional Choice," *American Political Science Review*, vol. 63, pp. 40–56.

Rousseau, J.-J. (1762) *du Contrat Social; ou, Principes du Droit Politique* reprinted: Paris : Garnier; 1966.

Schofield, N.J. (1971) "Ethical Decision Rules for Uncertain Voters," *British Journal of Political Science*, vol. 2, pp 193–207.

Schofield, N.J. (1972) "Is Majority Rule Special?" in *Probability Models of Collective Decision Making*, edited by R.G. Niemi and H.F. Weisberg, Merrill Publishing: Columbus Ohio.

Sosnowska, Honorata (2002) "Self-Stability for Weighted Voting Games," mimeo: Warsaw School of Economics.

Young, H.P. and A. Levenglick (1978) "A Consistent Extension of Condorcet's Election Principle" *SIAM Journal of Applied Mathematics, Part C*, vol. 35, pp. 283–300.

Appendix

Proof of Theorem 1:

Let $D_i(s) = U_i(s) - U_i(s-1)$. From equation (2) it follows that

$$D_i(s) = (1-p_i)P_i(s-1) - p_i P_i(s-2). \tag{5}$$

Thus,

$$D_i(s) = P_i(s-1)(1 - p_i \frac{P_i(s-1) + P_i(s-2)}{P_i(s-1)}).$$

Note that U_i is single-peaked if there exists $\widehat{s}_i$ (possibly equal to 1 or n) such that $D_i(s) > 0$ for every $\widehat{s}_i > s \geq 2$, $D_i(s) < 0$ for every $n \geq s > \widehat{s}_i$, and $D_i(s) \geq 0$ at $s = \widehat{s}_i$ (with equality holding only when there are twin peaks). Thus, if we can show that $D_i(s)$ has this form, then we will have shown that U_i is single-peaked.

Note that the sign of $D_i(s)$ depends only on the size of $\frac{P_i(s-1)+P_i(s-2)}{P_i(s-1)}$ relative to $\frac{1}{p_i}$. This means that showing that $\frac{P_i(s-1)+P_i(s-2)}{P_i(s-1)}$ is increasing in s for $n \geq s \geq 2$ establishes that $D_i(s)$ has the form specified above. Rewriting

$$\frac{P_i(s-1) + P_i(s-2)}{P_i(s-1)} = 1 + \frac{P_i(s-2)}{P_i(s-1)},$$

means that we need only show that $\frac{P_i(s-2)}{P_i(s-1)}$ is increasing[24] in s.

We follow a proof by induction on n. The case where $n = 2$ is trivial, since then there is only one s that satisfies $n \geq s \geq 2$. Now for the induction step. Suppose that $P_i(s-2)/P_i(s-1)$ is increasing for any $n' \geq s \geq 2$ for societies of size $n-1 \geq n'$. We show that $P_i(s-2)/P_i(s-1)$ is increasing for any $n \geq s \geq 2$.

Let

$$P_{i,j}(s) = \sum_{C \subset N \setminus \{i,j\} : |C| = s} \times_{\ell \in C} p_\ell \times_{k \notin C} (1-p_k).$$

$P_{i,j}(s)$ is the probability that exactly s of the voters other than i and j support the change.

$$\frac{P_i(s-2)}{P_i(s-1)} = \frac{p_j P_{i,j}(s-3) + (1-p_j)P_{i,j}(s-2)}{p_j P_{i,j}(s-2) + (1-p_j)P_{i,j}(s-1)},$$

where $P_{i,j}(s-3) = 0$ when $s = 2$. Rewrite the above equality as

$$\frac{P_i(s-2)}{P_i(s-1)} = \frac{p_j P_{i,j}(s-3)}{p_j P_{i,j}(s-2) + (1-p_j)P_{i,j}(s-1)} + \frac{(1-p_j)P_{i,j}(s-2)}{p_j P_{i,j}(s-2) + (1-p_j)P_{i,j}(s-1)}. \tag{6}$$

[24] When we say "increasing" we refer to the strict sense, and we use the term "non-decreasing" to refer to the weaker sense.

We show that each term on the right hand side of (6) is increasing in s for $n \geq s \geq 2$. Take the first term. It is clear that since $P_{i,j}(s-3) = 0$ when $s = 2$, that it is increasing from $s = 2$ to $s = 3$. So, we need only show that its inverse is decreasing in s for $n \geq s \geq 3$.

$$\frac{p_j P_{i,j}(s-2) + (1-p_j) P_{i,j}(s-1)}{p_j P_{i,j}(s-3)} = \frac{p_j P_{i,j}(s-2)}{p_j P_{i,j}(s-3)} + \frac{(1-p_j)}{p_j} \frac{P_{i,j}(s-1)}{P_{i,j}(s-3)}$$

$$= \frac{p_j P_{i,j}(s-2)}{p_j P_{i,j}(s-3)} + \frac{(1-p_j)}{p_j} \frac{P_{i,j}(s-1)}{P_{i,j}(s-2)} \frac{P_{i,j}(s-2)}{P_{i,j}(s-3)}. \tag{7}$$

Note that the induction step implies that $\frac{P_{i,j}(s-2)}{P_{i,j}(s-3)}$ is increasing in s for $n \geq s \geq 3$. So, each expression on the right hand side is decreasing in s for each $n-1 \geq s \geq 3$ by the induction step, and so the overall expression is. So we only have to worry about the case where $s = n$ and the expression $\frac{P_{i,j}(s-1)}{P_{i,j}(s-2)}$. Note that $P_{i,j}(n-1) = 0$, and so this follows as well.

Recall that the expression in (7) is the inverse of the first term on the right hand side of (6). A similar argument establishes that the second term on the right hand side of (6) is increasing in s. ∎

Proof of Theorem 2:

We first show that $U_i(s) \geq U_i(s-1)$ implies that $U_j(s) \geq U_j(s-1)$ for any j such that $p_j \geq p_i$. Recall that

$$U_i(s) - U_i(s-1) = (1-p_i)P_i(s-1) - p_i P_i(s-2).$$

So, we write

$$U_i(s) - U_i(s-1) = (1-p_i)(P_{i,j}(s-2)p_j + P_{i,j}(s-1)(1-p_j)) - p_i(P_{i,j}(s-3)p_j + P_{i,j}(s-2)(1-p_j)).$$

Likewise,

$$U_j(s) - U_j(s-1) = (1-p_j)(P_{i,j}(s-2)p_i + P_{i,j}(s-1)(1-p_i)) - p_j(P_{i,j}(s-3)p_i + P_{i,j}(s-2)(1-p_i)).$$

It follows that

$$[U_i(s) - U_i(s-1)] - [U_j(s) - U_j(s-1)] = 2(p_j - p_i)P_{i,j}(s-2).$$

Notice that the right hand side of the above equation is nonnegative, because $p_j \geq p_i$.

So, we have shown that

$$U_i(s) - U_i(s-1) \geq U_j(s) - U_j(s-1), \tag{8}$$

whenever $p_j \geq p_i$. Note that if $s > s'$, then

$$U_i(s) - U_i(s') = \sum_{k=s'+1}^{s} U_i(k) - U_i(k-1). \tag{9}$$

So, (9) implies that

$$U_i(s) - U_i(s') \geq U_j(s) - U_j(s'), \tag{10}$$

whenever $p_j \geq p_i$, provided $s > s'$. This establishes that preferences satisfy the single-crossing property.

We now show that this implies that preferences are intermediate.

First, consider the case where $U_i(s) \geq U_i(s')$, $U_j(s) \geq U_j(s')$, $p_j \geq p_k \geq p_i$ and $s > s'$. Since $U_i(s) \geq U_i(s')$ and $p_k \geq p_i$, (10) implies that $U_k(s) \geq U_k(s')$ (with strict inequality if the strict inequality holds for i). So, the desired conclusion of intermediate preferences is established for this case (and the corresponding strict inequality case). Next, consider the case where $U_i(s) \geq U_i(s')$, $U_j(s) \geq U_j(s')$, $p_j \geq p_k \geq p_i$ and $s < s'$. Suppose to the contrary that $U_k(s') > U_k(s)$. Then since $p_j \geq p_k$ it follows from (10) (applied with the roles of s and s' reversed) that $U_j(s') > U_j(s)$, which is a contradiction. Thus, our supposition was incorrect and so $U_k(s) \geq U_k(s')$. Finally, consider the case where $U_i(s) > U_i(s')$, $U_j(s) > U_j(s')$, $p_j \geq p_k \geq p_i$ and $s < s'$. Suppose to the contrary that $U_k(s') \geq U_k(s)$. Then since $p_j \geq p_k$ it follows from (10) (applied with the roles of s and s' reversed) that $U_j(s') \geq U_j(s)$, which is a contradiction. Thus, our supposition was incorrect and so $U_k(s) > U_k(s')$. We have shown that preferences are intermediate. ∎

Proof of Theorem 3: s^{maj} maximizes total societal welfare (see the concluding remarks), as it always results in the maximal number of voters who get a utility of 1. Consider the case where some voter j's peak is greater than s^{maj}. So, $U_j(s^{\text{maj}}+1) \geq U_j(s^{\text{maj}})$. As $\sum_i U_i(s^{\text{maj}}) > \sum_i U_i(s^{\text{maj}}+1)$, it follows that there exists some i with $U_i(s^{\text{maj}}+1) < U_i(s^{\text{maj}})$ which by single-peaked preferences (Theorem 1) implies that $s^{\text{maj}} \geq \widehat{s}_i$. We are left with the case where all voters' peaks are no more than s^{maj}. Suppose to the contrary of Theorem 3 that all the peaks are strictly less than s^{maj}. By the single-peakedness of preferences, this implies that $U_i(s^{\text{maj}}-1) > U_i(s^{\text{maj}})$ for all i, which contradicts the fact that s^{maj} maximizes $\sum_i U_i(s)$. Thus our supposition was incorrect and the Theorem is established. ∎

Proof of Theorem 5: Let $n^2 = \#N^2$ and $n^1 = \#N^1$, and without loss of generality take $p_1 \geq p_2$.

In the case where $p^1 = p^2$ it is easily checked that all preferences are identical with $\widehat{s}_i = n/2 + 1$ if n is even, and $\widehat{s}_i = (n+1)/2$ if n is odd. In that case, $\widehat{s}_i$ is self-stable. So, we consider the case where $p^1 > p^2$, $n_1 \geq 1$, and $n_2 \geq 1$.

Theorem 3 and Corollary 1 implies that $\widehat{s}_2 \geq s^{\mathrm{maj}} \geq \widehat{s}_1$, since $p_2 < p_1$. If $n_2 \geq s^{\mathrm{maj}}$, then it must be that $s^{\mathrm{maj}} > n_1$ and so $\widehat{s}_2$ is self-stable. Therefore, we need only examine the case where $n_1 \geq s^{\mathrm{maj}} > n_2$.

Suppose to the contrary that there is no self-stable voting rule. It must be that $n_2 \geq \widehat{s}_1$ and $n_1 \geq \widehat{s}_2$. Thus,

$$n_1 \geq \widehat{s}_2 \geq s^{\mathrm{maj}} > n_2 \geq \widehat{s}_1 \tag{11}$$

and $p_2 < p_1$.

For $k \in \{1, \ldots, n\}$, let $q_i^b(k)$ be the probability that a voter of type $i \in \{1, 2\}$ supports b *conditional* on knowing that k voters support b. Correspondingly, let $q_i^a(k)$ be the probability that a voter of type $i \in \{1, 2\}$ supports a *conditional* on knowing that k voters support a. By the definition of q_i^a and q_i^b it follows that

$$q_i^a(k) = 1 - q_i^b(n-k). \tag{12}$$

Note that i's peak is the largest s' such that $q_i^b(s') \geq 1/2$ and $1/2 \geq q_i^b(s)$ for $s < s'$.

Below we will establish that

$$\frac{q_2^a(k)}{k} > \frac{q_2^a(k+1)}{(k+1)}. \tag{13}$$

Before proving (13), let us argue that this will complete the proof. Since $q_2^b(\widehat{s}_2) \geq 1/2$ it follows that $1/2 \geq q_2^a(n - \widehat{s}_2)$. So, by (13) (applied iteratively) it follows that

$$1/2 > q_2^a(n - \widehat{s}_1)\frac{n - \widehat{s}_2}{n - \widehat{s}_1}.$$

From the inequality above, we then have

$$n^2 q_2^a(n - \widehat{s}_1) < \frac{n_2}{2}\frac{n - \widehat{s}_1}{n - \widehat{s}_2}.$$

Since it must be that $n_1 q_1^a(n-s) + n_2 q_2^a(n-s) = n - s$, it follows that

$$n_1 q_1^a(n - \widehat{s}_1) > n - \widehat{s}_1 - \frac{n_2}{2}\frac{n - \widehat{s}_1}{n - \widehat{s}_2}.$$

Noting that $n - \widehat{s}_1 \geq n_1$ (recall that $n_1 + n_2 = n$ and $n_2 \geq \widehat{s}_1$ from inequality (11)), the previous inequality requires that

$$q_1^a(n - \widehat{s}_1) > 1 - \frac{n_2}{2(n - \widehat{s}_2)}.$$

Since $n - \widehat{s}_2 \geq n_2$ (recall that $n_1 + n_2 = n$ and $n_1 \geq \widehat{s}_2$ from inequality (11)), the above inequality implies that $q_1^a(n - \widehat{s}_1) > 1/2$. By the definition of $\widehat{s}_1$ we know that $q_1^b(\widehat{s}_1) \geq 1/2$, but then $q_1^a(n - \widehat{s}_1) > 1/2$ contradicts equation (12).

Now, we complete the proof by showing that (13) holds. Let $P^a(n', k)$ denote the probability that, in a society with n^1 voters with p_1 and n' voters with p_2, exactly k of the voters support a. So in this calculation, the number of voters of type 1 is always fixed, but the number of voters of type 2 is given by n'. Writing in the expressions for q_2^a from Bayes' rule, we need to show that

$$\frac{1}{k}\left(\frac{(1-p_2)P^a(n^2-1,k-1)}{p_2P^a(n^2-1,k)+(1-p_2)P^a(n^2-1,k-1)}\right) >$$

$$\frac{1}{k+1}\left(\frac{(1-p_2)P^a(n^2-1,k)}{p_2P^a(n^2-1,k+1)+(1-p_2)P^a(n^2-1,k)}\right). \tag{14}$$

Note that

$$P^a(n',k) = p_2P^a(n'-1,k)+(1-p_2)P^a(n'-1,k-1). \tag{15}$$

So substituting from (15) and simplifying, we rewrite (14) as

$$(k+1)\frac{P^a(n^2-1,k-1)}{P^a(n^2,k)} > k\frac{P^a(n^2-1,k)}{P^a(n^2,k+1)}. \tag{16}$$

We show this by induction on n^2. A straightforward (but tedious) expansion of the expressions (that we leave to the reader) verifies that (16) holds for $n_2 = 1$ and any $k \geq 1$ (set $\frac{P^a(n^2,k)}{P^a(n^2,k+1)} = 0/0 = 1$ when $k > n^1 + n^2$). We now show that if (16) holds for each $n^2 < n'$ and $k \geq 1$, then it holds for n' and any $k \geq 1$. Rewriting (16) at n' and expanding using (15) in each expression we obtain,

$$(k+1)\Big(p_2^2P^a(n'-2,k-1)P^a(n'-1,k+1)+p_2(1-p_2)P^a(n'-2,k-1)P^a(n'-1,k)$$

$$+p_2(1-p_2)P^a(n'-2,k-2)P^a(n'-1,k+1)+(1-p_2)^2P^a(n'-2,k-2)P^a(n'-1,k)\Big)$$

$$> k\Big(p_2^2P^a(n'-2,k)P^a(n'-1,k)+p_2(1-p_2)P^a(n'-2,k)P^a(n'-1,k-1)$$

$$+p_2(1-p_2)P^a(n'-2,k-1)P^a(n'-1,k)+(1-p_2)^2P^a(n'-2,k-1)P^a(n'-1,k-1)\Big)$$

Using the induction hypothesis, we eliminate the first expression on each side of the inequality, and then collecting terms and simplifying we obtain

$$p_2P^a(n'-2,k-1)P^a(n'-1,k)+(k+1)p_2P^a(n'-2,k-2)P^a(n'-1,k+1)$$

$$+(1-p_2)P^a(n'-2,k-2)P^a(n'-1,k)$$

$$> kp_2P^a(n'-2,k)P^a(n'-1,k-1)+(1-p_2)P^a(n'-2,k-1)P^a(n'-1,k-1)$$

Now, substituting for $P^a(n'-1,\cdot)$ from (15), we rewrite the above as

$$p_2^2 P^a(n'-2,k-1)P^a(n'-2,k) + p_2(1-p_2)P^a(n'-2,k-1)P^a(n'-2,k-1)$$

$$+p_2(1-p_2)P^a(n'-2,k-2)P^a(n'-2,k) + (1-p_2)^2 P^a(n'-2,k-2)P^a(n'-2,k-1)$$

$$+(k+1)p_2^2 P^a(n'-2,k-2)P^a(n'-2,k+1) + (k+1)p_2(1-p_2)P^a(n'-2,k-2)P^a(n'-2,k)$$

$$> p_2(1-p_2)P^a(n'-2,k-1)P^a(n'-2,k-1) + (1-p_2)^2 P^a(n'-2,k-2)P^a(n'-2,k-1)$$

$$+kp_2^2 P^a(n'-2,k)P^a(n'-2,k-1) + kp_2(1-p_2)P^a(n'-2,k)P^a(n'-2,k-2)$$

Simplifying, we must only show the inequality

$$(k+1)p_2 P^a(n'-2,k-2)P^a(n'-2,k+1) + (k+1)(1-p_2)P^a(n'-2,k-2)P^a(n'-2,k)$$

$$> (k-1)p_2 P^a(n'-2,k)P^a(n'-2,k-1) + (k-1)(1-p_2)P^a(n'-2,k)P^a(n'-2,k-2)$$

Using (15) at $n'-1$ we rewrite this as

$$(k+1)P^a(n'-2,k-2)P^a(n'-1,k+1) > (k-1)P^a(n'-2,k)P^a(n'-1,k-1). \qquad (17)$$

So we need only show that (17) holds. By the induction hypothesis, we know that

$$(k+1)P^a(n'-2,k-1)P^a(n'-1,k+1) > kP^a(n'-2,k)P^a(n'-1,k),$$

and

$$kP^a(n'-2,k-2)P^a(n'-1,k) > (k-1)P^a(n'-2,k-1)P^a(n'-1,k-1),$$

or

$$kP^a(n'-1,k) > (k-1)P^a(n'-2,k-1)P^a(n'-1,k-1)/P^a(n'-2,k-2).$$

Combined, these imply that

$$(k+1)P^a(n'-2,k-1)P^a(n'-1,k+1)$$

$$> P^a(n'-2,k)(k-1)P^a(n'-2,k-1)P^a(n'-1,k-1)/P^a(n'-2,k-2),$$

which simplifies to

$$(k+1)P^a(n'-1,k+1)P^a(n'-2,k-2) > (k-1)P^a(n'-2,k)P^a(n'-1,k-1).$$

This verifies that (17) holds and completes the proof. ∎

Proof of Theorem 9: First, note that given the single-peaked preferences (accounting for the possibility of two peaks), (s^{maj}, S) is self-stable if and only if

$$|\{i : U_i(s^{\text{maj}}) \geq U_i(s^{\text{maj}} - 1)\}| > n - S \quad \text{and} \quad |\{i : U_i(s^{\text{maj}}) \geq U_i(s^{\text{maj}} + 1)\}| > n - S. \quad (18)$$

A sufficient condition for this is that

$$|\{i : U_i(s^{\text{maj}}) \geq U_i(s^{\text{maj}} - 1)\}| > n - S \quad \text{and} \quad |\{i : U_i(s^{\text{maj}} - 1) \geq U_i(s^{\text{maj}})\}| > n - S.$$

which is in turn guaranteed by

$$S > |\{i : U_i(s^{\text{maj}} - 1) \geq U_i(s^{\text{maj}})\}| > n - S. \quad (19)$$

Recall from (2) that

$$U_i(s) - U_i(s-1) = (1 - p_i)P_i(s-1) - p_i P_i(s-2).$$

Thus,

$$\{i : U_i(s^{\text{maj}} - 1) \geq U_i(s^{\text{maj}})\} = \left\{ i : \frac{p_i}{1 - p_i} = z_i \geq \frac{P_i(s^{\text{maj}} - 1)}{P_i(s^{\text{maj}} - 2)} \right\}. \quad (20)$$

From the definition of $P_i(s)$ it follows that

$$\frac{P_i(s)}{P_i(s-1)} = \frac{\sum_{C \subset N \setminus i, |C| = s} \left[\times_{j \in C} p_j \times_{j \notin C} (1 - p_j) \right]}{\sum_{C \subset N \setminus i, |C| = s-1} \left[\times_{j \in C} p_j \times_{j \notin C} (1 - p_j) \right]}$$

Dividing top and bottom by $\times_{j \neq i}(1 - p_j)$, this becomes

$$\frac{P_i(s)}{P_i(s-1)} = \frac{\sum_{C \subset N \setminus i, |C| = s} \times_{j \in C} z_j}{\sum_{C \subset N \setminus i, |C| = s-1} \times_{j \in C} z_j}$$

So, by the above equation and (20), we can rewrite (19) as

$$S > |\{i : z_i \geq \frac{\sum_{C \subset N \setminus i, |C| = \frac{n}{2}} \times_{j \in C} z_j}{\sum_{C \subset N \setminus i, |C| = \frac{n}{2} - 1} \times_{j \in C} z_j}\}| > n - S. \quad (21)$$

This can be rewritten as

$$S > |\{i : \sum_{C \subset N, i \in C, |C| = \frac{n}{2}} \times_{j \in C} z_j \geq \sum_{C \subset N \setminus i, |C| = \frac{n}{2}} \times_{j \in C} z_j\}| > n - S,$$

which is the claimed expression.

A direct rewriting of (21) leads to the claimed expression in (4):

$$S > |\{i : z_i \geq \sum_{k \neq i} \lambda_k^i z_k\}| > n - S,$$

where

$$\lambda_k^i = \frac{2}{n} \frac{\sum_{|C|=\frac{n}{2}-1;i,k\notin C}(\times_{j\in C} z_j)}{\sum_{|C|=\frac{n}{2}-1;i\notin C}(\times_{j\in C} z_j)}.$$

Direct inspection shows that $\sum_{k\neq i} \lambda_k^i = 1$ for all i. ∎

Proof of Vote Trading Claim: Given that $p_i \in (0,1)$ for each i, every C has a positive probability of being realized. For each C let π_C denote the probability that C is realized. Let $\underline{p} = \min_C \pi_C$.

If s is not the simple majority rule, identify the set of C's where a minority ends up having their alternative chosen. Let M denote this set of C's.

For each $C \in M$, select $p_C > 0$ so that the product $p_C\pi_C$ is the same across all C's in M, and set $p_C = 0$ for any $C \notin M$. This vote trading scheme upsets s. To see this, note that M is a symmetric set as it will be all C's of sizes that fall below s and are at least $\frac{n+1}{2}$ (or above s and no more than $\frac{n+1}{2}$ if s is smaller than $\frac{n+1}{2}$). Given this symmetry and the fact that for each of these C's, we have $\#C \neq \#N \setminus C$, it follows that there are more of these sets C for which any fixed voter i is in the majority than in the minority. By the choice of probabilities, the probabilities of vote trading are the same across all $C \in M$ and so any voter i expects to benefit since they are more often in the majority than the minority. ∎