Games and Economic Behavior 100 (2016) 376–389

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

On the self-(in)stability of weighted majority rules ☆

Yaron Azrieli [a], Semin Kim [b,*]

[a] Dept. of Economics, The Ohio State University, United States
[b] Dept. of Economics, Yonsei University, Republic of Korea

ARTICLE INFO

Article history:
Received 25 November 2015
Available online 5 November 2016

JEL classification:
D72

Keywords:
Voting rules
Weighted majority rules
Self-stability

ABSTRACT

A voting rule f is *self-stable* (Barberà and Jackson, 2004) if any alternative rule g does not have sufficient support in the society to replace f, where the decision between f and g is based on the rule f itself. While Barberà and Jackson focused on anonymous rules in which all agents have the same voting power, we consider here the larger class of weighted majority rules. Our main result is a characterization of self-stability in this setup, which shows that only few rules of a very particular form satisfy this criterion. This result provides a possible explanation for the tendency of societies to use more conservative rules when it comes to changing the voting rule. We discuss self-stability in this latter case, where a different rule F may be used to decide between f and g.

1. Introduction

The voting rule used by a society to choose between different alternatives impacts which alternative is eventually chosen. Therefore, agents' preferences over alternatives naturally translate into preferences over voting rules. Given that members of a society have preferences over voting rules, and that members can propose to replace the current rule by a different one, which rules are likely to survive in the long-run? And how does the answer depend on the characteristics of the society?

Barberà and Jackson (2004) (BJ henceforth) develop a theoretical framework to study these questions, a theory that is based on the endogenous preferences of agents over voting rules. The key concept, which characterizes voting rules that can withstand proposed changes to the rule, is called *self-stability*. Roughly speaking, a rule f is self-stable if, given any proposed alternative rule g, the coalition of agents who prefer g to f is not large enough to win the vote on replacing f by g, where the rule used for this latter decision is f itself. Thus, the idea underlying this concept is that the same voting rule (f) governs ordinary decisions as well as decisions about changes to the voting rule itself.[1]

In their paper, BJ analyze self-stability of voting rules that treat all the voters symmetrically. That is, BJ study the case of anonymous rules in which a reform passes if and only if the number of its supporters exceeds the threshold specified by the rule. They show that self-stable rules may not exist in this setup, and they establish conditions (on the characteristics of the society) that guarantee existence.

Our contribution in this paper is to extend the analysis of BJ to a larger class of voting rules, namely to the class of weighted majority rules. There are many examples of institutions that use rules in which different agents have different

☆ We thank two anonymous referees, an associate editor, and participants of several conferences and seminars for their comments. Kim gratefully acknowledges financial support from the National Science Council of Taiwan, grants NSC103-2410-H-002-246 and NSC104-2410-H-002-019.

* Corresponding author.
E-mail addresses: azrieli.2@osu.edu (Y. Azrieli), semn.kim@gmail.com (S. Kim).

[1] BJ also consider the case of constitutions, where a different rule F can be used for the decision between f and g. We discuss this possibility in Section 5.

http://dx.doi.org/10.1016/j.geb.2016.10.010
0899-8256/

voting weights: In the United Nations Security Council the permanent members have veto power, which effectively means that their voting power is larger than that of other members[2]; in the Council of the European Union the number of seats of each state depends on its size; and in the International Monetary Fund members' voting weights depend on the size of their economies. It is therefore important to understand what type of rules are self-stable when anonymity is not assumed.

In our formal model, a voting rule is any mapping from preference profiles over $\{Reform, Status\ quo\} = \{R, S\}$ to lotteries over this set. A weighted majority rule is a voting rule that can be described by assigning weights to the agents and setting a quota, such that R is chosen if the total weight of the agents that support R exceeds the quota, and S is chosen otherwise. Our main result (Theorem 1) is a characterization of the class of self-stable weighted majority rules. In other words, we characterize those weighted majority rules that cannot be defeated by any arbitrary voting rule.[3]

Our characterization shows that only few rules of a very particular form are self-stable. Namely, each self-stable rule partitions the society into at most three groups, where the weights of agents within each group are the same. The first group contains 'veto' players, i.e. agents that can single-handedly vote down any reform. The second group contains 'null' players with zero weight whose vote never affect the outcome. The last group contains the rest of the society, and we refer to these agents as 'normal' players. Put differently, according to a self-stable rule a Reform passes if and only if the coalition of agents who support it contains all the veto players and at least a certain number of normal players. There are additional constraints on the numbers of veto players and normal players that need to be satisfied in order for the rule to be self-stable. These constraints vary with the characteristics of the society, but in every society self-stability implies that the rule has the form described above.

An important point to notice is that the class of self-stable rules does not depend at all on the ex-ante distribution of agents' preferences. Specifically, we prove that if an incumbent rule f is defeated by a rule g under a particular distribution of preferences (hence f is not self-stable in this society), then for any other distribution of preferences it is possible to find a rule g' that defeats f in this alternative society. This property is a key step in the characterization of self-stability, and it stands in stark contrast to the situation in BJ, where the connection between self-stability and the distribution of preferences is the main research question.

Our characterization implies that self-stability is an extremely restrictive requirement, and that voting rules used in the real world almost never satisfy this criterion.[4] It is therefore not surprising that societies usually use different voting rules for everyday decisions than for decisions involving changes to the rules. Our results strengthen the conclusions of BJ, who argue that (Barberà and Jackson, 2004, page 1011) "...constitutions where the voting rule used to amend the constitution is the same as the voting rule used for ordinary business are dangerously simplistic". We therefore extend our analysis to the case of constitutions in which a rule F (possibly different than f) is used to decide whether the ordinary business rule f will be replaced. While we do not have a complete characterization of self-stability in this case, we do obtain several necessary conditions which suggest that even in this setup self-stability is quite restrictive. In particular, we show that self-stability implies that F is more conservative than f, in the sense that if a coalition T is not sufficiently large to pass a Reform according to f then T is also not sufficiently large to replace f by another rule (according to F).

We now discuss how this paper is related to previous literature. The model we use, in which agents' preferences over voting rules are endogenously determined from their assessments regarding their future preferences over alternatives, was first suggested in early papers by Rae (1969), Badger (1972), and Curtis (1972). These papers only consider anonymous voting rules with the same weight to all agents.

The theoretical investigation of weighted majority rules appears already in the seminal book of von-Neumann and Morgenstern (1944, Section 5), who are mainly interested in measures of the voting power of agents under the rule. A common scenario leading to heterogeneous voting weights is that of a representative democracy with heterogeneous district sizes. An early paper on this topic is Penrose (1946). More recent papers are Barberà and Jackson (2006) and Fleurbaey (2008), who point out the advantage of weighted majority rules from a utilitarian point of view. In our recent work (Azrieli and Kim, 2014) we show that, in a standard mechanism design setup, weighted majority rules naturally arise from considerations of efficiency and incentive compatibility.

Several papers extend the analysis of BJ's self-stability concept in various directions. The closest one to ours is Sosnowska (2002, 2003) who considers a model in which the voting weights of agents are fixed but the quota is subject to amendments. She provides examples that demonstrate the differences between the case of heterogeneous voting weights and the case of equal voting power as in BJ. Wakayama (2002) considers self-stability under the possibility that agents can abstain from voting. Kultti and Miettinen (2009) study self-stability in a model of constitutions that contain several layers of voting rules, where the voting rule in each layer is used to decide on changes to the voting rule of the previous one. The same authors consider in Kultti and Miettinen (2007) a setup with a continuum of agents and analyze stability of voting rules using the

[2] See Azrieli and Kim (2014, Section 8) for a representation of the voting rule used in the UN Security Council as a weighted majority rule.

[3] Note that there may be weighted majority rules that cannot be defeated by any other weighted majority rule, but are defeated by a rule that is not a weighted majority rule. Thus, in our analysis the set of rules that are candidates for self-stability is smaller than the set of rules that can be proposed as alternatives to the incumbent rule. We discuss this point in greater detail in the concluding section (Section 6). In particular, we argue that the gap between the two sets can be reduced significantly without affecting the results.

[4] In the UN Security Council the five permanent member states (China, France, Russia, the UK, USA) are veto players and the other ten states are normal players. To pass a resolution, the support of all veto players and at least four normal players is needed (this assumes no abstentions). However, there is no society for which this rule satisfies the constraints on the number of normal players implied by self-stability. See Theorem 1 for details.