Optimal Voting Rules

Alex Gershkov, Benny Moldovanu and Xianwen Shi*

July 31, 2016

Abstract

We derive the incentive compatible and ex-ante welfare maximizing (i.e., utilitarian) mechanism for settings with an arbitrary number of agents and alternatives where the privately informed agents have single-crossing and single-peaked preferences. The optimal outcome can be implemented by modifying a sequential voting scheme, due to Bowen (1943), and used in many legislatures and committees. The modification uses a flexible majority threshold for each of several alternatives, and allows us to replicate, via a single sequential procedure, the entire class of anonymous, unanimous and dominant strategy incentive compatible mechanisms. Our analysis relies on the elegant characterization of this class of mechanisms for single-peaked preferences by Moulin (1980) and, subsequently, for single-crossing preferences by Saporiti (2009).

*We wish to thank Dimitri Vayanos and anonymous referees for many insightful comments. We are grateful to Felix Bierbrauer, Andreas Kleiner, Ilya Segal, Martin Schmalz, Tymon Tatur, Tilman Börgers, Deniz Dizdar, Frank Rosar, Gabor Virag and various seminar participants for helpful discussion. Gershkov wishes to thank the Israel Science Foundation, the German-Israeli Foundation and the Maurice Falk Institute for Economic Research for financial support. Moldovanu wishes to thank the German Science Foundation, the German-Israeli Foundation and the European Research Council for financial support. Shi is grateful to the Social Sciences and Humanities Council of Canada for financial support. Gershkov: Department of Economics, Hebrew University of Jerusalem, Israel and School of Economics, University of Surrey, UK, alexg@huji.ac.il; Moldovanu: Department of Economics, University of Bonn, Germany, mold@uni-bonn.de; Shi: Department of Economics, University of Toronto, Canada, xianwen.shi@utoronto.ca.

1 Introduction

We derive the welfare maximizing (i.e., utilitarian) mechanism for settings with an arbitrary number of social alternatives where privately informed agents have single-crossing and single-peaked preferences. Our analysis takes into account the agents' strategic incentives in such situations.

The point of departure for our analysis is a classical "converse" to the *Median Voter Theorem* due to Moulin [1980]. Moulin's result says that, on the *full* domain of single-peaked preferences, all dominant strategy incentive compatible (DIC), Pareto efficient and anonymous mechanisms can be described as *generalized median* schemes that choose the median among the n real peaks of actual voters and additional $(n-1)$ fixed "phantom" voters' peaks.[1]

The characterization via phantoms is not very intuitive, and we first show that all generalized median schemes can be implemented by modifying the well-known *successive voting procedure* used in European legislatures (Rasch [2000]). In this voting procedure, alternatives are brought to the ballot in a pre-specified order, and at each step an alternative is either adopted (and voting stops), or eliminated from further consideration (and the next alternative is considered).[2] Usually the result at each stage is determined by a fixed, possibly qualified majority. It is well known that, under complete information, single-peaked preferences and simple majority, the "sophisticated" equilibrium outcome (reached by backward induction) of the successive voting procedures (and of many other binary voting schemes) is the Condorcet winner, i.e., the alternative preferred by the median voter.

For the case where the alternatives are clearly defined by a numerical magnitude, such as the level of a public good, an interest rate, a tax levy, or a minimum wage, the successive voting procedure is equivalent to another well-known procedure, devised by Bowen [1943] for the provision of a public good.[3] Starting from a status-quo quantity of public good, voters vote on successive increments (decrements) until no more increases (decreases) garner a sufficient majority. In our modification, the majority needed for successive increases or decreases is not constant; instead, this majority increases as the process gets further away from the status-quo. Such staggered hurdles are commonplace in the laws governing tax or expenditure increases (for financing various public goods such as education or transportation) in most U.S. states.[4] For example, the legislature of Nebraska can vote to increase property taxes reflecting changes in the Consumer Price Index by simple majority, while larger increases

[1]Several authors have extended Moulin's characterization by discarding the assumption that mechanisms can only depend on peaks: examples are Barbera, Gul and Stacchetti [1993], Sprumont [1991], Ching [1997], Schummer and Vohra [2002], and Chatterjee and Sen [2011].

[2]In Scandinavia and the Anglo Saxon world, where the alternative *amendment procedure* is used, voting occurs over pairs of alternatives, with the winner advancing to the next stage until all alternatives are exhausted. Apesteguia, Ballester and Masatlioglu [2014] offer a parallel axiomatic characterization of both procedures.

[3]Bowen's scheme has been extensively studied in the subsequent literature (see for example Chapter 14 in Green and Laffont [1979]).

[4]See Joyce et al. (1995) for a summary of local tax and expenditure limitations imposed by states.

up to 5% require a three-quarters majority. Increases above 5% require a referendum in the population. In Florida, the law sets a maximum rolled-back rate that a county can adopt to fund the county's general public hospital. A rate of up to 110% of the rolled-back rate needs a two-thirds majority of the county's governing body, while a rate in excess of 110% requires a three-fourths majority, or a referendum.[5]

As noted above, even if formal decisions are taken by a qualified majority in a legislature, additional indirect hurdles apply to various motions, according to their main properties/potential for disruption or harm to those in a minority.[6] These include quorum rules, referenda, mandatory intervening elections, double majorities in several houses of parliaments, or approval by additional public or federal bodies.[7] The prevalent use of supermajorities and other non-neutral measures suggests a utilitarian rationale where potential gains and losses are weighed against each other and where minorities get proper legal protection.[8]

The utilitarian principle uses cardinal information and dictates the choice of an alternative that is preferred by the average voter. At least since Galton's 1907 famous letter to *Nature* where he proposes the choice of the median alternative, it has been recognized that the average is not implementable if voters are strategic: each agent has an incentive to exaggerate her position in order to move the average towards her preferred outcome.[9]

Our optimization analysis reconciles these two conflicting principles—implementable median-like procedures versus non-implementable but ex-ante efficient average procedures—in the best possible way for any number of voters and alternatives: this is achieved by introducing flexible adoption thresholds that depend on the respective alternative, and optimizing over these thresholds. As a result, the optimal mechanism chooses the alternative favored by the "average" type, but the average is not the naive one: instead it is calculated from the coarse information that could be inferred solely from the agents' equilibrium voting behavior. In other words, to satisfy DIC, the mechanism needs to filter the agents' information, and then chooses the efficient alternative given this filtration.

Our main results are:

[5]See Florida Statutes (2015), http://dor.myflorida.com/dor/property/legislation/pdf/2015statutes.pdf.

[6]Public firms also use super-majority requirements when shareholders vote about major issues such as mergers. See Gompers et al. [2003] and Bebchuk et al. [2009].

[7]These bodies can be seen as "phantoms" in Moulin's analysis.

[8]We note that the Austrian, Belgian, Finnish, Dutch, German and US constitutions can be changed only if two-thirds of the parliaments' members are in favor. Three-fifths majorities for constitutional changes are used in France, Greece and Spain.

[9]Galton (1907) writes: "... (1) A jury has to assess damages. (2) The council of a society has to fix on a sum of money, suitable for some particular purpose. How can the right conclusion be reached, considering that there may be as many different estimates as there are members? That conclusion is clearly not the average of all the estimates, which would give a voting power to 'cranks' in proportion to their crankiness. One absurdly large or small estimate would leave a greater impress on the result than one of reasonable amount, and the more an estimate diverges from the bulk of the rest, the more influence would it exert. I wish to point out that the estimate to which least objection can be raised is the middlemost estimate, the number of votes that it is too high being exactly balanced by the number of votes that it is too low. Every other estimate is condemned by a majority of voters as being either too high or too low, the middlemost alone escaping this condemnation."

1. We show that, by varying the threshold requirement in the successive voting schemes, we can replicate the outcome of any anonymous, unanimous and DIC mechanism. The equilibrium notion is the very simple and robust (ex-post perfect and Markov) Nash equilibrium where, at each stage, agents use sincere strategies. Conversely, for any successive voting scheme with decreasing thresholds, there is an anonymous, unanimous and DIC mechanism that generates the same equilibrium outcome.

2. Under standard assumptions on preferences and their distribution (yielding symmetric, independent, private values), we compute the adoption thresholds that maximize ex ante expected welfare. In other words, we derive the incentive compatible optimal mechanism (second-best). We then extend this result to the correlated case with linear utilities, where types are independently drawn conditional on the realization of one out of several possible states of the world.

3. We show that, for large societies where types are independently drawn, our solution (that converges then to the first-best) can be approximated by a fixed adoption threshold that equals the statistical proportion of voters with peaks below the efficient alternative. In contrast, when types are correlated, the first-best cannot be approximated by a fixed threshold: flexible decreasing thresholds are then necessary to approximate the first-best even if the number of voters is very large.

Moulin's characterization does not directly apply to our setting because it requires a full domain of single-peaked preferences. Since the (standard) one-dimensional model of private information used here cannot generate this full domain, we rely instead on an analogous result for maximal domains of single-crossing preferences, due to Saporiti [2009].[10] To understand the logic of our results, let $m(k)$ be the number of phantom voters with peaks to the left of, and including alternative k, in a generalized median mechanism. By definition, this function is increasing. The outcome of the median mechanism with such phantom distribution can be replicated by the sincere equilibrium of the successive voting procedure among privately informed agents where the adoption threshold for alternative k is decreasing, and given by $\tau(k) = n - m(k)$.[11] The optimization task is then of combinatorial nature, to determine the optimal, decreasing function τ as a function of the agents' preferences and their distribution. Since the domain of maximization is finite, a solution always exists. The main difficulty is the analytic identification of the solution and its properties. Just to give an example, the June 1991 successive voting procedure that determined the new capital of the reunited Germany involved 658 members of parliament and 4 alternatives.[12] This yields 47 698 420 different

[10] A well-known application of a social choice framework with single-crossing preferences is voting over linear tax schedules—see for example, Roberts [1977], Romer [1975], and Meltzer and Richard [1981].

[11] Although they do not refer to the successive voting procedure, our argument is inspired by Barbera, Gul and Stacchetti's [1993] interpretation of generalized median mechanisms in terms of *coalitional systems*. See also the survey by Barbera [2001].

[12] Besides simple alternatives such as Bonn and Berlin, there were composite ones that involved different

anonymous, unanimous and incentive compatible mechanisms among which we look for the optimal one (see formula (7) below).

We also briefly sketch how our model could be used on actual data. Just to give an example, assume that income is the relevant defining "type" of voters. Then, for a large population with an income inequality measure in the observed range of Western democracies (where Gini coefficients are between 0.25 and 0.55) our approach suggests that a required majority of about two-thirds will be approximately optimal and relatively stable under changes of the underlying distribution.[13]

The rest of the paper is organized as follows. In the remainder of this Section we review the related literature. In Section 2 we describe the social choice model and the mechanism design problem. In Subsection 2.1 we illustrate the model and some implications of incentive compatibility in the simple special case where utilities are linear. In particular, we show that the welfare-maximizing rule (first-best) is not implementable although it is monotone. In Section 3 we first introduce the modification of the successive voting procedure and derive an ex-post Nash equilibrium where voters vote sincerely. Next, we prove that, for any unanimous and anonymous DIC mechanism, there exists a successive voting procedure with decreasing majority requirements that generates the same outcome, and vice versa. In Section 4 we use the equivalence result to derive the precise decreasing sequence of the majority thresholds associated with the ex-ante welfare maximizing DIC mechanism. In Section 4.1 we illustrate some of the insights, including comparative statics, for the case of linear utilities. Section 5 treats the case of decisions with large number of voters. In Section 6 we extend our results to the case of correlated types. Section 7 concludes. All omitted proofs are in Appendix A. Appendix B discusses the regularity condition used in the characterization of optimal mechanisms.

Related Literature

A large body of work has focused on the implementation of desirable social choice rules in abstract frameworks with ordinal preferences. Classical results include the Gibbard-Satterthwaite *Impossibility Theorem* (Gibbard [1973] and Satterthwaite [1975]) and the *Median Voter Theorem* for settings with single-peaked preferences (see Black [1948]). Preference intensities are not part of those models and maximization goals are not easily formulated within them.

The idea of comparing voting rules in terms of the ex-ante expected utility they generate goes back to Rae [1969].[14] That paper and almost the entire following literature focus on

locations of parliament and government.

[13] The two-thirds requirement probably stems from the rules for electing a new pope, devised by Pope Alexander III in 1179. Although only unanimity was thought to reveal the will of God, Pope Pius II summarized his own election in 1458: "What is done by two thirds of the sacred college, that is surely of the Holy Ghost, which may not be resisted" (in Gragg and Gabel, 1959: 88).

[14] A recent analysis of the median versus the mean mechanism is in Rosar [2012].

settings with two social alternatives (a reform and a status quo, say). Schmitz and Tröger [2012] identify qualified majority rules as ex-ante welfare maximizing in the class of DIC mechanisms with two alternatives—this can be seen as an implication of our main result.[15] Azrieli and Kim [2014] nicely complement this analysis for two alternatives by showing that any interim Pareto efficient, Bayesian incentive compatible (BIC) choice rule must be a qualified majority rule. Dekel and Piccione [2000], Callander [2007] and Ali and Kartik [2012] analyzed procedures where voters act sequentially, one after the other. In their settings, there are two alternatives and voters have common or interdependent values.

Drexl and Kleiner [2013] also confine attention to settings with two social alternatives and show that a principal who wishes to maximize the agents' welfare from the physical allocation minus potential transfers to outsiders will use a mechanism that does not involve any monetary transfers. In particular, for settings with two alternatives, their optimal mechanism coincides with the one derived in this paper.

The situation dramatically changes when there are three, or more alternatives: the DIC/BIC constraints and the mechanisms themselves are much more numerous and complex. Börgers and Postl [2009] study a setting with three alternatives where it is common knowledge that the top alternative for one agent is the bottom for the other, and vice-versa. The agents differ in the relative intensity of their preferences for a middle alternative (the *compromise*). In addition to a characterization of BIC mechanisms, Börgers and Postl conduct numerical simulations and show that the efficiency loss from second-best rules is often small.

Apesteguia, Ballester and Ferrer [2011] consider a general social choice model with cardinal utility. Strategic voting is not considered in their analysis—this would lead there to impossibility results—and the scoring rules that emerge as optimal in their analysis are known to be subject to strategic manipulation.

Flexible thresholds have been advocated with a clear utilitarian rationale in mind by Gersbach and Pachl [2009] in the context of the common European monetary policy: the size of the required majority should depend monotonically on the proposed change in interest rate. In this way, small shocks affecting only a few countries can be readily accommodated, while radical changes that affect the entire Euro area should only be implemented if they command a broad support. Interestingly, Riboni and Ruge-Murcia [2010] empirically show that (sincere) successive voting, augmented by a supermajority requirement, best explains the decision on interest rates by monetary committees at several major, independent central banks.

2 The Social Choice Model

We consider n agents who have to choose one out of K mutually exclusive alternatives. Let $\mathcal{K} = \{1, ..., K\}$ denote the set of alternatives. Agent $i \in \{1, ..., n\}$ has (cardinal) utility $u^k(x_i)$,

[15] These authors also perform an analysis for Bayesian mechanisms, which is not covered by our study.

where $k \in \mathcal{K}$ is the chosen alternative and where x_i is a parameter (or type) privately known to agent i only. We assume that $u^k(x_i)$ is continuous in x_i for any k. The types $x_1, ..., x_n$ are distributed on the interval $[\underline{x}, \overline{x}]^n$, $0 \leq \underline{x} < \overline{x} < \infty$, according to a commonly known, joint cumulative distribution function Ψ with density ψ having full support.[16] This is the one-dimensional, private values specification, the most common one in the vast literature on optimal mechanism design with monetary transfers. Monetary transfers, however, are not allowed here.

Agents' utilities are assumed to be single-crossing with respect to the order of alternatives $1, ..., K$. Formally, for any two alternatives k and l with $k < l$, we assume that there exists a unique cutoff type $x^{k,l}$ with $u^l(x^{k,l}) = u^k(x^{k,l})$ such that[17]

$$\begin{cases} u^k(x_i) > u^l(x_i) & \text{if} \quad x_i < x^{k,l} \\ u^k(x_i) < u^l(x_i) & \text{if} \quad x_i > x^{k,l} \end{cases}. \tag{1}$$

To simplify notation, we denote $x^k \equiv x^{k-1,k}$. We further assume that each alternative is the top alternative for some type of the agents.[18] That is, for any $k \in \mathcal{K}$, there exists $x_i \in [\underline{x}, \overline{x}]$ such that

$$u^k(x_i) > \max_{l \in \mathcal{K}, l \neq k} u^l(x_i). \tag{2}$$

We shall focus on the case of a utilitarian planner whose objective is to maximize the sum of the agents' expected utilities

$$\max_{k \in \mathcal{K}} E\left[\sum\nolimits_i u^k(x_i)\right].$$

Remark 1 *The single-crossing property (1), together with assumption (2), implies that the cutoffs x^k are well-ordered:*

$$\underline{x} \equiv x^1 < ... < x^K < x^{K+1} \equiv \overline{x}. \tag{3}$$

To see this, we note that, by definition of x^k and the single-crossing property (1),

$$u^k(x_i) < u^{k-1}(x_i) \text{ for all } x_i < x^k.$$

Similarly, by definition of x^{k+1} and the single-crossing property (1), we have

$$u^k(x_i) < u^{k+1}(x_i) \text{ for all } x_i > x^{k+1}.$$

If $x^k \geq x^{k+1}$, any type x_i satisfies either $x_i \leq x^k$ or $x_i \geq x^{k+1}$, and thus alternative k is (weakly) dominated either by alternative $k-1$ or by alternative $k+1$, which contradicts (2).

[16] We can allow $\overline{x} = \infty$ as long as $u^k(x_i)$ is bounded for all k.

[17] We also assume that the indifference types $x^{k,l}$ are different across pairs, which is a generic assumption.

[18] This assumption ensures that our single-crossing preferences are also single-peaked (see Remark 1). However, it rules out the setting of Börgers and Postl [2009] where the third alternative, *compromise*, is not the top alternative of any agent.

Therefore, we must have $x^k < x^{k+1}$ for all $k \in \mathcal{K}$, which proves (3). By the definition of x^k and by (3), agents with type x_i have k as their top alternative if and only if $x_i \in \left[x^k, x^{k+1}\right]$.

Note also that the agents' preferences are single-peaked. To see this, consider agent i with type $x_i \in \left(x^k, x^{k+1}\right)$. By definition of x^k, agent i prefers alternative k to any alternative $l < k$, and by definition of x^{k+1}, agent i prefers k over any $l > k$. Consider two alternatives l and m with $l < m < k$. Since $x^l < x^m < x^k$, we have $x_i > x^{l,m}$ and agent i prefers m to l. Similarly, agent i prefers m to l if $k < m < l$. Therefore, agent i's preferences are single-peaked. On the other hand, our preference domain is a strict subset of the full single-peaked preference domain whenever $K > 3$: not all single-peaked preferences are compatible with our environment (see below an explicit illustration in the linear environment).

A deterministic, direct mechanism asks agents to report their types, and, for any profile of reports, the mechanism chooses one alternative from $\mathcal{K}$. Formally, a deterministic direct mechanism is a function $g : [\underline{x}, \overline{x}]^n \to \mathcal{K} = \{1, ..., K\}$. A deterministic mechanism is *dominant strategy incentive compatible* (DIC) if for any player i and for any x_i, x_i' and x_{-i}:

$$u^{g(x_i, x_{-i})}(x_i) \geq u^{g\left(x_i', x_{-i}\right)}(x_i). \tag{4}$$

It is clear from the above definition that two types that have the same ordinal preferences must be treated in the same way by a DIC mechanism. Thus, an implication of the lack of monetary transfers is that deterministic DIC mechanisms cannot depend on preference intensities.

2.1 An Illustration: Linear Preferences

Suppose the utilities are linear: $u^k(x_i) = a_k + b_k x_i$. These preferences are necessarily single-crossing. We assume that $b_K > b_{K-1} > ... > b_1 \geq 0$ and $a_1 > a_2 > ... > a_K$. The cutoff type who is indifferent between two adjacent alternatives k and $k-1$ is given by

$$x^k \equiv x^{k-1,k} = \frac{a_{k-1} - a_k}{b_k - b_{k-1}}. \tag{5}$$

We impose further restrictions on b_k and a_k so that our previous assumption (2) is satisfied.[19] These restrictions, together with the definition of $x^{k,l}$, imply that $x^{k,l} \in (x^{k+1}, x^l)$ for $l > k+1$, because

$$x^{k,l} = \frac{a_k - a_l}{b_l - b_k} = \frac{(a_k - a_{k+1}) + ... + (a_{l-1} - a_l)}{(b_{k+1} - b_k) + ... + (b_l - b_{l-1})}.$$

Similarly to the general case, we assume that $x^{k,l}$ are different across pairs.

Our preference domain is a strict subset of the full single-peaked preference domain. Indeed, consider a setting with 4 different alternatives $(1, 2, 3 \text{ and } 4)$ with $x^{1,4} \in \left(x^{1,2}, x^{3,4}\right)$.

[19] That is, we assume that for all $k \geq 2$,

$$\frac{a_{k-1} - a_k}{b_k - b_{k-1}} < \frac{a_k - a_{k+1}}{b_{k+1} - b_k},$$

so cutoffs are ordered according to (3).

If $x^{1,4} \in \left(x^{2,3}, x^{3,4}\right)$, as shown in Figure 1, then the feasible single-peaked preferences that have alternative 2 on their top are $2 \succ 1 \succ 3 \succ 4$ and $2 \succ 3 \succ 1 \succ 4$. In particular, the preference $2 \succ 3 \succ 4 \succ 1$ is not compatible with the linear environment. Similarly, if $x^{1,4} \in \left(x^{1,2}, x^{2,3}\right)$, the feasible single-peaked preferences that have alternative 3 on their top are $3 \succ 2 \succ 4 \succ 1$ and $3 \succ 4 \succ 2 \succ 1$. Here the preference profile $3 \succ 2 \succ 1 \succ 4$ is not compatible with our structure.

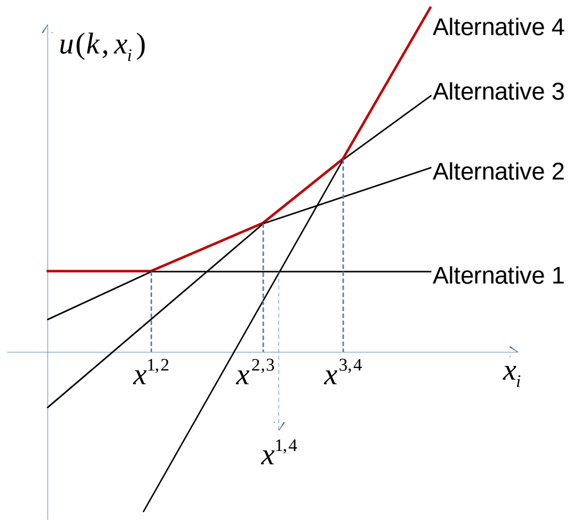

Figure 1: Not all single-peaked preferences are compatible with our linear structure.

Analogously to the classical framework with monetary transfers, a mechanism $g(x_i, x_{-i})$ is DIC if and only if (i) for all x_{-i} and for all i, $g(x_i, x_{-i})$ is increasing in x_i; and (ii) for any agent i, any $x_i \in [\underline{x}, \overline{x}]$ and $x_{-i} \in [\underline{x}, \overline{x}]^{n-1}$, the following envelope condition holds

$$u^{g(x_i, x_{-i})}(x_i) = u^{g(0, x_{-i})}(0) + \int_0^{x_i} b_{g(z, x_{-i})} dz. \tag{6}$$

When monetary transfers are feasible, any monotone decision rule $g(x_i, x_{-i})$ is incentive compatible since it is always possible to augment it with a transfer such that the equality required by (6) holds. Thus, with transfers, only monotonicity really matters for DIC. If monetary transfers were available, the welfare-maximizing allocation would be implementable via the well-known Vickrey-Clarke-Groves mechanisms. But, without monetary transfers, not all monotone decision rules $g(x_i, x_{-i})$ are implementable, and in particular, the welfare maximizing allocation need not be incentive compatible although it is monotone. This well-known phenomenon is illustrated in the next example.

Example 1 (First-best Rule Not Implementable) *Consider the linear environment with two alternatives $\{1, 2\}$, two agents $\{i, -i\}$ and $[\underline{x}, \overline{x}] = [0, 1]$. The designer is indifferent between alternatives 1 and 2 if*

$$2a_1 + b_1(x_i + x_{-i}) = 2a_2 + b_2(x_i + x_{-i}).$$

The first-best rule conditions on the value of the average type, and is given by

$$g(x_i, x_{-i}) = \begin{cases} 1 & \text{if } \frac{1}{2}(x_i + x_{-i}) \in [0, x^2) \\ 2 & \text{if } \frac{1}{2}(x_i + x_{-i}) \in [x^2, 1] \end{cases}$$

where the cutoff x^2 is defined in (5): $x^2 \equiv (a_1 - a_2) / (b_2 - b_1)$. The first-best rule is increasing in both x_i and x_{-i}. Yet, it is not implementable. To see this, consider the realization of types $(x_i, x_{-i}) = (x^2 + \epsilon, x^2 - 2\epsilon)$ where $\epsilon > 0$ is small. Then $\frac{1}{2}(x_i + x_{-i}) < x^2$, and hence $g(x_i, x_{-i}) = 1$, which is not optimal for player i. By misreporting any type above $x^2 + 2\epsilon$ player i can shift the decision to his preferred alternative 2. Therefore, this mechanism is not DIC.

3 The Successive Voting Procedure

In this procedure, alternatives are first arranged in some pre-determined voting order, say $1, 2, ..., K$. The first ballot determines whether there is a (qualified) majority for alternative 1. If so, alternative 1 is adopted and voting ends. If alternative 1 fails to command a majority, this alternative is removed from future consideration, and the parliament proceeds to vote on alternative 2. If a majority supports alternative 2, alternative 2 is adopted; otherwise, the agents proceeds to vote on alternative 3. Voting continues until one alternative gains majority. If no alternative gains majority in earlier stages, the last two alternatives $K-1$ and K are paired and the one with majority support is adopted. In most cases, the required majority for adoption is the same across alternatives, and the voting order is either suggested by the agenda setter or is pre-determined by custom. Without single-peaked preferences, the voting outcome (even under complete information) is sensitive to the voting order.[20]

In order to link the successive procedure to unanimous, anonymous and DIC mechanisms, we consider a modified successive procedure with two properties. First, the order of vote is according to the natural order $(1, 2, ..., K)$ under which the preferences are single-peaked (see Remark 1).[21] Second, the required majority for adoption is no longer kept constant across alternatives: instead, the adoption threshold for choosing alternative k, $\tau(k)$, is a decreasing function. That is, a more stringent majority requirement (which may be more or less than simple majority) is set for earlier alternatives, while a lower majority is required for adopting later alternatives. Equivalently put, it becomes increasingly difficult to keep the voting process in motion.

As mentioned in the Introduction, the modified successive procedure can also be seen as a simple variation of the classical Bowen's scheme for public good provision. There voters vote on successive increments (or decrements) to the status quo — the quantity of public good already provided. If a simple majority of voters are against the first increment, voting

[20] If voters vote sincerely, later alternatives have better chance to be adopted (Black [1958]), but if voters vote strategically earlier alternatives are more likely to be adopted (Farquharson [1969]).

[21] Alternatively, the successive voting procedure can be run in reverse order.

stops and the status quo is adopted. If the first increment gains a majority support, voting continues to another increment. If the second increment fails to gain a simple majority support, the voting stops and the first increment is implemented, otherwise, voters are asked to vote yet another increment, and so on. Therefore, if the underlying collective decision is numerical in nature and has a clear direction of change, the natural order of alternatives of our modified procedure is the same as in Bowen's scheme, but we have a *declining* threshold $\tau(k)$ for adopting alternative k, which corresponds to an *increasing* threshold for voting on further increments in Bowen's original formulation.

The results in this section are purely ordinal, and thus do not depend on the particular cardinal specification of utility (nor on the distribution of types) as long as the single-crossing assumption is satisfied and the domain of preferences is maximal with respect to single-crossing.[22]

Definition 1 1. *A voting strategy for an agent is sincere if, at each stage, the agent votes in favor of the respective alternative if and only if it is the best (among the remaining alternatives) given his preferences.*

2. *A voting strategy for an agent is monotone if it consists of a series of "No" in early stages (possible none), followed by a series of "Yes" in all later stages.*

Note that, with single-peaked preferences and with our natural voting order, a sincere strategy has a particular structure: the agent votes "No" for all alternatives that appear on the ballot before his most preferred one (he wants the voting to continue), and then votes "Yes" for his peak alternative and for all successive ones (he wants the voting to stop). Hence, under the successive voting rule with the natural order, monotone voting is a generalization of sincere voting.[23] The next result holds for the entire domain of single-peaked preferences.[24]

Proposition 1 *Consider the successive procedure with a decreasing threshold function $\tau(k)$, and assume that all agents except agent i use monotone voting strategies. Then, the sincere voting strategy is optimal for agent i. In particular, the strategy profile where all agents vote sincerely constitutes an ex-post perfect Nash equilibrium.*[25]

Proof. Assume that all agents other than i use monotone strategies, and let the peak of agent i be on alternative k. Consider first an alternative $k' < k$. The sincere voting strategy

[22] A domain of preferences is *maximal* with respect to single-crossing if one cannot add to it another ordinal preference profile without violating the single crossing property (see Saporiti [2009] for a formal definition).

[23] Monotone strategies are also *Markov*, i.e., they do not condition on the history of votes before the current one.

[24] The amendment procedure where alternatives are voted one against the other in the natural order (or its reverse) and the winner advances to the next stage given some (flexible) qualified majority need not possess an ex-post, sincere equilibrium.

[25] An n-tuple strategy profile is said to constitute an ex-post perfect equilibrium if at every stage, and for every realization of private information, the n-tuple of continuation strategies constitutes a Nash equilibrium of the subgame in which the realization of the agents' types is common knowledge.

calls for i to vote against k'. A deviation from sincere voting matters only if, by changing his strategy from "No" to "Yes" at this stage, alternative k' is chosen whereas it would not be chosen if i voted sincerely. But in this case, the number of "Yes" votes for alternative k' must be $\tau(k') - 1$. Since τ is decreasing, and since all other agents use monotone strategies, voting "No" on alternative k' implies that, in this case, agent i can ensure that the chosen alternative k'' satisfies $k' < k'' \leq k$ (either voting stops before reaching k, or i ensures the choice of k by voting "Yes" on it, and then joining at least $\tau(k') - 1 \geq \tau(k) - 1$ "Yes" votes). All alternatives k'' with $k' < k'' \leq k$ are preferred by i to k', so this deviation from sincere voting is not beneficial. Consider now $k' \geq k$. The sincere strategy calls for i to vote "Yes" at the relevant stage. By deviating to "No", the chosen alternative must satisfy $k'' \geq k'$. All these alternatives are dominated by k' from i's point of view, so a deviation is again not beneficial. This completes the proof of optimality of the sincere voting strategy for agent i. Since the argument applies to all agents, and since sincere voting is monotone, sincere voting constitutes an ex-post perfect Nash equilibrium.[26] ■

Remark 2 *The above result applies, as stated, for the case where the results of previous voting stages is not revealed: the only possible inference at a particular stage is that no earlier alternative has obtained a necessary majority. In practical applications one needs to consider more permissive information disclosure policies, such as revealing the margin of past decisions, or even the individual voting records. By the robust nature of the ex-post equilibrium, where agents do not regret their strategies even if all private information is revealed ex-post, sincere voting after each history remains an ex-post perfect equilibrium for **any** disclosure policy. But, disclosing more information adds more strategies and potential equilibria. Nevertheless, we show in Lemma 1 in Appendix A that, irrespective of the disclosure policy, sincere voting is the unique outcome that survives iterated elimination of (weakly) dominated strategies. This gives another strong rationale for the simplest, sincere equilibrium. The reader may also notice the close parallels to the implementation, via a dynamic auction, of the Vickrey-Clarke-Groves mechanism due to Ausubel [2004].*

We now uncover the connection between the outcome of any DIC mechanisms and the sincere equilibria of the successive procedures with decreasing thresholds. An influential paper by Moulin [1980] shows that, if each agent is restricted to report his top alternative only, then every DIC, Pareto efficient and anonymous voting scheme on the full domain of single-peaked preferences is equivalent to a generalized median voter scheme that is obtained by adding $(n-1)$ fixed peaks (phantoms) to the n voters' reported peaks and then choosing the median of this larger set of peaks.

[26] Bowen's procedure for public good provision is analyzed, under incomplete information, by Green and Laffont [1979]. They erroneously claim (Theorem 14.2) that sincere voting constitutes an equilibrium in dominant strategies. This holds only if agents are a-priori restricted to play monotone strategies, as defined here.

Moulin's characterization also holds in our setting although agents are not restricted to report only their peaks, and although the domain of preferences is a strict subset of the full domain of single-peaked preferences. The relevant result is due to Saporiti [2009]: he provides a characterization of unanimous, anonymous and DIC mechanisms for maximal domains of single-crossing preferences, in a spirit similar to Moulin [1980].

To get some intuition about Moulin's characterization, consider a mechanism that always picks the top alternative of the agent with the lowest type (an analogous intuition works for other order statistics). Such a mechanism is clearly DIC, Pareto efficient and anonymous. It can be replicated by a generalized median that places all $(n-1)$ phantoms at alternative 1, because the median of the n real votes and the $(n-1)$ phantoms is always the top alternative of the lowest type agent. The number of phantoms cannot be n because then generalized median is not uniquely defined. It cannot be $n+1$ or higher, because the resulting generalized median may not be Pareto efficient. For example, if there are $(n+1)$ phantoms all placed at alternative 1, then alternative 1 is the generalized median, which may not be Pareto efficient.[27]

To formally state the connection between DIC mechanisms and the successive procedures, we need several definitions:

Definition 2 *1. A mechanism g is unanimous if $x_i \in \left(x^k, x^{k+1}\right)$ for all i implies $g(x) = k$.*

2. A mechanism g is Pareto efficient if, for any profile of reports $(x_i, x_{-i}) \in [\underline{x}, \overline{x}]^n$, there is no alternative $k \in \mathcal{K}$ such that $u^k(x_i) \geq u^{g(x)}(x_i)$ for all i, with strict inequality for at least one agent.

3. A mechanism g is anonymous if, for any profile of reports $(x_i, x_{-i}) \in [\underline{x}, \overline{x}]^n$, $g(x_1, ..., x_n) = g\left(x_{\sigma(1)}, ..., x_{\sigma(n)}\right)$ where σ denotes any permutation of the set $\{1, ..., n\}$.

It is clear that a Pareto-efficient mechanism is unanimous. In the presence of dominant strategy incentive compatibility, an anonymous and unanimous mechanism is also Pareto efficient (Corollary 1 in Saporiti [2009]). We are now ready to state our first main result.

Theorem 1 *1. For any unanimous and anonymous DIC mechanism g, there exists a decreasing threshold function $\tau^g(k)$ with $\tau^g(k) \leq n$ for all $k \in \mathcal{K}$ and $\tau^g(K) = 1$ such that, for any realization of types, the outcome of g coincides with the outcome in the sincere equilibrium of successive procedure with thresholds $\tau^g(k)$.*

2. Conversely, for any decreasing threshold function $\tau(k)$ with $\tau(k) \leq n$ for all $k \in \mathcal{K}$ and $\tau(K) = 1$, there exists an anonymous, unanimous and DIC mechanism g^τ such that,

[27] Pareto efficiency implies that the chosen alternative must be between the top alternatives of the agents with the lowest and the highest type). Moulin shows that all DIC and anonymous (not necessarily Pareto efficient) mechanisms can be replicated by a generalized median with $(n+1)$ phantoms.

for any realization of types, the outcome of g^τ coincides with the outcome of the sincere equilibrium in successive procedure with thresholds $\tau(k)$.

Proof. 1. Our underlying domain of ordinal preferences is *maximal* with respect to single-crossing. To see this, note that every two alternatives, say k and l, induce a cutoff $x^{k,l}$, and each cutoff $x^{k,l}$ divides the set of types into two intervals where ordinal preferences differ with respect to the ordering of alternative k and l.[28] Since each alternative is top for some types, the interval of types is thus partitioned into $K(K-1)/2+1$ parts, each corresponding to a distinct ordinal preference. But this is also the maximum number of ordinal profiles in a maximal domain of single-crossing preferences on K alternatives.

Saporiti [2009] shows that, on a maximal domain of (ordinal) single-crossing preferences any anonymous, unanimous and DIC mechanism in an environment with n voters can be obtained as a generalized median voter mechanism with $n-1$ phantom voters.[29] In the dominant strategy equilibrium of such a generalized median voter scheme, all n voters truthfully report their top alternatives, and the outcome is the median of the n real peaks and the $n-1$ fixed phantom peaks.

Let $\ell_k \geq 0$ denote the number of phantom voters with peak on alternative k in the generalized median voter scheme corresponding to a DIC mechanism g. To construct an equivalent successive voting scheme, we define the thresholds $\tau^g(k) \equiv n - \sum_{m=1}^{k} \ell_m$, and note that $\tau^g(k)$ is decreasing and that $\tau^g(K) = 1$.

Alternative 1 is the generalized median only if the number of (real) agents who report this alternative as their top alternative exceeds $n - \ell_1 = \tau^g(1)$. Alternative 2 is the generalized median if the number of agents who report a peak on alternative 1 is less than $n - \ell_1$ and if the number of the agents who report either alternative 1 or alternative 2 as their top alternative is at least $n - \ell_1 - \ell_2 = \tau^g(2)$. In general, alternative k is the generalized median if, for any $k' < k$, the number of reported peaks on alternatives $1, 2, ..., k'$ was strictly less than $\tau^g(k')$ and if the number of agents who report their peak on alternative k or lower is at least $n - \sum_{m=1}^{k} \ell_m = \tau^g(k)$. Otherwise, alternative K is the generalized median. With this interpretation, it is now clear that the outcome of the sincere equilibrium under successive voting with threshold $\tau^g(k)$ coincides with the outcome of mechanism g.

2. Conversely, for a given successive procedure with decreasing cutoffs $\tau(k)$ such that $\tau(k) \leq n$ for any $k \in \mathcal{K}$ and $\tau(K) = 1$, we can define $\ell_1 \equiv n - \tau(1)$, and $\ell_k \equiv \tau(k-1) - \tau(k)$ for $k \geq 2$. Since $\tau(k)$ is decreasing, $\tau(k) \leq n$ for all $k \in \mathcal{K}$ and $\tau(K) = 1$, we have $\ell_k \geq 0$ for all $k \in \mathcal{K}$ and $\sum_{k=1}^{K} \ell_k = n - \tau(K) = n - 1$. The constructed phantom distribution $\{\ell_k\}$ is part of a generalized median voter scheme which corresponds to some unanimous and anonymous DIC mechanism g^τ. Moreover, it is easy to verify that the outcome of mechanism g^τ is the same as the sincere equilibrium outcome in successive voting with threshold $\tau(k)$. ■

[28] Recall the (generic) assumption that all $x^{k,l}$ are distinct.

[29] See Theorem 3 in the Appendix A for a formal statement of Saporiti's characterization.

The above theorem implies that the search for optimal mechanisms within the class of successive procedures with decreasing cutoffs $\tau(k)$ is without loss of generality. With many agents and alternatives, this remains a rather complex discrete optimization problem since there are many decreasing sequences $\tau(k)$: the number of DIC, anonymous and unanimous mechanisms for n agents and K alternatives is given by[30]

$$\frac{(n+K-2)!}{(K-1)!\,(n-1)!}. \tag{7}$$

We conclude this Section with a simple illustration of the welfare benefit of having flexible thresholds.

Example 2 *There are three alternatives, denoted by* $1,2,3$, *and preferences are single-peaked with respect to the order:* $1,2,3$. *Assume for simplicity and concreteness that all feasible ordinal single-peaked rankings occur with the same probability. Note that alternative* 2 *is never ranked as the bottom alternative in such a scenario.*[31] *Endow each agent with a simple cardinal preference where the top alternative yields utility* $\epsilon > 0$, *the middle alternative yields utility* 0, *and the bottom alternative yields utility* $\eta < 0$ *with* $|\eta| \gg \epsilon$.[32] *Whenever some agents rank alternative* 1 *at the top and other agents rank alternative* 3 *at the top, the only way to avoid a substantial utility loss is to choose alternative* 2. *This outcome cannot be achieved by any fixed threshold policy since, for any such policy (for example simple majority with* $\tau(k) = (n+1)/2$, $k = 1,2,3$), *there is a positive probability of choosing either alternative* 1 *or alternative* 3. *In contrast, the successive procedure with the decreasing threshold function* $\tau(1) = n$, $\tau(2) = \tau(3) = 1$ *always generates positive utility since: 1) it chooses alternatives* 1 *and* 3 *only when there is unanimity in their favor (yielding welfare* $n\epsilon > 0$), *and chooses alternative* 2 *otherwise (yielding welfare* $n_2\epsilon > 0$, *where* n_2 *is the number of agents who rank* 2 *at the top).*

4 The Optimal Mechanism

We now characterize the welfare maximizing allocations that respect the incentive constraints (constrained efficiency, or "second-best"). We first introduce two assumptions that put more structure on the optimization problem, allowing us to solve it analytically.

Assumption A *Agents' signals are distributed identically and independently of each other on the interval* $[\underline{x}, \overline{x}]$ *according to a cumulative distribution* F *with density* f.

[30] The problem is to partition $(n-1)$ phantoms into K alternatives, which can be represented by $(K-1)$ bars placed among $(n-1)$ balls. Hence, it is equivalent to choosing $(K-1)$ out of $(n+K-2)$ positions to place $(K-1)$ bars.

[31] This insight is more general and it applies irrespective of the number of alternatives: alternatives that are not extreme in the linear order determining single-peakedness cannot be ranked at the bottom of the preference list.

[32] We can approximate these utilities through our continuous, type-dependent function $u^k(x_i)$.

This assumption yields the standard symmetric, independent private values model (SIPV) that is widely used in the literature on trading mechanisms with transfers (where utility is usually linear). We need another joint requirement on the utility functions and on the distribution of types, ensuring that the optimal threshold function τ^*—which necessarily exists and is monotone—is identified by the necessary first order conditions, and thus amenable to analysis. We first need some notation. Let us define, for all $k \geq 2$ and $l \geq 1$,

$$u^l_{x<x^k} = E\left[u^l\left(x\right)|x < x^k\right]$$

as the expected utility from alternative l, conditional on the agent's type x being lower than the cutoff x^k. Similarly, we define

$$u^l_{x>x^k} = E\left[u^l\left(x\right)|x > x^k\right]$$

as the expected utility from alternative l conditional on the agent's type x being higher than the cutoff x^k. With single-crossing preferences, the entire (convex) interval of types below (above) x^k prefer alternative $k-1$ to k (alternative k to $k-1$). Finally, let us define

$$\beta\left(k\right) = \frac{\left(u^k_{x>x^k} - u^{k-1}_{x>x^k}\right)}{\left(u^{k-1}_{x<x^k} - u^k_{x<x^k}\right) + \left(u^k_{x>x^k} - u^{k-1}_{x>x^k}\right)}, \; k \geq 2. \tag{8}$$

By the definition of x^k and by the single-crossing property, $u^{k-1}_{x<x^k} > u^k_{x<x^k}$ and $u^k_{x>x^k} > u^{k-1}_{x>x^k}$. Therefore, $\beta\left(k\right) \in (0,1)$ for all $k \geq 2$.

Assumption B *The function β is decreasing.*

The function β plays a crucial rule in our analysis and is intimately related to the optimal threshold function τ^*. In order to better understand its definition, we can rewrite (8) as

$$\beta\left(k\right)\left(u^{k-1}_{x<x^k} - u^k_{x<x^k}\right) + \left[1-\beta\left(k\right)\right]\left(u^{k-1}_{x>x^k} - u^k_{x>x^k}\right) = 0.$$

Suppose that the chosen alternative changes from k to $k-1$. The expected gain for an agent with type below x^k is $u^{k-1}_{x<x^k} - u^k_{x<x^k}$, while the expected loss for an agent with type above x^k is $u^{k-1}_{x>x^k} - u^k_{x>x^k}$. The function β is defined such that the expected gain weighted by $\beta\left(k\right)$ and the expected loss weighted by $1-\beta\left(k\right)$ cancel out.

In Appendix B we derive sufficient conditions on the primitives of the social choice model (utility functions and the distribution of types) for the above assumption to hold. In Section 4.1, where we assume linear utility, we show how it reduces to simple and well-known conditions on the distribution function only.

Consider then an environment with n voters, and let τ^* be the optimal threshold function in the successive procedure. The analysis is based on the following simple observation. Fix any alternative k with $\tau^*\left(k-1\right) > \tau^*\left(k\right)$. If $\tau^*\left(k-1\right)$ and $\tau^*\left(k\right)$ are part of the optimal voting procedure, then increasing $\tau^*\left(k\right)$ by 1 or decreasing $\tau^*\left(k-1\right)$ by 1 should weakly reduce the total expected utility.[33] For instance, increasing $\tau^*(k)$ by 1 (while keeping $\tau^*(k')$ with $k' \neq k$ unchanged) has an impact only if it changes the chosen alternative. The proposed deviation will change the chosen alternative only if there are exactly $\tau^*(k)$ voters with values below x^{k+1}. These arguments generate the following bounds on the threshold function τ^*:

$$\begin{aligned} \tau^*(k-1) &\leq n\beta(k)+1, \text{ for all } k \geq 2, && (9) \\ \tau^*(k) &\geq n\beta(k+1), \text{ for all } k \leq K-1. && (10) \end{aligned}$$

The proposed deviation, however, is not feasible for alternative k with $\tau^*(k-1) = \tau^*(k)$, because the cutoff function $\tau^*(k)$ must be weakly decreasing. It turns out that, under Assumption B, the two derived bounds (9) and (10) remain valid also for alternatives k with $\tau^*(k-1) = \tau^*(k)$ (see Lemma 2 in Appendix A). Since $\tau^*(k)$ has to be integer, the above two bounds lead to an essentially unique threshold function.

Theorem 2 *Let $\lceil z \rceil$ denote the smallest integer that is above z. Under Assumptions A and B, the sincere equilibrium of the successive procedure with the threshold function*

$$\tau^*(k) = \begin{cases} \lceil n\beta(k+1) \rceil & \text{if } k < K \\ 1 & \text{if } k = K \end{cases}$$

implements the optimal anonymous, unanimous and DIC mechanism.

Proof. See Appendix A. ■

The above theorem reveals that adding or eliminating an alternative has only a local effect. That is, adding an alternative k_1 such that an interval $[x^k, x^{k+1}]$ is further divided into $[x^k, x^{k_1}]$ and $[x^{k_1}, x^{k+1}]$ changes only the threshold of alternative k. Similarly, the elimination of an alternative k changes the threshold of alternative $k-1$ only, without any effect on the other alternatives. This "locality-effect" follows from the single-peaked preferences: the social planner does not want to change the chosen alternative if the peak of the median voter does not change as a result of adding/eliminating alternatives.

The following corollary characterizes the optimal voting rule for the case of two alternatives by specifying the optimal qualified majority rule (or super-majority).[34] Note that Assumption B is not needed for the case of two alternatives.

Corollary 1 *Suppose there are n agents and only two alternatives, $K = 2$. Under Assumption A, the optimal rule is implemented through the sequential procedure where alternative* 1 *is chosen if and only if at least $\tau^*(1) = \lceil n\beta(2) \rceil$ voters voted in its favour.*

[33] In the language of phantoms, increasing $\tau^*(k)$ by 1 while keeping other cutoffs unchanged corresponds to moving one phantom voter from alternative k to alternative $k+1$ in the generalized median voter scheme. Similarly, decreasing $\tau^*(k-1)$ by 1 while keeping other cutoffs unchanged is equivalent to shifting one phantom from alternative k to alternative $k-1$.

[34] See Nehring [2004], Barbera and Jackson [1994] and Schmitz and Tröger [2012] for related results.

4.1 The Linear Case

We now illustrate our characterization of optimal mechanisms in the linear environment set out in Section 2.1. For this environment, we first derive a more intuitive assumption to replace Assumption B. Let X be the random variable governing the agents' type. We first recall two well-known concepts from the theory of reliability:

Definition 3 *1. The mean residual life (MRL) of a random variable $X \in [\underline{x}, \overline{x}]$ is defined as*

$$MRL(x) = \begin{cases} E[X - x | X > x] & \text{if} \quad x < \overline{x} \\ 0 & \text{if} \quad x = \overline{x} \end{cases}$$

A random variable X satisfies the decreasing mean residual life (DMRL) property if the function $MRL(x)$ is decreasing in x. This is equivalent to requiring that the function $\int_x^{\overline{x}} (1 - F(t))dt$ is log-convex where F is the CDF of X.

2. The reversed mean residual life (RMRL) of a random variable $X \in [\underline{x}, \overline{x}]$ is defined as

$$RMRL(x) = \begin{cases} E[x - X | X < x] & \text{if} \quad x > \underline{x} \\ 0 & \text{if} \quad x = \underline{x} \end{cases}$$

A random variable X satisfies the increasing reversed mean residual life (IRMRL) property if the function $RMRL(x)$ is increasing in x. This is equivalent to requiring that the function $\int_{\underline{x}}^{x} F(t)dt$ is log-concave where F is the CDF of X.

If we let X denote the life-time of a component, then $MRL(x)$ measures the mean remaining life of a component that has survived until time x: intuitively this should decrease as the component ages. The function $RMRL(x)$ measures the mean time since the failure of a component that has already failed by time x: intuitively this should increase as x increases. The DMRL and IRMRL properties hold for a large, non-parametric, class of distributions (in fact a lifetime with *decreasing* RMRL does not even exist on an unbounded interval). A simple sufficient condition for both properties to hold is the log-concavity of the density function f.[35]

In the linear setting, the function β that determines the optimal thresholds becomes

$$\beta(k) = \frac{E\left[X | X > x^k\right] - x^k}{E[X | X > x^k] - E[X | X < x^k]} = \frac{1}{1 + \frac{E\left[x^k - X | X < x^k\right]}{E\left[X - x^k | X > x^k\right]}}, \tag{11}$$

and therefore a sufficient condition for β to decrease is:

[35] The log-concavity of density is stronger than (and implies) increasing failure rate (IFR) which is equivalent to log-concavity of the reliability function $(1 - F)$. Moreover, it implies the logconcavity of F and $\int_{\underline{x}}^{x} F(t)dt$ which is equivalent to IRMRL. The family of log-concave densities is large and includes many commonly used distributions such as uniform, normal, exponential, logistic, extreme value etc. The power function distribution $(F(x) = (x)^s)$ has log-concave density if $s \geq 1$, but it does not if $s < 1$. However, one can easily verify that the two properties in Assumption B' still hold for $F(x) = (x)^s$ even with $s < 1$. Therefore, a log-concave density is not necessary. See Bagnoli and Bergstrom [2005] for an excellent discussion of these implications.

Assumption B' *The random variable X governing the distribution of types has both DMRL and IRMRL properties.*

Corollary 2 *Suppose utilities are linear and Assumptions A and B' hold. The sincere equilibrium of the successive procedure with the threshold function*

$$\tau^*(k) = \begin{cases} \lceil n\beta\,(k+1) \rceil & \text{if } k < K \\ 1 & \text{if } k = K \end{cases}$$

implements the optimal anonymous, unanimous and DIC mechanism.

To get a better intuition for the characterization, consider the optimal threshold $\tau^*(k)$ for alternative k. Ignoring the integer problem, the threshold is chosen such that, in case of pivotality, the inferred average type (given the information revealed by this event) is exactly the cutoff type x^{k+1} who is indifferent among alternatives k and $k+1$. To see this, recall that in case of pivotality there are exactly $\tau^*(k) = n\beta\,(k+1)$ agents with types in the interval $\left[\underline{x}, x^{k+1}\right]$, while all the remaining $n - n\beta\,(k+1)$ agents have types in the interval $\left[x^{k+1}, \overline{x}\right]$. Given this information, it follows from the definition of β (see equation 11) that the inferred average type is

$$\beta\,(k+1)\,E\left[X|X < x^{k+1}\right] + (1 - \beta\,(k+1))\,E\left[X|X > x^{k+1}\right] = x^{k+1}. \tag{12}$$

We can obtain immediate and intuitive comparative statics with respect to parameters of the linear utility function $\{a_k, b_k\}_{k=1}^{K}$. By the definition of x^k, increases in either a_k or b_k decrease x^k and increase x^{k+1}, which in turn leads to a higher threshold $\tau^*(k-1)$ for adopting alternative $k-1$, and a lower threshold $\tau^*(k)$ for adopting alternative k. That is, if the attractiveness of any alternative increases, the chances of adopting that alternative increase as well.

Our next proposition shows how the entire optimal threshold function τ^* changes with respect to the distribution of types. It uses the following well known stochastic orders (see Shaked and Shanthikumar [2007]). Let $\leq_{st}$ denote the standard first order stochastic dominance relation.

Definition 4

1. *A random variable Y dominates a random variable X in the hazard rate order (denoted as $X \leq_{hr} Y$) if $[X|X > x] \leq_{st} [Y|Y > x]$ for all x.*
2. *A random variable Y dominates a random variable X in the reverse hazard rate order (denoted as $X \leq_{rh} Y$) if $[X|X < x] \leq_{st} [Y|Y < x]$ for all x.*
3. *A random variable Y dominates a random variable X in the likelihood ratio order (denoted as $X \leq_{lr} Y$) if $[X|a \leq X \leq b] \leq_{st} [Y|a \leq Y \leq b]$ for all $a < b$.*

It is clear from the above definitions that $X \leq_{lr} Y$ implies both $X \leq_{hr} Y$ and $X \leq_{rh} Y$.

Proposition 2 *Consider two distinct type distributions F and $\widetilde{F}$. Let X and $\widetilde{X}$ be the random variables representing agent types associated with distribution F and $\widetilde{F}$, respectively. Assume that $X \leq_{hr} \widetilde{X}$ and $X \leq_{rh} \widetilde{X}$. Let $\widetilde{\tau}^*$ and τ^* be the optimal threshold function under $\widetilde{X}$ and X, respectively. Then, for any $k \in \{1, ..., K\}$, $\widetilde{\tau}^*(k) \geq \tau^*(k)$.*

Proof. See Appendix A. ■

For an intuition, recall the identity (12) with $k+1$ replaced by k:

$$\beta(k) E\left[X|X < x^k\right] + (1 - \beta(k)) E\left[X|X > x^k\right] = x^k.$$

When the distribution is improved by the likelihood ratio order, both conditional expectations go up. Thus, in order to keep the average constant at x^k, one needs to increase the weight on the smaller term $E\left[X|X < x^k\right]$. Thus, the weight of this term must increase and the optimal threshold is shifted upwards.

5 Large Societies

If the number of voters is large, then we can ignore the integer problem. We also normalize the threshold $\tau^*(k)$ by the size of voter population n, and write $\tau^*(k) = \beta(k+1)$, which is interpreted as the minimal *proportion* of voters required in order to undertake alternative k.

In a large society, the optimal (second-best) mechanism approximates the welfare maximizing mechanism (first-best), which, as illustrated in Example 1, is not directly implementable.[36] This is intuitive since the aggregate uncertainty vanishes in the limit. For a simple proof, assume that the ex-ante optimal alternative is l, and consider a fixed threshold mechanism that requires the support of a proportion t of voters where $F(x^l) < t < F(x^{l+1})$. This mechanism is anonymous, unanimous and DIC, and hence it must be welfare inferior to the optimal mechanism derived above. The per-capita welfare attained by this mechanism converges to the first-best when the population size tends to infinity since the welfare maximizing alternative is chosen with a probability that converges to unity.

With linear utility and a large number of voters, the maximization of average utility coincides with the maximization of the utility of the mean (average) voter. Thus, our optimal mechanism should pick the alternative favored by the mean voter with probability going to one. To see that this is indeed the case, assume that the mean voter's top alternative is k_μ, which implies that $\mu \in [x^{k_\mu}, x^{k_\mu+1}]$ and $F(x^{k_\mu}) \leq F(\mu) \leq F(x^{k_\mu+1})$. With a slight abuse of notation, we express the function β as a continuous function of type:

$$\beta(x) = \frac{E[X|X > x] - x}{E[X|X > x] - E[X|X < x]}. \tag{13}$$

It is easy to verify that

$$\beta(x) = \frac{F(x)(E[X|X > x] - x)}{E[X|X > x] F(x) - \mu + E[X|X > x][1 - F(x)]} = F(x)\frac{E[X|X > x] - x}{E[X|X > x] - \mu}.$$

[36] A result in the same spirit for settings with only two alternatives has been obtained by Ledyard and Palfrey [2002].

Thus $\beta(\mu) = F(\mu)$ and $\beta\left(x^{k_\mu}\right) \geq F(\mu) \geq \beta\left(x^{k_\mu+1}\right)$, because $\beta(x)$ is decreasing by Assumption B'. This implies that, under our optimal voting thresholds, any alternative $k < k_\mu$ is not likely to be chosen because $\mu \geq x^{k+1}$ and thus $\tau^*(k) = \beta\left(x^{k+1}\right) > F(x^{k+1})$, while the alternative k_μ, preferred by the mean voter, will be chosen with probability going to one because $\tau^*(k_\mu) = \beta\left(x^{k_\mu+1}\right) \leq F(\mu)$.

To conclude, for any distribution of preferences F satisfying our assumptions, a simple threshold mechanism which requires an $F(\mu)$-majority for adoption, where μ is the mean of F, is approximately efficient (and optimal) when the number of voters is sufficiently large, types are independent and utility is linear.

5.1 Voting over the Provision of a Public Good

In order to illustrate the above insight in a less abstract setting, we sketch below a very simple, textbook example about the provision of a public good subject to congestion. There are n agents, each endowed with an exogenous amount of a private good M_i. An agent i with type x_i has a utility function of the form $u_i = x_i G/\sqrt{n} + Z_i$ where G is the amount of public good, Z_i is the amount of the private good, and the factor of $1/\sqrt{n}$ captures the effect of congestion.[37] Types distribute identically and independently of each other according to distribution F. Producing G units of the public good costs $G^2/2$ units of the private good. The cost is equally shared among the agents, so that the only decision is about the level of public good provision.

Suppose endowments are sufficiently large. Individual utility maximization reveals that each individual i prefers a public good level of $G_i^* = \sqrt{n}x_i$. Preferences over the various levels of the public good are easily seen to be single-peaked and single-crossing. For each realization of types, the outcome of simple majority voting produces the Condorcet winner $G^{sm} = \sqrt{n}x_{sm}$ where x_{sm} is the sample median. In contrast, the efficient production level, that satisfies Samuelson's (or Bowen's !) well-known condition is such that the sum of individual marginal rate of substitution must be equal to the social marginal cost. Here we obtain $G^{s\mu} = \sqrt{n}x_{s\mu}$, where $x_{s\mu} = \frac{1}{n}\sum x_i$ is the sample mean. Given the linear structure, we can identify any level of public good G with the type $x_G = G/\sqrt{n}$ for which this level is optimal.

It is obvious, and has been noted by Bowen, that the outcome of majority voting is almost always inefficient if the distribution of types is skewed, so that the mean and the median do not coincide. If the number of voters is large so that the sample mean $x_{s\mu}$ and sample median x_{sm} approach the mean μ and median m respectively, too little (too much) public good is provided by majority voting if the the median m is lower (higher) than the mean μ. As shown above, the efficient outcome G^μ can be attained by an $F(\mu)$-majority rule.

If we normalize each individual's utility without the public good to be zero, we can express the inefficiency ratio (IR) of simple majority as the ratio between the welfare from the public

[37]For expositional simplicity we assume here that any level of the public good can be provided, so that the quantity is continuous.

good obtained under the simple majority rule and the first-best outcome as

$$IR = \frac{m}{\mu}(2 - \frac{m}{\mu})$$

where the ratio $\frac{m}{\mu}$ is a simple measure of skewness. This function is strictly less than 1 if $m \neq \mu$.

5.2 Gini Coefficients and the Efficient Supermajority

Let us apply the $F(\mu)$-majority rule to social decisions that depend on the distribution of income, where agents with higher income prefer higher decisions. As already seen above, the necessary correction versus the simple majority rule increases with the skewness of the distribution. Assume, as in a large number of empirical studies, that the distribution of income is given by a lognormal distribution with parameters μ and σ.[38] Such distributions are of course skewed, with the mean larger than the median. The Gini coefficient, which is readily available in practice, is given by $g_\sigma = 2\Phi(\sqrt{2}\sigma/2) - 1$ where Φ is the standard normal CDF. Note that this only depends on σ. The threshold at the mean as a function of the Gini coefficient can be computed by

$$F_\sigma = \Phi\left[\frac{\sqrt{2}}{2}\Phi^{-1}(\frac{g_\sigma + 1}{2})\right],$$

This threshold is increasing in g, and hence in σ. For the typical range of Gini coefficients found in (Western) democracies $g_\alpha \in [0.25, 0.55]$ the resulting optimal supermajority does not vary too much, ranging between 58% and 68%.[39]

6 Voting with Correlated Types

The optimality calculations in the previous parts were conducted under the basic assumptions that types are independent. But, there are numerous situations where the agents' preferences depend on some underlying state of the world, and are thus correlated. Consider, for example, the decision to change income taxes in an environment where the economy's fundamentals (and hence the government's fiscal needs) are not constant over time. Individual preferences depend then on the economy's fundamentals, and hence, presumably, the optimality conditions should also reflect this dependence. As a typical illustration, note that tax increases over \$70 million require a referendum in Missouri, but the legislature can raise itself such taxes by a two-thirds majority if the governor declares a state emergency (e.g., after a flood).[40]

[38] We use notational convention here. The parameters μ and σ of the lognormal distribution determine, but are not identical to, the mean and the standard deviation.

[39] As an example, the U.S. has a Gini of .45 and a required threshold of exactly two-thirds.

[40] See *National Conference of State Legislators,* http://www.ncsl.org/issues-research/budget/state-tax-and-expenditure-limits-2010.aspx

For another example, consider a monetary committee that votes on interest rates. Members usually agree that the optimal rate needs to balance the trade-off between inflation and unemployment, but have different signals or perceptions about the economy. Relatively pessimistic (optimistic) members will be more dovish (hawkish) regarding a raise in the interest rate. Perceptions depend of course on the prevailing market conditions, and thus types are again correlated.

Given the ubiquity of examples such as the above, it is of interest to establish whether the intuitions developed above for the case of independent types continue to hold with correlated types. Note that the class of anonymous, unanimous and DIC mechanisms is independent of assumptions about the probability distribution of voters' types. The same holds for the sincere equilibria of the successive voting procedure with a decreasing threshold. Thus, the changes induced by possible correlations will affect our results only via their effect on the threshold function itself and its calculation.

The analysis is complicated by the fact that the various events of pivotality—which determine the optimal thresholds—reveal now information about the underlying state of the world, and, moreover, the revealed information is influenced by the imposed thresholds themselves. Nevertheless, we are able to extend our main insights, albeit under more stringent sufficient conditions.

Assume that there are S states of the world:[41] if the state is $s \in \{1, ..., S\}$, then types x_s distribute identically and independently of each other according to a distribution F_s with mean μ_s. Let $p_s \geq 0$ denote the probability that the state of the world is s, with $\sum p_s = 1$. Utilities are assumed here to be linear, as in Section 4.1.

Let τ^* denote the optimal cutoff function and apply the same reasoning as in the case of independent types. Consider any alternative k with $\tau^*(k-1) > \tau^*(k)$ and $k \geq 2$. For τ^* to be optimal, "local" changes should weakly decrease the expected social welfare. In particular, decreasing $\tau^*(k-1)$ by 1 while keeping all other cutoffs unchanged will have any effect only if there are exactly $\tau^*(k-1) - 1$ agents with types in $\left[\underline{x}, x^k\right]$ and $n - \tau^*(k-1) + 1$ agents with types in $\left[x^k, \overline{x}\right]$. In this case, the chosen alternative becomes $k-1$ instead of k. With correlated types, however, additional complication arises because this pivotal event affects the implied probabilities of the states of the world: conditional on having exactly $\tau^*(k-1) - 1$ agents with types in $\left[\underline{x}, x^k\right]$ and $n - \tau^*(k-1) + 1$ agents with types in $\left[x^k, \overline{x}\right]$, the posterior probability of state s is

$$\mathbb{P}\left(s|\tau^*(k-1)\right) = \frac{p_s \binom{n}{\tau^*(k-1)-1} \left(F_s(x^k)\right)^{\tau^*(k-1)-1} \left(1 - F_s\left(x^k\right)\right)^{n-\tau^*(k-1)+1}}{\sum_{l=1}^{S} p_l \binom{n}{\tau^*(k-1)-1} \left(F_l\left(x^k\right)\right)^{\tau^*(k-1)-1} \left(1 - F_l\left(x^k\right)\right)^{n-\tau^*(k-1)+1}}.$$

Note that this probability depends on the number of agents, on the cutoff x^k, and on the assumed threshold $\tau^*(k-1)$.

[41] We assume here for simplicity that there is a finite number of states, but the analysis does not depend on this assumption.

The expected social utility from alternative $k-1$ in this case is given by

$$na_{k-1}+b_{k-1}\sum_{s=1}^{S}\mathbb{P}\left(s|\tau^{*}(k-1)\right)\left\{\begin{array}{c}\left(\tau^{*}\left(k-1\right)-1\right)E\left[X_{s}|X_{s}<x^{k}\right]\\ +\left(n-\tau^{*}\left(k-1\right)+1\right)E\left[X_{s}|X_{s}>x^{k}\right]\end{array}\right\}. \tag{14}$$

The expected social utility from alternative k is given by

$$na_{k}+b_{k}\sum_{s=1}^{S}\mathbb{P}\left(s|\tau^{*}(k-1)\right)\left\{\begin{array}{c}\left(\tau^{*}\left(k-1\right)-1\right)E\left[X_{s}|X_{s}<x^{k}\right]\\ +\left(n-\tau^{*}\left(k-1\right)+1\right)E\left[X_{s}|X_{s}>x^{k}\right]\end{array}\right\}. \tag{15}$$

Because $\tau^{*}\left(k-1\right)$ is part of the optimal voting procedure, decreasing $\tau^{*}\left(k-1\right)$ by 1 should weakly decrease the expected social welfare. This implies that expression (15) is weakly higher than expression (14). This yields

$$\sum_{s=1}^{S}H_{s}\left(k\right)\left(\frac{F_{s}\left(x^{k}\right)}{1-F_{s}\left(x^{k}\right)}\right)^{\tau^{*}(k-1)-1}\left(n\beta_{s}\left(k\right)-\tau^{*}\left(k-1\right)+1\right)\geq 0, \tag{16}$$

where the function β_{s} for state s is defined by

$$\beta_{s}\left(k\right)=\frac{E[X_{s}|X_{s}>x^{k}]-x^{k}}{E[X_{s}|X_{s}>x^{k}]-E[X_{s}|X_{s}<x^{k}]},$$

and where the function $H_{s}\left(k\right)$ is defined as

$$H_{s}\left(k\right)=p_{s}\left(1-F_{s}(x^{k})\right)^{n}\left(E[X_{s}|X_{s}>x^{k}]-E[X_{s}|X_{s}<x^{k}]\right). \tag{17}$$

Similarly, fix any alternative k with $\tau^{*}\left(k-1\right)>\tau^{*}\left(k\right)$ and $k\leq K-1$. Increasing $\tau^{*}\left(k\right)$ by 1 should also lead to a weakly lower expected social welfare, which yields

$$\sum_{s=1}^{S}H_{s}\left(k+1\right)\left(\frac{F_{s}\left(x^{k+1}\right)}{1-F_{s}\left(x^{k+1}\right)}\right)^{\tau^{*}(k)}\left(n\beta_{s}\left(k+1\right)-\tau^{*}\left(k\right)\right)\leq 0. \tag{18}$$

Now let us define functions $T_{k}\left(\tau\right)$, $k=1,2,...,K-1$, as

$$T_{k}\left(\tau\right)=\sum_{s=1}^{S}H_{s}\left(k+1\right)\left(\frac{F_{s}\left(x^{k+1}\right)}{1-F_{s}\left(x^{k+1}\right)}\right)^{\tau}\left(n\beta_{s}\left(k+1\right)-\tau\right). \tag{19}$$

Our candidate solution $\tau\left(k\right)$ is an integer that satisfies $T_{k}\left(\tau-1\right)\geq 0$ and $T_{k}\left(\tau\right)\leq 0$, i.e., that satisfies inequalities (16) and (18)). In order to ensure that $\tau\left(k\right)$ is well-defined it suffices to establish that each function T_{k} is single-crossing in τ, i.e., there is a unique τ such that $T_{k}\left(\tau\right)=0$.

The difficulty of proving the single-crossing property (SCP) lies in the fact that the terms of the form $n\beta_{s}\left(k+1\right)-\tau$ (which were the only relevant terms in the independent case) are now weighted by the updated probabilities that vary in both the chosen alternative and in the threshold imposed there. Nevertheless, SCP holds if the states of the world can be ordered stochastically according to the likelihood ratio order.[42] Its proof requires recently developed tools from the theory of monotone comparative statics (Quah and Strulovici [2012]) that allow us to aggregate SCP's relevant in each state.

[42] As in Proposition 2, the combination of the hazard and reversed hazard rate order also suffices here.

Proposition 3 *Assume that (possibly after reordering the states of the world) $X_1 \leq_{lr} X_2 \leq_{lr} ... \leq_{lr} X_S$ and that each X_s, $s = 1, ..., S$ satisfies Assumption B'. Then, for each k, the function T_k is single-crossing.*

Proof. See Appendix A. ■

Since we are optimizing a bounded function over a discrete and finite domain of monotone decreasing cutoff sequences, an optimal solution always exists. Similarly to the independent types case, in order to identify the solution by the necessary first order conditions we also need to ensure that the two derived inequalities (16) and (18) hold for all alternatives, including alternatives k with $\tau^*(k-1) = \tau^*(k)$. This is proved in Lemma 3 in the Appendix A where, in addition to the likelihood ratio stochastic dominance and Assumption B', we offer a simple sufficient condition under which these two inequalities hold. Finally, this yields

Proposition 4 *Assume that $X_1 \leq_{lr} X_2 \leq_{lr} ... \leq_{lr} X_S$ and that each X_s, $s = 1, ..., S$ satisfies Assumption B'. Suppose $\beta_S(k+1) \leq \beta_1(k)$ for all $k \leq K-1$. Let $\tau^*(k)$ be the unique integer that satisfies $T_k(\tau) \leq 0$ and $T_k(\tau - 1) \geq 0$. Then the threshold $\tau^*(k)$ is optimal and the obtained function τ^* is decreasing.*

Proof. See Appendix A. ■

The condition $\beta_S(k+1) \leq \beta_1(k)$ for all $k \leq K-1$, imposes restrictions on both agents' utilities and type distributions. It becomes more restrictive when there are more alternatives (and the interval of types is bounded) and when the distance between the distributions in the lowest and highest states of the world is large. But it is relatively mild, as illustrated by the following example.

Example 3 *There are two states of the world with distributions $F_1 = x$ and $F_2 = (x)^2$ on $[0,1]$ respectively. It is clear that $X_1 \leq_{lr} X_2$. Moreover, $\beta_1(k) = 1 - x^k$ and $\beta_2(k) = 1 - \frac{1}{2}(x^k + (x^k)^2)$. As should be the case by the above argument, these functions are decreasing in k and $\beta_1(k) \leq \beta_2(k)$. Then the condition in Proposition 4 requires $\beta_2(k+1) \leq \beta_1(k)$ for all $k \leq K-1$, which is equivalent to $x^{k+1} \geq \frac{1}{2}(\sqrt{8x^k+1} - 1)$. Since it is always the case that $x^{k+1} \geq x^k$, a mild, sufficient condition is $x^{k+1} - x^k \geq 1/8 = \max_{x \in [0,1]} \frac{1}{2}(\sqrt{8x+1} - 1) - x$.*

6.1 Large Societies with Correlated Types

The implementation of the optimal mechanism via a fixed threshold when the society is large (see Section 5 above) clearly hinges on the assumption that uncertainty about the efficient alternative vanishes in the limit when the number of agents gets large. It fails if there is uncertainty about the states of the world, and hence residual uncertainty about the best course of action.

Consider a large society with S states of the world where types x_s distribute identically and independently of each other according to a distribution F_s (with a mean μ_s) if the state is $s \in \{1, ..., S\}$. Denote by k_s the efficient alternative in state s, preferred by the mean voter μ_s. When states are stochastically ordered we obtain that $\mu_1 \leq \mu_2 \leq ... \leq \mu_S$ and therefore $k_1 \leq k_2 \leq ... \leq k_S$. A flexible threshold $\tau^*(k_s) = F_s(\mu_s)$ implements then the efficient outcome in each state of the world if $F_1(\mu_1) \geq F_2(\mu_2) \geq ... \geq F_S(\mu_S)$.[43] It is obvious that in general this outcome cannot be obtained by a fixed threshold policy.

It is interesting to note that for the special case where $F_s(x) = 1 - e^{-\lambda_s x}$ we obtain that $F_s(\mu_s) = F_s(\lambda_s^{-1}) = 1 - e^{-1} \approx 0.63$, so that a fixed two-thirds majority is approximately optimal in all states of the world. The same holds if the states of the world are governed by lognormal distributions with Gini coefficients in the actual range of Western democracies, so that the same two-thirds majority obtained above is relatively stable under the typical range of income inequality.

To conclude, the large-society setting with correlated types forcefully demonstrates the importance of flexible thresholds, and our results are consistent with observed practices that, for example, adjust legislative hurdles in times of financial distress or national emergency.

7 Concluding Remarks

We have characterized constrained efficient (i.e., second-best) dominant strategy incentive compatible and deterministic mechanisms in a setting where privately informed agents have single-crossing utility functions, but where monetary transfers are not feasible. Our approach allows a systematic choice among Pareto-efficient mechanisms based on the ex-ante utility they generate. We have also shown that the optimal mechanism can be implemented by a modification of a widely used voting procedure. This modification is an extension to several alternatives of the idea behind qualified majorities (or supermajorities) that are also widely used for binary decisions. In practice, one could use flexible thresholds in a simplified way (e.g., by using only one switching point from a high threshold to a low one) instead of changing the required threshold for each alternative. Such schemes are already welfare superior to those using a fixed threshold.

An open question is whether random mechanisms can yield a improvement over the deterministic mechanisms studied in this paper. The answer would be clearly negative if one could show that any probabilistic, DIC and anonymous mechanism is a lottery over deterministic, DIC and anonymous mechanisms. Peters et al. [2014] prove exactly that on single-peaked domains satisfying a minimally richness condition. But, their result is not immediately applicable here, mainly because their incentive compatibility concept is ordinal: a deviation from truth-telling must be disadvantageous for any cardinal utility representation of the ordinal single-peaked preferences; thus, their concept is stronger than the incentive compatibility concept for a specific and given cardinal utility function, and it potentially excludes more mechanisms.

[43] As an example, consider $F_s(x) = (x)^s$. Then $F_s(\mu_s) = (\frac{s}{s+1})^s$, which is decreasing in the parameter s. Because $F_s(\mu_s) = \beta_s(\mu_s)$, $s = 1, 2, ..., S$, the condition here is a special case of what was assumed in Proposition 4 for a finite number of voters.

Another open question is whether using the more permissive Bayesian incentive compatibility concept can improve the performance of constrained efficient mechanisms. It is instructing to note that in the standard setting with independent types, linear utility and with monetary transfers, a general welfare equivalence result between dominant strategy incentive compatible and Bayes-Nash incentive compatible mechanisms has been established by Gershkov et al. [2013].

Finally, recall that we studied a private value environment in the sense that a voter's payoff, conditional on the chosen alternative, depends on only his own private information. Given that in many applications voters' payoffs often depend on other voters' private information, it will be interesting to study optimal voting rules for information aggregation in an interdependent value environment.

Appendix A: Proofs

Lemma 1 *Sincere voting is the unique outcome surviving iterated elimination of (weakly) dominated strategies.*[44]

Proof. The final vote is between alternatives K and $K-1$: for any observed history of previous play, voters with peaks up to and including $K-1$ have a dominant action, to vote for $K-1$, while voters with peak on alternative K must vote for K. Thus, at the first stage of elimination, for all players, we can delete the strategies that prescribe an insincere action at the last vote. Consider now the vote to approve or reject alternative $K-2$. Voting insincerely is clearly dominated for all agents with peaks up to and including $K-2$ (since any outcome that can be obtained by voting No is worse than any outcome that can be obtained by voting Yes), and for all agents with peak on alternative K (vice-versa). Thus, we can eliminate all strategies that prescribe insincere voting at stage $K-2$ for these types of all agents. It remains to deal with types having a peak on alternative $K-1$. If such an agent prefers alternative K to alternative $K-2$, then the argument is the same as for an agent with peak on K, and insincere voting at $K-2$ is dominated. Look then at an agent with peak on alternative $K-1$, who prefers alternative $K-2$ to K. He may, theoretically, believe that, given the observed history of play, voting Yes (stopping at $K-2$) is better, as voting No will ultimately lead to K, a worse option. But note that the action of such an agent at stage $K-2$ makes any difference only in the particular instance where there are exactly $\tau(K-2)-1$ other agents that vote Yes at that stage. By the previous arguments, this set must consist of all agents with peak up to and including $K-2$ and, possibly, of some other agents with peak on $K-1$ who favor alternative $K-2$ over K. But then, because $\tau(K-1) \leq \tau(K-2)$, our

[44] If agents have strict ordinal preferences, then a simple condition ensuring that the order of elimination does not matter is trivially satisfied in our setting (for example the condition of *transference of decision maker indifference* in Marx and Swinkels [1997]). The model with a continuum of types is the limit of models with a finite number of types (representing the possible ordinal profiles) with strict preferences. In the limit, only a finite number of types (measure zero) do not have strict preferences.

agent must conclude that, in the only instance where he can affect the outcome, voting No at stage $K-2$ (and voting Yes at the next stage, as required) necessarily leads to the adoption of alternative $K-1$, his peak. Thus, voting insincerely at $K-2$ is also dominated for such agents, and thus for all possible types of all agents. The proof for all earlier stages continues analogously, and sincere voting after each possible history is the only outcome surviving the iterated elimination process. ■

The formal statement of the main theorem in Saporiti [2009] used in the proof of Theorem 1 is as follows:

Theorem 3 (Saporiti, 2009) *An unanimous, anonymous mechanism g is DIC if and only if there exists $(n-1)$ numbers $\alpha_1, ..., \alpha_{n-1} \in \mathcal{K}$ such that for any type profile $(x_1, ..., x_n) \in [\underline{x}, \overline{x}]^n$ with $x_i \in (x^{k_i}, x^{k_i+1})$ for all i, it holds that*

$$g(x_1, ..., x_n) = M(\alpha_1, ..., \alpha_{n-1}, k_1, ..., k_n),$$

where the function M returns the median of $(\alpha_1, ..., \alpha_{n-1}, k_1, ..., k_n)$.

Proof of Theorem 2. Consider the optimal mechanism with decreasing threshold function τ^*. Suppose the optimal threshold function satisfies $\tau^*(k-1) > \tau^*(k)$ for some $k \geq 2$. Suppose the planner decreases the cutoff $\tau^*(k-1)$ by 1 while keeping all other cutoffs unchanged. Because $\tau^*(k-1) > \tau^*(k)$, the alternative cutoff sequence is still monotone and thus feasible. This change matters only if there are exactly $\tau^*(k-1)-1$ voters with values below x^k and $(n+1-\tau^*(k-1))$ voters with values above x^k (recall that x^k is the cutoff type that is indifferent between alternative $k-1$ and alternative k). In this case, by decreasing the cutoff $\tau^*(k-1)$ by 1, the planner might change the allocation from k to $k-1$ given that $\tau^*(k-1)-1 \geq \tau^*(k)$. In this case, the total expected utility from alternative k is given by

$$[\tau^*(k-1)-1]\, u^k_{x<x^k} + [n+1-\tau^*(k-1)]\, u^k_{x>x^k}.$$

The total expected utility from alternative $k-1$ is given by

$$[\tau^*(k-1)-1]\, u^{k-1}_{x<x^k} + [n+1-\tau^*(k-1)]\, u^{k-1}_{x>x^k}.$$

Since the planner (weakly) prefers k to $k-1$, the total expected utility from alternative k must be higher than the total expected utility from alternative $k-1$. This gives us the following "first-order condition" for all $k \geq 2$ with $\tau^*(k-1) > \tau^*(k)$:

$$[\tau^*(k-1)-1]\left(u^k_{x<x^k} - u^{k-1}_{x<x^k}\right) + [n+1-\tau^*(k-1)]\left(u^k_{x>x^k} - u^{k-1}_{x>x^k}\right) \geq 0. \tag{20}$$

Similarly, suppose the optimal threshold function satisfies $\tau^*(k-1) > \tau^*(k)$ for some $k \leq K-1$. Now suppose the planner increases $\tau^*(k)$ by 1 while keeping all other cutoffs unchanged. Since $\tau^*(k-1) > \tau^*(k)$ and $k \leq K-1$, this alternative cutoff sequence is still monotone and thus feasible. This change matters only if only if there are exactly $\tau^*(k)$ voters with values below x^{k+1} and $(n-\tau^*(k))$ voters with values above x^{k+1}. In this case, by increasing the cutoff $\tau^*(k)$ by 1, the planner might change the allocation from k to $k+1$ given that $\tau^*(k+1) \leq \tau^*(k)$. In this case, the total expected utility from alternative k is given by

$$\tau^*(k)\, u^k_{x<x^{k+1}} + [n-\tau^*(k)]\, u^k_{x>x^{k+1}}.$$

The total expected utility from alternative $k+1$ is given by

$$\tau^*(k)\, u^{k+1}_{x<x^{k+1}} + [n-\tau^*(k)]\, u^{k+1}_{x>x^{k+1}}.$$

This yields another "first-order condition" for all $k \leq K-1$ with $\tau^*(k-1) > \tau^*(k)$:

$$\tau^*(k)\left(u^k_{x<x^{k+1}} - u^{k+1}_{x<x^{k+1}}\right) + [n-\tau^*(k)]\left(u^k_{x>x^{k+1}} - u^{k+1}_{x>x^{k+1}}\right) \geq 0. \tag{21}$$

These two first-order conditions can be rewritten as bounds on the cutoff functions $\tau^*(k)$:

$$\begin{aligned}
\tau^*(k-1) &\leq \frac{n\left(u^k_{x>x^k} - u^{k-1}_{x>x^k}\right)}{\left(u^{k-1}_{x<x^k} - u^k_{x<x^k}\right) + \left(u^k_{x>x^k} - u^{k-1}_{x>x^k}\right)} + 1, \\
\tau^*(k) &\geq \frac{n\left(u^k_{x>x^{k+1}} - u^{k+1}_{x>x^{k+1}}\right)}{\left(u^k_{x<x^{k+1}} - u^{k+1}_{x<x^{k+1}}\right) + \left(u^{k+1}_{x>x^{k+1}} - u^k_{x>x^{k+1}}\right)}.
\end{aligned}$$

We can use the definition (8) of $\beta(k)$ to rewrite it as

$$\tau^*(k-1) \leq n\beta(k) + 1, \text{ for all } k \geq 2, \tag{22}$$

$$\tau^*(k) \geq n\beta(k+1), \text{ for all } k \leq K-1, \tag{23}$$

which are exactly the inequalities (9) and (10) in the main text. Lemma 2 below shows that the above two conditions also hold for k with $\tau^*(k-1) = \tau^*(k)$.

Therefore, we can construct the (generically unique) optimal cutoff function $\tau^*(k)$ as follows. We first derive bounds for $\tau^*(1)$ by taking $k=2$ in (22) and $k=1$ in (23):

$$n\beta(2) \leq \tau^*(1) \leq n\beta(2) + 1.$$

Since the two bounds differ by 1 and $\tau^*(1)$ must be an integer, $\tau^*(1)$ is generically unique and must be equal to $\lceil n\beta(2) \rceil$,where $\lceil z \rceil$ denotes the smallest integer that is above z. Next, for all $2 \leq k \leq K-1$, conditions (22) and (23) imply that

$$n\beta(k+1) \leq \tau^*(k) \leq n\beta(k+1) + 1.$$

Hence, $\tau^*(k)$ is also generically unique and must be equal to $\lceil n\beta(k+1) \rceil$. Finally, for $k=K$, the cutoff $\tau^*(K)$ is fixed at 1.

Note that by Assumption B, β is decreasing in k, so the optimal cutoff $\tau^*(k) = \lceil n\beta(k+1) \rceil$ is indeed decreasing for all $k \leq K-1$. Further note that $\beta(K) > 0$, so we must have $\tau^*(K-1) \geq 1 = \tau^*(K)$. Therefore, the optimal cutoff function τ is decreasing.

To complete the proof, we need to argue that the cutoff function τ^* we constructed above is indeed optimal. Note that we are optimizing a bounded function over a finite domain of decreasing sequences $\tau(k)$ where $\tau(k) \leq n$ for all k and $\tau(K) = 1$. Thus an optimal solution always exists. Because the optimal solution has to satisfy the two necessary conditions (22) and (23), and because there is an essentially unique cutoff function that satisfies these two conditions, our candidate distribution τ^* must be optimal. ■

Lemma 2 *The bounds (22) and (23) hold for all $k \in \mathcal{K}$ with $\tau^*(k) = \tau^*(k-1)$.*

Proof. First let us define κ_1 and κ_2 as follows:

$$\begin{aligned} \kappa_1 &= \max\{m \in \mathcal{K} : \tau^*(m) = n\}, \\ \kappa_2 &= \min\{m \in \mathcal{K} \text{ and } m \leq K-1 : \tau^*(m) = 1\}. \end{aligned}$$

We need to consider four cases.

Case 1: Both κ_1 and κ_2 exist. Then by definition, we have $\tau^*(1) = ... = \tau^*(\kappa_1) = n$, $\tau^*(\kappa_1+1) < n$, $\tau^*(\kappa_2) = ... = \tau^*(K) = 1$ and $\tau^*(\kappa_2-1) > 1$. An alternative k with $\tau^*(k) = \tau^*(k-1)$ could belong to one of the following three possible scenarios:

(i) $k \leq \kappa_1$. Then $\tau^*(k) = n$ and condition (23) holds trivially. We only need to prove condition (22). By definition of κ_1, $\tau^*(\kappa_1) > \tau^*(\kappa_1+1)$. Thus, (22) must hold at κ_1+1:

$$\tau^*(\kappa_1) \leq n\beta(\kappa_1+1) + 1.$$

Therefore, we have

$$\tau^*(k-1) = \tau^*(k) = \tau^*(\kappa_1) \leq n\beta(\kappa_1+1) + 1 \leq n\beta(k) + 1,$$

where the second inequality follows because β is decreasing and $\kappa_1 + 1 > k$.

(ii) $k \geq \kappa_2 + 1$. Then $\tau^*(k) = \tau^*(\kappa_2) = 1$ and condition (22) is trivially satisfied, and we only need to prove condition (23). By definition of κ_2, $\tau^*(\kappa_2-1) > \tau^*(\kappa_2)$. So we have (23) hold at κ_2:

$$\tau^*(\kappa_2) \geq n\beta(\kappa_2+1).$$

Therefore,

$$\tau^*(k) = \tau^*(\kappa_2) \geq n\beta(\kappa_2+1) \geq n\beta(k+1).$$

Again, the last inequality follows from the monotonicity of $\beta(\cdot)$ and the fact that $k \geq \kappa_2 + 1$.

(iii) $k \in (\kappa_1, \kappa_2+1)$. Define k_1 and k_2 as follows:

$$\begin{aligned} k_1 &= \max\{m \in \mathcal{K} : \tau^*(m) > \tau^*(k)\}, \\ k_2 &= \min\{m \in \mathcal{K} : \tau^*(m) < \tau^*(k)\}. \end{aligned}$$

It's clear that both k_1 and k_2 are well defined for all $k \in (\kappa_1, \kappa_2)$, and $k_1 < k < k_2$. By definition of k_1 and k_2, we have

$$\tau^*(k_1) > \tau^*(k_1+1) = ... = \tau^*(k-1) = \tau^*(k) = ... = \tau^*(k_2-1) > \tau^*(k_2).$$

Therefore, both conditions (23) and (22) hold at $k_1 + 1$ and k_2. In particular,

$$\tau^* (k_1 + 1) \geq n\beta (k_1 + 2) \text{ and } \tau^* (k_2 - 1) \leq n\beta (k_2) + 1.$$

Since $\tau^* (k - 1) = \tau^* (k)$, we must have $k_1 \leq k - 1$ or equivalently $k \geq k_1 + 1$. It follows from the monotonicity of β that

$$\tau^* (k) = \tau^* (k_1 + 1) \geq n\beta (k_1 + 2) \geq n\beta (k + 1),$$

which is (23). Similarly, we can use the monotonicity of β and the fact that $k_2 > k$ to obtain

$$\tau^* (k - 1) = \tau^* (k_2 - 1) \leq n\beta (k_2) + 1 \leq n\beta (k) + 1,$$

which is (22).

Case 2: Neither κ_1 nor κ_2 exists. Then the argument of Case 1(iii) applies for all k with $\tau^* (k - 1) = \tau^* (k)$.

Case 3: κ_1 exists but κ_2 does not. Consider alternative k with $\tau^* (k - 1) = \tau^* (k)$. If $k \leq \kappa_1$, the argument of Case 1(i) applies. If $k > \kappa_1$, the argument of Case 1 (iii) applies.

Case 4: κ_2 exists but κ_1 does not. Consider alternative k with $\tau^* (k - 1) = \tau^* (k)$. If $k \geq \kappa_2 + 1$, the argument of Case 1(ii) applies. If $k < \kappa_2 + 1$, the argument of Case 1(iii) applies. ■

Proof of Proposition 2. Observe that, under distribution F,

$$\begin{aligned} \beta (k) &\equiv \frac{E\left[X|X > x^k\right] - x^k}{E\left[X|X > x^k\right] - E\left[X|X \leq x^k\right]} \\ &= \frac{E\left[X|X > x^k\right] - x^k}{E\left[X|X > x^k\right] - x^k + x^k - E\left[X|X \leq x^k\right]} \\ &= \frac{1}{1 + \frac{x^k - E\left[X|X \leq x^k\right]}{E\left[X|X > x^k\right] - x^k}}. \end{aligned}$$

Similarly, under distribution $\widetilde{F}$, we can obtain the corresponding $\widetilde{\beta} (k)$ as

$$\widetilde{\beta} (k) = \frac{1}{1 + \frac{x^k - E[\widetilde{X}|\widetilde{X} \leq x^k]}{E[\widetilde{X}|\widetilde{X} > x^k] - x^k}}.$$

Note that, for any $x \in [\underline{x}, \overline{x}]$,

$$\begin{aligned} X &\leq_{hr} \widetilde{X} \Rightarrow E\left[X|X > x\right] \leq E[\widetilde{X}|\widetilde{X} > x] \\ X &\leq_{rh} \widetilde{X} \Rightarrow E\left[X|X \leq x\right] \leq E[\widetilde{X}|\widetilde{X} > x] \end{aligned}$$

Therefore, we have $\widetilde{\beta} (k) \geq \beta (k)$ under the assumptions given in the proposition. As a result, we must have $\widetilde{\tau^*} (k) \geq \tau^* (k)$ for all k. ■

Proof of Proposition 3. Consider the function T_k and express each component in this summation as

$$Z_k(\tau, s) = H_s(k+1)\left(\alpha_s(k+1)\right)^{\tau}\left(n\beta_s(k+1) - \tau\right)$$

where $H_s(k+1)$ is defined in (17) and

$$\alpha_s(k+1) = \frac{F_s\left(x^{k+1}\right)}{1 - F_s\left(x^{k+1}\right)}.$$

Then, each Z_k is single-crossing in τ on $[0, n]$ since the function $\alpha^{\tau}(n\beta - \tau)$ is first positive, zero at $\tau = n\beta$, and then negative. The function H does not play here any role. Look now at s and s' such that $X_s \geq_{lr} X_{s'}$. Then by usual stochastic dominance and the monotonicity of function $\frac{x}{1-x}$ on $[0, 1]$, we have $\frac{\alpha_s(k+1)}{\alpha_{s'}(k+1)} < 1$. In addition, the proof of Proposition 2 shows that the likelihood ratio order implies $\beta_s(k+1) > \beta_{s'}(k+1)$. Look now at the ratio:

$$\frac{Z_k(\tau, s)}{Z_k(\tau, s')} = \frac{H_s(k+1)}{H_{s'}(k+1)}\left(\frac{\alpha_s(k+1)}{\alpha_{s'}(k+1)}\right)^{\tau}\frac{n\beta_s(k+1) - \tau}{n\beta_{s'}(k+1) - \tau}.$$

Since $\beta_s(k+1) > \beta_{s'}(k+1)$, we know that $n\beta_{s'}(k+1) - \tau < 0$ if $n\beta_s(k+1) - \tau < 0$. Thus the ratio $\frac{Z_k(\tau,s)}{Z_k(\tau,s')}$ is negative only on the interval $[n\beta_{s'}(k+1),\ n\beta_s(k+1)]$. By Theorem 1 of Quah and Strulovici [2012], the function $T_k(\tau) = \sum_s Z_k(\tau, s)$ will be single-crossing in τ if for each parameters s, s', the ratio $\frac{Z_k(\tau,s)}{Z_k(\tau,s')}$ is increasing in τ on the interval $[n\beta_{s'}(k+1),$ $n\beta_s(k+1)]$. Since τ is not a variable of the functions H_s and $H_{s'}$, it is sufficient to check the derivative of a function of the form $v^{\tau}(n\beta - \tau)/\left(n\beta' - \tau\right)$, where $v = \alpha/\alpha' < 1$ and $\beta > \beta'$, on the interval $[n\beta',\ n\beta]$. Its derivative with respect to τ is

$$\frac{v^{\tau}}{\left(n\beta' - \tau\right)^2}\left[n\beta - n\beta' + (\ln v)\left(n\beta - \tau\right)\left(n\beta' - \tau\right)\right] > 0$$

because $\ln v < 0$ and $n\beta' < \tau < n\beta$. ■

Lemma 3 *Assume that $X_1 \leq_{lr} X_2 \leq_{lr} ... \leq_{lr} X_S$ and that each X_s, $s = 1, ..., S$ satisfies Assumption B'. Assume also that $\beta_S(k+1) \leq \beta_1(k)$ for all $k \leq K-1$. Then conditions (16) and (18) hold for all $k \in \mathcal{K}$ with $\tau^*(k) = \tau^*(k-1)$.*

Proof. This lemma extends Lemma 2 to the case with correlated types, and the proof follows the same steps as before. Define κ_1 and κ_2 as follows:

$$\begin{aligned}
\kappa_1 &= \max\left\{m \in \mathcal{K} : \tau^*(m) = n\right\}, \\
\kappa_2 &= \min\left\{m \in \mathcal{K} \text{ and } m \leq K-1 : \tau^*(m) = 1\right\}.
\end{aligned}$$

We need to consider four cases.

Case 1: Both κ_1 and κ_2 exist. Then, by definition, we have $\tau^*(1) = ... = \tau^*(\kappa_1) = n$, $\tau^*(\kappa_1 + 1) < n$, $\tau^*(\kappa_2) = ... = \tau^*(K) = 1$ and $\tau^*(\kappa_2 - 1) > 1$. An alternative k with $\tau^*(k) = \tau^*(k-1)$ could belong to one of the following three possible scenarios:

(i) $k \leq \kappa_1$. Then $\tau^*(k) = n$ and condition (18) holds trivially. We only need to prove condition (16). By definition of κ_1, $\tau^*(\kappa_1) > \tau^*(\kappa_1 + 1)$. Thus, it is feasible to decrease $\tau^*(\kappa_1)$ by 1, so condition (16) must hold at $\kappa_1 + 1$. That is,

$$\sum_{s=1}^{S} H_s(\kappa_1 + 1) \left(\frac{F_s(x^{\kappa_1} + 1)}{1 - F_s(x^{\kappa_1 + 1})} \right)^{\tau^*(\kappa_1) - 1} (n\beta_s(\kappa_1 + 1) - \tau^*(\kappa_1) + 1) \geq 0.$$

Note that by likelihood ratio order, $\beta_1(\kappa_1 + 1) \leq \beta_2(\kappa_1 + 1) \leq ... \leq \beta_S(\kappa_1 + 1)$. Therefore, in order for the above inequality to hold, it must be the case that

$$n\beta_S(\kappa_1 + 1) - \tau^*(\kappa_1) + 1 \geq 0.$$

By assumption $\beta_S(\kappa_1 + 1) \leq \beta_1(\kappa_1)$, and by monotonicity of β, we have $\beta_1(\kappa_1) \leq \beta_1(k)$. These two inequalities imply that

$$\beta_S(\kappa_1 + 1) \leq \beta_1(k).$$

As a result,

$$n\beta_1(k) - \tau^*(\kappa_1) + 1 \geq 0$$

Again by likelihood ratio order, $\beta_1(k) \leq \beta_2(k) \leq ... \leq \beta_S(k)$ for all k, so we must have

$$n\beta_s(k) - \tau^*(\kappa_1) + 1 \geq 0, \text{ for all } s = 1, ..., S.$$

By definition of κ_1, we have $\tau^*(k-1) = \tau^*(\kappa_1)$, so

$$n\beta_s(k) - \tau^*(k-1) + 1 \geq 0, \text{ for all } s = 1, ..., S.$$

But this implies that condition (16) holds at k.

(ii) $k \geq \kappa_2 + 1$. Then $\tau^*(k) = \tau^*(\kappa_2) = 1$ and condition (16) is trivially satisfied, and we only need to prove condition (18). By definition of κ_2, $\tau^*(\kappa_2 - 1) > \tau^*(\kappa_2)$. So it is feasible to increase $\tau^*(\kappa_2)$ by 1, which indicates that (18) must hold at κ_2:

$$\sum_{s=1}^{S} H_s(\kappa_2 + 1) \left(\frac{F_s\left(x^{\kappa_2 + 1}\right)}{1 - F_s(x^{\kappa_2 + 1})} \right)^{\tau^*(\kappa_2)} (n\beta_s(\kappa_2 + 1) - \tau^*(\kappa_2)) \leq 0.$$

Since $\beta_1(\kappa_2 + 1) \leq ... \leq \beta_S(\kappa_2 + 1)$ by the likelihood ratio order, it must be the case that

$$n\beta_1(\kappa_2 + 1) - \tau^*(\kappa_2) \leq 0.$$

Since $\beta_1(\kappa_2 + 1) \geq \beta_S(\kappa_2 + 2)$ by assumption, and $\beta_S(\kappa_2 + 2) \geq \beta_S(k+1)$ by monotonicity of β_S, we have

$$n\beta_S(k+1) - \tau^*(\kappa_2) \leq 0.$$

Again by the likelihood ratio order, $\beta_1(k+1) \leq \beta_2(k+1) \leq ... \leq \beta_S(k+1)$, so we must have

$$n\beta_s(k+1) - \tau^*(\kappa_2) \leq 0, \text{ for all } s = 1, ..., S.$$

By definition of κ_2, we know $\tau^*(k) = \tau^*(\kappa_2)$. Hence,

$$n\beta_s(k+1) - \tau^*(k) \leq 0, \text{ for all } s = 1, ..., S.$$

Therefore, condition (18) must hold at k.

(iii) $k \in (\kappa_1, \kappa_2 + 1)$. Define k_1 and k_2 as follows:

$$\begin{aligned} k_1 &= \max\{m \in \mathcal{K} : \tau^*(m) > \tau^*(k)\}, \\ k_2 &= \min\{m \in \mathcal{K} : \tau^*(m) < \tau^*(k)\}. \end{aligned}$$

It's clear that both k_1 and k_2 are well defined for all $k \in (\kappa_1, \kappa_2 + 1)$, and $k_1 < k < k_2$. By definition of k_1 and k_2, we have

$$\tau^*(k_1) > \tau^*(k_1+1) = ... = \tau^*(k-1) = \tau^*(k) = ... = \tau^*(k_2-1) > \tau^*(k_2).$$

Therefore, both conditions (16) and (18) hold at $k_1 + 1$ and k_2. In particular,

$$\sum_{s=1}^{S} H_s(k_2) \left(\frac{F_s(x^{k_2})}{1 - F_s(x^{k_2})} \right)^{\tau^*(k_2-1)-1} (n\beta_s(k_2) - \tau^*(k_2-1) + 1) \geq 0,$$

and

$$\sum_{s=1}^{S} H_s(k_1+2) \left(\frac{F_s(x^{k_1+2})}{1 - F_s(x^{k_1+2})} \right)^{\tau^*(k_1+1)} (n\beta_s(k_1+2) - \tau^*(k_1+1)) \leq 0.$$

Similar to what we argued previously for Case 1(i) and (ii), in order for these two inequalities to hold, we must have

$$n\beta_S(k_2) - \tau^*(k_2-1) + 1 \geq 0, \text{ and } n\beta_1(k_1+2) - \tau^*(k_1+1) \leq 0.$$

By definition of k_1 and k_2, we have $\tau^*(k_1+1) = \tau^*(k-1) = \tau^*(k) = \tau^*(k_2-1)$. Hence we can rewrite the above two inequalities as

$$n\beta_S(k_2) - \tau^*(k-1) + 1 \geq 0, \text{ and } n\beta_1(k_1+2) - \tau^*(k) \leq 0.$$

Furthermore, by a similar argument as above, we can show that

$$\beta_S(k_2) \leq \beta_1(k_2-1) \leq \beta_1(k) \leq \beta_2(k) \leq ... \leq \beta_S(k)$$

and

$$\beta_1(k_1+2) \geq \beta_S(k_1+1) \geq \beta_S(k+1) \geq \beta_{S-1}(k+1) \geq ... \geq \beta_1(k+1).$$

Therefore, we must have

$$n\beta_s(k) - \tau^*(k-1) + 1 \geq 0, \text{ and } n\beta_s(k+1) - \tau^*(k) \leq 0, \text{ for all } s = 1, ..., S.$$

Therefore, conditions (16) and (18) must hold at k.

Case 2: Neither κ_1 nor κ_2 exists. Then the argument of Case 1(iii) applies for all k with $\tau^* (k-1) = \tau^* (k)$.

Case 3: κ_1 exists but κ_2 does not. Consider alternative k with $\tau^* (k-1) = \tau^* (k)$. If $k \leq \kappa_1$, the argument of Case 1(i) applies. If $k > \kappa_1$, the argument of Case 1 (iii) applies.

Case 4: κ_2 exists but κ_1 does not. Consider alternative k with $\tau^* (k-1) = \tau^* (k)$. If $k \geq \kappa_2 + 1$, the argument of Case 1(ii) applies. If $k < \kappa_2 + 1$, the argument of Case 1(iii) applies. ■

Proof of Proposition 4. It remains to verify that the constructed τ is monotone. Note that each function β_s, $s = 1, 2, ..., S$ is decreasing in k by Assumption B'. Moreover, it holds that $\beta_1 (k) \leq \beta_2 (k) \leq ... \leq \beta_S (k)$ for all k by the likelihood ratio ordering of the types across states (see the proof of Proposition 2). Consider again the function

$$T_k (\tau) = \sum_{s=1}^{S} H_s (k+1) \left(\frac{F_s (x^{k+1})}{1 - F_s (x^{k+1})} \right)^{\tau} (n\beta_s (k+1) - \tau).$$

Let θ^k be the (possible real valued) solution to the equation $T_k (\tau) = 0$. Define $\theta^k_s = n\beta_s(k+1)$ for each $s = 1, 2, ..., S$. Then $\theta^k_1 \leq \theta^k_2 \leq ... \leq \theta^k_S$, and it must hold that $\theta^k_1 \leq \theta^k \leq \theta^k_S$. Similarly, we must have $\theta^{k-1}_1 \leq \theta^{k-1} \leq \theta^{k-1}_S$. The assumption $\beta_S (k+1) \leq \beta_1(k)$ in the proposition then implies that $\theta^{k-1}_1 \geq \theta^k_S$. It immediately follows that $\theta^k \leq \theta^{k-1}$. Therefore, we must have $\tau^* (k) \leq \tau^* (k-1)$ for all $k \leq K-1$, given that $\tau^*(k)$ is the smallest integer such that $\tau^*(k) \geq \theta^k$. ■

Appendix B: Sufficient Conditions for Assumption B

Note first that requiring of $\beta (k)$ to be decreasing in k is equivalent to requiring

$$\frac{u^{k-1}_{x<x^k} - u^k_{x<x^k}}{u^k_{x>x^k} - u^{k-1}_{x>x^k}}$$

to be increasing in k. To derive sufficient conditions for Assumption B, we let $h_k (x)$ denote the utility difference for a type-x agent from two adjacent alternatives k and $k-1$:

$$h_k(x) = u^{k-1} (x) - u^k (x).$$

We claim that if the random variables $\{h_k (x)\}_{k \in \mathcal{K}}$ are ordered in terms of both hazard rate order and reverse hazard rate order, that is, $h_k \leq_{hr} h_{k+1}$ and $h_k \leq_{rh} h_{k+1}$, then Assumption B holds.[45] To see this, note that we can write

$$u^{k-1}_{x<x^k} - u^k_{x<x^k} = E[h_k(x) \mid x < x^k] = E[h_k(x) \mid h_k(x) > 0]$$

[45] Note that conditions $h_k \leq_{hr} h_{k+1}$ and $h_k \leq_{rh} h_{k+1}$ impose restrictions on the shapes of both the distribution F and the utility function u. Alternatively, if we assume F is uniform, we could explicitly derive the required conditions for Assumption B only on function u. On the other hand, if we assume that the utility function u is linear as in Section 4.1, the required conditions for Assumption B impose restrictions only on the distribution F (see Assumption B').

where the second equality follows from the definition of cutoff x^k and the single-crossing property. By rewriting $u^k_{x>x^k} - u^{k-1}_{x>x^k}$ analogously, we obtain

$$\frac{u^{k-1}_{x<x^k} - u^k_{x<x^k}}{u^k_{x>x^k} - u^{k-1}_{x>x^k}} = -\frac{E[h_k(x) \mid h_k(x) > 0]}{E[h_k(x) \mid h_k(x) < 0]}.$$

Note that $h_k \leq_{hr} h_{k+1}$ implies that $E[h_k(x) \mid h_k(x) > 0]$ is increasing in k, and $h_k \leq_{rh} h_{k+1}$ implies that $E[h_k(x) \mid h_k(x) < 0]$ is increasing in k. Therefore, $\beta(k)$ is decreasing in k.

References

[2012] Ali, N., and Kartik, N. (2012), "Herding with Collective Preferences," *Economic Theory* **51**, 601-626.

[2011] Apesteguia, J, Ballester, M., and Ferrer, R. (2011), "On the Justice of Decision Rules," *Review of Economic Studies* **78**(1), 1-16.

[2014] Apesteguia, J, Ballester, M., and Y. Masatlioglu (2014), "A Foundation for Strategic Agenda Voting," *Games and Economic Behavior* **87**: 91-99.

[2004] Ausubel, L. (2004), "An Efficient Ascending-Bid Auction for Multiple Objects," *American Economic Review* **94**(5), 1452-1475

[2014] Azrieli, Y. and Kim, S. (2014), "Pareto Efficiency and Weighted Majority Rules," *International Economic Review*, forthcoming.

[2005] Bagnoli, M. and Bergstrom, T. (2005), "Log-Concave Probability and Its Applications," *Economic Theory* **26**(2), 445-469.

[2001] Barbera, S (2001), "An Introduction to Strategy-Proof Social Choice Functions," *Social Choice and Welfare* **18**, 619-653.

[1993] Barbera, S., Gul, F. and Stacchetti, E. (1993), "Generalized Median Voter Schemes and Committees," *Journal of Economic Theory* **61**(2): 262-289

[1994] Barbera, S., and Jackson, M. (1994), "A Characterization of Strategy Proof Mechanisms for Economies with Pure Public Goods," *Social Choice and Welfare* **11**, 241-252.

[1990] Barbera, S. and Peleg, B. (1990), "Strategy-proof Voting Schemes with Continuous Preferences," *Social Choice and Welfare* **7**, 31-38.

[2009] Bebchuk, L., Cohen, A., and Ferrell, A. (2009), "What Matters in Corporate Governance," *The Review of Financial Studies* **22**, 783-827.

[1948] Black, D. (1948), "On the Rationale of Group Decision-Making," *Journal of Political Economy* **56,** 23-34

[1958] Black, D. (1958), *The Theory of Committees and Elections*, Cambridge University Press, Cambridge, U.K.

[2009] Börgers, T., and Postl, P. (2009), "Efficient Compromising," *Journal of Economic Theory* **144**, 2057-2076.

[1943] Bowen, H. (1943), "The Interpretation of Voting in the Allocation of Economic Resources," *The Quarterly Journal of Economics* **58**(1), 27-48.

[2007] Callander, S. (2007), "Bandwagons and Momentum in Sequential Voting," *Review of Economic Studies* **74**(3), 653-684.

[2011] Chatterji, S. and Sen, A. (2011), "Tops-only Domains," *Economic Theory* **46**, 255-282.

[1997] Ching, S. (1997), "Strategy-Proofness and 'Median Voters', " *International Journal of Game Theory* **26,** 473-490.

[2000] Dekel, E., and Piccione, M, (2000), "Sequential Voting Procedures in Symmetric Binary Elections," *Journal of Political Economy* **108**(1), 34-55

[2013] Drexl, M. and Kleiner, A. (2013), "Why Voting? A Welfare Analysis," discussion paper, University of Bonn.

[2002] Ehlers, L., Peters, H., and Storcken, T. (2002), "Strategy-Proof Probabilistic Decision Schemes for One-Dimensional Single-Peaked Preferences," *Journal of Economic Theory* **105**, 408-434.

[1969] Farquharson, R. (1969), *The Theory of Voting*, Yale University Press, New Haven, U.S.A.

[1907] Galton, F. (1907), "One Vote, One Value," reprinted in *Nature* 75, Issue 1948, 414.

[2009] Gersbach, H. and Pachl, B. (2009), "Flexible Majority Rules for Central Banks," *Journal of Money, Credit and Banking* **41** (2-3), 507-516.

[2013] Gershkov, A., Goeree J., Kushnir, A., Moldovanu, B. and Shi, X. (2013), "On the Equivalence of Bayesian and Dominant Strategy Implementation," *Econometrica* **81**(1), 197-220.

[1973] Gibbard, A. (1973), "Manipulation of Voting Schemes: A General Result," *Econometrica* **41**(4), 587-601.

[2003] Gompers, P., Ishi., J. and Metrick, A (2003), "Corporate Governance and Equity Prices," *Quarterly Journal of Economics* **118**(1), 107-155.

[1979] Green J., and Laffont, J. (1979), *Incentives in Public Decision-Making,* Amsterdam: North Holland.

[1995] Joyce, M. et al. (1995), *Tax and Expenditure Limits on Local Governments*, Advisory Commission on Intergovernmental Relations, Center for Urban Policy and the Environment, Indiana University.

[2002] Ledyard, T., and Palfrey, T. (2002), "The Approximation of Efficient Public Good Mechanisms by Simple Voting Schemes," *Journal of Public Economics,* **83** (2), 151-171.

[1997] Marx, L.M., Swinkels, J.M. (1997), "Order Independence for Iterated Weak Dominance," *Games and Economic Behavior* **18**, 219-245.

[1981] Meltzer A., and Richard, S. (1981), "A Rational Theory of the Size of Government," *Journal of Political Economy* **89**, 914-927.

[1980] Moulin, H. (1980), "On Strategy-Proofness and Single Peakedness," *Public Choice* **35**, 437-455.

[2004] Nehring, K. (2004), "The Veil of Public Ignorance," *Journal of Economic Theory* **119**, 247-270

[2014] Peters, H., Roy, S., Sen, A., and Storcken, T. (2014), "Probabilistic Strategy-Proof Rules over Single-Peaked Domains," *Journal of Mathematical Economics* **52**, 123-127.

[2012] Quah, J., and Strulovici, B. (2012), "Aggregating the Single Crossing Property," *Econometrica* **80**(5), 2333-2348.

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

[2000] Rasch, B. E. (2000), "Parliamentary Floor Voting Procedures and Agenda Setting in Europe," *Legislative Studies Quarterly* **25**(1), 3-23.

[2010] Riboni. A., and Ruge-Murcia, F. (2010), "Monetary Policy By Committee: Consensus, Chairman Dominance, or Simple Majority?" *The Quarterly Journal of Economics* **125** (1), 363-414.

[1977] Roberts, K. (1977), "Voting over Income Tax Schedules," *Journal of Public Economics* **8**, 329-340.

[1975] Romer, T. (1975), "Individual Welfare, Majority Voting and the Properties of a Linear Income Tax," *Journal of Public Economics* **7**, 163-168.

[2012] Rosar, F. (2015), "Continuous Decisions by a Committee: Median versus Average Mechanisms," *Journal of Economic Theory* **159**, 15-65.

[2009] Saporiti, A. (2009), "Strategy-proofness and Single-crossing," *Theoretical Economics* **4**, 127-163.

[1975] Satterthwaite, M. (1975), "Strategy-Proofness and Arrow's Conditions. Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory* **10**, 187-217.

[2012] Schmitz, P. and Tröger T. (2012), "The (Sub-) Optimality of the Majority Rule," *Games and Economic Behavior* **74**(2), 651-665.

[2002] Schummer, J., and Vohra, R. (2002), "Strategy-proof Location on a Network," *Journal of Economic Theory* **104**, 405-428.

[2007] Shaked, M., and Shanthikumar, J. (2007), *Stochastic Orders*, Springer, New York.

[1991] Sprumont, Y. (1991), "The Division Problem with Single-Peaked Preferences: A Characterization of the Uniform Allocation Rule," *Econometrica* **59,** 509-519.