An optimal Voting System when Voting is costly

Bognar, Katalin and Börgers, Tilman and Meyer-ter-Vehn, Moritz

University of Michigan (Ann Arbor)

October 2010

An Optimal Voting System When Voting is Costly*

Katalin Bognar[†], Tilman Börgers[‡], Moritz Meyer-ter-Vehn[§]

October 2010

Abstract

We consider the design of an optimal voting system when voting is costly. For a private values model with two alternatives we show the optimality of a voting system that combines three elements: (i) there is an arbitrarily chosen default decision and non-participation is interpreted as a vote in favor of the default; (ii) voting is sequential; (iii) not all voters are invited to participate in the vote. We show the optimality of such a voting system by first arguing that it is first best, that is, it maximizes welfare when incentive compatibility constraints are ignored, and then showing that individual incentives and social welfare are sufficiently aligned to make the first best system incentive compatible. The analysis in this paper involves some methods that are new to the theory of mechanism design, and it is also a purpose of this paper to explore these new methods.

*This paper supersedes a previous draft entitled "Dynamic Mechanism Design with Costly Participation." We are grateful to Arthur Lupia, Romans Pancs, Alessandro Pavan, Marek Pycia, Ilya Segal, and Ricky Vohra for helpful comments. Tilman Börgers thanks the Center for Mathematical Studies in Economics and Management Science at Northwestern University for the hospitality that it offered while he was working on this paper. Moritz Meyer-ter-Vehn thanks the Hausdorff Research Institute for Mathematics in Bonn for its hospitality, which he enjoyed while working on this paper.

[†]Department of Economics, UCLA, kbognar@umich.edu.

[‡]Department of Economics, University of Michigan, tborgers@umich.edu.

[§]Department of Economics, UCLA, mtv@econ.ucla.edu.

1 Introduction

When participation in votes is costly, and welfare is defined to include the cost of participation, then the design of an expected welfare-maximizing voting system must trade off the expected welfare loss that results from sub-optimal decisions against the expected cost of voting. In this paper, we examine this tradeoff, and determine a voting system that resolves the tradeoff optimally. For this optimal voting system we also show that individual voters' participation incentives are sufficiently aligned with the social welfare objective to make the optimal voting system incentive compatible. The analysis presented in this paper involves some methods that are new to the theory of mechanism design. It is also a purpose of this paper to explore these new methods.

We consider the following setting. A set of voters has to choose one candidate from a set of two possible candidates for a public office. Each voter has a strict preference over candidates, and knows his or her own preference, but not other voters' preferences. Voters' preferences are independent, and each voter is equally likely to prefer either candidate. Participation in the voting procedure has a known positive cost that is common to all voters. The system that we find to be optimal declares arbitrarily one of the two candidates to be the *default candidate*. Voters are asked sequentially whether they want to participate in the voting procedure or abstain. A voter who abstains incurs no participation cost. Non-participation is counted as a vote in favor of the default candidate. A voter who participates incurs the cost of voting, and is automatically counted as a vote against the default. The candidate who receives the majority of votes is elected. Ties can be broken arbitrarily. The process is terminated at the latest when a majority has been established that cannot be overturned by the remaining voters. The process may be terminated earlier.

The optimal procedure thus combines three features of voting procedures, all of which one observes in practice, although not always together. The first feature, the use of an arbitrarily chosen default candidate, is reminiscent of a department chair informing department members that she will take one particular course of action unless a majority of members objects. The second feature is that voting is sequential so that voters who come late in the sequence can form an assessment of the likelihood that their vote is pivotal based on earlier votes. This is reminiscent of members of the US congress watching roll-call votes on a screen in their offices, and basing the decision whether to leave their offices and to enter the chamber on the votes of those who come before them. The third feature of the optimal voting system is that participation may be artificially constrained, so that some voters are prevented from participating even if they want to. This is reminiscent of the formation of executive committees for decision making. All three features of voting systems may also in practice be motivated by voting cost considerations.

An important assumption on which our construction is based is that voters can observe the actions of other voters without incurring participation cost. Observation only requires watching a TV or computer screen, which is essentially costless, whereas participation requires actual presence in a room. We also assume that the participation cost is independent of the mechanism. In some contexts different assumptions about exogenous participation cost will be more appropriate. Participation costs may depend on the mechanism. They may also depend on the extent to which an agent observes actions by other agents before participating. Finally, in some contexts it may make sense to consider a mechanism designer who can subsidize voters' participation cost. All this is ruled out in our model. Our model is thus special, but the set-up that we choose yields results that seem realistic in the voting context.

Our study is also of methodological interest because it would not be sufficient in our set-up to consider static games, i.e. games in which all players choose simultaneously. Rather we have to consider extensive games with sequences of moves and non-trivial information sets. Formally, the choice set of the designer of the voting system in this paper will be the set of all finite extensive game forms with endogenous participation choices. Choosing optimally from this set is a much more complex problem than choosing the optimal mechanism from static games, which is traditionally considered in mechanism design, where in addition typically one can assume that the agents' message spaces are identical to their type spaces.[1]

The same issue arises in a recent paper by Gershkov and Szentes (2009). They, too, study the optimal design of a voting scheme. However, whereas we consider a private value setting, they consider a setting with common values. While in our setting participation is costly, in their setting information acquisition is costly. As is the case in our paper, they cannot restrict attention

[1]We explain in Sections 4.1 and 4.2 below why the consideration of static revelation games is not useful for our analysis.

to static revelation games but instead need to consider all extensive games. Their paper therefore deals with a similar methodological issue as ours.

It is interesting to compare Gershkov and Szentes' analytical approach to ours. We proceed by first identifying a canonical class of mechanisms to which we can restrict attention. We then show that within this class the mechanisms that maximize ex ante expected welfare when incentive constraints are relaxed, i.e. "first best" mechanisms, are also incentive compatible. Gershkov and Szentes, too, find a class of canonical mechanisms to which they can restrict attention, and then optimize within this class. However, first-best mechanisms are not incentive compatible in their setting.

The reason why the difference between private and social interests has more severe consequences in Gershkov and Szentes' setting than in ours is that in our set-up, by interpreting non-participation as a signal, the mechanism designer can effectively choose which action is costly, and which action is free. So, in particular, he can always make whatever action has positive expected externalities freely available to agents. Therefore, if in the first best mechanism the mechanism designer asks an agent to take a costly action, we can conclude that asking the agent to take this action is not based on the positive externalities of this action, but on the benefits that the mechanism designer expects to accrue to this particular agent. The agent will therefore also find that it is in his interest to take the costly action, and the mechanism designer's request will be incentive compatible.

In Gershkov and Szentes' set-up, by contrast, the costs of actions are intrinsic to those actions. They cannot be chosen by the mechanism designer. If information acquisition has positive externalities, then the mechanism designer does not have any way of making information acquisition free. If the mechanism designer in the first best mechanism asks an agent to acquire information, this request may well be based on the positive externalities from information acquisition rather than on that agent's own interests. The request may therefore not be incentive compatible.

Gershkov and Szentes' canonical mechanisms provide minimal information to voters to relax voters' incentive constraints. Even with this construction, they do not obtain incentive compatibility of the first best mechanism. By contrast, we consider mechanisms in which all information about previous votes is revealed to agents and show that even though such a mechanism maximizes the opportunities for deviations, the optimal canonical mechanism is incentive compatible. We view this maximal informativeness as an attractive feature of our mechanism because in practice it may be difficult to conceal information from voters.

Some authors have considered welfare properties of particular voting schemes either in a private or in a common value setting, without examining the general mechanism design problem. For example, Börgers (2004) showed in a model similar to ours the superiority of voluntary voting participation over required voting participation when voting is costly. Ghosal and Lockwood (2009) and Krasa and Polborn (2009) describe models in which the opposite conclusion can be reached. Gershkov and Szentes (2009) reference a number of papers that study particular voting institutions, and in which the emphasis is not on participation cost but on information acquisition cost.

Issues analogous to the ones investigated in this paper arise in other mechanism design problems whenever participation is costly and it is possible that participants observe other players' actions without incurring all participation cost. Procurement auctions seem a realistic example. Participants in procurement auctions might first observe the participation decision of other bidders before incurring the cost of preparing their own bidding material. The existing literature has studied auction design problems in the case in which all participation decisions have to be made simultaneously. Examples of relevant papers are Stegeman (1996) who considers welfare maximizing auction design in this environment, and Celik and Yilankaya (2009), who consider profit maximizing auction design. Our paper indicates how one can set up the auction designer's problem if one wants to consider the possibility of reducing participation cost, or perhaps of encouraging participation, through sequential participation decisions.[2]

Considering information acquisition cost rather than participation cost, Bergemann and Välimäki (2002) find for a general environment with transferable utility that Vickrey-Clarke-Groves mechanisms provide socially optimal incentives for information acquisition when agents have independent private values. Our main result is similar in that it also proves the incentive compatibility of a first best solution, but the underlying intuition is different: the transfer payments in a Vickrey-Clarke-Groves mechanism induce perfectly aligned individual and social incentives, whereas in our setting individual and social incentives potentially diverge, yet equilibrium participation decisions are socially optimal. Further results in the literature on information

[2] Celik and Yilankaya (2009, p. 3) note that dynamic mechanisms might be of advantage in their setting.

acquisition in environments with transferable utility are surveyed in Bergemann and Välimäki (2006). A recent paper by Pancs (2010) considers a setting with two potential buyers of a single, indivisible object, and finds a sequential mechanism with partial information disclosure to be maximizing the expected value of a weighted average of welfare and revenue. This paper is related to our work because it also considers an optimization problem where the objects among which the mechanism designer chooses are extensive game forms.

A different type of cost, communication cost, are considered in Fadel and Segal (2009). The communication cost of a decision rule equals the minimal number of bits of information agents must transmit in an incentive compatible mechanism that computes this decision rule. Our participation cost can be interpreted as a fixed cost of communication, whereas Fadel and Segal's communication costs are proportional to the number of bits communicated. Fadel and Segal fix a decision rule and study the minimal communication costs to implement it. By contrast, we also study the tradeoff between suboptimal decisions and participation costs.

We present our framework in Section 2. In Section 3 we state and prove the main result. Section 4 contains some concluding remarks.

2 Set-Up

There are $n \in \mathbb{N}$ voters. We denote the voters by: $i \in N \equiv \{1, 2, \ldots, n\}$. The voters have to pick one of two candidates: $k \in \{A, B\}$. Each voter $i \in N$ has a type $\theta_i \in \{A, B\}$. The type of voter i indicates which candidate voter i prefers. Types are random, and different voters' types are independent. Each voter is equally likely to prefer candidate A and candidate B. Each voter observes his or her own type, but not the other voters' types. Voter i's Bernoulli utility equals $1-c$ if i's preferred candidate wins and i participated in the decision process, $-c$ if i's preferred candidate loses and i participated in the decision process, 1 if i's preferred candidate wins and i did not participate in the decision process, and 0 if i's preferred candidate loses and i did not participate in the decision process. Here, $c > 0$ is a constant, the cost of participation in the decision making process. The distribution of types, what each voter observes, the Bernoulli utility function, and the value of c are common knowledge among voters and the mechanism designer. The mechanism designer does not observe any voter's type. The mechanism designer seeks to maximize the sum of all voters' ex ante expected utilities.

The setting described so far is obviously very special. We study this simple setting to focus on methodological issues without distraction and because elementary modifications of the set-up make the question that we examine analytically significantly harder, as we explain in the last section of the paper. As anticipated in the Introduction, despite its simplicity, the setting that we study yields results that are of some real world plausibility.

The mechanism designer chooses firstly an extensive game form that the voters then use to pick one of the two candidates, and secondly a sequential equilibrium of the game implied by the game form, the voters' Bernoulli utilities, and the voters' information structure. It is assumed that the voters play the sequential equilibrium that the mechanism designer chooses.

We begin by describing the set of extensive game forms that the mechanism designer can choose from. We have adapted the following definition of *extensive game forms* from Osborne and Rubinstein (1994, pp. 200-201).

Definition 1. *An* extensive game form *consists of:*

1. *The set N of* players *(identical to the set of voters).*

2. *A finite set H of finite sequences with the following properties:*

   (a) *The empty sequence $\emptyset$ is an element of H;*

   (b) *If $\left(a^k\right)_{k=1,\ldots,K} \in H$ and $L < K$ then $\left(a^k\right)_{k=1,\ldots,L} \in H$.*

   (Each element of H is a history*, or, equivalently, a* node*. Each component of a history is an* action *taken by a player, or a chance move. A history $\left(a^k\right)_{k=1,\ldots,K} \in H$ is* terminal *if there is no a^{K+1} such that $(a^k)_{k=1,\ldots,K+1} \in H$. The set of terminal histories is denoted by Z. The set of actions available after a nonterminal history $h \in H \backslash Z$ is denoted by $\mathcal{A}(h) \equiv \{a : (h; a) \in H\}$.)*

3. *A function P that assigns to each nonterminal history $h \in H \setminus Z$ an element of $N \cup \{C\}$. (P is the* player function*, $P(h)$ being the player who takes an action after history h. If $P(h) = C$ then chance determines the action taken after history h.)*

4. *A function f_C that associates with every history h for which $P(h) = C$ a probability measure $f_C(\cdot|h)$ on $\mathcal{A}(h)$. ($f_C(a|h)$ is the probability that a occurs after history h.)*

5. *For each player $i \in N$ a partition $\mathcal{I}_i$ of $\{h \in H : P(h) = i\}$ with the property that $\mathcal{A}(h) = \mathcal{A}(h')$ whenever h and h' are in the same element of the partition. ($\mathcal{I}_i$ is the* information partition *of player i. Any set $I_i \in \mathcal{I}_i$ is an* information set *of player i.)*

6. *A function D that assigns to each terminal history $h \in Z$ a decision $D(h) \in [0,1]$. ($D(h)$ as the probability that A is chosen after history h. Candidate B is chosen with the remaining probability $1 - D(h)$. We also write $D(h) = A$ if $D(h) = 1$, and $D(h) = B$ if $D(h) = 0$.)*

We next single out particular extensive game forms, namely those that can be interpreted as decision making procedures that include voters' decisions about participation in the procedure. We call such extensive game forms *mechanisms with participation decisions.*

Definition 2. *An extensive game form is a* mechanism with participation decisions *if:*

1. *For every nonterminal history $h \in H \setminus Z$ we either have:*

   (a) *$P(h) \in N$ and $\mathcal{A}(h) \subseteq \{0,1\}$, or*

   (b) *$\mathcal{A}(h) \cap \{0,1\} = \emptyset$.*

   (Histories that satisfy (a) will be called participation nodes *of player $P(h)$. For such histories we interpret the action 1 as participation by voter $P(h)$, and the action 0 as non-participation by voter $P(h)$.)*

2. *For every terminal history $\left(a^k\right)_{k=1,\dots,K} \in Z$ and every $i \in N$ we either have:*

   (a) *if $\ell < K$ and $P(a^1, \dots, a^\ell) = i$ then $a^{\ell+1} = 0$*

   or there is a unique $L < K$ such that:

   (b) *if $\ell < L$ and $P(a^1, \dots, a^\ell) = i$ then $a^{\ell+1} = 0$;*

   (c) *$P(a^1, \dots, a^L) = i$ and $a^{L+1} = 1$;*

   (d) *if $\ell > L$ and $P(a^1, \dots, a^\ell) = i$ then $a^{\ell+1} \notin \{0,1\}$.*

(Either player i does not participate in this terminal history, or he participates at the $L+1$st action node of the history. In the latter case, all of i's earlier action nodes are participation nodes at which i does not participate, but none of the later action nodes are participation nodes. This definition implies that before doing anything else players have to choose to participate.)

For every mechanism with participation decisions there is an associated game of incomplete information in which voters first learn privately their types, and then play the mechanism. Voters evaluate outcomes according to their Bernoulli utility function. The cost c is incurred by a player i if and only if a participation history of that player i has been reached and the player made the choice to participate, that is, chose action 1 at that point. We denote strategies of player i by σ_i. The solution concept that we shall use is sequential equilibrium. The mechanism designer's problem is then as follows:

The Mechanism Designer's Problem. *Choose a mechanism with participation decisions m and a strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ for the incomplete information game associated with m to maximize the sum of ex ante expected utilities of all voters, subject to the constraint that σ is a sequential equilibrium of the incomplete information game associated with m.*

3 Result

Proposition 1. *A solution (m, σ) to the mechanism designer's problem exists. Moreover, there is at least one solution with the following properties:*

(i) There is no $h \in H \setminus Z$ such that $P(h) = C$. (There are no chance moves.)

(ii) For every $i \in N$ and every $I_i \in \mathcal{I}_i$: $\#I_i = 1$. (The extensive game form is of perfect information.)

(iii) For every $h = (a^1, \ldots, a^K) \in Z$ and every $i \in N$ there is at most one $\ell < K$ such that $P(a^1, \ldots, a^\ell) = i$. Moreover then: $A(a^1, \ldots, a^\ell) = \{0, 1\}$. (Every voter makes at most one decision. This decision is a participation decision. Participation is voluntary.)

(iv) *There is an* $x \in \{A, B\}$ *such that for every* $h = (a^1, \ldots, a^K) \in Z$*:* $D(h) = x$ *if* $\#\{k|a^k = 1\} < K/2$*, and* $D(h) \neq x$ *if* $\#\{k|a^k = 1\} > K/2$*. (Candidate* x *is the* default *candidate. If a majority of voters who make a choice decide not to participate, then* x *is chosen. If a majority of voters who make a choice decide to participate, then the default is overturned.)*

(v) *If* $P(h) = i$ *and* $\mathcal{A}(h) = \{0, 1\}$*, then* $\sigma(h, \theta_i) = 1$ *if and only if* $\theta_i \neq x$*. (A voter participates if and only if she opposes the default.)*

(vi) *If* $h = (a^1, \ldots, a^K)$ *is such that* $\#\{k|a^k = 0\} \geq n/2$ *or* $\#\{k|a^k = 1\} \geq n/2$ *then* $h \in Z$*. (The decision process ends at the latest when a majority of potential voters has indicated a preference for one candidate.)*

Proof. The proof of Proposition 1 has two parts. In the first part we consider a relaxed version of the mechanism designer's problem in which the constraint that σ has to be a sequential equilibrium is dropped. We show that a pair (m^*, σ^*) exists that solves the relaxed problem, and that has the properties described in Proposition 1. In the second part of the proof we show for any solution (m^*, σ^*) that has the properties described in Proposition 1 that σ^* is indeed a sequential equilibrium of the extensive game with incomplete information corresponding to m^*. Therefore such a solution of the relaxed maximization problem is also a solution of the mechanism designer's problem.

Part 1. We consider the relaxed maximization problem. Let m be any mechanism with participation decisions, and let σ be a strategy combination for the corresponding extensive game with incomplete information. We shall explain how one can transform m into a mechanism with participation decisions $\hat{m}$, and σ into a strategy combination $\hat{\sigma}$ for the extensive game with incomplete information corresponding to $\hat{m}$, such that $(\hat{m}, \hat{\sigma})$ have the properties described in Proposition 1, and expected welfare resulting from $(\hat{m}, \hat{\sigma})$ is at least as large as expected welfare from (m, σ). Existence of an optimal solution with the properties described in Proposition 1 then follows immediately, as there are only finitely many mechanisms and strategy combinations with the properties described in Proposition 1. One of them needs to be optimal.

We now describe the transformation. If m has information sets with multiple elements, we begin by modifying m so that all information sets are singletons. We adjust σ so that each player's strategy assigns to every non-terminal history at which a player moves in the new extensive game form the same action that σ assigned previously to that history. In this step we thus provide players with additional information, but assume that they do not make use of this additional information. This step leaves expected welfare unchanged.

Next, if either there are chance moves in the extensive game form, or the decision rule D of the extensive game form involves randomization, or if the strategies σ involve any randomization, we consider the set of combinations of actions by the chance player C, non-randomized decision rules D, and pure strategies σ that are assigned positive probability, and we pick from this (finite) set one combination that yields highest expected welfare. We then replace the chance moves, decision rule and the strategies by this combination of non-randomized rules. This step either leaves expected welfare unchanged, or increases it. We next remove all actions that are not taken according to σ, and we remove all nodes corresponding to chance player C. This step leaves expected welfare unchanged.

We now pick one of the two candidates, denoted by x, as the "default candidate." We denote the other candidate by y. It does not matter which candidate we pick to be the default candidate. Proceeding in an arbitrary order of players, we then successively for each player i make the following changes to the extensive game form: We first consider the earliest nodes h at which player i moves. These must be participation nodes. If one type of player i participates at h, and another one does not participate, then we label the choice that type $\theta_i = x$ makes as "0" (i.e. don't participate), and the choice that type $\theta_i = y$ makes as "1" (i.e. participate). If the game tree that follows the choice 0 contains further choices of player i, then we remove the choices that type $\theta_i = y$ would have made at those nodes, and then we remove these nodes from the game tree. Similarly, if the game tree that follows the choice 1 contains further choices of player i, then we remove the choices that type $\theta_i = x$ would have made at those nodes, and then we remove these nodes from the game tree. If both types of player i participate at h, then we remove h from the game tree, assuming that player i does not participate at this history, and label the next decision nodes of player i as participation nodes. If both types of player i don't participate at h, we remove h from the game tree, assuming that player i does not participate at this node. We iterate this operation for player i until there are no further nodes left to consider. Then we proceed to the next player until there are no further players left to consider.

This step either leaves expected welfare unchanged or it increases it. This is because we don't change for each vector of realized types of all voters the candidate chosen. Moreover, in each step, if at the node that we are considering exactly one type of a player participates, then we don't change the expected participation cost or we reduce them. We may have switched which type participates, but because both types are equally likely, and both types have the same participation cost, this is inconsequential. Moreover, for the type that does not participate, we have fixed that this type does not participate at future nodes either. This potentially reduces expected participation cost. If both types participate then we have clearly either left participation cost unchanged, or reduced them. Moreover, if both types did not participate at the node that we were considering then we have left participation cost unchanged.

Our next step is to modify the candidate chosen after each history to be the candidate to maximize expected welfare conditional on the information revealed by the voters' choices. This obviously requires that the candidate is chosen who is preferred by the majority of players who have made a choice, where we count non-participation as an expression of a preference for x, and participation as an expression of a preference for y.

The final step is to remove all decision nodes that don't affect the final decision, starting at the end of the game tree and moving iteratively to the beginning. This step leaves the collective decision unchanged, and either leaves expected participation cost unchanged or reduces them. After this step we have obtained a mechanism with participation decisions, and a strategy vector for the extensive game with incomplete information corresponding to this mechanism, which have all the properties described in Proposition 1.

Part 2. Let (m^*, σ^*) be a combination of mechanism with participation decisions and strategy vector that has the properties described in Proposition 1, and that maximizes expected welfare among all such combinations. We now show that σ^* is a sequential equilibrium of the incomplete information game corresponding to m^*. We first note that all information sets in the extensive game with incomplete information corresponding to m^* are reached with positive probability. Therefore, beliefs are given by Bayesian updating. The strategies form a sequential equilibrium if and only if they are sequentially rational given these beliefs.

Consider any information set. Let i be the voter choosing at that information set. Denote by Δ the difference between the probability of the default candidate x winning if voter i abstains and the probability of the default candidate winning if voter i participates. We calculate these two probabilities conditional on the information set that we are considering. The prescribed actions of the two types of voter i at this information set are sequentially rational if and only if:

$$\Delta \geq c. \tag{1}$$

To see this note that for voter i the benefits from not-participating are Δ and the cost is zero if he prefers the default candidate, and the gains from participating are Δ and the cost of participating are c if he opposes the default candidate. The benefits are at least as large as the costs in both cases if and only if (1) is true.

We derive (1) from the fact that the given mechanism and strategy combination maximize expected welfare. Let z_1 and z_2 both be elements of the set of candidates $\{A, B\}$. Denote by "$w_{z_1 z_2}$" expected welfare conditional on the decision node that we are considering, conditional on voter i being type z_1, and conditional on voter i choosing at the decision node the action that the strategy σ^* prescribes for the case that voter i is of type z_2. Here, we calculate expected welfare taking everything into account *except voter i's participation cost*. We obtain four numbers: $w_{xx}, w_{xy}, w_{yx}, w_{yy}$, where x is the default candidate, and y is the other candidate.

The expected welfare optimality of the strategy that we are considering implies that expected welfare is at least as large when voter i's types take the actions prescribed as it is when the two types of voter i take the same action. Note that when the two types take the same action then the mechanism designer can label that action "non-participation," and thus incur zero participation cost for voter i. By contrast, if the two types take opposite actions, then with probability 0.5 the mechanism designer incurs cost of c. This implies the following two inequalities:

$$0.5w_{xx} + 0.5w_{yy} - 0.5c \geq 0.5w_{xx} + 0.5w_{yx}, \tag{2}$$
$$0.5w_{xx} + 0.5w_{yy} - 0.5c \geq 0.5w_{xy} + 0.5w_{yy}. \tag{3}$$

These two inequalities are equivalent to:

$$w_{yy} - w_{yx} \geq c, \tag{4}$$
$$w_{xx} - w_{xy} \geq c. \tag{5}$$

Now we decompose the expected welfare differences on the left hand sides of these inequalities into differences in voter i's expected welfare, and differences in all other voters' expected welfare. Denote all other voters' expected welfare by w_z^{-i} if voter i chooses the action that is prescribed for the case that he favors candidate z. Note that this expected welfare does not depend on voter i's true preference. Using this notation we can re-write the above inequalities as:

$$\Delta + w_y^{-i} - w_x^{-i} \geq c \tag{6}$$
$$\Delta + w_x^{-i} - w_y^{-i} \geq c \tag{7}$$

or, equivalently, as:

$$\Delta \geq c + \left(w_x^{-i} - w_y^{-i}\right) \tag{8}$$
$$\Delta \geq c - \left(w_x^{-i} - w_y^{-i}\right) \tag{9}$$

The expression $\left(w_x^{-i} - w_y^{-i}\right)$ must be positive, zero, or negative. In either case, one of the last two inequalities implies (1). □

4 Discussion

4.1 Sub-optimality of Static Mechanisms

Proposition 1 describes features that *at least one* optimal pair of mechanism and its equilibrium have. The result does not assert that *all* optimal pairs of mechanisms and their corresponding equilibria have these features. We argue in this subsection that when n is at least 3 and the voting cost c is sufficiently small *no* optimal mechanism will be static. Here, we call a mechanism with participation decisions *static* if all moves, including participation decisions, are made simultaneously.

We begin with the observation that a static mechanism with participation decisions and its equilibrium cannot be optimal if in equilibrium there is a positive probability that at least $n/2+1$ (if n is even) or $n/2+1.5$ (if n is odd) voters participate. This is an implication of the fact that the procedure described in part 1 of the proof of Proposition 1, when applied to such a static mechanism and its equilibrium, strictly increases expected welfare. It strictly increases expected welfare because it allows voting to be terminated when sufficiently many agents have participated to establish a majority which cannot be overturned. This economizes on voting cost.

Now suppose that voting cost c are sufficiently small, and that there are at least 3 voters. Then it is easy to see that the mechanism designer, if restricted to a static mechanism, will choose a mechanism and a corresponding equilibrium where there is a positive probability that more voters than given by the above thresholds participate.[3] Therefore, no optimal mechanism is static if n is at least 3 and the voting cost c are sufficiently small.

4.2 The Revelation Principle

We have not used the revelation principle in our analysis, although this principle plays a central role in many studies of mechanism design problems. The revelation principle, applied to our model, says that for any sequential equilibrium of a mechanism with participation decisions one can construct a direct mechanism in which voters reveal their type to the mechanism designer who then implements the outcome that would have resulted in the original mechanism if the voters had played their equilibrium strategies. Moreover, in this direct mechanism, all voters reporting their types truthfully is a Bayesian Nash equilibrium, that is, the direct mechanism is incentive compatible. In our context, the outcome that the mechanism designer implements once types have been revealed consists of a selection of one of the two candidates, and also participation decisions for each voter. Reporting one's type in the direct mechanism is costless. A voter incurs participation cost only when the direct mechanism's outcome specifies that the voter participates. With this notion of a direct mechanism, the standard proof of the revelation principle can be used to prove the revelation principle also for our model.

Note that the direct mechanisms described in the previous paragraph are *not* mechanisms with participation decisions in the sense of Definition 2. According to Definition 2 an agent has to incur participation cost before taking any other action. In particular, in Definition 2, reporting one's type is not possible without incurring participation cost. When solving the mechanism designer's problem, we cannot therefore proceed in the usual way and maximize expected welfare among all incentive compatible direct mechanisms, and then conclude that the optimal direct mechanism is also a solution to

[3] We omit the formal proof of this assertion.

the mechanism designer's original problem. Direct mechanisms are not in the feasible set of the mechanism designer's original problem.

One might try instead to first maximize expected welfare among all incentive compatible direct mechanisms that are equivalent[4] to some mechanism with participation decisions and a corresponding sequential equilibrium, reconstructing the underlying mechanism with participation decisions and its sequential equilibrium only in a second step. However, there is no obvious characterization of the incentive compatible direct mechanisms that the designer can choose from. Clearly, not all incentive compatible direct mechanisms can be allowed. For example, any direct mechanism where no voter ever participates, yet the collective decision depends on voters' types, must be ruled out. Which restrictions exactly describe the admissible set of incentive compatible direct mechanisms seems difficult to determine.

A possibly more useful version of the revelation principle in our model considers mechanisms with participation decisions where any voter who has decided to participate reports her type subsequently, and then does not get to move again.[5] One might call such mechanisms *direct mechanisms with participation decisions*. The standard logic of the revelation principle implies that for every mechanism with participation decisions, and corresponding sequential equilibrium, there is an equivalent direct mechanism with participation decisions and a sequential equilibrium such that any voter who is asked to reveal his type reveals that type truthfully. This result is true in more general models than ours, and in such more general models may simplify the search for optimal mechanisms. In our model there is no need for participating voters to reveal their types because the types can be inferred from the participation decision. Therefore, our analysis in this paper seems best conducted without explicit appeal to this version of the revelation principle.

[4]We call a direct mechanism "equivalent" to a mechanism with participation decisions and a corresponding sequential equilibrium if in this sequential equilibrium for every vector of voters' types, the same candidate is chosen, and the same voters participate, as in the direct mechanism if voters reveal their types truthfully in that mechanism.

[5]This revelation principle is related to the revelation principle invoked in Myerson (1986) where players in a multi-stage mechanism report in each period their privately observed information to the mechanism designer. However, participation decisions play no role in Myerson's set-up, and the rationale for reporting private information not initially but later is in Myerson's paper not that a player did not participate earlier, but that the information was not available to the player at an earlier stage.

4.3 Correlated Types

A natural extension of our model is a model in which agents' types are correlated. For concreteness suppose there are two equally likely possible states of the world, a and b, where conditional on state a individual voters' preferences are i.i.d. with the probability that a voter prefers candidate A being $p > 0.5$ and conditional on state b individual voters' preferences are i.i.d. with the probability that a voter prefers candidate B being p. Voters observe the state and their own preferences. This structure is common knowledge.

If the mechanism designer does not observe the state of the world before designing the mechanism, then, unlike in our setting, it may be that no first best mechanism is incentive compatible. To see this suppose for simplicity that preferences are perfectly correlated, i.e. $p = 1$. Then any first best mechanism will invite at most one voter to participate. Non-participation will be interpreted as a preference for one candidate, and participation will be interpreted as a preference for the other candidate. This one voter's preference will determine the collective choice. However, this one voter's individual incentives to participate do not reflect the positive externality that he exerts on other voters, and therefore he may individually prefer not to participate even if the first best mechanism requires him to participate.[6]

If the mechanism designer does observe the state of the world before designing the mechanism, then the first best mechanism can be analyzed as in this paper except that the choice of default candidate is no longer arbitrary, but the candidate who is more likely to be preferred by voters must be made the default candidate. A difficulty is that the argument of this paper for incentive compatibility of the first best mechanism no longer applies. The intuition is not obvious from the proof that we presented in Section 3. It is as follows: When constructing the first best mechanism in the case that both candidates are equally likely to be preferred, the mechanism designer is at every information set indifferent between labeling either of the two actions available to an agent as "non-participation," and thereby making it available for free. Therefore, in particular, even if in our construction he

[6]Note the analogy with the analysis of Gershkov and Szentes (2009) that we cited in the Introduction. In their paper, too, first best mechanisms are not incentive compatible because information acquisition by one voter exerts a positive externality on other voters, and therefore individual incentives for information acquisition need not be as large as the social benefit from information acquisition.

doesn't do so, the mechanism designer could have labelled the action with positive externalities as free, and his request to the agent to sometimes take the costly action reveals that the participation cost are not larger than the benefits accruing to this agent. This implies incentive compatibility of the first best mechanism. But if the voters' types are not equally likely, then the mechanism designer is no longer indifferent between labeling either of the two actions as "non-participation." He will always want to label that action as "participation" that the less likely type takes. This action might have positive externalities. The mechanism designer's request to the agent to take the costly action may be based in parts on such a positive externality of that action. The agent's individual incentives need not be sufficient to cover the cost of participation. The first best mechanism need not be incentive compatible. This case thus requires a separate analysis.

4.4 Privately Observed Voting Costs

Another modification of our model assumes that not only voters' preferences but also their voting cost c are privately observed random variables. The reason why the analysis of Section 3 does not apply to this modified model is as follows. When analyzing the incentive issues raised in the second part of the proof we may find that voters with large voting cost who oppose the default have an incentive to participate even if the mechanism designer prefers them not to participate. For voters who oppose the default the argument in the second part of our proof shows that their individual incentives to participate are at least as large as c whenever the designer wants them to participate. But the argument allows the possibility that the individual incentives are strictly larger than c. Thus, when the planner's first best mechanism requests a voter to abstain when her cost realization is high, the voter may find it privately beneficial to mimic a low cost type and participate anyway. If c is privately observed the planner cannot prevent her from doing so. Therefore, this modification of our model, too, requires a separate analysis.

References

Dirk Bergemann and Juuso Välimäki (2002), "Information Acquisition and Efficient Mechanism Design," *Econometrica* 70, 1007-1033.

Dirk Bergemann and Juuso Välimäki (2006), "Information in Mechanism

Design," in: Richard Blundell, Whitney Newey and Torsten Persson (eds.), *Proceedings of the 9th World Congress of the Econometric Society*, Cambridge: Cambridge University Press, 186-221.

Tilman Börgers (2004), "Costly Voting," *American Economic Review* 94, 57-66.

Gorkem Celik and Okan Yilankaya (2009), "Optimal Auctions With Simultaneous and Costly Participation," *The B. E. Journal of Theoretical Economics (Advances)* 9, Article 24.

Ronald Fadel and Ilya Segal (2009), "The Communication Costs of Selfishness," *Journal of Economic Theory* 144, 1895-1920.

Alex Gershkov and Balázs Szentes (2009), "Optimal Voting Schemes With Costly Information Acquisition," *Journal of Economic Theory* 144, 36-68.

Sayantan Ghosal and Ben Lockwood (2009), "Costly Voting When Both Information And Preferences Differ: Is Turnout too High or too Low?," *Social Choice and Welfare* 33, 25-50.

Stefan Krasa and Mattias Polborn (2009), "Is Mandatory Voting Better Than Voluntary Voting?," *Games and Economic Behavior* 66, 275-291.

Roger B. Myerson (1986), "Multistage Games With Communication," *Econometrica* 56, 323-358.

Martin Osborne and Ariel Rubinstein (1994), *A Course in Game Theory*, Cambridge (MA): The MIT Press.

Roman Pancs (2010), "Optimal Information Disclosure in Sequential Negotiations with Costly Information Acquisition," University of Rochester.

Mark Stegeman (1996), "Participation Costs and Efficient Auctions," *Journal of Economic Theory* 71, 228-259.