Preference Intensities in Repeated Collective Decision-Making*

Moritz Drexl Andreas Kleiner

14th February 2014

Abstract

We study welfare-optimal decision rules for committees that repeatedly take a binary decision. Committee members are privately informed about their payoffs and monetary transfers are not feasible. In static environments, the only strategy-proof mechanisms are voting rules which are inefficient as they do not condition on preference intensities. The dynamic structure of repeated decision-making allows for richer decision rules that overcome this inefficiency. Nonetheless, we show that often simple voting is optimal for two-person committees. This holds for many prior type distributions and irrespective of the agents' patience.

JEL classification: D72; D82; C61

Keywords: Dynamic mechanism design; Voting; Collective decision making

1 Introduction

Simple voting rules are known to be inefficient when a majority with weak preferences outvotes a minority with strong preferences. For instance, if ten out of one hundred citicens of a village are willing to pay $20 for changing a law, but the rest has a willingness-to-pay of $1 for keeping the old one, votes would be 90 to 10 against the new law, althought it would be efficient to pass it.

Money could be used as a tool to elicit preference intensities and thereby to implement the efficient allocation, but in many situations there are moral or other considerations that prevent the use of monetary means. Instead, this paper examines the possibilities of using the dynamic structure of environments, where group decisions have to be made repeatedly, in order to condition on preference intensities. In fact, repeated decision problems are ubiquitous in everyday life, ranging from examples in parliament to hiring committees. As Buchanan and Tullock (1962) emphasize, "any rule must be analyzed in terms of the results it will produce, not on a single issue, but on the whole set of issues."

Consider the following example, which illustrates the possibility of increasing sensitivity to preference intensities: Assume that the decision rule prescribes to accept if at

*Drexl, Kleiner: Bonn Graduate School of Economics, University of Bonn, Kaiserstr. 1, 53113 Bonn, Germany; drexl@uni-bonn.de, akleiner@uni-bonn.de

least one of two agents is in favor of the proposal, unless the other agent uses one of his limited possibilities to exercise a veto. In this situation, agents are faced with a trade-off between the current and future periods. If an agent exercises a veto now, the decision rule decides in her favor, but at the cost of fewer possibilities to use a veto in the future, which reduces the agent's continuation value. Intuitively, agents will use their veto right only if their preference against the proposal exceeds some threshold. This has the effect that more refined information about the agents' preferences is elicited and potentially a more efficient allocation can be implemented.

Given these ideas, the question is why we see so many decision rules that use simple majority voting in every period, and, more generally, which decision rule is the best in terms of providing the highest welfare to the agents. In this paper, we tackle the latter question and show that, surprisingly, voting rules are optimal among many reasonable decision rules. This provides a hint to the answer for the former question on why voting is used so universally.

To be specific, we analyze a model with two agents who are repeatedly presented a proposal that they need to either accept or reject. Each agent is privately informed about her willingness-to-pay (or value) for accepting the proposal; she can be in favor or against the proposal and the intensity of preference is unknown to the designer. Values are distributed independently across time and agents, and agents are symmetric ex-ante. In a related setting, Jackson and Sonnenschein (2007) establish the strong result that the first-best allocation can be approximated arbitrarily well in Bayes-Nash equilbrium. Implementation in Bayes-Nash equilibrium imposes strong assumption on the information available to the agents; specifically, it is required that each agent's belief about the other agents equals the prior belief. Moreover, the mechanism they propose hinges stronly on the assumption that there are many identical decision problems and that agents are patient. In contrast, we consider mechanisms that are more robust and hence easier to implement. We focus on mechanisms that implement decision rules in *periodic ex-post equilibrium*, where the strategies of the agents are optimal even if they come to know the current values of the other agents. This requirement renders incentives robust to uncontrolled changes in the information structure as well as some deviations of the other player.

We provide a characterization of incentive compatible decision rules in terms of the allocation in a given period and the continuation values the rule promises, which shows that the class of decision rules compatible with our equilibrium concept is relatively large. Methodologically, we view continuation values as substitutes for money, which enables us to relate any given decision rule to a static mechanism with money. We show that if preference distributions satisfy an increasing hazard rate condition, then voting rules are optimal within two classes of mechanisms. First, they are optimal among decision rules that satisfy *unanimity*, i.e., rules that never contradict the decision that both agents would unanimously agree on. This is a reasonable robustness requirement since one could expect that agents will not adhere to the decision rule if they unanimously agree to do something else. Second, if type distributions are *neutral across alternatives*, i.e., densities are symmetric around zero, then voting rules are also optimal among all deterministic decision rules. Therefore, if the type distributions are neutral across alternatives, we get the summarizing result that any decision rule yielding higher welfare than every voting rule has both weaknesses of not satisfying unanimity and not being deterministic. This provides a rationale for voting rules in the setting we consider. Our results are not only robust with respect to the exact specification of the information structure, but also does not depend on the number of decision problems, or the discount factor, and the results also hold if decision problems differ over time.

Relation to the Literature

Buchanan and Tullock (1962) provide an early analysis of collective decision making in dynamic settings. They restrict their examination to standard voting rules, but consider non-sincere equilibria. In a seminal contribution, Casella (2005) designs a specific decision rule for a dynamic setting comparable to ours. He proposes the concept of storable votes: In each period, each agent receives an additional vote and can use some of his votes for the current decision or, alternatively, he can store his additional vote for future usage. By shifting their votes inter-temporally, agents can concentrate their votes on decisions for which they have a strong preference. This decision rule thereby takes preference intensities into account, and increases welfare if there are only two agents. Hortala-Vallve (2012) analyzes a similar proposal (called qualitative voting) for a static setting (meaning that agents are completely informed about their preferences in all decision problems when making the first decision), in which agents face a number of binary decisions.

Going one step further, one can systematically look for the "best" decision rule. Jackson and Sonnenschein (2007) take a mechanism design approach and show that for a static setting the efficient outcome can be approximated even in the absence of money, by linking a large number of independent copies of the decision problem. This result extends to dynamic settings, as long as individuals are arbitrarily patient. This surprising result hinges critically on a number of strong assumptions: each decision problem has to be an identical copy, the designer is required to have the correct prior belief, agents need to be arbitrarily patient and their beliefs about other agents have to be identical to the common prior. In an attempt to find more robust decision rules, Hortala-Vallve (2010) characterizes the set of strategy-proof decision rules for a static problem. Given that strategy-proofness is a strong requirement in multi-dimensional settings, it is not too surprising that voting rules are the only decision rules that satisfy this restriction.

In contrast, our focus on periodic ex-post equilibrium implies that on the one hand, the set of implementable decision rules is very rich, but on the other hand our results are robust and the optimal mechanism is bounded away from attaining the first-best.

The paper is structured as follows: In Section 2 we present our model in detail. The results are presented in Section 3 and discussed in Section 4. Some proofs are omitted from the main text and relegated to the appendix.

2 Model

There are two agents who are repeatedly faced with a proposal and have to accept or reject each proposal. Periods are indexed by $t = 0, 1, ... \in T = \mathbb{N}$. The type of an agent i in a given period t is denoted by θ_{it} and indicates his willingness-to-pay for the proposal. Type spaces and distribution functions are the same for each period and each agent, denoted by Θ_i and F respectively, and types are drawn independently across time and agents. We denote by $\tilde{\theta}_{it}$ the random variable corresponding to the type of agent i, and by θ_t a type profile which is an element of the product type space Θ.

In each period, a decision $x_t \in \{0, 1\}$ has to be made. We denote the sequence of decisions up to period t by x^t, and similarly for a sequence of types θ_i^t. Accordingly, for an infinite sequence we write x^T.

Mechanisms

In this model a dynamic version of the revelation principle holds (Myerson (1986), for similar arguments see Pavan, Segal and Toikka (2008)), hence we can focus on truthfully implementable direct revelation mechanisms.

Definition 1. *A* mechanism χ *is a sequence of decision rules* $\{\chi_t\}_{t\in T}$ *that map past decisions and type profiles into a distribution over decisions in the current period:*

$$\chi_t : \Theta^t \times \{0, 1\}^{t-1} \to [0, 1].$$

Preferences

Agents have linear von-Neumann-Morgenstern utility functions and there are no monetary payments. Given a period t and a decision x_t for this period, the utility of agent i with type θ_{it} is $v_{it}(\theta_{it}, x_t) = \theta_{it}x_t$. Agents discount the future with the common discount factor δ. Consequently, utility of agent i with type sequence θ_i^T is

$$V_i(\theta_i^T, x^T) = \sum_{t\in T} \delta^t \theta_{it} x_t$$

for the decision sequence x^T.

Equilibrium Concept and Incentive Compatibility

We consider implementation in *periodic ex-post equilibrium* (for a formal definition, see Miller 2012), which is equivalent to agents playing an ex-post equilibrium in each period. In every period t, agent i learns about his preference type θ_{it}, which is his private information, and then sends a report r_{it}. The history known to the designer in period t, $h^t = (x^{t-1}, r^{t-1})$, consists of past decisions and past reports.

Given a mechanism χ, we can write the value function for agent i:

$$W_i(h^t, \theta_t) = \sup_{r_{it}\in\Theta_i} \theta_{it}\chi_t(h^t, r_{it}, \theta_{-it}) + \delta \mathbf{E}_{\Theta_{t+1}} W_i(h^{t+1}, \tilde{\theta}_{t+1}) \tag{1}$$

Here, h^{t+1} is the history in the next period, consisting of $\chi_t(h^t, r_{it}, \theta_{-it})$ and (r_{it}, θ_{-it}) appended to h^t. The valuation function specifies, given any history h^t, and the current type profile θ_t, the highest utility the agent can possibly obtain for some report r_{it}, assuming that she reports optimally in the future and the other agents report truthfully. Given a specific history h^t, the mechanism χ induces an allocation rule and continuation functions which we will denote

$$x_t(\theta_t) = \chi_t(h^t, \theta_t) \quad \text{and}$$
$$w_{it}(\theta_t) = \delta \mathbf{E}_{\Theta_{t+1}} W_i(h^{t+1}, \tilde{\theta}_{t+1}).$$

If the current period is clear from the context, we will also drop the subscript t. The pair (x_t, w_t) is called the *stage mechanism after history* h_t and we say that w_t is *generated* by the mechanism χ. A stage mechanism is *admissible* if it is generated by some mechanism χ.

Definition 2. *A mechanism is* periodic ex-post incentive compatible (IC) *if for every period t and for all histories h^t the following holds: For every θ_{-i} and every θ_i we have that*

$$\theta_{it}x(\theta_{it}, \theta_{-it}) + w_{it}(\theta_{it}, \theta_{-it}) \geq \theta_{it}x(r_{it}, \theta_{-it}) + w_{it}(r_{it}, \theta_{-it}) \tag{2}$$

for all reports $r_i \in \Theta_i$.

See, e.g., Athey and Miller (2007), Bergemann and Välimäki (2010). The definition in particular states that if a mechanism is incentive compatible, then every stage mechanism for all histories is incentive compatible. The following lemma can be proved using the Envelope Theorem (which is a standard exercise in mechanism design).

Lemma 1. *A mechanism is IC if and only if for each agent i the following two conditions hold:*

1. *Monotonicity of x: $x(\theta_i, \theta_{-i}) \leq x(\theta_i', \theta_{-i})$ for $\theta_i \leq \theta_i'$.*
2. *Payoff equivalence: Fix $\hat{\theta}_i \in \Theta_i$. Then for all θ*

$$\theta_i x(\theta_i, \theta_{-i}) + w_i(\theta_i, \theta_{-i}) = \hat{\theta}_i x(\hat{\theta}_i, \theta_{-i}) + w_i(\hat{\theta}_i, \theta_{-i}) + \int_{\hat{\theta}_i}^{\theta_i} x(\beta, \theta_{-i})d\beta. \tag{3}$$

Since the term $\hat{\theta}_i x(\hat{\theta}_i, \theta_{-i}) + w_i(\hat{\theta}_i, \theta_{-i})$ is independent of θ_i, we will write $h_i(\theta_{-i})$ for it. Note, however, that $h_i(\theta_{-i})$ does depend on the particular choice of $\hat{\theta}_i$.

Objective

For a given stage mechanism we can write down the expected welfare going forward from period t as

$$U_{h^t}(\chi) = U_{h^t}(x, w) := \mathbf{E}_{\Theta_t}\big[(\theta_1 + \theta_2)x_t(\theta) + w_{1t}(\theta) + w_{2t}(\theta)\big].$$

This is the period-t expected discounted welfare that the agents receive after history h^t. The aim of this paper is to identify welfare-optimal mechanisms, that is, mechanisms χ that solve

$$\max_{\chi} \; U(\chi) := U_{h^0}(\chi), \quad \text{s.t.} \quad \chi \text{ is IC}.$$

Lemma 2 in the appendix provides a useful way to rewrite the objective function in terms of the allocation rule x and $h_i(\theta_{-i})$.

3 Results

The aim of this section is to identify mechanisms that are optimal in the above stated sense. The following conditions on F which we need to derive our results are standard in the mechanism design literature.

Condition 1 (Monotone Hazard Rates)**.** *The hazard rate* $\frac{f(\theta_i)}{1-F(\theta_i)}$ *is non-decreasing in* θ_i *and the reversed hazard rate* $\frac{f(\theta_i)}{F(\theta_i)}$ *is non-increasing in* θ_i*.*

A *voting rule* x is a rule where $x(\theta)$ only depends on $\{sgn(\theta_i)\}_{i=1,2}$. A voting mechanism is a mechanism where the allocation rule after all histories is a voting rule. In each of the two subsections below we will present a setting in which the welfare-maximizing dynamic decision rule is a voting mechanism.

The proofs in each part will proceed as follows: First, we show that under the appropriate assumptions stage mechanisms consisting of a voting rule and promising the same continuation payoffs for all type profiles are weakly welfare-superior to all other stage mechanisms. Then we make use of the following proposition to deduce that also the best dynamic mechanism uses a voting rule in every period. For this step to work it is helpful that optimal stage mechanisms are of as simple a form as voting mechanisms.

Proposition 1. *Assume that for every history* h^t *and admissible stage mechanism* (x_t, w_t) *in period* t*, there exists an admissible stage mechanism* $(\hat{x}_t, \hat{w}_t)$*, where* $\hat{x}_t$ *is a voting rule and* $\hat{w}_t$ *is constant, and such that*

$$U_{h^t}(x_t, w_t) \leq U_{h^t}(\hat{x}_t, \hat{w}_t).$$

Then a voting mechanism is among the optimal mechanisms.

Proof. We start with any dynamic mechanism χ and transform it into a mechanism that uses a voting rule in every period and such that U weakly increases. Start with $t = 0$. The assumption states that there exists a voting stage mechanism $(\hat{x}_0, \hat{w}_0)$ with constant $\hat{w}_0$ and such that $U(\hat{x}_0, \hat{w}_0) \geq U(x_0, w_0)$. Since the voting stage mechanism is admissible and promises constant continuations, these continuations can be generated by a mechanism that is independent of h^1. Denote by χ' this new dynamic mechanism. Since x'_1 and w'_1 are independent of h^1, we know (again by the assumption) that there exists a voting stage mechanism $(\hat{x}_1, \hat{w}_1)$ with constant $\hat{w}_1$ and such that $U_{h^1}(\hat{x}_1, \hat{w}_1) \geq U_{h^1}(x'_1, w'_1)$ for all h^1. Again, $\hat{w}_1$ can be generated by a mechanism that does not condition on histories h^2. Now if we let χ'' be the mechanism that arises if one exchanges the stage mechanism (x'_1, w'_1) in χ' for $(\hat{x}_1, \hat{w}_1)$, we know that χ'' is still incentive compatible: All promised continuations in period 0 change by the same amount, independent of the history h^1 and in particular independent of θ_0. Repeating this argument inductively for $t \geq 2$ completes the proof. □

Unanimity

Unanimity requires the mechanism to always adhere to a decision to which both agents agree. For example, if both types in some period are positive the mechanism has to choose $x_t = 1$ for sure. Formally, the condition is defined as follows:

Definition 3. *A mechanism is called* unanimous *if, for every period and all possible histories,* $x(\theta) = 1$ *if* $\theta > 0$ *and* $x(\theta) = 0$ *if* $\theta < 0$.

Note that mechanisms not satisfying this requirement will probably have legitimacy problems: Although all parties involved in the decision process opt in favor of the proposal, the mechanism forces its rejection. Furthermore, if agents are not able to collectively commit to the decision prescribed by the mechanism, then mechanisms satisfying unanimity are the only feasible mechanisms. Also note that mechanisms proposed in the literature are not excluded by this assumption (see, e. g., Jackson and Sonnenschein 2007, Casella 2005). In the next subsection we will see that even when relaxing this restriction, for certain distribution functions only non-deterministic decision rules can yield a higher expected welfare than voting rules.

Theorem 1. *Suppose that* F *satisfies Condition 1. Then a voting mechanism is optimal among all unanimous mechanisms.*

Proof. The proof consists of establishing the preconditions of Proposition 1. So let (x, w) be a stage mechanism after some history h^t (since we are only concerned with unanimous mechanisms, x satisfies unanimity). Set $(\hat{\theta}_1, \hat{\theta}_2) = (0, 0)$ and let h_i be the resulting redistribution functions implied by Lemma 1. Let $\theta^* \in \arg\max_{\theta \in \Theta_i} h_1(\theta) + h_2(\theta)$. We first show that setting $h_1(\theta_2) = h_1(\theta^*)$ for all θ_2 and $h_2(\theta_1) = h_2(\theta^*)$ for all θ_1 does not decrease $U_{h^t}(x, w)$.

Since so far we have not changed x, by Lemma 2 it is enough to show that the terms involving the redistribution functions do not decrease in this step. But this follows from

$$\begin{aligned}\int_{\Theta_1} h_2(\theta_1) dF(\theta_1) + \int_{\Theta_2} h_1(\theta_2) dF(\theta_2) &= \int_{\underline{\theta}}^{\overline{\theta}} \big[h_2(\beta) + h_1(\beta)\big] dF(\beta) \\ &\le \int_{\underline{\theta}}^{\overline{\theta}} \big[h_2(\theta^*) + h_1(\theta^*)\big] dF(\beta).\end{aligned}$$

Next we show that changing x to a voting rule does not decrease welfare. It is enough to consider the regions where $\theta_1 \le 0, \theta_2 \ge 0$ and $\theta_1 \ge 0, \theta_2 \le 0$ because the mechanism is unanimous. By Lemma 3 and the choice of $(\hat{\theta}_1, \hat{\theta}_2)$, we know that the first term in (4), which for the region $\theta_1 \le 0, \theta_2 \ge 0$ amounts to

$$\int_{\underline{\theta}}^{0} \int_{0}^{\overline{\theta}} \left[\frac{-F(\theta_1)}{f(\theta_1)} + \frac{1 - F(\theta_2)}{f(\theta_2)}\right] x(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1),$$

is maximized by setting x to 1, as soon as Condition 1 holds. Since the same is true for the region where $\theta_1 \ge 0, \theta_2 \le 0$, we have constructed a voting stage mechanism that is weakly welfare superior to the old stage mechanism.

Let (x', w') denote the new stage mechanism. The proof is complete if we can show that w' is constant and can be generated. Constancy of w' holds for any stage mechanism where x' is a voting rule and the functions h'_i are constant. More specifically, w'_i is equal to $h_i(\theta^*)$. Since the old mechanism was unanimous, $w_i(\theta^*, \theta^*) = h_i(\theta^*)$. Because $w_i(\theta^*, \theta^*)$ could be generated, it follows that w' can be generated. □

Neutrality of Alternatives

In this section, we show that in some situations we can derive optimality of voting mechanisms even if unanimity does not hold. This shows that the restriction imposed in the previous section does in many cases not reduce welfare.

We assume that the distribution of types is *neutral across alternatives*, i.e., it is symmetric around 0. This is an important special case of our general model and has been analyzed, among others, by Carrasco and Fuchs (2011). For instance, this assumption is satisfied if a committee has to decide among two proposals that are valued equally ex ante. Specifying one alternative as the default, the distribution of valuations for changing from the default to the alternative proposal is symmetric around 0.

Theorem 2. *Suppose F satisfies Condition 1 and is neutral across alternatives. Then a voting mechanism is optimal among all deterministic mechanisms.*

The proof of Theorem 2 is presented in the appendix. Similar arguments as in the last subsection can be given for restricting attention to deterministic mechanisms: First, stochastic mechanisms are difficult to implement and face legitimacy problems in practice. It is barely conceivable that a parliament would introduce decision protocols that involve random elements. Second, all proposed mechanisms in the literature and mechanisms observed in practice are usually deterministic and therefore not excluded from our analysis. Numerical simulation also suggests that expected welfare can be improved only slightly using stochastic mechanisms. The following corollary combines Theorem 1 and Theorem 2 and summarizes all properties one has to give up in order to improve upon voting rules.

Corollary 1. *Assume F satisfies Condition 1 and is neutral across alternatives. Then every decision rule that is strictly welfare-superior to any voting rule is stochastic and does not satisfy unanimity.*

4 Discussion

We have seen that despite the absence of money as a means for implementing rules other than majority voting, the possibility to condition decision rules on the past gives us the possibility to design dynamic decision rules that take preference intensities into account. However, we have shown that for committees consisting of two players the welfare maximizing dynamic decision rule nonetheless consists of simple majority voting in every period. This holds unless desirable properties of the decision rule are given up. We therefore provide a possible explanation for why majority voting is used almost universally in practice.

One extension of our model is to allow for correlation of agent types over time. However, this restricts the class of incentive compatible mechanisms since the quasi-linear separation of continuation payoffs from the payoff in the current period disappears. While voting rules would still be optimal in this restricted class, our model without correlation shows that voting rules are also optimal in the larger class.

A major open problem is the question as to what extent our results generalize to more than two agents. We believe that a substantial difficulty towards progress in this direction is to understand in how far continuation values can be redistributed among the agents.

Appendix

Helpful Lemmata

The following shows how the welfare of every incentive compatible mechanism can be expressed in terms of the allocation function and the functions h_i defined following Lemma 1.

Lemma 2. *Let χ be an incentive compatible mechanism and define*

$$\psi(\theta_i) = \begin{cases} \frac{-F(\theta_i)}{f(\theta_i)} & \text{if } \theta_i \le \hat{\theta}_i, \\ \frac{1-F(\theta_i)}{f(\theta_i)} & \text{otherwise.} \end{cases}$$

Then for every history h^t we have

$$U_{h^t}(\chi) = \int_\Theta \big[\psi(\theta_1) + \psi(\theta_2)\big] x(\theta) dF(\theta) + \int_{\Theta_1} h_2(\theta_1) dF(\theta_1) + \int_{\Theta_2} h_1(\theta_2) dF(\theta_2). \tag{4}$$

Proof. First note that

$$U_{h^t}(\chi) = \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\overline{\theta}} \Big[\theta_1 x(\theta) + \theta_2 x(\theta) + w_1(\theta) + w_2(\theta)\Big] dF(\theta_2) dF(\theta_1), \tag{5}$$

and by Lemma 1

$$w_i(\theta) = \int_{\hat{\theta}_i}^{\theta_i} x(\beta, \theta_{-i}) d\beta - \theta_i x(\theta) + h_i(\theta_{-i}). \tag{6}$$

Using integration by parts, we first rewrite the term

$$\begin{aligned} &\int_{\underline{\theta}}^{\overline{\theta}} \left[\int_{\hat{\theta}_i}^{\theta_i} x(\beta, \theta_{-i}) d\beta\right] f(\theta_i) d\theta_i \\ &= \left[\int_{\hat{\theta}_i}^{\overline{\theta}} x(\beta, \theta_{-i}) d\beta \underbrace{F(\overline{\theta})}_{=1} - \int_{\hat{\theta}_i}^{\underline{\theta}} x(\beta, \theta_{-i}) d\beta \underbrace{F(\underline{\theta})}_{=0}\right] - \int_{\underline{\theta}}^{\overline{\theta}} x(\theta_i, \theta_{-i}) F(\theta_i) d\theta_i \\ &= \int_{\hat{\theta}_i}^{\overline{\theta}} \frac{1 - F(\theta_i)}{f(\theta_i)} x(\theta) dF(\theta_i) + \int_{\underline{\theta}}^{\hat{\theta}_i} \frac{-F(\theta_i)}{f(\theta_i)} x(\theta) dF(\theta_i). \end{aligned} \tag{7}$$

Now plug (6) into (5) and use (7) to complete the proof. □

The next lemma implies, together with Condition 1, that the first part of (4) is maximized by a constant allocation function whenever only one part of the function ψ is considered.

Lemma 3. *Suppose that $\psi(\theta_1, \theta_2)$ is non-increasing in θ_1 and θ_2, and that $\int \psi(\theta) dF(\theta) < \infty$. Then the problem*

$$\begin{aligned} \max_x \; & \int_a^b \int_c^d \psi(\theta_1, \theta_2) \cdot x(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) \\ \text{s. t. } & x \text{ is non-decreasing in } \theta \\ & 0 \le x(\theta) \le 1 \end{aligned}$$

is solved optimally either by setting $x^*(\theta) = 1$ or $x^*(\theta) = 0$.

Proof. Suppose to the contrary that there exists a function $\hat{x}(\theta)$ that achieves a strictly higher value. Define $x'(\theta_1, \theta_2) := \frac{1}{F(d)-F(c)} \int_c^d \hat{x}(\theta_1, \beta) dF(\beta)$. This function is feasible for the above problem given that $\hat{x}$ is feasible and, by Chebyshev's inequality, for all θ_1,

$$\begin{aligned}\int_c^d & \psi(\theta_1, \theta_2)\hat{x}(\theta_1, \theta_2) dF(\theta_2) \\ & \leq \int_c^d \psi(\theta_1, \theta_2) dF(\theta_2) \frac{1}{F(d) - F(c)} \int_c^d \hat{x}(\theta_1, \theta_2) dF(\theta_2) \\ & = \int_c^d \psi(\theta_1, \theta_2) x'(\theta_1, \theta_2) dF(\theta_2).\end{aligned}$$

Since this inequality holds for every θ_1, we also have

$$\int_a^b \int_c^d \psi(\theta_1, \theta_2)\hat{x}(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) \leq \int_a^b \int_c^d \psi(\theta_1, \theta_2) x'(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1).$$

Defining $x''(\theta_1, \theta_2) = \frac{1}{F(b)-F(a)} \int_a^b x'(\theta_1, \theta_2) dF(\theta_1)$ and again applying Chebyshev's inequality as above, we get that

$$\int_a^b \int_c^d \psi(\theta_1, \theta_2) x'(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) \leq \int_a^b \int_c^d \psi(\theta_1, \theta_2) x''(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1).$$

Since the objective function is linear in x, the constant function x'' is weakly dominated by either $x \equiv 1$ or $x \equiv 0$, contradicting the initial claim. □

Proof of Theorem 2

Proof. We establish the preconditions of Proposition 1. Fix an arbitrary history h_t and consider the stage mechanism (x, w) employed after this history. Let $\overline{w} := max_\theta\{w_1(\theta) + w_2(\theta)\}$ and let θ_w be an optimizer. We normalize w such that $w_1(\theta_w) = w_2(\theta_w) = 0$ by decreasing w_i by $w_i(\theta_w)$ for all i. This does not affect incentive compatibility. After the normalization we have

$$w_1(\theta) + w_2(\theta) \leq 0.$$

We start with some preliminaries where we derive a set of inequalities that are satisfied by every incentive compatible stage mechanism for which the above inequality holds.

Preliminaries:

Set $(\hat{\theta}_1, \hat{\theta}_2) := (\overline{\theta}, \underline{\theta})$, let h_i denote the resulting redistribution functions implied by Lemma 1 and define $g_i(\theta) := \theta_i x(\theta) - \int_{\hat{\theta}_i}^{\theta_i} x(\beta, \theta_{-i}) d\beta$. It follows from Lemma 1 that $w_i(\theta) = -g_i(\theta) + h_i(\theta_{-i})$. Let $h^* := \max_\theta\{h_1(\theta) + h_2(-\theta)\} - \overline{\theta}$ and θ^* be a maximizer. Normalize h such that $h_1(\theta^*) = h^* + \overline{\theta}$ and $h_2(-\theta^*) = 0$ by increasing $h_1(x_2)$ and decreasing $h_2(x_1)$ by $h_2(-\theta^*)$. The definition of h^* implies

$$h_1(\theta) + h_2(-\theta) \leq h^* + \overline{\theta} \qquad \text{for all } \theta, \tag{8}$$

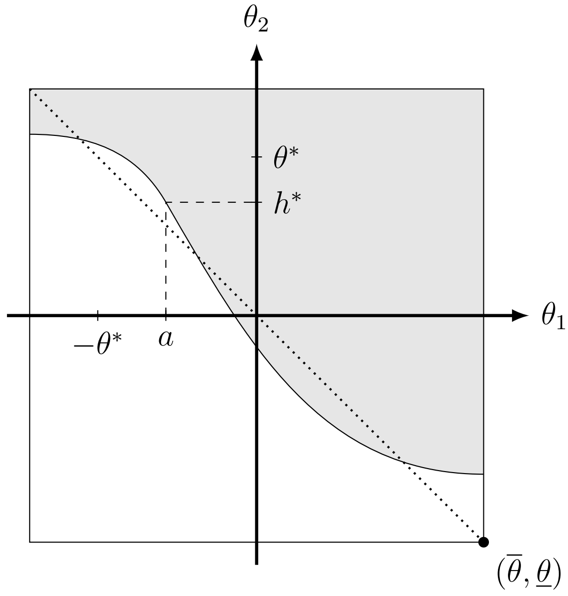

Figure 1: Proof of Theorem 2. The shaded area indicates the profiles θ where $x(\theta) = 1$.

and $w_1(\theta, -\theta) + w_2(\theta, -\theta) \leq 0$ implies

$$\begin{aligned} h_1(\theta) + h_2(-\theta) &\leq g_1(-\theta, \theta) + g_2(-\theta, \theta) \\ &= -\int_{\overline{\theta}}^{-\theta} x(\beta, \theta) d\beta - \int_{\underline{\theta}}^{\theta} x(-\theta, \beta) d\beta \\ &\leq \int_{-\theta}^{\overline{\theta}} x(\beta, \theta) d\beta \leq \overline{\theta} + \theta. \end{aligned} \tag{9}$$

By plugging θ^* into (9) and using the definition of h^*, it follows that $h^* \leq \theta^*$.

Define $a := \inf\{\theta_1 \mid x(\theta_1, h^*) = 1\}$. If there does not exist θ_1 such that $x(\theta_1, h^*) = 1$, set $a := \overline{\theta}$. Without loss we can assume that $a \geq -h^*$, since otherwise we can "mirror" the mechanism on the dotted line shown in Figure 1.[1] Let $\theta_1 \geq a$. Then expanding and rearranging $w_1(\theta_1, \theta^*) + w_2(\theta_1, \theta^*) \leq 0$ yields

$$\begin{aligned} h_2(\theta_1) &\leq -(h^* + \overline{\theta}) + g_1(\theta_1, \theta^*) + g_2(\theta_1, \theta^*) \\ &= -h^* - \overline{\theta} + \theta_1 - \int_{\overline{\theta}}^{\theta_1} x(\beta, \theta^*) d\beta + \theta^* - \int_{\underline{\theta}}^{\theta^*} x(\theta_1, \beta) d\beta \\ &= -h^* + \theta^* - \theta^* + h^* - \int_{\underline{\theta}}^{h^*} x(\theta_1, \beta) d\beta \\ &= -\int_{\underline{\theta}}^{h^*} x(\theta_1, \beta) d\beta, \end{aligned} \tag{10}$$

where in the second equality we made use of the fact that $x(\beta, \theta^*) = 1$ for $\beta \geq a$ and $x(\theta_1, \beta) = 1$ for $\theta_1 \geq a$, $h^* \leq \beta \leq \theta^*$ (see Figure 1). Similar arguments will be used more often in the equalities below.

[1] Let $(x^{\#}, w^{\#})$ be the mirrored mechanism, then $x^{\#}(\theta_1, \theta_2) = 1 - x(-\theta_2, -\theta_1)$, $w_i^{\#}(\theta_1, \theta_2) = w_{-i}(-\theta_2, -\theta_1)$. The new mechanism is IC iff. the old mechanism is IC and by our symmetry assumptions the mirrored mechanism yields the same welfare. Also, h^* and θ^* will not be changed by this operation.

Define $b := \inf\{\theta_2 \mid x(-h^*, \theta_2) = 1\}$ (if there is no θ_2 such that $x(-h^*, \theta_2) = 1$, set $b := \overline{\theta}$) and let $\theta_2 \leq b$. Then $w_1(-\theta^*, \theta_2) + w_2(-\theta^*, \theta_2) \leq 0$ implies

$$\begin{aligned} h_1(\theta_2) &\leq g_1(\theta^*, \theta_2) + g_2(\theta^*, \theta_2) \\ &= 0 - \int_{\overline{\theta}}^{-\theta^*} x(\beta, \theta_2) d\beta - \int_{\underline{\theta}}^{\theta_2} x(-\theta^*, \beta) d\beta \\ &= \int_{-\theta^*}^{\overline{\theta}} x(\beta, \theta_2) d\beta. \end{aligned} \tag{11}$$

Since by Lemma 1 an incentive compatible stage mechanism is completely determined by x and h, we will in the following change x and h in a number of consecutive steps while making sure that x stays monotone and we never decrease the welfare $U^{h_t}(x, h) := U^{h_t}(x, w)$. At the end of the proof we will make sure that the resulting mechanism is admissible. First, we increase $h_2(\theta_1)$ for $\theta_1 \geq a$ and $h_1(\theta_2)$ for $\theta_2 \leq b$ until (10) and (11) hold with equality since this trivially weakly increases welfare.

Step 1:

In this step we will change the variables $x(\theta)$ with $\theta \in A := \{(\theta_1, \theta_2) \mid \theta_1 \geq a, \theta_2 \leq h^*\}$, $h_2(\theta_1)$ with $\theta_1 \geq a$ and $h_1(\theta_2)$ with $\theta_2 \leq h^*$. If we change h_1 and h_2 such that (11) and (10) continue to hold with equality, we can express changes of all the variables in terms of changes of x. Making use of the fact that for $\theta_2 \leq h^*$, (11) is equivalent to

$$h_1(\theta_2) = \int_a^{\overline{\theta}} x(\beta, \theta_2) d\beta,$$

and by substituting (11) and (10), we can rewrite the the part of U_{h^t} that depends on changes of the variables $x(\theta)$ for $\theta \in A$ as

$$\begin{aligned} &\int_a^{\overline{\theta}} \int_{\underline{\theta}}^{h^*} \left[\frac{-F(\theta_1)}{f(\theta_1)} + \frac{1 - F(\theta_2)}{f(\theta_2)} \right] x(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1) \\ &\qquad + \int_{\underline{\theta}}^{h^*} \int_a^{\overline{\theta}} x(\beta, \theta_2) d\beta \, dF(\theta_2) - \int_a^{\overline{\theta}} \int_{\underline{\theta}}^{h^*} x(\theta_1, \beta) d\beta \, dF(\theta_1) \\ &= \int_a^{\overline{\theta}} \int_{\underline{\theta}}^{h^*} \left[\frac{1 - F(\theta_1)}{f(\theta_1)} + \frac{-F(\theta_2)}{f(\theta_2)} \right] x(\theta_1, \theta_2) dF(\theta_2) dF(\theta_1). \end{aligned}$$

Lemma 3 implies that this term is maximized by setting $x(\theta) = 0$ or 1 for $\theta \in A$. To see that we cannot gain by setting $x(\theta) = 1$ we bound

$$\begin{aligned} U_{h^t}(1, h) &= \int_a^{\overline{\theta}} \int_{\underline{\theta}}^{h^*} \left[\frac{1 - F(\theta_1)}{f(\theta_1)} + \frac{-F(\theta_2)}{f(\theta_2)} \right] dF(\theta_2) dF(\theta_1) \\ &= \int_{\underline{\theta}}^{-a} \int_{\underline{\theta}}^{h^*} \left[\frac{F(\theta_1)}{f(\theta_1)} + \frac{-F(\theta_2)}{f(\theta_2)} \right] dF(\theta_2) dF(\theta_1) \\ &= \int_{\underline{\theta}}^{-a} \int_{-a}^{h^*} \left[\frac{F(\theta_1)}{f(\theta_1)} + \frac{-F(\theta_2)}{f(\theta_2)} \right] dF(\theta_2) dF(\theta_1) \\ &\leq 0 = U_{h^t}(0, h). \end{aligned}$$

Here, the second equality is due to the symmetry of F around zero, the third equality is because the integral over $[\underline{\theta}, -a] \times [\underline{\theta}, -a]$ vanishes, and the inequality is due to log-concavity of F and the fact that $-a \leq h^*$. Hence, we weakly increase welfare by setting x to 0 or 1 in A and h_1 and h_2 according to (11) and (10), respectively.

Step 2:

For this step define the set $B = \{\theta_1 > -h^*, \theta_2 > h^* \mid x(\theta_1, \theta_2) = 0\}$. Set $x(\theta) = 1$ for $\theta \in B$ and $h_1(\theta_2) = h^* + \overline{\theta}$ for all θ_2 for which there is a θ_1 such that $(\theta_1, \theta_2) \in B$. We claim that this does not decrease U_{h^t}. Since allocative efficiency improved in this step, we only need to check that the sum of promised continuations increased. First, let $(\theta_1, \theta_2) \in B$. Then (11) is equivalent to

$$h_1(\theta_2) = \int_{-h^*}^{\overline{\theta}} x(\beta, \theta_2) d\beta.$$

Continuations before this change are given by

$$h_2(\theta_1) + h_1(\theta_2) + \int_{\overline{\theta}}^{\theta_1} x(\beta, \theta_2) d\beta = h_2(\theta_1) + \int_{-h^*}^{\overline{\theta}} x(\beta, \theta_2) d\beta + \int_{\overline{\theta}}^{\theta_1} x(\beta, \theta_2) d\beta = h_2(\theta_1).$$

After the change we get

$$h_2(\theta_1) + h^* + \overline{\theta} - \theta_1 + \int_{\overline{\theta}}^{\theta_1} \underbrace{x(\beta, \theta_2)}_{=1} d\beta - \theta_2 + \int_{h^*}^{\theta_2} \underbrace{x(\theta_1, \beta)}_{=1} d\beta = h_2(\theta_1).$$

Fixing $(\theta_1, \theta_2) \in B$, the claim can similarly be shown for points of the form (θ_1', θ_2) and (θ_1, θ_2') where $\theta_2' > \theta_2$.

Step 3:

We claim that setting $x(\theta) = 1$ or $x(\theta) = 0$ for $\theta \in [\underline{\theta}, -h^*] \times [h^*, \overline{\theta}]$ increases U_{h^*}. This follows from the fact that, since, ignoring the part which depends on h_i, the objective function in the area where we change x has the form required by Lemma 3. Symmetry implies that $x(\theta) = 0$ gives the same welfare as $x(\theta) = 1$.

Step 4:

Note that the original mechanism satisfied

$$h_1(-\theta) + h_2(\theta) \leq h^* + \overline{\theta}.$$

Therefore, welfare is not decreased by setting $h_2(\theta) := 0$ and $h_1(-\theta) = h^* + \overline{\theta}$ for $\theta \leq -b$.

Note that the changed mechanism satisfies $w_1(\theta, -\theta) + w_2(\theta, -\theta) \leq 0$: For $a \leq \theta$ this holds as we assumed (10) and (11) to be binding in Step 1, hence $g_1(\theta, -\theta) = g_2(\theta, -\theta) = h_1(-\theta) = h_2(\theta) = 0$. For $-h^* \leq \theta \leq a$, this holds as continuations weren't changed for these values (changed Pivot payments were offset by changes in the h functions, as (11) was assumed to hold with equality in Step 1). For $-b \leq \theta \leq -h^*$ this holds as constraints were assumed to bind in Step 2. For $\underline{\theta} \leq \theta \leq -b$ this holds as $h_1(-\theta) + h_2(\theta) \leq h^* + \overline{\theta} = g_1(\theta, -\theta) + g_2(\theta, -\theta)$.

The fact that $w_1(\theta, -\theta) + w_2(\theta, -\theta) \leq 0$ implies that $h_1(-\theta) + h_2(\theta) \leq g_1(\theta, -\theta) + g_2(\theta, -\theta)$. We can increase h so that equality holds, thereby again improving the mechanism, ending up with the following stage mechanism:

$$x(\theta) = \begin{cases} 1 & \text{if } \theta_2 \geq h^* \\ 0 & \text{else,} \end{cases}$$
$$h_1(\theta_2) = \begin{cases} 0 & \text{if } \theta_2 \leq h^* \\ h^* + \overline{\theta} & \text{else,} \end{cases}$$
$$h_2(\theta_1) = 0.$$

We call this class of mechanisms *phantom dictatorship with parameter h^**.

Step 5:

So far we have shown that every stage mechanism can be modified until it is a phantom dictatorship while weakly improving welfare. To prove that for every stage mechanism there is a simple voting stage mechanism with weakly higher welfare, we show that simple voting weakly welfare-dominates every phantom dictatorship: Indeed, the optimal phantom dictatorship is given by the parameter $h^* = \mathbf{E}[\theta]$. Therefore, symmetry of F around 0 implies that the optimal phantom dictatorship is characterized by $h^* = 0$, which has the same aggregate welfare as unanimity voting.

The voting stage mechanism we have constructed so far has the continuations profile $w_1(\theta) = w_2(\theta) = 0$ for all θ. It remains to show that this mechanism is admissible. But this follows from the fact that $(0, 0)$ was an implementable continuation profile of the original mechanism (namely, at the type profile θ_w). We therefore established the conditions for Proposition 1, which completes the proof of the theorem. □

References

Athey, S. and Miller, D. A. (2007). Efficiency in repeated trade with hidden valuations, *Theoretical Economics* **2**(3): 299–354.

Bergemann, D. and Välimäki, J. (2010). The dynamic pivot mechanism, *Econometrica* **78**(2): 771–789.

Buchanan, J. M. and Tullock, G. (1962). *The calculus of consent: logical foundations of constitutional democracy*, University of Michigan Press.

Carrasco, V. and Fuchs, W. (2011). From equals to despots: The dynamics of repeated decision taking in partnerships with private information, *Technical report*.

Casella, A. (2005). Storable votes, *Games and Economic Behavior* **51**(2): 391–419.

Hortala-Vallve, R. (2010). Inefficiencies on linking decisions, *Social Coice and Welfare* **34**: 471–486.

Hortala-Vallve, R. (2012). Qualitative voting, *Journal of Theoretical Politics* **24**: 526–554.

Jackson, M. O. and Sonnenschein, H. F. (2007). Overcoming incentive constraints by linking decisions, *Econometrica* **75**(1): 241–257.

Miller, D. (2012). Robust collusion with private information, *Review of Economic Studies* **79**(2): 778–811.

Myerson, R. B. (1986). Multistage games with communication, *Econometrica* **54**(2): 323–358.

Pavan, A., Segal, I. and Toikka, J. (2008). Dynamic mechanism design: Incentive compatibility, profit maximization and information disclosure, *Carlo Alberto Notebooks 84*, Collegio Carlo Alberto.