UNIVERSITY OF Southampton
School of Social Sciences

Economics Division
University of Southampton
Southampton SO17 1BJ, UK

Discussion Papers in
Economics and Econometrics

Title: Efficient Voting with Penalties

Author : Maksymilian Kwiek (University of Southampton)

No. 1419

ISSN 0966-4246

Efficient Voting with Penalties

Maksymilian Kwiek*

2014

Abstract

Simple majority does not reflect the intensity of voters' preferences. This paper presents an efficient collective choice mechanism when the choice is binary and the designer may use non-trasferable punishments to persuade agents to reveal their private information. The designer faces a dilemma – a punishment may induce a more correct choice, but its cost is socially wasteful. The efficient mechanism is a weighted majority. Weight of each individual is known ex ante and no punishments applied if preferences are relatively homogenous. Eliciting types through punishments in order to construct type-specific weights should occur if preference intensity is relatively heterogeneous, or if voters preferences represent a larger population.

JEL classification: D71, D82

1 Introduction

Simple majority is the benchmark of voting systems. The seminal contribution of Rae (1969) indicates that in case of a binary choice, simple majority has good normative properties. The key assumption in their work is that the intensity of voters' preferences in favor of an alternative is the same. This paper revisits the question of normative performance of various voting systems under the assumption that voters may differ in the intensity of preferences, and that interpersonal comparisons are possible.

As in any voting mechanism, monetary transfers are not allowed. The main premise of this study is that voters may be punished, and thus punishments can be a part of the mechanism's design. A benevolent designer

*Department of Economics, Univeristy of Southampton, SO17 1BJ, UK, M.Kwiek@soton.ac.uk

faces a dilemma – punishments can motivate voters to reveal their preferences, facilitating a correct assignment, but at the same time they represent an unrecoverable welfare loss.

There are real-life examples in which preference intensity can be expressed in a wasteful way in a collective choice context. Consider repeated voting with a supermajority – voters are locked until the support for one of the alternatives reaches some supermajority, or the support for the other one falls below some minimal fraction of voters. A form of a War of Attrition ensues, in which waiting time imposes a cost on participating voters, and this cost is a pure welfare loss. A conclave with supermajority 2/3 is used to select a Pope – the leader of the Roman Catholic Church. Many hiring committees vote repeatedly until one of the candidates gathers enough support. Jury trials use a similar procedure of repeated voting until all jurors agree on a verdict. Influencing political decisions through lobbying also falls into the category of environments studied here: non-cooperative lobbying efforts, wasteful as such, translate into a higher probability of winning. Once it is recognized that imposing wasteful efforts is available in the designer's tool-box, one may contemplate very exotic voting rules, furnished with completely artificial punishments.

This paper characterizes the efficient mechanism for both net efficiency, which includes the cost of penalties in welfare calculation, and allocative efficiency, which does not (and any convex combination between the two). The main result is that regardless of whether the voting mechanism should seek to reveal and use voters' preferences, an efficient mechanism always takes a form of *weighted majority*. The main dilemma faced by the designer revolves around the question of whether these weights should be type-dependent, and thus whether voters should be incentivised to reveal their types.

The solution to this dilemma depends on the level of dispersion, or heterogeneity in voters' preferences. At this stage, it is useful to distinguish between the *ex post heterogeneity*, which refers to the dispersion of the realized preferences, and the *ex ante heterogeneity*, which is concerned with how likely those randomly chosen preferences exhibit high dispersion. For example, the classical environment studied by Rae (1969) is bipolar; namely, some agents support one alternative while others support the other one, but the distribution of preferences, conditional on supporting one alternative, is degenerate at one point. There in no heterogeneity, ex ante or ex post. However, if this distribution is non-degenerate, and thus there is some ex ante heterogeneity, the actually realized preferences may be dispersed or not, generating high or low ex post heterogeneity.

If net efficiency is the designer's objective, then the level of ex ante heterogeneity in preferences is the key determinant of whether extracting information from the voters is desirable.

For example, if intensity of preferences is not ex ante heterogeneous enough, in the sense of the reciprocal hazard rate being decreasing, then improving the likelihood of selecting a better alternative is not worth the associated social cost of screening. Consequently, the efficient mechanism is a weighted majority with *weights known ex ante*. It does not extract any information and does not apply any penalties. In addition to this, if the environment is symmetric, then these weights are equal, and the mechanism is the old-fashioned simple majority. On the other hand, if intensity of preferences is ex ante heterogeneous enough, meaning that the reciprocal hazard rate is increasing, the efficient mechanism extracts information from the voters and uses it to construct *type-dependent weights*. Penalties are needed to create incentives for the voters to reveal their types. A simple mechanism of a Vickrey–Clarke–Groves type implementing the net efficient outcome in dominant strategies is easy to characterize.

If the designer concentrates only on allocative efficiency, for example when voters in the committee represent a wider population, then the cost of the penalty does not matter in welfare calculation. Weights ought to be type-dependent.

This paper also explores the role of ex post realization of preference intensities. Firstly, in any efficient mechanism *more agents* supporting one alternative increases the chances of selecting this alternative.[1] This is a strong result as it does not depend on any extra conditions. Secondly, if the reciprocal hazard rate is increasing, so that ex ante heterogeneity is high enough for the preference intensity to matter, then *higher preference intensity* leads to higher chances of this alternative being selected. Thirdly, if the reciprocal hazard rate is increasing, then ex post heterogeneity of preferences also matters, but the relationship is difficult to capture intuitively as it depends on the curvature of the reciprocal hazard rate. Having voters with *more dispersed realized preferences* supporting one alternative increases the likelihood of selecting their alternative, if the reciprocal hazard rate is convex. Having voters with *less dispersed realized preferences* supporting one alternative increases the likelihood of selecting their alternative, if the reciprocal hazard rate is concave.

Also, this paper proposes that some mechanisms with penalties which are not efficient but are realistic, may perform better than majority with no penalties. An example of a simple conclave presented in the last section achieves a higher welfare than simple majority, even if the cost of penalties is fully taken into welfare account.

[1]In the following weak sense: sometimes increases, and never decreases

Literature, further results and assumptions

Multiplayer decision mechanisms with penalties are used in many real-life applications, and there is literature that analyzes their positive or normative properties. These mechanisms are generally not efficient, but it is worth knowing the magnitude of the welfare loss associated with using these mechanisms relative to the efficient ones. In fact, this question motivated the current study, and an example is presented in the last section. Ponsati and Sakovics (1996) are concerned with properties of equilibria in multiplayer Wars of Attrition with supermajority. Kwiek (2014) is focused on welfare performance of these kind of mechanisms. In contrast to Kwiek (2014), the model below considers setups with private information, all incentive compatible mechanisms, and it looks for efficient mechanisms that take into account the penalty cost. Although Kwiek (2014) and the model below are different, the gist of comparative statics is similar, namely in both models simple majority is inefficient, provided that preference intensities are ex ante heterogeneous enough, or voters' decision has an externality on a wider population. Kwiek et al. (2015) experimentally study positive and welfare properties of conclaves.

The literature on efficient social choice with privately known preferences can be divided into two branches: those discussing environments permitting transfers and those where transfers are not allowed. The present study ties together those two strands in the following sense. On one hand, the efficient mechanism will take a form of a generalized majority rule, close to the results known from the no-transfers literature. On the other hand, agents may be incentivised by facing a prospect of suffering penalties, very much like in the case when transfers are allowed. For example, incentive compatibility constraints are similar to those in auction literature, and many of these techniques can be adopted to the present case. The rest of this section elaborates on how these two strands of literature connect with the results and assumptions of the model.

The analysis below draws heavily from optimal auction theory (e.g. Vickrey–Clarke–Groves, Myerson (1981), McAfee and McMillan (1992)). This literature postulates that individuals have single-dimensional private information about the values of the alternatives and their utilities are quasilinear in money. One benchmark assumption is that private information is statistically independent across individuals. We adopt this assumption too, keeping in mind that it is questionable in many real-life cases of committee voting. Committees may be viewed as a platforms for deliberation, designed to allow pooling dispersed information. However, in the model below, an agent cannot learn anything about her values from other agents.

McAfee and McMillan (1992), followed by a series of more recent papers by Hartline and Roughgarden (2008), Yoon (2011), Condorelli (2012) and Chakravarty and Kaplan (2013) are closer to the analysis below. They share one feature, namely that the participants' bids, or efforts, are a social waste. Thus a benevolent mechanism designer faces a trade-off. They can either elicit information to assign the good correctly and waste those efforts, or forgo the former to save the latter. From the perspective of the current study, the following specific points ought to be mentioned. Yoon (2011) allows an intermediate case between allocative and net efficiency, which, in the context of a voting environment, can be interpreted naturally as a type of representative democracy; the paper also provides comparative statics in which ex ante heterogeneity takes a center stage. Condorelli (2012) considers an interesting problem of interdependent values (although statistically independent signals), and identifies the whole Pareto frontier rather than just utilitarian welfare. Chakravarty and Kaplan (2013) assume that agents' private information may also affect their individual marginal cost of penalty.

The key difference is that these papers look at an assignment problem – how to allocate a rivalrous good among competing individuals. In contrast, the study below considers a collective choice problem, giving rise to different forms of externalities among the agents. For example, in the spirit of the classical Samuelson condition for public goods, the efficiency condition involves a sum of individual benefits over all agents supporting an alternative, although – since this is an incentive problem – these individual benefits are net of costs of incentive provision.

Another difference is that private information is multidimensional in the model below. Apart from the preference intensity, individual voters are assumed to have private information about two other aspects: which alternative they prefer, and the marginal cost of punishment. As far as the the latter is concerned, Chakravarty and Kaplan (2013) also assume that agents may differ in their individual marginal costs of effort, but here we assume that value and marginal cost are two different random variables, possibly statistically independent. The environment is still relatively tractable so incentive compatibility can be framed within the usual integrability condition.

The second strand of literature – the one emerging from political economy and voting without transfers – starts with the aforementioned seminal paper by Rae (1969).[2] More recently, some articles investigate how preference intensities can be reflected in voting mechanisms. Casella (2005) proposes a procedure in which voters can store their votes for future use in a sequence of elections, thus creating an opportunity to use accumulated votes on issues

[2] See also May (1952).

of a particular importance. A similar idea of linking different social choice problems to enhance incentives appears in Jackson and Sonnenschein (2007).

The literature closer to the model below analyzes environments with one independent voting problem. The paper by Azrieli and Kim (2014) asks a question what mechanism is efficient in the binary choice environment without transfers. The only difference with the current paper is that they do not permit the use of penalties. Their efficient mechanism is a weighted majority. In the current paper, when ex ante heterogeneity of preference intensities is low, the penalties should not be used and therefore the mechanism is exactly the same as the one of Azrieli and Kim (2014). Otherwise, penalties ought to be used to extract information about types, and consequently the weights are type-dependent.

Drexl and Kleiner (2013) is a study that occupies the gap between transfer and no-transfer literature. They assume that transfers are allowed, but due to anonymity and other assumptions, monies collected cannot be redistributed back to the voters. Consequently, potential payments of the voters have to be interpreted as penalties like in the model below. They focus on the case low ex ante heterogeneity, thus characterizing the efficient mechanism as weighted majority without employing any penalties. In contrast to this, the current study attempts to identify the cases under which nontrivial penalties are used efficiently, such as in high ex ante heterogeneity case, or when voters' decision has an externality on a wider population. Penalties are not interpreted as monetary payments, and thus the marginal cost is not necessary one, or even commonly known. The current model generalizes in many aspects a result reported in Kwiek and Zhang (2013).

There are other papers which deal with efficient mechanisms when transfers are not possible. For example, Schmitz and Tröger (2012) consider correlated signals, while Apesteguia at al. (2011) and Gershkov at al. (2014) analyze the case of more than two alternatives.

2 Environment and welfare criteria

Physical environment consist of the following elements. One *alternative* is to be selected from the set $\{A, B\}$, and k will denote a generic alternative. There are $n \geq 2$ *voters*. The key postulate is that there exists a way to penalize individuals. Namely, assume that an individual voter i can be forced to suffer c_i units of a *penalty*. One may hypothesize that there are different types of penalties (waiting, tedious tasks, electric shocks, etc.), and the mechanism designer may have an option to select one, but for now the penalty type is fixed.

Preferences

Preferences of each voter are represented by three components. The first element, $a_i \in \{A, B\}$, is called *direction* of preferences and describes which alternative voter i prefers. Sometimes we will refer to the set of players supporting k as party k. The second element describes relative *intensity* of this preference. Intensity is denoted by $x_i \geq 0$, interpreted as how many units of penalty this voter is willing to endure at most, and still select her preferred alternative over the other alternative if it is received without penalty. The third element is z_i, the *marginal cost* of penalty to individual i.

The ultimate net utility of voter i who is asked to suffer c_i units of penalty is

$$\begin{cases} z_i (x_i - c_i) & \text{if her preferred alternative is selected} \\ -z_i c_i & \text{otherwise} \end{cases}$$

Notice that the top line can also be written more conventionally as $v_i - z_i c_i$, where $v_i = z_i x_i$ could be called an absolute intensity of preferences. As it will become clear later, x_i and z_i will be values of some random variables which may or not be statistically independent. It is also assumed that the expected value of v_i conditional on x_i is increasing in x_i.

Information

The final element that has to be specified in order to complete the description of the model is the information available to the voters and to the mechanism designer. We assume that all three elements of voter's preferences are her private information. That is, voter i knows the realization of random variables (a_i, x_i, z_i) called the *type* of voter i, but other voters and the mechanism designer know only its joint distribution.

Types are independent across individuals. The intensity and marginal cost might not be independent for a given individual; we assume that the p.d.f. and c.d.f. of x_i conditional on z_i is $f_i(x_i|z_i)$ and $F_i(x_i|z_i)$, while the marginal p.d.f. of x_i is $f_{X_i}(x_i)$. For any z_i, function $F_i(x_i|z_i)$ is continuous and strictly increasing for $x_i \in [0, \bar{x}_i(z_i)]$, where $\bar{x}_i(z_i)$ could be infinity. It is assumed that a_i and (x_i, z_i) are statistically independent, but this is only for transparency of notation.

Welfare criteria

In the leading case, this paper will adopt utilitarian welfare criterion as a benchmark capturing interpersonal comparisons of intensity of preferences.

That is, social welfare is the sum of individual utilities. However, checking any point on the Pareto Frontier will not pose any difficulty, see footnote 3.

There will be two special cases of interest: net welfare and allocative (gross) welfare. Allocative welfare achieved by a mechanism is equal to $\sum_{i \in k} v_i$, if alternative k is selected. According to this criterion, A should be selected if and only if $\sum_{i \in A} v_i > \sum_{i \in B} v_i$, regardless of the penalties endured by the individuals. Thus, the type of the penalty applied to individuals is inconsequential. On the other hand, net welfare is defined as

$$\sum_{i \in k} v_i - \sum_{i=1}^{n} z_i c_i$$

if alternative k is selected, and $c_1, ..., c_n$ units of penalty are applied. This notion takes into account the costs of voting suffered by the participants.

Net welfare seems to be a more correct measure, since it takes into account all social benefits and costs of selecting an alternative. However, there could be cases when the cost of penalty should be ignored. Arguably, the active voters may be just representatives of *districts* of voters who do not participate directly. Consider a model in which (a_i, x_i, z_i) is a district-specific type. That is, districts are internally homogeneous, but they differ from each other. Each district i sends a representative with type (a_i, x_i, z_i) to sit in the committee. In particular, let the proportion of district's inhabitants who go to the committee be λ, and let $1-\lambda$ of them stay home. In other words, only fraction λ of the population has to pay the penalty associated with voting, while the rest free rides on those who are subjected to this penalty. The correct welfare is $\sum_{i \in k} v_i - \lambda \sum_{j=1}^{n} z_j c_j$. If $\lambda = 1$, then this expression results in net welfare, and if the fraction of representatives in the district is negligible, $\lambda = 0$, then this expression becomes allocative welfare as the other polar case.[3]

The interest below will be in incentive compatible mechanisms maximizing ex-ante welfare for a given λ.

3 Incentive compatible mechanisms

The mechanism designer specifies a Bayesian game, in which the voters play a Bayesian Nash equilibrium. By revelation principle any mechanism can be

[3] Instead of assuming that each individual or district is of the same size, one may consider a more general formulation, for example the one in which the size of the population in district i is m_i, and the fraction of representatives is a district-specific λ_i. In what follows, the weights $(\lambda, 1-\lambda)$ would be replaced by $(\lambda_i m_i, (1-\lambda_i) m_i)$. Vector $(m_i, ..., m_n)$ could also be interpreted as weights attached to different districts in a welfare function, tracing different points on the Pareto Frontier.

mimicked by a direct revelation and incentive compatible mechanism. In such a mechanism each player simultaneously reports their type, having incentives to report truthfully, and then the mechanism executes the outcome.

Let $a = (a_1, ..., a_n)$ be the profile of true preference directions, let $x = (x_1, ..., x_n)$ be the profile of true intensities and $z = (z_1, ..., z_n)$ the profile of true marginal costs. Let $\bar{a} = (\bar{a}_1, ..., \bar{a}_n)$ be the profile of reported preference directions, and let $\bar{x}$ and $\bar{z}$ be similar vectors of reported intensities and marginal costs. Reports may be different from true values. To shorten the notation, we will denote the triplet of types as $r_i = (a_i, x_i, z_i)$; likewise for reports and profiles: $\bar{r}_i = (\bar{a}_i, \bar{x}_i, \bar{z}_i)$, $r = (a, x, z)$, $\bar{r} = (\bar{a}, \bar{x}, \bar{z})$.

The direct revelation mechanism is a collection $(p_A, p_B, c_1, ..., c_n)$, where all elements are functions of the report profile $\bar{r}$. The mechanism works as follows. After the agents have reported their preferences to the mechanism, alternative k is eventually selected with probability $p_k(\bar{r}) \in [0, 1]$, where obviously $p_A(\bar{r}) + p_B(\bar{r}) = 1$. Then, agent i has to pay a non-negative and non-transferable expected penalty, $c_i(\bar{r}) \geq 0$. The expected utility of voter i, whose preference type is r_i while the report profile of types was $\bar{r}$, is

$$z_i \left(x_i p_{a_i}(\bar{r}) - c_i(\bar{r})\right)$$

Just to repeat, $p_{a_i}(\cdot)$ is the probability of selecting the alternative that voter i prefers.

Define $P_{a_i}(\bar{r}_i) = E_{r_{-i}} p_{a_i}(\bar{r}_i, r_{-i})$ to be the expected probability of selecting i's favorite alternative if her report is $\bar{r}_i$ (not necessarily truthful) when all other voters report truthfully, and $r_{-i} = (r_1, ..., r_{i-1}, r_{i+1}, ..., r_n)$; define $C_i(\bar{r}_i) = E_{r_{-i}} c_i(\bar{r}_i, r_{-i})$ to be the expected penalty in this situation; finally define

$$\tilde{\pi}_i(\bar{r}_i, r_i) = x_i P_{a_i}(\bar{r}_i) - C_i(\bar{r}_i) \tag{1}$$

and

$$\pi_i(r_i) = \tilde{\pi}_i(r_i, r_i) \tag{2}$$

Clearly, the expected utility of voter i, if all voters other than i report truthfully, is simply $z_i \tilde{\pi}_i(\bar{r}_i, r_i)$, and the expected utility when voter i reports truthfully along with everyone else is $z_i \pi_i(r_i)$.

Proposition 1. *The direct revelation mechanism is incentive compatible if and only if all conditions hold:*

1. $\frac{\partial}{\partial x_i} P_{a_i}(a_i, x_i, z_i) \geq 0$
2. $\frac{\partial}{\partial x_i} \pi_i(a_i, x_i, z_i) = P_{a_i}(a_i, x_i, z_i)$
3. $P_{a_i}(a_i, 0, z_i) \geq P_{a_i}(-a_i, 0, z_i)$ *and* $C_i(a_i, 0, z_i) = C_i(-a_i, 0, z_i)$

4. $\pi_i(a_i, x_i, \cdot)$ *and* $P_{a_i}(a_i, x_i, \cdot)$ *are constant functions.*

Points 1 and 2 of this Proposition ensures that individual i does not have incentives to misreport her intensity x_i and they are standard in mechanism design literature. Point 3 makes certain that voters have no incentives to misreport their directions. Condition 4 guarantees that voters will not lie about their marginal cost of penalty; it will also allow to simplify notation and write $\pi_i(a_i, x_i)$, $P_{a_i}(a_i, x_i)$, $C_i(a_i, x_i)$ and $p_{a_i}(a, x)$.

4 Welfare maximizing mechanisms

The initial focus of the analysis will be the case of net efficiency, $\lambda = 1$. It turns out that, as in McAfee and McMillan (1992), the core criterion for net efficiency involves the reciprocal hazard rate associated with the distribution of x_i. Specifically, let $H_i(\cdot)$ be defined as

$$H_i(x_i) = E_{Z_i} z_i \frac{1 - F_i(x_i|z_i)}{f_{X_i}(x_i)}.$$

For example, if z_i and x_i are statistically independent, then $H_i(\cdot)$ is the conventional reciprocal hazard ratio of x_i multiplied by a constant $z_i^e = E_{Z_i} z_i$.

One can express the payoff of member i of the committee only in terms of the allocation probability and the payoff of the indifferent type. This will be asserted in the following series of Lemmas.

Lemma 1. *Net expected payoff of voter i in an incentive compatible mechanism is*

$$N_i = z_i^e E_{a_i} \pi_i(a_i, 0) + E_{a_i, X_i} P_{a_i}(a_i, x_i) H_i(x_i) \tag{3}$$

Equation 3 may be altered to allow for a stronger allocative efficiency concern. That is, one can consider the case of $\lambda < 1$. Recall that each i may be interpreted as a district sending representatives who would experience the penalties employed by the mechanism, for the good of the population who stays at home. Clearly, the expected payoff of the representative of district i is given by equation 3. The payoff of each constituent who stays home and does not suffer the penalty is $z_i x_i P_{a_i}(a_i, x_i)$. Define the expected marginal penalty cost of agent i, conditional on x_i, as

$$Z_i(x_i) = \int_0^\infty z_i f_{Z_i|X_i}(z_i|x_i)\, dz_i$$

The expected payoff of a stay-at-home agent is therefore $E_{a,X} x_i Z_i(x_i) p_{a_i}(a, x)$.

Adding together the payoff of the representative, weighted by λ, and the payoff of the stay-at-home agent, weighted by $1-\lambda$, leads to a weighted average

$$H_{i\lambda}\left(x_i\right) = \lambda H_i\left(x_i\right) + \left(1-\lambda\right) x_i Z_i\left(x_i\right) \tag{4}$$

The expected payoff of district i has a very similar form to the one in equation 3.

Lemma 2. *Fix any $0 \leq \lambda \leq 1$. The expected payoff of district i in an incentive compatible mechanism is*

$$\lambda z_i^e E_{a_i} \pi_i\left(a_i, 0\right) + E_{a,X} p_{a_i}\left(a, x\right) H_{i\lambda}\left(x_i\right) \tag{5}$$

A number of specific cases can be described using just function $H_{i\lambda}\left(\cdot\right)$. But the general characterization requires further modifications. Thus, for every $q \in [0,1]$, define $\phi_i\left(q\right) = H_{i\lambda}\left(\left(F_{X_i}\right)^{-1}\left(q\right)\right)$, where $\left(F_{X_i}\right)^{-1}$ is the inverse of the marginal c.d.f. of x_i. Furthermore, let $\Phi_i\left(q\right) = \int_0^q \phi_i\left(s\right) ds$, and let

$$\Gamma_i\left(q\right) = conv\left(\Phi_i\left(q\right)\right) \tag{6}$$

be the convexification of Φ_i, and let $\gamma_i\left(q\right) = \Gamma_i'\left(q\right)$. Obviously, all these functions are derived from the primitives of the environment, hence exogenous, and known to the designer.

The next result also gives the payoff of district i, and therefore is another version of the previous Lemmata. However, function $H_{i\lambda}\left(\cdot\right)$, which may be decreasing, is replaced by $\gamma_i\left(F_{X_i}\left(\cdot\right)\right)$, which never is.

Lemma 3. *The expected payoff of district i can be written as*

$$N_i = \lambda z_i^e E_{a_i} \pi_i\left(a_i, 0\right) - \Lambda_i + E_{a_i,x_i} P_{a_i}\left(a_i, x_i\right) \gamma_i\left(F_{X_i}\left(x_i\right)\right) \tag{7}$$

where $\Lambda_i \geq 0$ is defined as

$$\Lambda_i = E_{a_i} \int_0^1 \frac{\partial P_{a_i}\left(a_i, x_i\right)}{\partial x_i} \left(\Phi_i\left(F_{X_i}\left(x_i\right)\right) - \Gamma_i\left(F_{X_i}\left(x_i\right)\right)\right) dx_i \tag{8}$$

The main result follows

Theorem 1. *Fix any $0 \leq \lambda \leq 1$. The mechanism that selects alternative A with probability 1 if and only if*

$$\sum_{i:a_i=A} \gamma_i\left(F_{X_i}\left(x_i\right)\right) \geq \sum_{j:a_j=B} \gamma_j\left(F_{X_j}\left(x_j\right)\right)$$

is efficient for this λ.

Theorem 1 looks obscure. The rest of this section will outline two special cases that lead to qualitatively different results. Broader ramifications will be discussed in the next section.

Decreasing $H_{i\lambda}(\cdot)$

It turns out that whether function $H_{i\lambda}(\cdot)$ is increasing or decreasing is of key importance to the type of the mechanism that should be selected by the efficiency-motivated designer.

Assume that function $H_{i\lambda}(\cdot)$ is decreasing. In this case, function ϕ_i is also decreasing, and thus Φ_i is concave. Its convexification in expression 6, Γ_i, is a straight line, and so γ_i is a constant equal to $\Phi_i(1)$. This, in turn, is equal to the expectation $E_{x_i} H_{i\lambda}(x_i)$, which, finally, is equal to the expectation of absolute intensity $v_i^e = E_{x_i} x_i Z_i(x_i)$ for any $0 \leq \lambda \leq 1$.[4]

Corollary 1. *Fix any $0 \leq \lambda \leq 1$ and suppose that $H_{i\lambda}(\cdot)$ is decreasing for every i. The mechanism that selects alternative A with probability 1 if and only if*

$$\sum_{i:a_i=A} v_i^e \geq \sum_{j:a_j=B} v_j^e$$

is efficient for this λ.

This condition can be assessed regardless of the reported intensities. Only reported directions matter. The mechanism gives potentially different weights to voters, but these weights $v_1^e, ..., v_n^e$ are known ex ante. Obviously, no penalty is needed.[5] If values v_i^e are the same for all voters, then the efficient mechanism is a classical simple majority.

Corollary 2. *Fix any $0 \leq \lambda \leq 1$ and suppose that $H_{i\lambda}(\cdot)$ is decreasing for every i. If agents' expected values v_i^e are constant for every i, then unweighted simple majority is efficient for this λ.*

Increasing $H_{i\lambda}(\cdot)$

On the other hand, if $H_{i\lambda}(\cdot)$ is increasing, then convexification in expression 6 is trivial and yields the original reciprocal hazard rate itself, $\gamma_i(F_{X_i}(x_i)) = H_{i\lambda}(x_i)$. We obtain

[4]This is a variant of a known result that the expectation of the reciprocal hazard rate is equal to the expectation of the random variable itself. Observe that

$$\begin{aligned} V_i^e &= \int_0^\infty x_i \int_0^\infty z_i f_{Z_i|X_i}(z_i|x_i)\, dz f(x)\, dx_i = \int_0^\infty x_i \left(E_{z_i} z_i f_i(x_i|z_i)\right) dx_i = \\ &= -x_i E_{z_i} z_i \left(1 - F_i(x_i|z_i)\right)|_0^\infty + \int_0^\infty \left(E_{z_i} z_i \left(1 - F_i(x_i|z_i)\right)\right) dx_i = H_i^e \end{aligned}$$

where integration by parts was used. Then $H_{i\lambda}^e = \lambda H_i^e + (1-\lambda) V_i^e$.

[5]Azrieli and Kim (2014) obtain the same criterion, but they assume that penalties cannot be used.

Corollary 3. *Fix any* $0 \leq \lambda \leq 1$ *and suppose* $H_{i\lambda}(\cdot)$ *is increasing for every* i. *The mechanism that selects alternative* A *with probability 1 if and only if*

$$\sum_{i:a_i=A} H_{i\lambda}(x_i) \geq \sum_{j:a_j=B} H_{j\lambda}(x_j) \tag{9}$$

is efficient for this λ.

This class of mechanisms asks agents to reveal not only their preference directions but also their relative preference intensity. There are many mechanisms that can realize allocation function in Corollary 3. Among them, there is a version of Vickrey–Clarke–Groves mechanism, henceforth VCG_λ, which implements outcomes in dominant strategies. This mechanism works by having the agents report their types, and the mechanism then selecting alternative k if and only if

$$\sum_{i:\bar{a}_i=k} H_{i\lambda}(\bar{x}_i) \geq \sum_{j:\bar{a}_j=-k} H_{j\lambda}(\bar{x}_j)$$

Those who supported the losing alternative are not subjected to any penalty. Those who supported the winning alternative k are subjected to penalty $c_{i\lambda}(\bar{a},\bar{x})$ defined by

$$H_{i\lambda}(c_{i\lambda}(\bar{a},\bar{x})) = \max\left\{H_{i\lambda}(0), \sum_{j:\bar{a}_j=-k} H_{j\lambda}(\bar{x}_j) - \sum_{\{j:\bar{a}_i=k, j\neq i\}} H_{j\lambda}(\bar{x}_j)\right\} \tag{10}$$

Proposition 2. *The mechanism* VCG_λ *is dominance solvable and it achieves the allocation function defined in Corollary 3.*

These two special cases of decreasing and increasing $H_{i\lambda}$ do not cover all possibilities. When $H_{i\lambda}$ is not monotone, then the VCG_λ constructed from γ_i could be used.

5 Comparative statics

This section will consider only a symmetric case, where probability distributions are independent of i.

Recall the main object of the analysis in expression 4. The focus now is in identifying factors that affect the shape of $H_\lambda(\cdot)$, which – as was seen above – is the key determinant of which mechanisms are efficient.

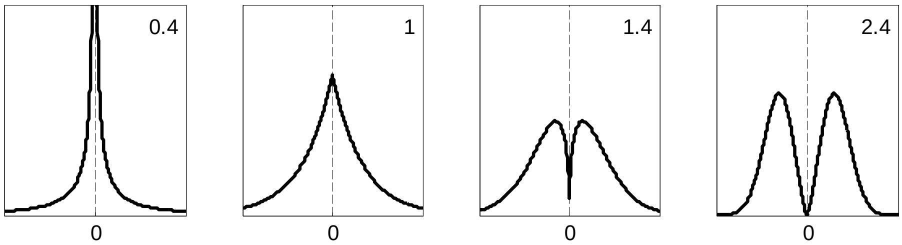

Figure 1: Weibull distribution for different shape parameters η

Statistical independence

Suppose that variables x_i and z_i are statistically independent, and without loss of generality, let $Ez_i = 1$.

Start with the case of net welfare, $\lambda = 1$. In other words, $H_\lambda(x_i) = H(x_i) = (1 - F_X(x_i))/f_X(x_i)$. Certainly, many distributions commonly used in examples have a decreasing H and therefore fall under the remit of Corollaries 1 and 2. It is sometimes claimed that this is the more likely case[6]. However, as soon as we establish a link between the slope of H and the notion of dispersion of values, this becomes an object of interpretable economic quality, which cannot be assumed away, a point made by Yoon (2011).

The argument will be illustrated with the following parametric example.

Example 1. Assume that voter's direction of preferences is equally likely to be A and B; intensity x_i (and v_i) is distributed according to Weibull distribution, the same for all i. That is, let $F_X(x_i) = 1 - \exp(-(x_i/\mu)^\eta)$, and therefore $H(x_i) = (\mu^\eta/\eta)\, x_i^{1-\eta}$, where $\eta > 0$ is the shape parameter, and μ is the scale parameter set up so that the mean is equal to one, $\mu = 1/\Gamma(1 + 1/\eta)$.

Weibull density function for four different shape parameters is presented in Figure 1. Shape parameter η controls ex ante heterogeneity of individual preferences. The greater this parameter, the more homogeneous within each party the intensity of preferences becomes, and the closer to the classical bipolar framework of Rae (1969) this environment is. Conversely, if this parameter is getting closer to zero, then the more heterogeneous each party is likely to become. In other words, the more likely it is that the individual

[6]For example, McAfee and McMillan (1992) write that "$H' \leq 0$ can be thought of as the more likely case."

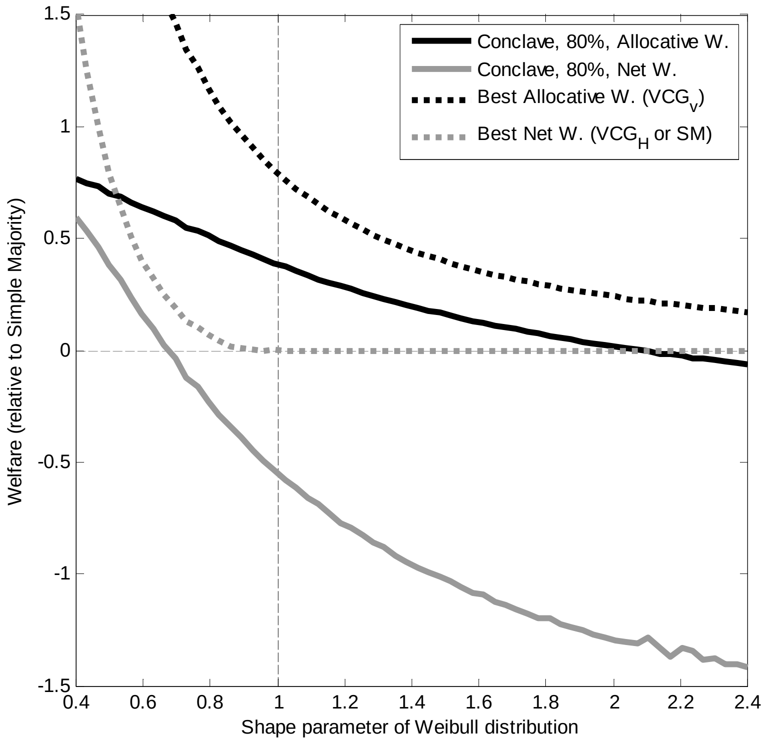

Figure 2: Best allocative welfare (black dots) and best net welfare (gray dots). Committee $N = 24$.

voters in a party are either almost indifferent, concentrated around zero, or if they are extreme then they are very extreme. The slope of function $H(x_i)$ depends on η. If $\eta > 1$, then the reciprocal hazard ratio is a decreasing function and simple majority is net efficient. If $\eta < 1$, then $H(\cdot)$ is increasing, and screening for types along the lines of Corollary 3 is net efficient.

This can be presented graphically, as in Figure 2. The curves show welfare performance (relative to simple majority) of three different mechanisms for different levels of heterogeneity of preferences, in a committee consisting of 24 voters. The gray dot curve illustrates the level of net welfare of a mechanism achieving the best net welfare. This mechanism could be VCG_1 if $\eta < 1$, and simple majority if $\eta > 1$.

A word of caution is in order here. The above statements, that ex ante heterogeneity of preference intensity is a key parameter, could be misunderstood, as not just any type of dispersion measure of X_i is important. It is certainly possible to give example of two distributions such that the first one has a greater variance and negative H', and the other one has lower variance and positive H'. By high heterogeneity we mean a particular relationship between central/indifferent values and tail/extreme values.

Formally, suppose that X has a c.d.f. F, and Y has an exponential distribution with a c.d.f. G and any mean. The first known result is the following: X has an increasing reciprocal hazard rate if and only if is smaller than Y in the convex transform order, $Y \leq_c X$ (i.e. $G^{-1}(F(x))$ is convex in x). Another result is that $Y \leq_c X$ together with $EX = EY$ implies that X second order stochastically dominates Y. These two results prove the following statement.

Proposition 3. *X has an increasing reciprocal hazard rate, then it second order stochastically dominates Y which has an exponential distribution with the same mean. X has a decreasing reciprocal hazard rate, then it is second order stochastically dominated by such Y.*

Generalizing this analysis to any $\lambda \in [0,1]$ is not a major complication. Now, $H_\lambda(x_i) = \lambda(1 - F(x_i))/f(x_i) + (1-\lambda)x_i$. For example, the black dot curve in Figure 2 shows the allocative welfare of a mechanism achieving the best allocative welfare, $\lambda = 0$. These welfare levels are higher than for the best net welfare, or any other mechanism. However, as η increases and the preferences converge to the classical bipolar Rae (1969) case, the best allocative welfare mechanism converges in its performance to simple majority.

5.1 Correlation

Suppose that z_i and x_i are not independent. Again, consider first a more involved case of net welfare, $\lambda = 1$. Function $H_\lambda(x_i)$ can be written as

$$H_\lambda(x_i) = H(x_i) = \int_0^\infty z_i \frac{1 - F(x_i|z_i)}{f_X(x_i)} f_Z(z_i)\, dz_i.$$

To present a clear comparative statics, one may follow Chakravarty and Kaplan (2013) by assuming a non-stochastic relationship between z_i and x_i. Namely, suppose that intensity of preferences is a random variable with a c.d.f. $F_X(x_i)$, but also assume that the marginal cost is determined by some monotonic function, $z_i = \beta(x_i)$, where $E\beta(x_i) = 1$ without loss of generality.[7] With this, the conditional c.d.f $F(x_i|z_i)$ is zero if $z_i > \beta(x_i)$ or

[7]If indeed such is the relationship between x_i and z_i then $Z_i(x_i) = \beta(x_i)$. Recall also that absolute value $x_i\beta(x_i)$ is assumed to be increasing in relative value x_i.

one otherwise, and therefore

$$\begin{aligned} H(x_i) &= \frac{1}{f_X(x_i)} \int_{\beta(x_i)}^{\infty} z_i f_Z(z_i)\, dz_i \\ &= \frac{1}{f_X(x_i)} \int_{x_i}^{\infty} \beta(s) f_X(s)\, ds \end{aligned}$$

For example, if z does not vary and so $\beta(\cdot)$ is a constant function, then $H(\cdot)$ becomes the usual reciprocal hazard rate.

Suppose that the association between x_i and z_i is positive, meaning that β is an increasing function. How does this assumption affects the shape of $H(\cdot)$ in comparison to the conventional reciprocal hazard rate? Notice that

$$\int_0^{\infty} \beta(s) f_X(s)\, ds = E\beta(s) = 1$$

and hence $f_X^{\beta}(x) = \beta(x) f_X(x)$ itself is a density function, rotated counter-clockwise relative to $f_X(x)$. An immediate implication is that in the interior of the support $F_X^{\beta} < F_X$, where F_X^{β} is the c.d.f. associated with $f_X^{\beta}(x)$. This proves the first part of the following result; the second part is obtained in the same way.

Proposition 4. *Assume that x is the interior of the support. If β is increasing then $H(x_i) > (1 - F_X(x_i))/f_X(x_i)$; if β is decreasing then $H(x_i) < (1 - F_X(x_i))/f_X(x_i)$.*

Next proposition is also easy to obtain and its proof is in the appendix.

Proposition 5. *Suppose that β is increasing; if $(1 - F_X(x_i))/f_X(x_i)$ is increasing, then so is $H(x_i)$. Suppose that β is decreasing; if $(1 - F_X(x_i))/f_X(x_i)$ is decreasing, then so is $H(x_i)$.*

If mechanism designer's objective is allocative welfare then $H_\lambda(\cdot)$ is affected by the stochastic relation between x_i and z_i through the expression of absolute value $x_i Z(x_i)$. Recall that $x_i Z(x_i)$ is assumed to be increasing in x_i, even if $Z(\cdot)$ may be mildly decreasing. Thus, Corollary 3 applies.

6 Ex post realizations

A concept of dispersion for vectors is needed in this section. Formally, vector $x \in R^n$ is smaller in the majorization order than $\tilde{x} \in R^n$, denoted $x \prec \tilde{x}$, if $\sum_{i=1}^n x_i = \sum_{i=1}^n \tilde{x}_i$ and $\sum_{i=1}^j x_{[i]} \leq \sum_{i=1}^j \tilde{x}_{[i]}$ for $j = 1, ..., n-1$, where $x_{[i]}$ denotes the ith largest element of vector x.

Consider the following thought experiment. Compare two different realizations of preference intensities for party k with the given number of supporters, denoted $x = \{x_i\}_{i:a_i=k}$ and $\tilde{x} = \{\tilde{x}_i\}_{i:a_i=k}$. Suppose that $x \prec \tilde{x}$. In other words, both realizations represented by these preference intensities are equivalent from the allocative efficiency perspective as they generate the same welfare for party k, but that the latter is more dispersed than the former, representing a realization that is more ex post heterogeneous. Inspecting condition 9, we see that the curvature of $H_\lambda(\cdot)$ plays a role. Namely,

Proposition 6. *Suppose that $H_\lambda(\cdot)$ is increasing, and suppose that $x, \tilde{x}$ are two different realizations of preference intensities in party k. If $H_\lambda(\cdot)$ is concave, then $x \prec \tilde{x}$ implies that $\sum_{i:a_i=k} H_\lambda(x_i) \geq \sum_{i:a_i=k} H_\lambda(\tilde{x}_i)$. Thus replacing x with $\tilde{x}$ weakly decreases the chances of selecting alternative k. If $H_\lambda(\cdot)$ is convex, then replacing this x with this $\tilde{x}$ increases those chances.*

How can the curvature of $H_\lambda(\cdot)$ be linked to some economically significant qualities? The answer to this question is much less clear-cut than in the previous discussion about the monotonicity of $H_\lambda(\cdot)$. Consider only the polar case of net efficiency and independence: this question boils down to when the reciprocal hazard rate of preference intensities, $(1-F)/f$, is convex or concave. In the Weibull example presented above this function is concave. Another parametric example is generalized Pareto distribution which has a linear reciprocal hazard rate:

Example 2. Generalized Pareto distribution with a c.d.f. $F(x) = 1 - (1+\xi x/(1-\xi))^{-1/\xi}$; assume that $0 \leq x$ for $0 < \xi < 1$, and $0 \leq x \leq (\xi - 1)/\xi$ for $\xi < 0$ (the mean is always 1). The corresponding reciprocal hazard rate is $H(x) = (1-\xi) + \xi x$. The comparative statics depend on the shape parameter ξ; if it is positive, then the reciprocal hazard rate is increasing, and if it is negative then it is decreasing.

Intuitively, the curvature of the reciprocal hazard rate appears to be linked to dispersion measure of the distribution, in that that the convex $(1-F)/f$ is associated with more dispersed values. To my knowledge, however, there is no established stochastic order that captures this.

In another polar case, the one of full allocative efficiency, we have $H_0(x_i) = x_i Z(x_i)$ and the curvature of this absolute value as a function of relative value x_i is of interest. If x_i and z_i are independent then $H_0(\cdot)$ is linear. If $Z(\cdot)$ is decreasing (and linear), then $H_0(\cdot)$ is concave; if $Z(\cdot)$ is increasing (and linear), then $H_0(\cdot)$ is convex.

The above analysis suggests three layers of the argument relating to ex post realization of preference intensities in an efficient mechanism. Informally, they can be stated as:

1. Is *more agents* in party k a good news for the likelihood of selecting its alternative? The answer is affirmative, and it does not depend on any extra conditions.

2. Suppose that the number of members in party k is given. Is having members with *higher preference intensity* a good news for the likelihood of selecting their alternative? The answer is affirmative, if H_λ is increasing. Intensity is irrelevant if this function is decreasing.

3. Suppose that the number of members in party k is given, and they have a given total intensity, $\sum_{i:a_i=k} x_i$. Is it true that having voters with more *concentrated preference intensities in party k* (in the majorization order) is a good news for the likelihood of selecting their alternative ? The answer is affirmative, if, in addition to H_λ being increasing, it is a concave function. Likewise, *less concentrated preference intensities* is a good news, if H_λ is a convex function.

Azrieli and Kim (2014) make an observation similar to point 1 above. They say that "only the ordinal ranking of the two alternatives as reported by the agents matters for the outcome". In the current paper, the word "only" should be removed, as reporting preference intensity matters under some circumstances. But the gist is that the efficient mechanism always responds positively to the number of supporters. Another interpretation of this point invokes the Samuelson condition for optimal allocation of public goods. Namely, one needs to sum the benefits to all agents together in order to calculate the correct rank of each of the two alternatives. The only twist in the present paper is that those benefits are net of the costs of a nontrivial incentive provision.

The analysis of McAfee and McMillan (1992), Yoon (2011), Condorelli (2012) and Chakravarty and Kaplan (2013) underline the importance of the the monotonicity of function the reciprocal hazard rate and the role of ex ante heterogeneity of preferences, exactly along the lines of point 2 in the above list.

However, the observations made in point 3 do not arise in those studies. There is no reporting of preference intensities in Azrieli and Kim (2014) or Drexl and Kleiner (2013), and no Samuelson condition in the papers in the tradition of McAfee and McMillan (1992).

7 Concluding remarks

One conclusion of this paper, relevant for real-life designers, can be intuitively summarized as follows: if all voters are likely to have relatively similar stakes, then asking them for ordinal preferences is efficient, as in weighted majority. If, however, there is likely to be a lot of fairly indifferent voters and a few, but extreme ones (ex ante heterogeneity) or if representatives' decision has a strong externality on a wider population, then the designer ought to consider more complicated rules, that in particular involve incentives.

One of the aims of this paper is to assess the welfare performance of some versions of real-life decision mechanisms that employ penalties, especially relative the best possible mechanisms. Hence consider simple conclaves, as one such example of a mechanism that is not efficient in general. Conclave is a mechanism in which voting occur repeatedly until sufficient supermajority is reached. The following is one possible way to formalize this in the context of two alternatives.

Assume that time is continuous and all voters are equally likely to support A and B. All voters start supporting their preferred alternative by pressing the relevant button A or B. As time passes, they can decide to irreversibly withdraw their support. Supermajority required is $n - m + 1$ where a key parameter $m \in \{1, ..., (n-1)/2\}$ is a minimal blocking minority (n odd). If m voters, or more, still vote for an alternative then the other alternative does not yet reach sufficient supermajority. As soon as the support for one of the alternatives falls below m, the other alternative is declared the winner and the game ends.

To complete the description of the extensive form game, one has to specify what is observable as the game progresses. For example, in many real-life situations, players could observe how many voters supported an alternative at any point in the game. This creates a very complicated game in which the shape of the allocation function, $p_k(a, x)$, is unclear. Since this function is the key tool in the analysis, a hard result seems difficult to derive.

Consider however a version of conclave that forms a simple timing game. Suppose that voters are locked in individual rooms and do not get to observe initial profile of support or anything about its evolution. This environment is symmetric, and so there is an equilibrium with symmetric strategies. Higher report translates into higher value. It turns out that in this case we can express the equilibrium allocation function quite easily.

Define $\tilde{x}_i = (1 - 2 \times 1_{i \in A}) x_i$, i.e. $\tilde{x}_i$ it is equal to x_i, except that it has a negative sign if and only if $i \in A$. Sort voters from the most negative $\tilde{x}_i$ to the most positive. Define two pivotal voters, the mth one from the left and the mth one from the right, with values $\tilde{x}_{(m)}$ and $\tilde{x}_{(n-m+1)}$, respectively, where $\tilde{x}_{(m)}$ is the mth order statistic of vector $(\tilde{x}_1, ..., \tilde{x}_n)$.

The allocation function then is

$$p_A(a,x) = \begin{cases} 1 & \text{if } \tilde{x}_{(m)} + \tilde{x}_{(n-m+1)} < 0 \\ 0 & \text{otherwise} \end{cases}$$

In particular, there are three possible cases. Either both $\tilde{x}_{(m)}$ and $\tilde{x}_{(n-m+1)}$ are negative, indicating that A has a supermajority support from the beginning; or both are positive, indicating that B has a supermajority support from the beginning; or $\tilde{x}_{(m)}$ is negative and $\tilde{x}_{(n-m+1)}$ is positive. In this last case, A is selected if and only if $\tilde{x}_{(m)}$ is greater in absolute terms than $\tilde{x}_{(n-m+1)}$. All three cases are captured by inequality $\tilde{x}_{(m)} + \tilde{x}_{(n-m+1)} < 0$.

In addition to the best net and the best allocative mechanisms, Figure 2 illustrates the welfare performance of conclave with supermajority 80%. Its allocative welfare level is depicted as a solid black curve, while the net welfare performance is depicted as a solid gray curve. One take-home message from this example is that for high enough ex ante heterogeneity even this unrefined version of conclave outperforms simple majority, even if the cost of penalty is taken into welfare account. This example also suggests that there is a great deal of research that should be done to study particular rules that may be used in similar repeated voting mechanisms.

8 Proofs

8.1 Proof of Proposition 1

Necessity. Suppose that the mechanism induces truth telling.

1. $\frac{\partial}{\partial x_i} P_{a_i}(a_i, x_i, z_i) \geq 0$.

   Truth telling implies that (where the scaling factor z_i in front is not included because it does not affect anything)

$$\begin{aligned}\pi_i(r_i) &\geq \tilde{\pi}_i(\bar{r}_i, r_i) = x_i P_{a_i}(\bar{r}_i) - C_i(\bar{r}_i) \\ &= (x_i - \bar{x}_i) P_{a_i}(\bar{r}_i) + \bar{x}_i P_{a_i}(\bar{r}_i) - C_i(\bar{r}_i) \\ &= (x_i - \bar{x}_i) P_{a_i}(\bar{r}_i) + \pi_i(\bar{r}_i)\end{aligned}$$

   Hence, if $x_i > \bar{x}_i$ we obtain the inequality on the left in

$$P_{a_i}(\bar{r}_i) \leq \frac{\pi_i(r_i) - \pi_i(\bar{r}_i)}{x_i - \bar{x}_i} \leq P_{a_i}(r_i).$$

   Similarly, using $\pi_i(\bar{r}_i) \geq \tilde{\pi}_i(r_i, \bar{r}_i)$ we obtain the inequality on the right. Since this is true for arbitrary reports of direction and marginal cost, even true ones, $\bar{a}_i = a_i$ and $\bar{z}_i = z_i$, the function $P_{a_i}(a_i, \cdot, z_i)$ is non-decreasing.

2. $\frac{\partial}{\partial x_i} \pi_i(a_i, x_i, z_i) = P_{a_i}(a_i, x_i, z_i)$

   A non-decreasing function $P_{a_i}(a_i, \cdot, z_i)$ is differentiable almost everywhere, hence continuous. We obtain the result by taking a limit $\bar{x}_i \to x_i$.

3. $P_{a_i}(a_i, 0, z_i) \geq P_{a_i}(-a_i, 0, z_i)$ and $C_i(a_i, 0, z_i) = C_i(-a_i, 0, z_i)$

   Suppose that $C_i(k, 0, z_i) > C_i(-k, 0, z_i)$. Then a voter whose true type is $(k, 0, z_i)$ would have incentives to misreport the direction of her preferences, as their payoff would be more negative if they reported k. Hence $C_i(a_i, 0, z_i) = C_i(-a_i, 0, z_i)$. Finally, consider a voter whose preferences are directed towards a_i. Telling the truth must be better than stating a different party and misreporting the intensity as zero:

$$x_i P_{a_i}(a_i, x_i, z_i) - C_i(a_i, x_i, z_i) \geq x_i P_{a_i}(-a_i, 0, z_i) - C_i(-a_i, 0, z_i) \quad (11)$$

   Since penalty $C_i(a_i, \cdot, z_i)$ is non-decreasing[8], we have $C_i(a_i, x_i, z_i) \geq C_i(a_i, 0, z_i) = C_i(-a_i, 0, z_i)$, and thus inequality (3) becomes

$$P_{a_i}(a_i, x_i, z_i) \geq P_{a_i}(-a_i, 0, z_i)$$

[8]This can be seen by taking the derivative of both sides in equation (2) with respect to x_i and using already established condition 2.

This is true for any x_i and so it must be true for $x_i = 0$.

4. $\pi_i(a_i, x_i, \cdot)$ and $P_{a_i}(a_i, x_i, \cdot)$ are constant.

   Firstly, $\tilde{\pi}_i(a_i, x_i, \bar{z}_i, a_i, x_i, z_i)$ is independent of z_i by definition in equation (1). Secondly, it is independent of the report $\bar{z}_i$ too. (To see this, fix a_i, x_i and note that the set of reports $\bar{z}_i$ that maximize the expected utility $z_i\tilde{\pi}_i(a_i, x_i, \bar{z}_i, a_i, x_i, z_i)$ does not depend on z_i, by previous point. If a certain report does not belong to this set then there is an incentive to lie, which cannot be a part of equilibrium. Hence all reports are maximizers and the function must be constant). The claim comes from the fact that $\pi_i(a_i, x_i, z_i) = \tilde{\pi}_i(a_i, x_i, z_i, a_i, x_i, z_i)$, and that the right-hand side is independent of z_i.

   Since $\pi_i(a_i, x_i, z_i)$ is constant over z_i and $\frac{\partial}{\partial x_i}\pi_i(a_i, x_i, z_i) = P_{a_i}(a_i, x_i, z_i)$, it must be that $P_{a_i}(a_i, x_i, z_i)$ is constant over z_i.

Sufficiency. Suppose that conditions 1-4 hold.

1. Reporting true direction a_i, but misreporting x_i can never improve the expected payoff (regardless of the report $\bar{z}_i$).

   Consider any $\bar{x}_i > x_i$ and any $\bar{z}_i$. Then the condition $\frac{\partial}{\partial s}\pi_i(a_i, s, \bar{z}_i) = P_{a_i}(a_i, s, \bar{z}_i)$ implies that

$$\pi_i(a_i, \bar{x}_i, \bar{z}_i) - \pi_i(a_i, x_i, \bar{z}_i) = \int_{x_i}^{\bar{x}_i} P_{a_i}(a_i, s, \bar{z}_i)\, ds$$

   Since $P_{a_i}(a_i, \cdot, \bar{z}_i)$ is non-decreasing, we have

$$\int_{x_i}^{\bar{x}_i} P_{a_i}(a_i, s, \bar{z}_i)\, ds \leq (\bar{x}_i - x_i)\, P_{a_i}(a_i, \bar{x}_i, \bar{z}_i)$$

   and therefore together

$$\pi_i(a_i, \bar{x}_i, \bar{z}_i) - \pi_i(a_i, x_i, \bar{z}_i) \leq (\bar{x}_i - x_i)\, P_{a_i}(a_i, \bar{x}_i, \bar{z}_i)$$

   Substituting in $\pi_i(a_i, \bar{x}_i, \bar{z}_i) = \bar{x}_i P_{a_i}(a_i, \bar{x}_i, \bar{z}_i) - C_i(a_i, \bar{x}_i, \bar{z}_i)$ implies

$$z_i\left(x_i P_{a_i}(a_i, \bar{x}_i, \bar{z}_i) - C_i(a_i, \bar{x}_i, \bar{z}_i)\right) \leq z_i \pi_i(a_i, x_i, \bar{z}_i)$$

   One can show a similar inequality for $\bar{x}_i < x_i$. This means that reporting true direction a_i, but misreporting x_i can never improve the expected payoff (regardless of the report of $\bar{z}_i$).

2. Misreporting the direction of support does not increase payoff.

   Observe that $P_{a_i}(-a_i, \cdot, z_i)$ is non-increasing. This follows from the fact that probabilities add up to one, that is, for any a_i, x_i, z_i

$$1 = P_{a_i}(a_i, x_i, z_i) + P_{-a_i}(a_i, x_i, z_i)$$

   where $-a_i$ in the subscript indicates the alternative that voter i does not prefer. Thus, if $P_{a_i}(a_i, \cdot, z_i)$ is non-decreasing by condition 1, then $P_{-a_i}(a_i, \cdot, z_i)$ is non-increasing, so is $P_{a_i}(-a_i, \cdot, z_i)$. As noted in footnote8, function $C_i(a_i, \cdot, z_i)$ is a non-decreasing, so the function

$$\tilde{\pi}_i(-a_i, \cdot, \bar{z}_i, r_i) = x_i P_{a_i}(-a_i, \cdot, \bar{z}_i) - C_i(-a_i, \cdot, \bar{z}_i)$$

   is non-increasing.

   In other words, if a voter masquerades herself as a member of a different party, then her payoff will be at least as high as if she also misrepresented her intensity as zero. This is the first inequality of the chain below. The second and third lines follow from condition 3, and the final line comes from step one above.

$$\begin{aligned}
\tilde{\pi}_i(-a_i, \bar{x}_i, \bar{z}_i, r_i) &\leq x_i P_{a_i}(-a_i, 0, \bar{z}_i) - C_i(-a_i, 0, \bar{z}_i) \\
&= x_i P_{a_i}(-a_i, 0, \bar{z}_i) - C_i(a_i, 0, \bar{z}_i) \\
&\leq x_i P_{a_i}(a_i, 0, \bar{z}_i) - C_i(a_i, 0, \bar{z}_i) \\
&\leq x_i P_{a_i}(a_i, x_i, \bar{z}_i) - C_i(a_i, x_i, \bar{z}_i) = \tilde{\pi}_i(a_i, x_i, \bar{z}_i, r_i)
\end{aligned}$$

   This means that $z_i\tilde{\pi}_i(-a_i, \bar{x}_i, \bar{z}_i, r_i) \leq z_i\tilde{\pi}_i(a_i, x_i, \bar{z}_i, r_i)$, or that misreporting a direction is never better than stating the direction and value correctly (regardless of report $\bar{z}_i$).

3. Finally, note that if $\pi_{a_i}(\bar{a}_i, \bar{x}_i, \cdot)$ and $P_{a_i}(\bar{a}_i, \bar{x}_i, \cdot)$ are constant then so is $C_i(\bar{a}_i, \bar{x}_i, \cdot)$; misreporting z_i does not improve utility.

8.2 Proof of Lemma 1

Since the payoff of voter (or district) i in an incentive compatibility mechanism is $z_i\pi_i(a_i, x_i)$ the expected payoff is

$$N_i = E_{a_i,z_i}\int_0^\infty z_i\pi_i(a_i, x_i) f_i(x_i|z_i)\, dx_i$$

Integrate by parts

$$N_i = -E_{a_i,z_i} z_i\pi_i(a_i, x_i)(1 - F_i(x_i|z_i))\,|_0^\infty + E_{a_i,z_i}\int_0^\infty \frac{\partial \pi_i(a_i, x_i)}{\partial x_i} z_i(1 - F_i(x_i|z_i))\, dx_i$$

Use condition 2 of Proposition 1

$$\begin{aligned} N_i &= E_{a_i,z_i} z_i \pi_i (a_i, 0) + E_{a_i,z_i} \int_0^\infty P_{a_i} (a_i, x_i) z_i \frac{1 - F_i (x_i|z_i)}{f_{X_i} (x_i)} f_{X_i} (x_i) dx_i \\ &= z_i^e E_{a_i} \pi_i (a_i, 0) + E_{a_i} \int_0^\infty P_{a_i} (a_i, x_i) \left(E_{z_i} z_i \frac{1 - F_i (x_i|z_i)}{f_{X_i} (x_i)} \right) f_{X_i} (x_i) dx_i \end{aligned}$$

Substitute $H_i (x_i) = E_{z_i} z_i \frac{1 - F_i(x_i|z_i)}{f_{X_i}(x_i)}$

$$N_i = z_i^e E_{a_i} \pi_i (a_i, 0) + E_{a_i,x_i} P_{a_i} (a_i, x_i) H_i (x_i)$$

8.3 Proof of Lemma 3

By adding and subtracting $E_{a_i,x_i} P_{a_i} (a_i, x_i) \gamma_i (F_{X_i} (x_i))$, payoff in equation 3 can be written as

$$\begin{aligned} N_i &= \lambda z_i^e E_{a_i} \pi_i (a_i, 0) + E_{a_i,x_i} P_{a_i} (a_i, x_i) \gamma_i (F_{X_i} (x_i)) \\ &\quad + E_{a_i,x_i} P_{a_i} (a_i, x_i) (\phi_i (F_{X_i} (x_i)) - \gamma_i (F_{X_i} (x_i))) \end{aligned}$$

The last component on the right can be integrated by parts and written as

$$\begin{aligned} & E_{a_i} P_{a_i} (a_i, x_i) (\Phi_i (F_{X_i} (x_i)) - \Gamma_i (F_{X_i} (x_i))) |_0^1 \\ & - E_{a_i} \int_0^1 \frac{\partial P_{a_i} (a_i, x_i)}{\partial x_i} (\Phi_i (F_{X_i} (x_i)) - \Gamma_i (F_{X_i} (x_i))) dx_i \end{aligned}$$

Note that because of the convexification, we have $\Phi_i (0) = \Gamma_i (0)$ and $\Phi_i (1) = \Gamma_i (1)$ at the end points; so the first term is equal to zero. This establishes payoff in equation 7, and the definition of Λ_i.

To show that Λ_i is non-negative, notice that $\partial P_{a_i} (a_i, x_i) / \partial x_i \geq 0$, and, because of convexification, we have $\Phi_i (q) \geq \Gamma_i (q)$ in the entire domain.

8.4 Proof of Theorem 1

By Lemma 1, total welfare is bounded

$$\sum_i N_i \leq E_{a,x} \sum_i p_{a_i} (a, x) \gamma_i (F_{X_i} (x_i))$$

The proof is conducted in two steps. Firstly, we will find $p_{a_i} (a, x)$ that maximizes this upper bound. Secondly, we will observe that this optimal function also guarantees that the upper bound is reached with equality, and thus it also maximizes total welfare itself.

Step 1. Take the sum under the expectation and write it as two separate terms for individuals who support A and B, respectively,

$$\begin{aligned}\sum_i N_i &\leq E_{a,x}\left(\sum_{i:a_i=A} p_A\left(a_i, a_{-i}, x\right)\gamma_i\left(F_{X_i}\left(x_i\right)\right) + \sum_{j:a_j=B} p_B\left(a_j, a_{-j}, x\right)\gamma_j\left(F_{X_j}\left(x_j\right)\right)\right)\\ &= E_{a,x}\left(p_A\left(a,x\right)\sum_{i:a_i=A}\gamma_i\left(F_{X_i}\left(x_i\right)\right) + p_B\left(a,x\right)\sum_{j:a_j=B}\gamma_j\left(F_{X_j}\left(x_j\right)\right)\right)\end{aligned}$$

Notice that $p_B\left(a,x\right) = 1 - p_A\left(a,x\right)$, and hence

$$\sum_i N_i \leq E_{a,x}\left(p_A\left(a,x\right)\left(\sum_{i:a_i=A}\gamma_i\left(F_{X_i}\left(x_i\right)\right) - \sum_{j:a_j=B}\gamma_j\left(F_{X_j}\left(x_j\right)\right)\right) + \sum_{j:a_j=B}\gamma_j\left(F_{X_j}\left(x_j\right)\right)\right)$$

The bound is maximized if the following rule is applied: Select A with probability 1 if and only if

$$\sum_{i:a_i=A}\gamma_i\left(F_{X_i}\left(x_i\right)\right) \geq \sum_{j:a_j=B}\gamma_j\left(F_{X_j}\left(x_j\right)\right)$$

This establishes a candidate solution $p_k\left(a,x\right)$. Notice that this function generates an non-decreasing $P_k\left(a_i, x_i\right)$; this is true because γ_i is monotonically increasing by construction.

Step 2. The last thing is to show that the proposed mechanism achieves this (maximal) bound. That is, we want to show that if the candidate solution is used then $\pi_i\left(a_i, 0\right) = 0$ and $\Lambda_i = 0$ for all individuals.

First, notice that the indifferent type is not expected to pay anything, hence $\pi_i\left(a_i, 0\right) = 0$.

Secondly, we show that if there exists an x_i such that $\Phi_i\left(F_{X_i}\left(x_i\right)\right) > \Gamma_i\left(F_{X_i}\left(x_i\right)\right)$, then $\partial P_{a_i}\left(a_i, x_i\right)/\partial x_i = 0$ for all a_i, and hence the second part of Λ_i is equal to zero too, proving the claim. Suppose that there is an x_i for which $\Phi_i\left(F_{X_i}\left(x_i\right)\right) > \Gamma_i\left(F_{X_i}\left(x_i\right)\right)$. But in this situation, the convexification Γ_i is linear, and hence its derivative, γ_i, is constant in its neighborhood. If this is the case, then a small change of such x_i in condition $\sum_{i:a_i=A}\gamma_i\left(F_{X_i}\left(x_i\right)\right) \geq \sum_{j:a_j=B}\gamma_j\left(F_{X_j}\left(x_j\right)\right)$ will not change the allocation probability $p_k\left(a,x\right)$ and so $\partial P_{a_i}\left(a_i, x_i\right)/\partial x_i = 0$.

8.5 Proof of Proposition 2

The mechanism does not make any use of report $\bar{z}_i$ so misreporting it does not improve the payoff regardless of the behavior of other voters.

Consider now reporting true direction but misreporting intensity. Notice that the level of payment $c_i(\bar{a}, \bar{x})$ of voter i, conditional on winning, is independent of her report. This is because the RHS of equation 10 does not depend on the report of voter i's intensity. So, the only element of the outcome that voter i can affect by choosing different reports is her winning-losing status.

Suppose that other voters vote such that

$$H_i(0) \leq \sum_{j:\bar{a}_i=-k} H_j(\bar{x}_j) - \sum_{\{j:\bar{a}_i=k, j\neq i\}} H_j(\bar{x}_j)$$

The payment is set so that

$$H_i(c_i(\bar{a}, \bar{x})) + \sum_{\{j:\bar{a}_i=k, j\neq i\}} H_j(\bar{x}_j) = \sum_{j:\bar{a}_i=-k} H_j(\bar{x}_j)$$

That is, if $x_i \geq c_i(\bar{a}, \bar{x})$ then reporting truthfully will guarantee that the alternative preferred by voter i will be selected, because $H_i(\cdot)$ is assumed to be increasing. However, since in this case the individual payoff, $z_i(x_i - c_i(\bar{a}, \bar{x}))$, is positive too, voter i wants his alternative to be selected at this price. Alternatively, if $x_i < c_i(\bar{a}, \bar{x})$ then reporting truthfully will guarantee that the alternative preferred by voter i will not be selected. However, since in this case the individual payoff is negative, voter does not want to change this situation.

If on the other hand

$$H_i(0) > \sum_{j:\bar{a}_i=-k} H_j(\bar{x}_j) - \sum_{\{j:\bar{a}_i=k, j\neq i\}} H_j(\bar{x}_j)$$

then payment is zero $c_i(\bar{a}, \bar{x}) = 0$. If voter i reports truthfully, then the condition

$$H_i(x_i) > \sum_{j:\bar{a}_i=-k} H_j(\bar{x}_j) - \sum_{\{j:\bar{a}_i=k, j\neq i\}} H_j(\bar{x}_j)$$

holds and his alternative is selected. Hence regardless of what other voters do, voter i has no incentives to misreport her intensity.

Misreporting direction of preferences.

8.6 Proof of Proposition 5

Observe that if $(1 - F_X(x))/f_X(x)$ is increasing then $-1/\int_x^\infty f_X(s)\,ds > (f_X'/f_X^2)$, simply by taking the derivative.

Now, suppose that β is increasing. Thus

$$\int_x^\infty \beta(x) f_X(s)\,ds < \int_x^\infty \beta(s) f_X(s)\,ds$$

or

$$-\frac{\beta(x)}{\int_x^\infty \beta(s) f_X(s)\, ds} > -\frac{1}{\int_x^\infty f_X(s)\, ds}$$

This implies that

$$-\frac{\beta(x)}{\int_x^\infty \beta(s) f_X(s)\, ds} > \left(f_X'/f_X^2\right)$$

But this implies that $H'(x) = -\beta(x) - (f_X'/f_X^2)\int_x^\infty \beta(s) f_X(s)\, ds$ is positive. The second part of the proposition can be obtained in a similar way.

9 Literature

Apesteguia, J., Ballester M.A. , Ferrer R., (2011), On the Justice of Decision Rules, *Review of Economic Studies*, 78, 116.

Azrieli Y. and S. Kim, (2014) Pareto Efficiency And Weighted Majority Rules, *International Economic Review* Vol. 55, No. 4, November.

Casella, Alessandra, (2005), Storable votes, *Games and Economic Behavior*, 51, 391-419.

Chakravarty, S., Kaplan T.R., (2013), Optimal allocation without transfer payments, *Games and Economic Behavior*, 77, 1-20.

Condorelli, D., (2012), What money can't buy: Efficient mechanism design with costly signals, *Games and Economic Behavior*, 75 (2), 613-624.

Drexl M., and Kleiner, A., (2013), Why voting, A Welfare Analysis, unpublished manuscript.

Gershkov, Alex, Moldovanu, Benny and Shi, Xianwen, (2014), Optimal Voting Rules, unpublished manuscript.

Hartline, Jason D. and Roughgarden, Tim, (2008), Optimal Mechanism Design and Money Burning, in: STOC'08 Conference Proceedings.

Jackson and Sonnenschein (2007)

Kwiek M. (2014), Conclave, *European Economic Review*,

Kwiek, M. and Zhang, Q., (2013), Majority Rule and Mechanism Design, unpublished manuscript.

Kwiek, M. Marreiros, Helia, and Vlassopoulos, Michael, (2015), Experimental study of conclaves. work in progress.

McAfee, R.P., McMillan, J. (1992), Bidding rings, *American Economic Review*, 82 (3), 579-599.

Myerson, R.B., (1981), Optimal auction design, *Mathematics of Operations Research*, 6, No. 1, 58-73.

May, Kenneth O., (1952), A Set of Independent, Necessary and Sufficient Conditions for Simple Majority Decision, *Econometrica*, 20, 680-4.

Ponsati, Clara and Sakovics, Jozsef, (1996), "Multiperson Bargaining over Two Alternatives," *Games and Economic Behavior, Elsevier*, vol. 12(2), pages 226-244, February.

Rae, Douglas W. (1969), Decision-Rules and Individual Values in Constitutional Choice, *The American Political Science Review*, Vol. 63, No. 1, pp. 40-56.

Schmitz, Patrick W. and Tröger, Thomas, (2012), The (sub-)optimality of the majority rule, “Games and Economic Behavior” 74, 651–665.

Yoon, K. (2011), Optimal mechanism design when both allocative inefficiency and expenditure inefficiency matter, Journal of Mathematical Economics, 670-676.