The Dimensions of Consensus

Alex Gershkov, Benny Moldovanu and Xianwen Shi*

December 22, 2016

Abstract

We study a multi-dimensional collective decision under incomplete information. Agents have Euclidean preferences and vote by simple majority on each issue (dimension), yielding the coordinate-wise median. Judicious rotations of the orthogonal axes – the issues that are voted upon – lead to welfare improvements. If the agents' types are drawn from a distribution with independent marginals then, under weak conditions, voting on the original issues is not optimal. If the marginals are identical (but not necessarily independent), then voting first on the total sum and next on the differences is often welfare superior to voting on the original issues. We also provide various lower bounds on incentive efficiency: in particular, if agents' types are drawn from a log-concave density with symmetric marginals, a second-best voting mechanism attains at least 88% of the first-best efficiency.

1 Introduction

In 1974 the U.S. Congress changed its budgeting process: instead of considering appropriations requests that were voted upon one at a time, which resulted in a gradually determined total level of spending, the *Congressional Budget and Impoundment Control Act* required voting first on an overall level of spending, before the determination of budgets for individual programs in subsequent votes. A large literature

*We wish to thank Yeon-Koo Che, Rahul Deb, Hans-Peter Grüner, Philippe Jehiel, Andreas Kleiner, Robert J. McCann, Konrad Mierendorff, Marcin Peski, Rob Ready, Colin Stewart, Thomas Tröger, Jürgen von Hagen and Cedric Wasser for helpful comments, and various seminar and conference participants for helpful discussion. Gershkov: Department of Economics, Hebrew University of Jerusalem, Israel and School of Economics, University of Surrey, UK, alexg@huji.ac.il; Moldovanu: Department of Economics, University of Bonn, Germany, mold@uni-bonn.de; Shi: Department of Economics, University of Toronto, Canada, xianwen.shi@utoronto.ca.

in the area of public finance (see for example the review articles in Poterba and von Hagen [1999]) has debated the costs and benefits of such procedural changes, with particular attention to the size of the expected budget deficit.[1]

In this paper we analyze the general problem of redefining (or bundling) the issues brought to vote in a multi-dimensional collective decision problem. Our interest in this topic stems from the potential of such methods to increase the welfare of the involved decision makers by allowing them to reach, in an incentive compatible way, a consensus that was not possible on the original issues. We are thus looking for the dimensions on which the best consensus can be found, or, put differently, the dimensions where the least cleavage among voters is present.

We study a multi-dimensional collective decision that is resolved via simple majority voting: an example is a legislature that needs to decide on individual budgets for public goods such as, say, education and defence. Another example is the decision on the geographical location of a desirable facility. But even "mundane" decisions such as hiring or project adoption based on multi-dimensional attributes can be viewed through our lens.

We adopt the standard spatial model of voting widely used in the political science literature (see for example, Chapter 5 in Austen-Smith and Banks [2005]), where voters have preferences characterized by ideal points in each dimension and by a quadratic loss caused by deviations from the ideal point. The main text deals with the two-dimensional case, while the generalization to more than two dimensions is in an Appendix.

The voters' ideal points are private information, and we look at the outcomes of voting by simple majority on each dimension separately — as we shall see below, the focus on simple majority voting yields, in combination with a decision over the dimensions that are the subject of voting, an analysis of more generality than immediately apparent.

With votes taken by simple majority in each separate dimension, the outcome is the coordinate-wise median of the voters' ideal points. This easily follows from Black's [1948] famous theorem because the induced preferences are single peaked on each one-dimensional issue. In general, this outcome does not coincide with the first-best, given here by the alternative that minimizes the overall distance from the individual ideal points. It is of course well-known that the first-best outcome is simply the coordinate-wise average (or mean) of the ideal points, and thus first-best welfare is given here by the corresponding variance (with a minus sign).

[1] There was a widespread belief that the new rules would lead to smaller deficits, and the act was passed almost unanimously in both House and Senate.

The first-best outcome is not implementable: each agent has an incentive to try to move the average closer to his/her ideal point by exaggerating his/her position on one or more issues. This phenomenon has been first documented by Galton [1907], who was also the first to recommended the use of the median as a robust and non-manipulable aggregator of opinions.[2]

Given the tension between first-best on one hand and implementable outcomes on the other, how well does voting by simple majority perform in terms of welfare? Using a classical inequality due to Hotelling and Solomons [1932], it can be shown that, for any distribution of preferences, voting by simple majority on any given issues achieves at least 50% of the welfare achievable in the first best.

The main insight of the present paper is that a judicious choice of the issues that are actually put to vote (while maintaining voting by simple majority, with its desirable incentive properties) can significantly improve welfare.[3] For example, instead of voting on two separate issues such as education and defense, the legislature could vote on a total budget, and then on a division of that budget between the two issues – just as Congress started to do in 1974. Ferejohn and Krehbiel [1987] have shown that the change adopted by Congress can be mathematically represented by a 45-degree rotation of the coordinates (or issues) on which voting takes place, and we analyze here the general issue of determining an optimal rotation.

Our main results are:

1) If the agents' ideal points in one dimension are independently distributed from the ideal points in the other dimension then, under very weak conditions on the distribution of preferences, voting on the original issues is sub-optimal; that is, a re-packaging of the issues brought to vote via rotation (this creates some correlation among the ideal points) increases welfare.

2) If the marginals of the distribution of agents' ideal points are identically distributed (but not necessarily independent), we provide sufficient conditions under which the 45-degree rotation is welfare superior to no rotation. The sufficient conditions are satisfied by commonly used distributions if marginals are identically and independently distributed (I.I.D.). We show analytically that, with I.I.D. marginals, the 45-degree rotation is always a critical point, and numerically that it is a global welfare maximum for many standard distributions. A key observation for these results is that, under the symmetry of the marginals, the 45-degree rotation entirely eliminates the conflict arising between efficiency and majority voting in one dimension – all remaining conflict is concentrated in the other, orthogonal dimension.

[2] His insights have been sharpened and much generalized in the literature on *robust estimation*.

[3] The idea of comparing voting rules in terms of their expected welfare goes back to Rae[1969].

3) We provide various lower bounds on incentive efficiency for large, non-parametric families of distribution of ideal points (such as unimodal distributions, distributions with an increasing hazard rate, etc.). For example, if agents' ideal points are drawn from a log-concave density with I.I.D. marginals, a voting mechanism that involves a 45 degrees rotation of the original dimensions attains at least 88% of the first-best efficiency. This should be compared to the universal lower bound of 50% that obtains without any assumption on the distribution, and without using rotations.

Technically and conceptually, our contribution builds upon and relates to several important and elegant contributions due to Moulin [1980], Border and Jordan [1983], Kim and Rousch [1984] and Peters, van der Stel and Storcken [1992]. In a one-dimensional setting with single-peaked preferences, Moulin considered mechanisms that depend on reported peaks, and characterized the set of dominant strategy incentive compatible (DIC), anonymous and Pareto efficient mechanisms: each mechanism in the class is obtained by choosing the median among the n reported peaks of the real voters and the peaks of a set of $n-1$ "phantom" voters (these are fixed by the mechanism, and do not vary with the reports).[4] Border and Jordan [1983] removed Moulin's assumption whereby mechanisms were allowed to only depend on peaks, and generalized Moulin's finding to a multi-dimensional setting with separable and quadratic preferences: each DIC mechanism was shown to be decomposable into a collection of one-dimensional DIC mechanisms, each described by the location of the phantom voters in the respective dimension.[5]

Gershkov, Moldovanu and Shi [2016] analyzed welfare maximization in a one-dimensional setting with cardinal utilities, and derived the ex-ante welfare maximizing placement of phantoms as a function of utilities and of the distribution of types. They also showed how to avoid the phantom interpretation by implementing Moulin's mechanisms (including the welfare optimal one) via a sequential, binary voting procedure together with a flexible qualified majority schedule needed for the adoption of various alternatives.[6] Combining their result with the Border-Jordan decomposition yields the welfare maximizing mechanism for multidimensional settings with separable and quadratic preferences. But, the ensuing solution, described by an optimal placement of phantoms in each dimension, is not satisfactory from a practical point of view: it implies that each issue (dimension) in each multi-dimensional problem must be voted upon according to a particular institution. This theoretically needed flexibility may be difficult, if not impossible, to achieve in practice and we do not

[4] Relaxing Pareto efficiency yields the same characterization, but with $n+1$ phantoms.

[5] See also Barbera, Gul and Stacchetti [1993].

[6] See also Kleiner and Moldovanu [2016] for a derivation of sufficient conditions under which sequential, binary voting procedures possess desirable properties.

expect to observe its deployment.

Instead, we take here a different approach to welfare improvement: we fix an ubiquitous institution – voting by simple majority on each issue – but we allow flexibility in the design of the issues that are actually put to vote. Such a limited form of agenda design is very common in practice, and, as we shall see, has important welfare consequences.

The simplest multidimensional setting for studying issue design and repackaging is the one with Euclidean preferences: intuitively, the presence of spherically symmetric preferences does not a-priori determine the dimensions of the Border and Jordan decomposition into one-dimensional mechanisms. Indeed, Kim and Rousch [1984] showed that the set of continuous, anonymous and DIC mechanisms can be described by performing the Border-Jordan analysis subsequent to any translation of the origin and any rotation of the orthogonal axes. Peters, van der Stel and Storcken [1992] showed that, for two dimensions, voting by simple majority in each dimension (after any translation/rotation of the plane) is also Pareto optimal. This is the unique anonymous and DIC mechanism with this property, and Pareto efficiency is generally not consistent with DIC in more than two dimensions.

Since both median and mean are translation equivariant, translations of the origin cannot improve welfare, and it is therefore without loss of generality to restrict attention to rotations of the axes followed by simple majority voting on each newly defined dimension.

A key observation, well known in the theory of spatial statistics, is that the mean is rotation equivariant (i.e., the mean after rotation is obtained by rotating the original mean) but the coordinate-wise median is not (see Haldane [1948], or the literature on spatial voting, e.g., Feld and Grofman [1988]). As a consequence, a rotation of the axes may decrease the distance between the coordinate-wise mean (first-best) and the coordinate-wise median (outcome of majority voting), thus increasing welfare in our framework.

The basic feature behind the welfare increasing properties of rotations is the non-linearity of the median function, i.e. the median of a sum of random variables is not equal to the sum of the medians. Note that the distributions of ideal points after rotation can be represented as convolutions of the original distributions, which explains here the appearance of sums of random variables.

On the one hand, this non-linearity is the driving force behind our results; on the other hand, it also implies that the analysis becomes relatively complex.[7] In order to

[7] This is true even for common distributions, such as the Gamma, Poisson, lognormal, etc. Some of our results are based on insights that go back to conjectures by Ramanujan (see Szegö [1928]).

use calculus and probabilistic/statistical techniques, we focus on the limit case where the number of voters is infinite. In particular, we employ methods from Fourier analysis to deal with convolutions, and various concentration inequalities that relate statistics such as the mean, median, mode and variance of distributions. We sketch in an Appendix how our analysis can be generalized to more than two dimensions.

It is well-known that, in multi-dimensional models of voting the existence of a Condorcet winner is rare (Kramer [1973]), and that dictatorship is often the only strategy-proof mechanism (Zhou [1991]). Kramer [1972] observed that, however, voting in a variety of institutions is often sequential, issue by issue, and he established that there exists an issue-by-issue sophisticated voting equilibrium if voters' preferences are continuous, convex and separable. Shepsle [1979] forcefully argued that the division of a complex, collective decision into several different jurisdictions, each jurisdiction being responsible for one aspect only (*germaneness*), creates stable equilibria that would not be possible in the general, unconstrained decision model. His main examples were the various legislative committees in the U.S. congress. Viewed in this context, the coordinate-wise median analyzed in our paper – obtained by simple-majority voting in each dimension – constitutes a basic instance of a *structure induced equilibrium* in the spirit of Shepsle [1979], and our goal is to endogenize the choice of jurisdictions in order to improve welfare, an issue that has not received much attention in formal studies. Of course, it is possible to perform an analysis similar to ours for different underlying goals, e.g., define jurisdictions that serve other purposes, such as the self-interest of an agenda setter or of a coalition of voters.

It is also instructive to compare our results to those in the classical papers by Caplin and Nalebuff ([1988], [1991]).[8] Again motivated by the instability of multi-dimensional voting, they considered instead the effect of super-majority requirements on the stability of the spatial mean. For a log-concave density governing the distribution of types (and also for other, more general forms of concavity), Caplin and Nalebuff showed that, once established as status-quo, the mean cannot be displaced by another alternative if the selection of that alternative requires a super-majority of at least 64% (or $1 - \frac{1}{e}$). In other words, given the distributional assumptions and a large population of voters, any coalition that prefers an alternative over the mean contains less than 64% of the voters, and is thus not effective given the super-majority requirement. Caplin and Nalebuff did not consider incomplete information and incentive constraints: recall that, for any finite number of voters, there is in fact no incentive compatible voting mechanism that would actually achieve the mean as an

[8] These authors were also the first to use modern concentration inequalities in the Economics literature.

outcome. Thus, it is not entirely clear how the first-best (and stable) status quo can be reached by a voting process in the first place. As mentioned above, for the log-concave case with independent marginals, our results display an incentive compatible mechanism that achieves at least 88% of the first-best utility. Thus, issue by issue voting by simple majority on appropriately defined dimensions constitutes an intuitive and incentive compatible institutional arrangement that is almost efficient in this case. Moreover, the relative efficiency of this mechanism increases, and tends to 100% when we increase the number of dimensions of the underlying problem.

The remaining part of the paper is organized as follows: In Section 2 we present the two dimensional voting model by simple majority, and connect incentive compatible mechanisms to the special orthogonal group of rotations in the plane. We also discuss the equi-invariance properties of means and medians. In Section 3 we focus on the setting with a large number of agents, for which we can use various available statistical techniques. We first show that voting on independent issues is always sub-optimal: a rotation that bundles independent issues is always beneficial. We next show that if the issues are identically distributed, a rotation by 45 degrees (corresponding, for example, to a vote on a total budget for the two issues and its division among issues) can improve welfare over the zero rotation. In particular, we give sufficient conditions that hold for large, non-parametric families of distributions under which the welfare under such a rotation is higher than in the original, no-rotation case. Numerical simulations suggest that the 45 degrees rotation is indeed an optimum in this case. In Section 4 we offer bounds on the relative efficiency of voting by simple majority complemented by rotations. The effect of rotations is shown to be substantial. Section 5 concludes. Several proofs are gathered in Appendix A, and generalizations to higher dimensions are sketched in Appendix B.

2 The Model

We consider an odd number of agents, n, who collectively decide about two issues, X and Y, on a convex region $D \subseteq \mathbb{R}^2$. Each agent's ideal position on these two issues is given by a peak $\mathbf{t}_i = (x_i, y_i)$, $i = 1, 2, ..., n$. The peak $\mathbf{t}_i$ is agent i's private information. Each agent i has a utility function of the form

$$-\left|\left|\mathbf{t}_i - \mathbf{v}\right|\right|^2$$

where $\mathbf{v} = (x, y)$ is a fixed point in D and where $||\cdot||$ is the standard Euclidean (L_2) norm. The peaks $\mathbf{t}_i = (x_i, y_i)$ are independently, identically distributed (I.I.D.) across agents, according to a joint distribution $F(x_i, y_i)$, with density f. Throughout the paper, we assume that $\mathbb{E}\,||\mathbf{t}_i||^2 < \infty$ for all $\mathbf{t}_i \in D$.

We consider a utilitarian planner who would like to choose $\mathbf{v} \in D$ to maximize the average of the agents' ex ante utilities, or equivalently, minimize the average of expected squared distance:

$$\min_{\mathbf{v}\in D} \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} ||\mathbf{t}_i - \mathbf{v}||^2\right],$$

subject to agents' incentive constraints. Ignoring agents' incentives, the planner can choose a point $\mathbf{u}$ that minimizes the average of ex post distance:

$$\mathbf{u} \in \arg\min_{\mathbf{v}\in D} \frac{1}{n}\sum_{i=1}^{n} ||\mathbf{t}_i - \mathbf{v}||^2,$$

which we will refer to as the *first-best solution*. For each fixed realization $(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)$, it is well known that the first-best solution is simply the mean of the ideal points

$$\mathbf{u} = \bar{\mathbf{t}} \equiv \frac{1}{n}\sum_{i=1}^{n} \mathbf{t}_i.$$

Hence, the first-best (per capita) expected utility is the variance (with negative sign) $-\frac{1}{n}\sum_{i=1}^{n}\left|\left|\mathbf{t}_i - \bar{\mathbf{t}}\right|\right|^2$. However, the first-best is clearly not implementable: each agent can advantageously move the mean towards her ideal point by reporting a false peak.

Remark 1 *The Euclidean norm adopted here, where first best choice and welfare are given by mean and variance (with negative sign), respectively, greatly facilitates our analysis because we can then use various statistical results relating to the mean, median and variance. We can incorporate other convex norms to address similar questions, but the technical analysis needs to change accordingly.*

2.1 Voting by Simple Majority

We consider voting by simple majority on each separate dimension. This is easily seen to be an incentive compatible scheme: each agent has a (weakly) dominant strategy, to state his true ideal point in each dimension. Our focus on simple majority voting stems from its wide applicability and its actual use in practice. We do not a-priori restrict the issues on the ballot to be X and Y. Instead, new issues can be created through "re-packaging and bundling" the basic issues X and Y. The main theme of the paper is, indeed, the analysis of the problem of optimal bundling of issues X and Y, i.e. finding what we call *the optimal dimensions of consensus.*

2.2 Rotations in the Plane

We model packaging and bundling of issues through rotations in the plane. Recall that, for fixed Cartesian coordinates, rotating a point $(x, y) \in \mathbb{R}^2$ counter-clockwise by an angle of θ can be represented by the multiplication of the vector (x, y) with a rotation matrix $R(\theta)$. The resulting, rotated point (z_-, z_+) is given then by

$$\begin{pmatrix} z_- \\ z_+ \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}}_{R(\theta)} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}.$$

Equivalently, one can obtain (z_-, z_+) by rotating the original Cartesian coordinates clockwise around the fixed origin by an angle of θ to obtain new orthogonal coordinates, and then projecting (x, y) to the new coordinates.

Let (Z_-, Z_+) denote the random variables obtained from rotating the random vector (X, Y) by an angle of θ. Then we have

$$\begin{aligned} Z_-(\theta) &= X\cos\theta - Y\sin\theta, \\ Z_+(\theta) &= X\sin\theta + Y\cos\theta. \end{aligned}$$

The voters vote then on the new issues Z_- and Z_+, instead of the original issues X and Y. By the simple majority rule, the voting outcome will be $(m_-(\theta, \mathbf{t}_1, ..., \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, ..., \mathbf{t}_n))$ where

$$\begin{aligned} m_-(\theta, \mathbf{t}_1, ..., \mathbf{t}_n) &= \text{median}\,(x_1\cos\theta - y_1\sin\theta, ..., x_n\cos\theta - x_n\sin\theta), && (1) \\ m_+(\theta, \mathbf{t}_1, ..., \mathbf{t}_n) &= \text{median}\,(x_1\sin\theta + y_1\cos\theta, ..., x_n\sin\theta + x_n\cos\theta), && (2) \end{aligned}$$

are the *marginal medians* after the rotation.

It is easy to verify that the mean $\bar{\mathbf{t}}$ of $\mathbf{t}_1, ..., \mathbf{t}_n$ is invariant to rotations (or *rotation equ-invariant*), i.e. the mean of rotated peaks is simply the rotated mean of the original peaks. In marked contrast, the marginal medians $(m_-(\theta, \mathbf{t}_1, ..., \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, ..., \mathbf{t}_n))$ **are not rotation equivariant**, i.e., rotating and taking medians is not the same as taking medians and rotating. Therefore, rotations are instruments by which the planner may try to influence welfare.

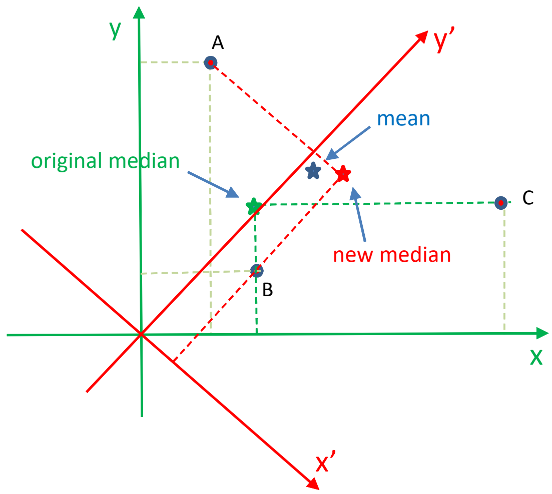


Figure 1. The mean is rotation equ-invariant, the median is not

The reason for the complex behavior of the median is the non-linearity of the median under convolutions.

Example 1 *Consider a discrete random variable X with three possible realizations, $a \le b \le c$, and $p_a = p_c = \frac{2}{5}$ and $p_b = \frac{1}{5}$.*[9] *The median of X is $m_X = b$ and its mean is $\mu_X = \frac{1}{5}(2a + b + 2c)$. Let σ_X^2 denote the variance of X. We have*

$$m_X \le \mu_X \Leftrightarrow a + c \ge 2b. \tag{3}$$

Consider another I.I.D. variable Y. The expected per capita utility with large number of voters by choosing marginal medians for each coordinate is

$$\begin{aligned} U(0) &= -\mathbb{E}[(X - m_X)^2 + (Y - m_Y)^2] \\ &= -2\sigma_X^2 - 2(\mu_X - m_X)^2 \\ &= -2\sigma_X^2 - \frac{8}{25}(a + c - 2b)^2 . \end{aligned}$$

Now suppose we rotate clockwise the two coordinates by $\frac{\pi}{4}$, and then project (X, Y) to the new coordinates. We obtain then two new random variables $\frac{\sqrt{2}}{2}X - \frac{\sqrt{2}}{2}Y$ and $\frac{\sqrt{2}}{2}X + \frac{\sqrt{2}}{2}Y$. It is easily seen that $\frac{\sqrt{2}}{2}X - \frac{\sqrt{2}}{2}Y$ is symmetric, so its median m_- and mean μ_- are both equal to zero. The random variable $\frac{\sqrt{2}}{2}X + \frac{\sqrt{2}}{2}Y$ has the following distribution:

value	$\sqrt{2}a$	$\frac{\sqrt{2}}{2}(a+b)$	$\sqrt{2}b$	$\frac{\sqrt{2}}{2}(a+c)$	$\frac{\sqrt{2}}{2}(b+c)$	$\sqrt{2}c$
probability	$\frac{4}{25}$	$\frac{4}{25}$	$\frac{1}{25}$	$\frac{8}{25}$	$\frac{4}{25}$	$\frac{4}{25}$

[9] Any other probabilities with $p_a = p_c > \frac{1}{3}$ will do as well.

Therefore, it has the mean $\mu_+ = \sqrt{2}\mu_X$ *which is the sum of the means of* $\frac{\sqrt{2}}{2}X$ *and* $\frac{\sqrt{2}}{2}Y$*, and its median is* $m_+ = \frac{\sqrt{2}}{2}(a+c)$*. Hence,*

$$m_+ \geq \mu_+ \Leftrightarrow a + c \geq 2b. \tag{4}$$

The sum of the medians of $\frac{\sqrt{2}}{2}X$ *and* $\frac{\sqrt{2}}{2}Y$ *is* $\sqrt{2}b$*, so the median of sum is* ***not*** *the same as the sum of the medians! In fact,*

$$m_+ \geq m_{\frac{\sqrt{2}}{2}X} + m_{\frac{\sqrt{2}}{2}Y} \Leftrightarrow a + c \geq 2b. \tag{5}$$

The expected utility associated with the $\frac{\pi}{4}$ *rotation is:*

$$\begin{aligned}
U(\frac{\pi}{4}) &= -\mathbb{E}\left[\left((X+Y)/\sqrt{2} - m_+\right)^2 + \left((X-Y)/\sqrt{2} - m_-\right)^2\right] \\
&= -2\sigma_X^2 - (\mu_+ - m_+)^2 \\
&= -2\sigma_X^2 - \frac{1}{50}(a+c-2b)^2 \\
&\geq U(0)
\end{aligned}$$

The inequality is strict if $a + c \neq 2b$*, i.e. if* $\mu_X \neq m_X$*. Therefore, the* $\frac{\pi}{4}$*-rotation generates higher social welfare than the* 0*-rotation.*

More generally, we could also consider an additional translation of the origin, say by a vector **w,** to obtain new orthogonal coordinates (and create new issues). The joint operation of rotation and translation can also be represented by a linear matrix.[10] But, medians (and means) are translation equ-invariant, and thus there is no extra welfare advantage from such translations. Therefore, we focus below on the family of rotations of coordinates around a fixed origin, described by the angle of rotation θ relative to standard Cartesian coordinates.

2.3 The Set of Voting Mechanisms

Kim and Roush [1984] and Peters et al. [1992] provided a complementary justification for our focus on simple, majority voting mechanisms. For any rotation angle $\theta \in [0, 2\pi]$, we can define the direct *marginal median mechanism* φ_θ as

$$\psi_\theta(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n) = (m_-(\theta, \mathbf{t}_1, .., \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, .., \mathbf{t}_n)),$$

[10] This set of general transformation matrices (rotation and translation) is called the *special orthogonal group* for the plane, and is denoted by $SO(2)$. Each matrix in $SO(2)$ is an orthogonal matrix. It is special because the determinant of each matrix is $+1$, whereas the determinant could be -1 for other orthogonal transformations such as reflections.

where $(m_-(\theta, \mathbf{t}_1, .., \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, .., \mathbf{t}_n))$ is the marginal median with respect to rotation θ and reported peaks $\mathbf{t}_i$ as defined in (1) and (2). Note that the function $\psi_\theta(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)$ is continuous in θ and in all its other arguments since both rotations and medians are continuous functions. It is also easy to see that ψ_θ is anonymous[11] and dominant-strategy incentive compatible (DIC).

Surprisingly enough, it turns out that the set of marginal median mechanisms (for all possible rotations) coincides with the entire class of anonymous, Pareto optimal[12] and DIC mechanisms.

Theorem 1 *(Kim and Roush [1984] and Peters et al. [1992]) A mechanism $\psi_\theta(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)$ is anonymous, Pareto optimal and DIC if and only if it is a marginal median mechanism $\psi_\theta(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)$ for some angle $\theta \in [0, 2\pi]$.*

It is worth noting that the characterization in Theorem 1 fails for higher dimensions because anonymous, Pareto optimal and DIC mechanisms need not exist: the vector of marginal medians need not be in the convex hull of the agents' peaks. Hence, our analysis can be extended to higher dimensional problems, but the solution need not be ex-post Pareto optimal.

The next result shows that a simple and intuitive indirect method to implement the entire class of mechanisms described in the above theorem is to define the issues (via rotations) and then sequentially vote by simply majority, one issue at a time, using a binary, sequential voting procedure with a convex agenda.[13] This defines then a structure induced equilibrium à là Shepsle [1979].

Theorem 2 *Assume that agents decide one issue at a time on the orthogonal dimensions $Z_+(\theta)$ and $Z_-(\theta)$ that are obtained by rotating original issues X and Y. Assume also that the vote on each issue is by simple majority according to a convex, binary sequential procedure. Then sincere voting is an ex-post equilibrium and the outcome is $(m_-(\theta, \mathbf{t}_1, ..., \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, ..., \mathbf{t}_n))$, independently of the order in which the issues are put up to vote.*[14]

[11] A mechanism ψ is *anonymous* if, for any profile of reports $(\mathbf{t}_i, \mathbf{t}_{-i})$, $\psi(\mathbf{t}_1, ..., \mathbf{t}_i, ..., \mathbf{t}_n) = \psi\left(\mathbf{t}_{p(1)}, ..., \mathbf{t}_{p(i)}, ..., \mathbf{t}_{p(n)}\right)$, where p is any permutation of the set $\{1, ..., n\}$.

[12] A mechanism ψ is *Pareto optimal* (or Pareto efficient) if, for any profile of reports $(\mathbf{t}_i, \mathbf{t}_{-i})$, there is no alternative $\mathbf{v}$ such that $||\mathbf{t}_i - \mathbf{v}||^2 \leq ||\mathbf{t}_i - \psi(\mathbf{t}_i, \mathbf{t}_{-i})||^2$ for all i, with strict inequality for at least one agent. Pareto optimality requires $\psi(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)$ to be in the convex hull $conv(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)$ for every type profile $(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)$ (see Lemma 2.2 in Peters et al. [1992]).

[13] At each stage of convex, sequential procedure on a fixed dimension, a binary decision is collectively taken among two ideologically coherent sets of alternatives that create a clear left-right divide. For details see Gershkov, Moldovanu and Shi (2016) and Kleiner and Moldovanu (2016).

[14] Sincere voting means that, at each binary decision note, an agent votes for the subset of alternatives containing his/her preferred alternative among those that are still relevant.

Proof. Assume that voters decide first on dimension $Z_+(\theta)$, and then on dimension $Z_-(\theta)$, and recall that these are orthogonal. Denote the first decision by $k_+(\theta, \mathbf{t}_1, ..., \mathbf{t}_n)$. This fixes the first coordinate of the final decision. In other words, at the second stage the agents choose only among alternatives of the form$(k_+(\theta, \mathbf{t}_1, ..., \mathbf{t}_n), z_-)$. This is a one-dimensional problem, on which agents have single peaked preferences. For any $k_+(\theta, \mathbf{t}_1, ..., \mathbf{t}_n)$, the ex-post equilibrium outcome of any binary, sequential voting with a convex agenda is sincere voting, and the outcome is the Condorcet winner $z_- = m_-(\theta, \mathbf{t}_1, ..., \mathbf{t}_n)$. Given this outcome, the first decision is a choice among alternatives of the form $(z_+, m_-(\theta, \mathbf{t}_1, ..., \mathbf{t}_n))$. Since this is again a one-dimensional problem, the outcome is the Condorcet winner, and the final outcome is $((m_+(\theta, \mathbf{t}_1, ..., \mathbf{t}_n), m_-(\theta, \mathbf{t}_1, ..., \mathbf{t}_n))$. An analogous reasoning yields the result for the other order of votes on the two issues. ■

3 The Limit Case when the Number of Agents Is Large

The full probabilistic optimization problem can be rewritten as

$$(\mathcal{P}) \qquad \min_{\theta \in [0,2\pi]} \int_D ... \int_D \left(\frac{1}{n} \sum_{i=1}^{n} ||R(\theta)\mathbf{t}_i - \varphi_\theta(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)||^2 \right) f(\mathbf{t}_1)...f(\mathbf{t}_n) d\mathbf{t}_1...d\mathbf{t}_n.$$

We focus here on the solution to problem $(\mathcal{P})$ when the number of agents is large. But, note that the resulting optimal mechanism will be incentive compatible, Pareto optimal and anonymous for any number of voters.[15] For a random variable X with finite mean μ_X and variance σ_X^2, we know from the central limit theorem that

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu_X\right) \to N(0, \sigma_X^2).$$

Bahadur (1966) showed that the quantiles of large samples display a similar behavior. In particular,

$$\sqrt{n}(X_{(n+1)/2:n} - m_X) \to N\left(0, \frac{1}{4f^2(m_X)}\right)$$

where

$$X_{(n+1)/2:n} = \text{median } (X_1, ..., X_n)$$

[15] This contrasts trivial incentive compatible mechanisms such as always choosing a fixed alternative, which may yield "catastrophic" results for a finite number of agents and particular realizations of types.

and where m_X is the median of the distribution. Thus, as n goes to infinity, the sample median converges to the median of the underlying distribution (and, of course, the sample mean converges to the mean).

By applying the above limit results to our setting, we obtain that, as $n \to \infty$,

$$\left(\begin{array}{c} m_{-}(\theta, \mathbf{t}_1, .., \mathbf{t}_n) \\ m_{+}(\theta, \mathbf{t}_1, .., \mathbf{t}_n) \end{array} \right) \longrightarrow \mathbf{m}(\theta) \equiv \left(\begin{array}{c} m_{-}(\theta) \\ m_{+}(\theta) \end{array} \right) \equiv \left(\begin{array}{c} \text{median } (X\cos\theta - Y\sin\theta) \\ \text{median } (X\sin\theta + Y\cos\theta) \end{array} \right)$$

Furthermore, since the norm operation $||\cdot||$ is continuous, we obtain that, as $n \to \infty$,

$$\begin{aligned} & \frac{1}{n}\sum_{i=1}^{n} ||R(\theta)\mathbf{t}_i - \varphi_\theta(\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_n)||^2 \\ = \; & \frac{1}{n}\sum_{i=1}^{n} \left[(x_i\cos\theta - y_i\sin\theta - m_{-}(\theta, \mathbf{t}_1, .., \mathbf{t}_n))^2 + (x_i\sin\theta + y_i\cos\theta - m_{+}(\theta, \mathbf{t}_1, .., \mathbf{t}_n))^2 \right] \\ \to \; & \mathbb{E}\,||X\cos\theta - Y\sin\theta - m_{-}(\theta), X\sin\theta + Y\cos\theta - m_{+}(\theta)||^2 \\ = \; & \sigma_X^2 + \sigma_Y^2 + \big(\mu_{-}(\theta) - m_{-}(\theta)\big)^2 + \big(\mu_{+}(\theta) - m_{+}(\theta)\big)^2 \end{aligned}$$

where

$$\mu_{-}(\theta) = \mu_X\cos\theta - \mu_Y\sin\theta, \text{ and } \mu_{+}(\theta) = \mu_X\sin\theta + \mu_Y\cos\theta$$

Therefore, in the limit where n is very large, our problem becomes

$$(\mathcal{P}) \qquad \min_{\theta\in[0,2\pi]} \big(\mu_{-}(\theta) - m_{-}(\theta)\big)^2 + \big(\mu_{+}(\theta) - m_{+}(\theta)\big)^2 + \sigma_X^2 + \sigma_Y^2.$$

In other words, we look for the rotation that creates the marginal median vector with the minimum distance from the mean.

For most parts of the analysis below, it will be convenient to normalize the means of X and Y to be zero - such a normalization is without loss of generality because of the translational equ-invariance of both mean and median. Let us define the normalized random variables $\tilde{X}$ and $\tilde{Y}$ as

$$\tilde{X} = X - \mu_X \text{ and } \tilde{Y} = Y - \mu_Y.$$

The corresponding normalized marginal medians $(\tilde{m}_{-}(\theta), \tilde{m}_{+}(\theta))$ are

$$\tilde{m}_{-}(\theta) = m_{-}(\theta) - \mu_{-}(\theta) \quad \text{and} \quad \tilde{m}_{+}(\theta) = m_{+}(\theta) - \mu_{+}(\theta).$$

Hence, the planner's problem becomes

$$(\mathcal{P}) \qquad \min_{\theta\in[0,2\pi]} \tilde{m}_{-}^2(\theta) + \tilde{m}_{+}^2(\theta) + \sigma_X^2 + \sigma_Y^2.$$

Since variances are fixed, the planner's goal under this normalization is simply to find the rotation resulting in a marginal median vector with minimum norm. To simplify notation, we shall drop the tilde symbol for normalized random variables where no confusion can arise.

The first basic Lemma shows that it is without loss of generality to restrict attention to rotations in the interval $[0, \pi/2]$.

Lemma 1 *For any $\theta \in [\pi/2, 2\pi]$ that minimizes the planner's objective, there exists $\theta' \in [0, \pi/2]$ that attains the same minimum.*

Proof. See the Appendix. ■

Consider the normalized planner's problem. The first order condition for the optimal rotation θ is then:

$$FOC : m_-(\theta)m'_-(\theta) + m_+(\theta)m'_+(\theta) = 0 \Leftrightarrow \langle \mathbf{m}(\theta), \mathbf{m}'(\theta) \rangle = 0 \tag{6}$$

In words, the vector of marginal medians and the vector of its derivatives must be orthogonal. The second order condition for the local optimality of θ is

$$m''_-(\theta)m_-(\theta) + (m'_-(\theta))^2 + m''_+(\theta)m_+(\theta) + (m'_+(\theta))^2 > 0 \tag{7}$$

and for a critical value θ to be locally sub-optimal we need

$$SOC : m''_-(\theta)m_-(\theta) + (m'_-(\theta))^2 + m''_+(\theta)m_+(\theta) + (m'_+(\theta))^2 < 0. \tag{8}$$

3.1 Sub-Optimality of Voting on Independent Issues

In this subsection, we assume that the marginals X and Y are independent. We work on the normalized version of the planner's problem. The zero-angle rotation corresponds then to votes on independent issues X and Y. Our goal is to show that the zero-angle rotation yields a local maximum of norm of the normalized marginal median, or in other words, it leads to a local utility minimum, and is thus sub-optimal.

Theorem 3 *Assume that X and Y are independent. Then the following hold:*

1. *The rotation with angle $\theta = 0$ is a critical point, i.e., it satisfies the first-order condition.*

2. *The rotation with angle $\theta = 0$ is a local utility minimum if*

$$m_X f'_X(m_X) \geq 0, m_Y f'_Y(m_Y) \geq 0, m_X^2 + m_Y^2 \neq 0.$$

Proof. See Appendix A. ■

Corollary 1 *Assume that X and Y are unimodal and independent.*[16] *Suppose also that X and Y satisfy*

$$\begin{aligned} M_X &\leq m_X \leq \mu_X \quad or \quad \mu_X \leq m_X \leq M_X \\ M_Y &\leq m_Y \leq \mu_Y \quad or \quad \mu_Y \leq m_Y \leq M_Y \end{aligned}$$

where M, m, μ are mode, median and mean, respectively. Then the rotation with angle $\theta = 0$ is a local utility minimum.

Proof. If $M_X \leq m_X \leq \mu_X = 0$ (where the last equality holds by normalization), then $m_X \leq 0$ and $f'(m_X) \leq 0$ because m_X is to the right of the mode. Hence $m_X f'_X(m_X) \geq 0$. If $0 = \mu_X \leq m_X \leq M_X$, then $m_X \geq 0$ and $f'(m_X) \geq 0$ because m_X is to the left of the mode. Hence $m_X f'_X(m_X) \geq 0$, and analogously for Y. ■

The sufficient condition stated in the above Corollary has the advantage that it is very intuitive: there are elegant, general characterizations of distributions where such orders of the mode, median, mean hold (see for example, Dharmadhikari and Joag-Dev [1988], Basu and DasGupta [1997]).

The next figure geometrically illustrates the intuition of suboptimality of voting on independent issues. Assume that $0 = \mu_X \leq m_X$ and $0 = \mu_Y \leq m_Y$. We want to show that a small rotation improves welfare if $f'_X(m_X) \geq 0$ and $f'_Y(m_Y) \geq 0$.

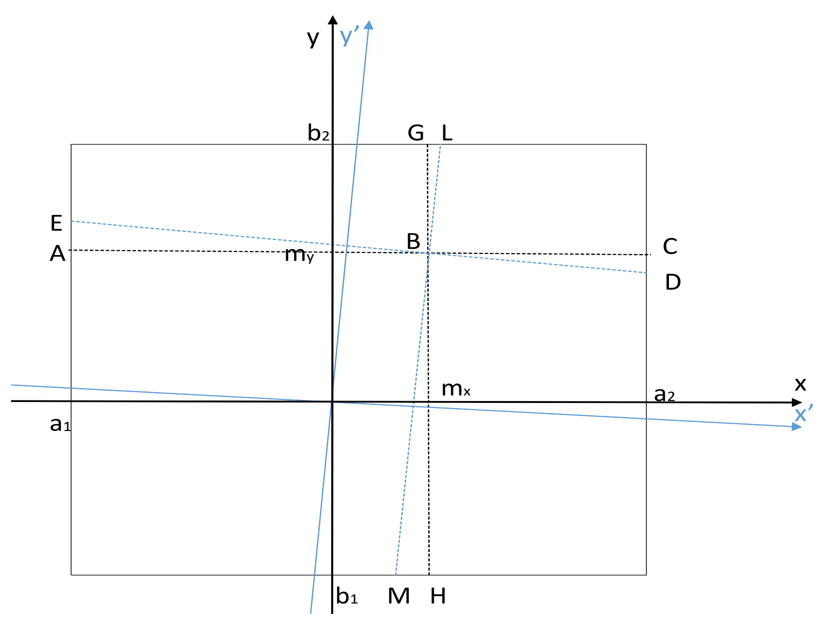


Figure 2. Small rotation improves welfare if $f'_X(m_X) \geq 0$ and $f'_Y(m_Y) \geq 0$

[16] A random variable Z is *unimodal* if its density $f(z)$ has a single mode (or peak).

Assume that the unrotated median is B. Therefore, by independence, there is a mass of 50% above the AC line and a mass of 50% to the right of GH line. Consider a small rotation with angle $\theta > 0$, so that new axes are x' and y'. We want to show that this moves the new median towards the mean $(0,0)$. That is, we want to show that the median moves towards the south-west. Consider the projection of B onto the new, rotated axes: the result will follow if the mass above DE and the mass to the right of LM are both below 50%. If the area of ABE is larger than the one of BCD, we obtain that the mass above ED is indeed smaller than 0.5 (the comparison for the other dimension is analogous).

For illustration purposes, let us assume that X and Y distribute on bounded intervals $[a_1, a_2]$ and $[b_1, b_2]$, respectively. The line DE passing through point B is given by $y = m_Y - (x - m_X)\tan\theta$. Therefore, the difference between the areas ABE and BCD is

$$ABE - BCD = \int_{a_1}^{a_2} \left[F_Y\left(m_Y - (x - m_X)\tan\theta\right) - F_Y\left(m_Y\right)\right] f_X(x)\, dx$$

Since $f'_Y(m_Y) \geq 0$, F_Y is locally convex at m_Y. Therefore, for sufficiently small θ, the curve $F_Y\left(m_Y - (x - m_X)\tan\theta\right)$ with $x \in [a_1, a_2]$ lies above the tangent line $F_Y(m_Y) + f_Y(m_Y)(m_X - x)\tan\theta$. As a result, for sufficiently small θ, we have

$$ABE - BCD \geq \int_{a_1}^{a_2} f_Y(m_Y)(m_X - x)\tan\theta f_X(x)\, dx = f_Y(m_Y)\, m_X \tan\theta > 0$$

as desired. The argument for the other dimension is analogous.

The main arguments in the rigorous proof of Theorem 3 are as follows: the first-order condition (6) evaluated at $\theta = 0$ is

$$m_-(0)\, m'_-(0) + m_+(0)\, m'_+(0) = 0 \tag{9}$$

and the second order condition (8) for sub-optimality, evaluated at $\theta = 0$, is

$$\left(m'_-(0)\right)^2 + m_-(0)\, m''_-(0) + \left(m'_+(0)\right)^2 + m_+(0)\, m''_+(0) < 0. \tag{10}$$

We first show that $m'_-(0) = m'_+(0) = 0$, which implies that condition (9) is fulfilled (recall that $m_X = m_-(0)$ and $m_Y = m_+(0)$). Condition (10) is then reduced to

$$m_X m''_-(0) + m_Y m''_+(0) < 0.$$

The main thrust of the proof is an application of the characteristic function approach (or inverse Fourier Transform) in order to show that the first-order condition (6), and the second-order condition (8) hold at $\theta = 0$. The characteristic function of a random variable Z is given by

$$\varphi_Z(t) = \mathbb{E}\left(e^{itZ}\right)$$

where i is the imaginary unit. Its main convolution property – used heavily in the proof – is

$$\varphi_{aX+bY}\left(t\right)=\mathbb{E}\left(e^{itX}\right)\mathbb{E}\left(e^{itY}\right)=\varphi_{X}\left(at\right)\varphi_{Y}\left(bt\right)$$

for independent random variables X, Y and real constants a, b. Therefore, by setting $a=\sin\theta$ and $b=\cos\theta$, we obtain

$$\begin{aligned}\varphi_{X\sin\theta+Y\cos\theta}\left(t\right) &= \varphi\left(t\sin\theta\right)\varphi\left(t\cos\theta\right),\\ \varphi_{X\cos\theta-Y\sin\theta}\left(t\right) &= \varphi\left(t\cos\theta\right)\varphi\left(-t\sin\theta\right).\end{aligned}$$

From the Fourier Inversion Theorem (see Gil-Pelaez [1951] or Shephard [1991]) we know that, for any random variable $Z(\theta)=X\cos\theta+Y\sin\theta$, we can uniquely recover its distribution from its characteristic function by the formula:

$$F_{Z(\theta)}(z)=\frac{1}{2}-\frac{1}{2\pi}\int_{0}^{\infty}\frac{\varphi_{Z(\theta)}\left(t\right)e^{-itz}-\varphi_{Z(\theta)}\left(-t\right)e^{itz}}{it}dt$$

Since by the definition of the median $m_{Z(\theta)}$, $F_{Z(\theta)}(m_{Z(\theta)})=\frac{1}{2}$, we obtain that

$$\int_{0}^{\infty}\frac{\varphi_{Z(\theta)}\left(t\right)e^{-itm_{Z(\theta)}}-\varphi_{Z(\theta)}\left(-t\right)e^{itm_{Z(\theta)}}}{it}dt=0.$$

This condition can be then implicitly differentiated to obtain information about how the median varies with rotation θ.

3.2 The $\pi/4$ Rotation for Distributions with Identical Marginals

We have shown above that a zero-angle rotation is sub-optimal if the issues brought to vote are such that the distribution of ideal points has independent marginals. In other words, it is not optimal to vote on independent issues, and, locally around zero, some rotation is always welfare improving. What is the optimal way to structure new issues? Formally, what is the optimal rotation that maximizes the per capita voters' expected utility?

In this subsection, we assume that the marginals X and Y are identically (but not necessarily independently) distributed. Then, by symmetry, the $\pi/4$ rotation is a natural candidate for the optimal rotation. For $\theta=\pi/4$ we have

$$\begin{pmatrix} m_{-}\left(\theta\right)\\ m_{+}\left(\theta\right)\end{pmatrix}=\begin{pmatrix}\text{median }\left(\frac{\sqrt{2}}{2}\left(X-Y\right)\right)\\ \text{median }\left(\frac{\sqrt{2}}{2}\left(X+Y\right)\right)\end{pmatrix}=\frac{\sqrt{2}}{2}\begin{pmatrix}0\\ \text{median }\ \left(X+Y\right)\end{pmatrix}$$

The last equality follows because median$(\lambda Z)=\lambda$median(Z) for any random variable Z, and because $X-Y$ is a symmetric random variable, where the median equals the mean (which, recall, is normalized to be zero). For our purposes, this implies that the $\pi/4$ rotation completely eliminates the conflict arising between efficiency and incentive compatibility along one dimension – all remaining such conflict is concentrated in the other dimension, as illustrated in the following figure (assuming $m_X > \mu_X = 0$):

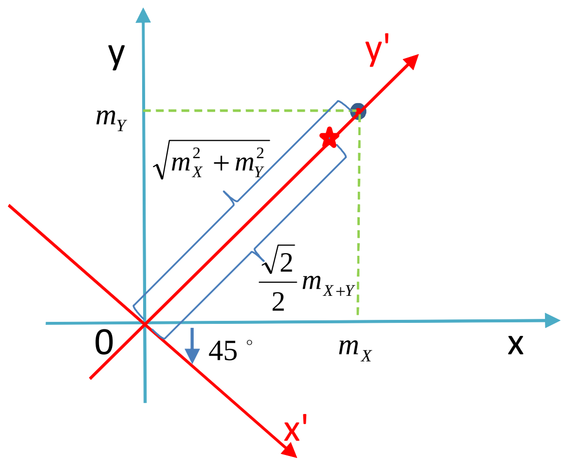


Figure 3. The $\pi/4$ rotation with symmetric marginals

The $\pi/4$ rotation has the following interpretation: Instead of voting X and Y separately, the vote is on issues $X+Y$ and $X-Y$, and the outcome is determined by the simple majority on each issue. Once voters have decided on $X+Y$ and $X-Y$, the planner can then obviously recover X and Y. The two-steps voting procedure associated with $\pi/4$ rotation resembles the "top-down" budgeting procedure widely used in practice: first a total budget is determined, and then it is allocated among several items.

We first compare the expected utility under the $\frac{\pi}{4}$ rotation with that under the 0 rotation when X and Y are identically distributed. As is also apparent from Figure 3, this amounts to check whether the original coordinate-wise median vector (m_X, m_Y) is closer to the origin than the new coordinate-wise median vector $(m_{X+Y}/2, m_{X+Y}/2)$, or vice-versa.

Theorem 4 *Suppose X and Y are identically distributed and $m_X \neq \mu_X$. In addition, suppose that convolution $X+Y$ maintains the relative magnitude of the mean and median, that is,*

$$m_X < (>)\, \mu_X \Rightarrow m_{X+Y} < (>)\, \mu_{X+Y}.$$

If $m_X < (>)\, \mu_X$, and if the median function is super-additive (sub-additive)

$$m_X + m_Y < (>)\, m_{X+Y}, \tag{11}$$

then the expected utility at $\theta = \frac{\pi}{4}$ exceeds the expected utility at $\theta = 0$.

Proof. Suppose that $m_X < \mu_X$ and that $\mu_X = 0$. The proof for the other case is completely analogous. By assumption, $m_X = m_Y \leq 0$ and $m_+(\frac{\pi}{4}) \leq 0$. The expected utility at $\theta = 0$ is

$$U(0) = -2\sigma_X^2 - 2m_X^2$$

and the expected utility at $\theta = \frac{\pi}{4}$ is

$$U(\frac{\pi}{4}) = -2\sigma_X^2 - m_+^2(\frac{\pi}{4})$$

Given our assumptions, we have

$$U(\frac{\pi}{4}) > U(0) \Leftrightarrow m_+(\frac{\pi}{4}) > \sqrt{2}m_X \Leftrightarrow m_{\frac{\sqrt{2}}{2}(X+Y)} > \sqrt{2}m_X \Leftrightarrow m_{X+Y} > 2m_X,$$

where we use the fact that for any random variable Z, it holds that $\lambda m_Z = m_{\lambda Z}$. ■

Remark 2 *The above sufficient conditions can also be directly applied to compare the level of total budget between the "bottom-up" and "top-down" budgeting procedures mentioned in the Introduction.*[17] *Whenever the median function is super (sub)-additive, the top-down procedure where a total budget is determined first leads to a higher (lower) overall budget than the bottom-up procedure where votes are item by item and the total budget is gradually determined.*

The super-additivity (or sub-additivity) condition on the median function, though elegant, may not be easily verified directly since it involves the computation of the convolution and its median. Assuming that X and Y are I.I.D. , we can present a simple, directly verifiable, sufficient condition that simultaneously guarantees $m_X < (>)\, \mu_X$ and $m_X + m_Y < (>)\, m_{X+Y}$.

Proposition 1 *Suppose that X and Y are I.I.D. and that $m_X \neq \mu_X$. If*

$$F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \leq (\geq)\, 1 \textit{ for all } \varepsilon > 0, \tag{12}$$

then

$$m_X < (>)\, \mu_X \textit{ and } m_X + m_Y < (>)\, m_{X+Y}.$$

Proof. See Appendix A. ■

It is worth noting that van Zwet [1979] shows that the same condition (12) implies that $\mu_X < m_X < M_X$ ($\mu_X > m_X > M_X$). It follows from Corollary 1 that condition (12) also implies the sufficient condition in Theorem 3 for zero rotation to be suboptimal.

[17] Note that this question is not identical to the question of utility comparisons.

We can apply Proposition 1 to show that, if F is either strictly convex or concave, then the $\pi/4$ rotation is strictly better than zero rotation.[18]

Corollary 2 *Suppose that X and Y are I.I.D. and that $\mu_X \neq m_X$. If $F(x)$ is strictly convex or strictly concave, then the expected utility at $\theta = \pi/4$ is strictly higher than the expected utility at $\theta = 0$.*

Proof. Note that $F(X)$ is uniformly distributed random variable , so that $E[F(X)] = 1/2$. Suppose that F is strictly convex. The concave case can be proved analogously. By Jensen's inequality

$$F(m_X) = \frac{1}{2} = E[F(X)] > F(E[X]) = F(\mu_X).$$

Hence, $m_X > \mu_X$. In order to show that $m_X + m_Y > m_{X+Y}$, it is sufficient to show that

$$F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \geq 1 \text{ for all } \varepsilon > 0.$$

Note that $f_X(m_X + \varepsilon) - f_X(m_X - \varepsilon) > 0$ by strict convexity of F, so $F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon)$ is increasing in ε and reaches a minimum at $\varepsilon = 0$. Since $F_X(m_X) + F_X(m_X) = 1$, we must have $F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \geq 1$ for all $\varepsilon > 0$. ■

It follows from Corollary 2 that the $\pi/4$ rotation strictly dominates the zero rotation for the exponential distribution because it is strictly concave. The domination also holds for the class of power function distribution $F(x) = x^k$ with $k > 0$ and $k \neq 1$, because $F(x)$ is strictly concave when $k < 1$ and strictly convex when $k > 1$.

We now show how the super-additivity condition in Theorem 4 is satisfied for two well-known families of distributions where condition (12) is not easily checked, or does not hold.[19] This requires a few definitions and some results that use majorization and Schur-convexity arguments.

Definition 1 *A vector (a, b) is said to majorize (a', b'), written as $(a, b) \succ (a', b')$, if $a + b = a' + b'$ and if $\max(a, b) > \max\{a', b'\}$. A function $h(a, b)$ is said to be Schur-convex in (a, b) if $h(a'', b'') \geq h(a', b')$ whenever $(a'', b'') \succ (a', b')$, and Schur-concave in (a, b) if $h(a'', b'') \leq h(a', b')$ whenever $(a'', b'') \succ (a', b')$.*

[18] Note that, if X has a bounded support (a, b), a sufficient condition for the case of $\mu_X < m_X$ is

$$F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \geq 1 \text{ for all } \varepsilon \in (0, b - m_X).$$

and a sufficient condition for the other case is

$$F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \leq 1 \text{ for all } \varepsilon \in (0, m_X - a).$$

[19] Although the super-additivity (or sub-additivity) condition is derived for normalized distributions, it is straightforward to verify that it is also sufficient for original distributions.

Consider first the large and important family of Gamma distributions with density

$$f_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \text{ for } x > 0.$$

This family contains the Exponential (that can be obtained by setting $\alpha = 1$) and many other well known distributions. For any constant $c > 0$, the random variable cX is also Gamma with parameters α and β/c. If X and Y are independent Gamma with parameters (α_X, β) and (α_Y, β), respectively, then $X + Y$ is also Gamma with parameters $(\alpha_X + \alpha_Y, \beta)$. Thus, the Gamma family is closed under scaling and under convolution. In a classic study, Bock et al. [1987] showed that $\Pr(aX + bY \leq t)$, $0 \leq a, b \leq 1$, is Schur-convex in (a, b) for all $t \leq \mu_X$. Since $(1, 0) \succ \left(\frac{1}{2}, \frac{1}{2}\right)$, we have $F_{\frac{1}{2}X+\frac{1}{2}Y}(t) \leq F_X(t)$ for all $t \leq m_X$. This implies $m_{\frac{1}{2}X+\frac{1}{2}Y} \geq m_X$ as desired.[20]

A second family is the Rayleigh distribution with cumulative distribution

$$F(x) = 1 - e^{-x^2} \text{ for } x \geq 0.$$

Suppose X, Y are I.I.D. distributed according to Rayleigh.[21] Then, according to Lemma 4 in Hu and Lin [2000], we have

$$\Pr(X\cos\theta + Y\sin\theta \leq z) = 1 - \int_0^{\pi/2} \sin(2\tau) \left(1 + \phi^2(\theta, \tau, z)\right) e^{-\phi^2(\theta,\tau,z)} d\tau$$

[20] Alternatively, let $m(\alpha, \beta)$ denote the median of Gamma random variable X with parameters α and β. Then $m(\alpha, \beta) = m(\alpha, 1)/\beta$. Note that

$$\begin{aligned} U(\tfrac{\pi}{4}) &= -2\sigma^2(\alpha, \beta) - \left(\mu_+ - m_+\right)^2 \\ &= -2\sigma^2(\alpha, \beta) - \left(\frac{\sqrt{2}\alpha}{\beta} - \frac{\sqrt{2}}{2\beta} m(2\alpha, 1)\right)^2 \\ &= -2\sigma^2(\alpha, \beta) - \frac{1}{2\beta^2}(2\alpha - m(2\alpha, 1))^2 \end{aligned}$$

and

$$U(0) = -2\sigma^2(\alpha, \beta) - 2\left(\mu_X - m_X\right)^2 = -2\sigma^2(\alpha, \beta) - \frac{2}{\beta^2}(\alpha - m(\alpha, 1))^2$$

Therefore,

$$\begin{aligned} U(\tfrac{\pi}{4}) > U(0) &\Leftrightarrow \frac{1}{2\beta^2}(2\alpha - m(2\alpha, 1))^2 < \frac{2}{\beta^2}(\alpha - m(\alpha, 1))^2 \\ &\Leftrightarrow (2\alpha - m(2\alpha, 1))^2 < 4(\alpha - m(\alpha, 1))^2 \\ &\Leftrightarrow m^2(2\alpha, 1) - 4\alpha m(2\alpha, 1) < 4m^2(\alpha, 1) - 8\alpha m(\alpha, 1) \\ &\Leftrightarrow m(2\alpha, 1) > 2m(\alpha, 1) \end{aligned}$$

The last inequality holds because, as shown in Berg and Pedersen [2008], $m(\alpha, 1)$ is convex in α.

[21] If Z_1, Z_2 is a random sample of size 2 from a normal distribution $N(0, 1)$ then the distribution of $X = \sqrt{Z_1^2 + Z_2^2}$ is Rayleigh. In other words, the Rayleigh is the distribution of the norm of a two-dimensional random vector whose coordinates are normally distributed.

where $\phi(\theta, \tau, z) = z/\cos(\theta - \tau)$. The medians of X and of Y are $m_X = m_Y = \sqrt{\ln 2}$. It can be (numerically) verified that

$$\begin{aligned}
\Pr\left((X+Y)/\sqrt{2} \leq \sqrt{2} m_X\right) &= 1 - \int_0^{\pi/2} \sin(2\tau)\left(1 + \phi^2(\frac{\pi}{4}, \tau, \sqrt{2\ln 2})\right) e^{-\phi^2(\frac{\pi}{4}, \tau, \sqrt{2\ln 2})} d\tau \\
&\approx 0.4658 \\
&< 0.5 \\
&= \Pr\left((X+Y)/\sqrt{2} \leq m_+(\frac{\pi}{4})\right)
\end{aligned}$$

where the last equality follows from the definition of $m_+(\frac{\pi}{4})$. Hence, $m_+(\frac{\pi}{4}) > \sqrt{2} m_X$ as desired.

By assuming independence between X and Y, we were able to derive operational, sufficient conditions for the $\pi/4$ rotation to dominate the zero rotation, but independence is not necessary in general. We now present an example where, even though X and Y are correlated, the median function is super-additive (sub-additive) so the $\pi/4$ rotation is welfare superior to the zero rotation. The standard tool we use to model correlation between X and Y for given marginals is the copula (see Nelson [2006] for an introduction).

Example 2 *Suppose that X and Y are identically distributed on $[0,1]$ with marginals $F_X(x) = x^2$ and $F_Y(y) = y^2$. To model correlation between X and Y, we consider here the Farlie-Gumbel-Morgenstern (FGM) copula*

$$C_\delta(p, q) = pq + \delta pq(1-p)(1-q)$$

with $p, q \in [0,1]$ and $\delta \in [-1,1]$. The correlation coefficient for FGM copula is $\rho = \delta/3 \in [-1/3, 1/3]$. It follows from the Sklar theorem that we can write the joint distribution $F(x, y)$ in terms of its marginals and a copula $C(p, q)$:

$$F(x, y) = C(F_X(x), F_Y(y)).$$

With some algebra, we can derive the joint density as

$$f(x, y) = 4xy + 4\delta xy\left(2x^2 - 1\right)\left(2y^2 - 1\right).$$

Therefore, as in the proof of Proposition 1, we can write $\Pr(X+Y<m_X+m_Y)$ *as*

$$
\begin{aligned}
& 2\int_{m_Y}^{1}\int_{0}^{m_X+m_Y-y} f(x,y)\,dxdy+\int_{0}^{m_Y}\int_{0}^{m_X} f(x,y)\,dxdy \\
= \; & 2\int_{\sqrt{2}/2}^{1}\int_{0}^{\sqrt{2}-y}\left(4xy+4\delta xy\left(2x^2-1\right)\left(2y^2-1\right)\right)dxdy \\
& +\int_{0}^{\sqrt{2}/2}\int_{0}^{\sqrt{2}/2}\left(4xy+4\delta xy\left(2x^2-1\right)\left(2y^2-1\right)\right)dxdy \\
= \; & \left(\frac{146}{35}-\frac{104}{35}\sqrt{2}\right)\delta-\frac{8}{3}\sqrt{2}+\frac{13}{3} \\
> \; & 0.5
\end{aligned}
$$

for all $\delta\in[-1,1]$. *Consequently, we have* $m_{X+Y}<m_X+m_Y$. *Since* $F_X(x)=x^2$ *is convex,* $\mu_X<m_X$. *Hence, the sufficient condition (11) in Theorem 4 is fulfilled. Alternatively, suppose* $F_X(x)=\sqrt{x}$ *and* $F_Y(y)=\sqrt{y}$. *If we again restrict attention to the FGM copula, we can follow the same procedure to show that* $m_{X+Y}>m_X+m_Y$ *and* $\mu_X>m_X$.

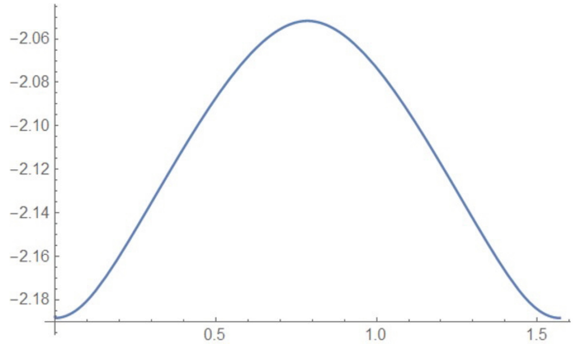
So far we have shown that the $\pi/4$ rotation is often better than the zero rotation. Given the symmetry of X and Y and the sub-optimality of the zero rotation, the $\pi/4$ rotation is the natural candidate for the optimal rotation. Even though all our numerical simulations clearly suggest it, we were unable to analytically prove that the $\pi/4$ rotation is fully optimal. But, if X and Y are I.I.D., we can analytically show that the $\pi/4$ rotation is a critical point, i.e., the first-order condition (6) is satisfied when $\theta=\pi/4$, and numerically show that the $\pi/4$ rotation is indeed optimal for several standard families of distributions.

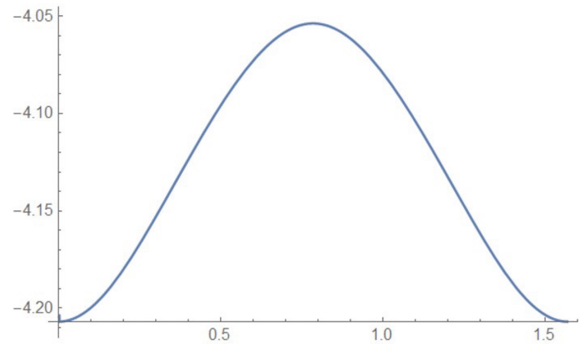
Proposition 2 *For any I.I.D. marginals* X *and* Y, $\theta=\pi/4$ *is a critical point, i.e., it satisfies the first order condition.*

Proof. See Appendix A. ■

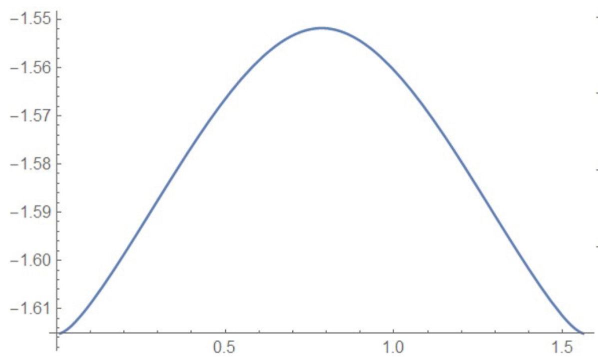
If we can verify second-order conditions either locally or globally, then Proposition 2 can tell us whether $\theta=\pi/4$ is local or global utility maximum. Unfortunately, the second order conditions, evaluated at $\theta=\pi/4$, turn out to be very elusive.

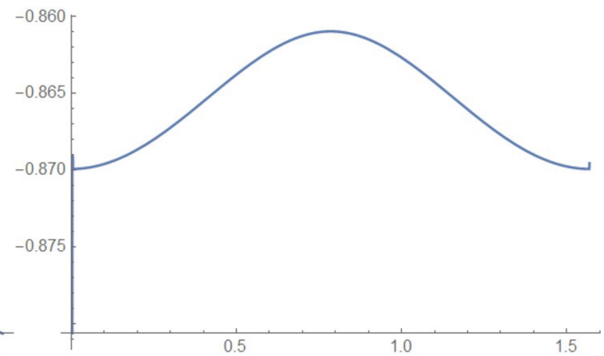
We conclude this section with several numerical simulations. In all these simulations, X and Y are I.I.D., and the original distributions (rather than the normalized ones) are used as inputs. We then use Mathematica to plot the aggregate expected welfare as a function of the rotation angle $\theta\in[0,\pi/2]$. All simulations suggest that a $\pi/4$-rotation (i.e., $\theta\approx 0.785$) is globally optimal. We have also varied the parameter values for the distributions, and the resulting graphs are similar. In fact, we were not able to find any standard distribution where the $\pi/4$-rotation is not globally optimal.





(a) Exponential: $f(x) = e^{-x}$, $x \geq 0$ (b) Gamma: $f(x) = xe^{-x}$, $x \geq 0$





(c) Pareto: $f(x) = x^{-2}$, $x \geq 1$ (d) Rayleigh: $f(x) = 2xe^{-x^2}$, $x \geq 0$

Figure 4. Simulation with I.I.D. marginals

4 Bounds on Relative Efficiency

In this section we provide several lower bounds on the (relative) efficiency loss of the marginal median mechanisms. In particular, for the logconcave case studied by Caplin and Nalebuff ([1988], [1991]), the lower bound is 88% of the first-best utility[22]. Various other bounds are obtained under other assumptions on the distributions governing the distribution of voter's ideal points. The proofs use several classical statistical inequalities, and some more recent concentration inequalities.

Assume that ideal points are distributed such that the marginals are given by random variables (X, Y) where X and Y are not necessarily identical, and are potentially correlated. Since the results heavily use statistical results that establish relations between the mean, median and variance, we work here with the **non-normalized**

[22]Recall that the first best is obtained by choosing the vector of marginal means, which, for any finite number of agents, is not incentive compatible

variables (so that the role of the mean and its relations to the other statistics does not get obscured by the normalization we used above). The first-best expected utility, attained by choosing the mean in each coordinate is given by

$$-\mathbb{E}(X-\mu_X)^2-\mathbb{E}(Y-\mu_Y)^2=-\sigma_X^2-\sigma_Y^2.$$

Note that the first best utility decreases as the variances increase. The expected utility of rotated medians with angle θ is given by

$$U(\theta)=-\sigma_X^2-\sigma_Y^2-\left(\mu_-(\theta)-m_-(\theta)\right)^2-\left(\mu_+(\theta)-m_+(\theta)\right)^2.$$

Thus, the relative efficiency of the rotation with angle θ (relative to first best) is given by:

$$EF(\theta)=\frac{\sigma_X^2+\sigma_Y^2}{\sigma_X^2+\sigma_Y^2+\left(\mu_-(\theta)-m_-(\theta)\right)^2+\left(\mu_+(\theta)-m_+(\theta)\right)^2}\leq 1$$

Observe that two forces play here a role: on the one hand, a distribution that is concentrated around a central location (such as the mean or the median) will have a small difference between mean and median, which tends to increase the relative efficiency. On the other hand, such a distribution also has a low variance so that the difference between mean and median plays a bigger overall role.

It is interesting to note that the covariance of X and Y does not play a direct role in the efficiency calculations: it only enters in the way that the medians of the convolutions are calculated. We define the maximal relative efficiency as

$$EF\equiv\max_{\theta}EF(\theta).$$

The first-best outcome can be attained by majority voting (in the limit with a large number of agents) if the distributions of both X and Y are symmetric around their respective means. In this case we have $\mu_-(\theta)=m_-(\theta)$ and $\mu_+(\theta)=m_+(\theta)$.

Example 3 (Normal Distribution) *Let X and Y be independently distributed normal random variables with zero mean. Then $X\cos\theta-Y\sin\theta$ and $X\sin\theta+Y\cos\theta$ are also normally distributed with mean and also median equal to zero. Thus, the first-best is implementable, and all rotations are welfare equivalent. This proves a conjecture about the normal distribution due to Kim and Roush [1984].*

We now obtain various lower bounds on the attained efficiency for various classes of distributions. We say a random variable X has *increasing failure rate* (IFR) if its hazard rate $f(x)/(1-F(x))$ is increasing in x.

Theorem 5 *The following relative efficiency bounds hold:*

1. *For any random variables X and Y, $EF \geq \frac{1}{2}$;*

2. *If both X and Y are unimodal, then $EF > \frac{5}{8}$;*

3. *If both X and Y have an increasing failure rate (IFR) such that $\mu_X \leq m_X$ and $\mu_Y \leq m_Y$, then $EF > \frac{3}{5}$;*

4. *If X and Y are identically distributed, then $EF \geq \frac{2\sigma_X^2}{3\sigma_X^2 + Cov(X,Y)}$. Thus, if X and Y are independent, $EF \geq \frac{2}{3}$, and in the co-monotonic scenario expected utility cannot be improved by rotation.*

5. *If X and Y are I.I.D., have an increasing failure rate (IFR), and $\mu_X \leq m_X$, then $EF \geq 0.754$;*

6. *If X and Y are I.I.D. and each has a log-concave density, then $EF \geq 0.876$.*

Proof. 1. A classical inequality due to Hotelling and Solomons [1932] says that the square distance between the mean and median of any random variable is always less than variance:

$$(\mu - m)^2 \leq \sigma^2.$$

Therefore,

$$\begin{aligned}
\left(\mu_-(\theta) - m_-(\theta)\right)^2 &\leq \sigma_-^2(\theta) \leq \sigma_X^2 \cos^2\theta + \sigma_Y^2 \sin^2\theta - 2\sin\theta\cos\theta Cov(X,Y) \\
\left(\mu_+(\theta) - m_+(\theta)\right)^2 &\leq \sigma_+^2(\theta) \leq \sigma_X^2 \sin^2\theta + \sigma_Y^2 \cos^2\theta + 2\sin\theta\cos\theta Cov(X,Y)
\end{aligned}$$

Hence we obtain the universal bound:

$$EF(\theta) \geq \frac{\sigma_X^2 + \sigma_Y^2}{2\sigma_X^2 + 2\sigma_Y^2} = \frac{1}{2}$$

2. For the class of unimodal distributions it can be shown that the squared distance between mean and median is at most $\frac{3}{5}$ variance (see Basu and DasGupta [1997]). Thus, for such distributions we get:

$$EF > EF(0) \geq \frac{\sigma_X^2 + \sigma_Y^2}{(\sigma_X^2 + \sigma_Y^2) + \frac{3}{5}(\sigma_X^2 + \sigma_Y^2)} = \frac{5}{8}$$

3. For the class of distributions with an increasing failure rate (IFR), we assume $\mu_X \leq m_X$ and then obtain from Rychlik [2000] that

$$\frac{(\mu_X - m_X)^2}{\sigma^2} \leq \frac{(-\log(\frac{1}{2}) - \frac{1}{2})^2}{\frac{3}{4} + \log(\frac{1}{2})} = 0.656$$

and hence an efficiency rate of

$$EF \geq EF(0) \geq \frac{\sigma_X^2 + \sigma_Y^2}{(\sigma_X^2 + \sigma_Y^2) + 0.656(\sigma_X^2 + \sigma_Y^2)} = \frac{1}{1 + 0.656} \simeq \frac{3}{5}$$

4. If X distributes as Y (not necessarily independent), we know that $X - Y$ is symmetric and hence that $m_-\left(\frac{\pi}{4}\right) = \mu_-\left(\frac{\pi}{4}\right) = 0$. This yields:

$$EF \geq EF(\frac{\pi}{4}) = \frac{2\sigma_X^2}{2\sigma_X^2 + \left(\mu_+\left(\frac{\pi}{4}\right) - m_+\left(\frac{\pi}{4}\right)\right)^2} \geq \frac{2\sigma_X^2}{3\sigma_X^2 + Cov(X, Y)}$$

Assume that (X_1, Y_1) and (X_2, Y_2) belong to the same Frechet class $M(F_1, F_2)$ of bivariate distributions with fixed marginals F_1 and F_2. Moreover, assume that (X_1, Y_1) $\leq_{PQD}$ (X_2, Y_2) where PQD stands for the *positive quadrant order* (see Lehmann [1966]). This stochastic order measures the amount of positive dependence of the underlying random vectors[23]. We obtain that all one-dimensional variances are identical, but that $Cov(X_1, Y_1) \leq Cov(X_2, Y_2)$. Thus, the worst case efficiency bound is higher when the variates are less positive dependent. In particular, for given marginals, the highest worst-case efficiency of the $\frac{\pi}{4}$ rotation is achieved for the I.I.D. case where $Cov(X, Y) = 0$, and where:

$$EF \geq EF(\frac{\pi}{4}) \geq \frac{2\sigma_X^2}{3\sigma_X^2} \geq \frac{2}{3}$$

The polar case to independence is the case where X and Y are *co-monotonic:* then, their covariance is maximized for given marginals, and, moreover, their convolution is quantile-additive (see Kaas et al. [2002]). In other words, quantiles and thus medians (the 50% quantile) are linear functions. Hence we obtain for the median that $m_+(\frac{\pi}{4}) = \sqrt{2}m_X$, that $m_+^2(\frac{\pi}{4}) \leq 2\sigma_X^2$ and hence that

$$\forall\theta,\ EF = EF(\frac{\pi}{4}) \geq \frac{2\sigma_X^2}{4\sigma_X^2} = \frac{1}{2}$$

In the co-monotonic scenario expected utility cannot be improved by rotation.

5. Consider the I.I.D. and IFR case with $\mu_X \leq m_X$.[24] Then the convolution of two such variables is again IFR (see Barlow and Proschan [1965]) and we obtain

$$EF \geq EF(\frac{\pi}{4}) \geq \frac{2\sigma_X^2}{2\sigma_X^2 + 0.65\sigma_X^2} = 0.754$$

6. Consider now the I.I.D. case with log-concave densities.[25] Then X and Y are unimodal. Their convolution is log-concave (Prekopa [1973]), and hence also

[23]It is implied, for example, by the supermodular order.

[24]For example, this holds for convex and IFR distributions.

[25]Note that any log-concave distribution on the plane yields log-concave marginals (Prekopa [1973]).

unimodal[26]. By applying the bound in (2) for unimodal distributions we would obtain that

$$EF \geq EF(\frac{\pi}{4}) \geq \frac{2\sigma_X^2}{2\sigma_X^2 + \frac{3}{5}\sigma_X^2} = \frac{10}{13} \approx 0.77$$

But, a better bound can be obtained by exploiting concentration inequalities that explicitly use the properties of log-concave densities. Denote by $f_X = f_Y$ the respective logconcave densities. Bobkov and Ledoux [2014] prove that[27]

$$\frac{1}{12\sigma_X^2} \leq f_X^2(m_X) \leq \frac{1}{2\sigma_X^2}$$

On the other hand, Ball and Böröczky [2010] prove that:

$$f_X(m_X) \cdot \mid m_X - \mu_X \mid \leq \ln\left(\sqrt{\frac{e}{2}}\right)$$

Combining the two inequalities above yields:

$$(m_X - \mu_X)^2 \leq \frac{1}{f_X^2(m_X)} \ln^2\left(\sqrt{\frac{e}{2}}\right) \leq 12\sigma_X^2 \ln^2\left(\sqrt{\frac{e}{2}}\right)$$

The efficiency bound in the log-concave case becomes then:

$$EF \geq EF(\frac{\pi}{4}) \geq \frac{2\sigma_X^2}{2\sigma_X^2 + 12\sigma_X^2 \ln^2\left(\sqrt{\frac{e}{2}}\right)} = \frac{1}{1 + 6\ln^2\left(\sqrt{\frac{e}{2}}\right)} = 0.876$$

■

It is important to note that the above calculations also show that the improvement obtained by rotation may be significant. Just to give one example, consider the distribution for which the Hotelling-Solomons bound is achieved with equality.[28] Then, the second-best welfare in the I.I.D. case without rotation is exactly half of the first-best welfare, while the welfare following the 45 degree rotation is at least two-thirds of the original first best, yielding an improvement of at least 30%.

In the Appendix we show how the above bounds can be obtained for the case of more dimensions. The bound derived for the I.I.D. log-concave case, for example, increases in the number of dimensions, and tends to 100% when the number of dimensions becomes infinite.

[26]The convolution of unimodal densities need not be unimodal ! But, the convolution of X and Y is unimodal for any Y iff X is log-concave (see Ibragimov [1956])

[27]Interestingly enough, the left hand side of the inequality applies to all probabiliy densities on the real line.

[28]This is a discrete distribution concentrated on two points.. But, it can be easily approximated by continuous distribution that satisfy the bound with almost equality, for any needed degree of precision.

To illustrate the above efficiency bounds, consider again the Gamma distribution with parameters α and β. The mean and variance are given by

$$\mu(\alpha,\beta)=\frac{\alpha}{\beta};\ \sigma^2(\alpha,\beta)=\frac{\alpha}{\beta^2}$$

The relative efficiency in the I.I.D. case is:

$$\begin{aligned}
EF(\frac{\pi}{4}) &= \frac{2\sigma_X^2}{2\sigma_X^2+\left(\mu_+\left(\frac{\pi}{4}\right)-m_+\left(\frac{\pi}{4}\right)\right)^2} \\
&= \frac{\frac{2\alpha}{\beta^2}}{\frac{2\alpha}{\beta^2}+\frac{1}{2\beta^2}(2\alpha-m(2\alpha,1))^2} \\
&= \frac{1}{1+\frac{1}{4a}(2\alpha-m(2\alpha,1))^2}\approx\frac{1}{1+\frac{1}{36\alpha}}=\frac{36\alpha}{36\alpha+1}
\end{aligned}$$

for $\alpha>\frac{1}{2}$ and

$$EF(\frac{\pi}{4})\approx\frac{1}{1+\frac{1}{4\alpha}(\ln 2)^2}=\frac{4\alpha}{4\alpha+(\ln 2)^2}$$

for $\alpha\le\frac{1}{2}$.[29] In contrast, the relative efficiency in the unrotated case is given by

$$\begin{aligned}
EF(0) &= \frac{\frac{2\alpha}{\beta^2}}{\frac{2\alpha}{\beta^2}+\frac{2}{\beta^2}(\alpha-m(\alpha,1))^2}=\frac{1}{1+\frac{1}{\alpha}(\alpha-m(\alpha,1))^2} \\
&\approx \frac{1}{1+\frac{1}{9\alpha}}=\frac{9\alpha}{9\alpha+1}
\end{aligned}$$

for $\alpha>1$ and by

$$EF(0)=\frac{1}{1+\frac{1}{\alpha}(\alpha-m(\alpha,1))^2}\approx\frac{1}{1+\frac{1}{\alpha}(\ln 2)^2}=\frac{\alpha}{\alpha+(\ln 2)^2}$$

for $\alpha\le 1$. As $\alpha\to\infty$ we obtain that $\lim EF(\frac{\pi}{4})=1$. This is due to the fact that the mean-median squared distance stays bounded while the variance increases without bounds. In particular, if $\alpha=1$ so that the distribution is exponential, we obtain

$$EF(\frac{\pi}{4})=\frac{36}{37}=0.97>EF(0)=0.9.$$

5 Concluding Remarks

We have shown that voting by simple majority on each dimension becomes a highly efficient aggregation mechanism when combined with a judicious choice of the issues that are put up for vote. Our study endogenizes the process by which a "structure

[29] The efficiency is between these bounds for $0<\alpha<\frac{1}{2}$.

induced equilibrium" can be reached in a complex multi-dimensional collective decision problem with incomplete information about preferences. As we have shown, a re-definition of issues facilitates the search for an optimal consensus among ex-ante conflicting interests. While we have focused on welfare maximization, other goals (such as the maximizing the utility of an agenda setter) can be analyzed by the same methods. A companion paper will explore in more detail the case of a finite number of voters.

6 Appendix A: Omitted Proofs

6.1 Proof of Lemma 1

Recall that we normalize the means of X and Y to be zero. First, suppose $\theta \in [\pi, 2\pi]$. If we let $\theta' = \theta - \pi$, then $\theta' \in [0, \pi]$. Furthermore,

$$\begin{aligned} m_-(\theta) &= \text{median } (X\cos\theta - Y\sin\theta) \\ &= \text{median } (-X\cos(\theta-\pi) + Y\sin(\theta-\pi)) \\ &= -m_-(\theta') \end{aligned}$$

and

$$\begin{aligned} m_+(\theta) &= \text{median } (X\sin\theta + Y\cos\theta) \\ &= \text{median } (-X\sin(\theta-\pi) - Y\cos(\theta-\pi)) \\ &= -m_+(\theta') \end{aligned}$$

As a result,

$$m_-^2(\theta) + m_+^2(\theta) = m_-^2(\theta') + m_+^2(\theta').$$

Next, suppose $\theta \in [\pi/2, \pi]$. If we let $\theta' = \theta - \pi/2$, then $\theta' \in [0, \pi/2]$. Furthermore,

$$\begin{aligned} m_-(\theta) &= \text{median } (X\cos\theta - Y\sin\theta) \\ &= \text{median } \left(-X\sin\left(\theta - \frac{\pi}{2}\right) - Y\cos\left(\theta - \frac{\pi}{2}\right)\right) \\ &= -\text{ median } (X\sin\theta' + Y\cos\theta') \\ &= -m_+(\theta') \end{aligned}$$

and

$$\begin{aligned} m_+(\theta) &= \text{median } (X\sin\theta + Y\cos\theta) \\ &= \text{median } \left(X\cos\left(\theta - \frac{\pi}{2}\right) - Y\sin\left(\theta - \frac{\pi}{2}\right)\right) \\ &= \text{median } (X\cos\theta' - Y\sin\theta') \\ &= m_-(\theta') \end{aligned}$$

Again, we have

$$m_-^2(\theta) + m_+^2(\theta) = m_-^2(\theta') + m_+^2(\theta').$$

Therefore, for any $\theta \in [\pi/2, 2\pi]$ that minimizes $m_-^2(\theta) + m_+^2(\theta)$, there exists $\theta' \in [0, \pi/2]$ that attains the same minimum.

6.2 Proof of Theorem 3

Recall our discussion in the text after the statement of Theorem 3: in order to show that $\theta = 0$ is suboptimal, it is sufficient to show that

$$m_-'(0) = m_+'(0) = 0,$$

and that

$$m_X m_-''(0) + m_Y m_+''(0) < 0.$$

We divide the proof in three steps. First, we derive expressions for $m_+'(\theta)$ and $m_+''(\theta)$, and verify $m_+'(0) = 0$. Second, we derive expressions of $m_-'(\theta)$ and $m_-''(\theta)$ and verify $m_-'(0) = 0$. The last step, that completes the proof, shows that $m_X m_-''(0) + m_Y m_+''(0) < 0$ if

$$m_X f_X'(m_X) \geq 0, m_Y f_Y'(m_Y) \geq 0, m_X^2 + m_Y^2 \neq 0.$$

Step 1: Compute $m_+'(\theta)$ and $m_+''(\theta)$, and verify $m_+'(0) = 0$.

By the inversion formula, the distribution of the convolution $Z = X\sin\theta + Y\cos\theta$ is

$$F_{X\sin\theta+Y\cos\theta}(z) = \frac{1}{2} - \frac{1}{2\pi}\int_0^\infty \frac{\varphi_X(t\sin\theta)\varphi_Y(t\cos\theta)e^{-itz} - \varphi_X(-t\sin\theta)\varphi_Y(-t\cos\theta)e^{itz}}{it}dt$$

where

$$\varphi_X(t\sin\theta) = \int_{-\infty}^\infty e^{itx\sin\theta} f_X(x)\,dx \text{ and } \varphi_Y(t\cos\theta) = \int_{-\infty}^\infty e^{ity\cos\theta} f_Y(y)\,dy$$

Since $F_{X\sin\theta+Y\cos\theta}(m_+(\theta)) = 1/2$, we must have

$$\int_0^\infty \frac{\varphi_X(t\sin\theta)\varphi_Y(t\cos\theta)e^{-itm_+} - \varphi_X(-t\sin\theta)\varphi_Y(-t\cos\theta)e^{itm_+}}{it}dt = 0$$

Let us define

$$G(m_+, \theta) = \int_0^\infty \frac{\varphi_X(t\sin\theta)\varphi_Y(t\cos\theta)e^{-itm_+} - \varphi_X(-t\sin\theta)\varphi_Y(-t\cos\theta)e^{itm_+}}{it}dt$$

Then we have

$$m_+'(\theta) = -\frac{\partial G/\partial\theta}{\partial G/\partial m_+} \tag{13}$$

and

$$
\begin{aligned}
m_+''(\theta) &= -\frac{\left(\frac{\partial^2 G}{\partial \theta^2} + \frac{\partial^2 G}{\partial m_+ \partial \theta} m_+'(\theta)\right)\frac{\partial G}{\partial m_+} - \left(\frac{\partial^2 G}{\partial m_+^2} m_+'(\theta) + \frac{\partial^2 G}{\partial m_+ \partial \theta}\right)\frac{\partial G}{\partial \theta}}{\left(\frac{\partial G}{\partial m_+}\right)^2} \\
&= -\frac{\frac{\partial^2 G}{\partial \theta^2}\left(\frac{\partial G}{\partial m_+}\right)^2 + \left(\frac{\partial^2 G}{\partial m_+^2}\frac{\partial G}{\partial \theta} - 2\frac{\partial^2 G}{\partial m_+ \partial \theta}\frac{\partial G}{\partial m_+}\right)\frac{\partial G}{\partial \theta}}{\left(\frac{\partial G}{\partial m_+}\right)^3} \qquad (14)
\end{aligned}
$$

By definition of φ_X and φ_Y, we can compute

$$
\begin{aligned}
\frac{\partial \varphi_X(t\sin\theta)}{\partial \theta} &= \int_{-\infty}^{\infty} itx\cos\theta e^{itx\sin\theta} f_X(x)\,dx \quad \text{so } \frac{\partial \varphi_X(t\sin\theta)}{\partial \theta}|_{\theta=0} = it\mu_X \\
\frac{\partial \varphi_X(-t\sin\theta)}{\partial \theta} &= \int_{-\infty}^{\infty} -itx\cos\theta e^{-itx\sin\theta} f_X(x)\,dx \quad \text{so } \frac{\partial \varphi_X(-t\sin\theta)}{\partial \theta}|_{\theta=0} = -it\mu_X \\
\frac{\partial \varphi_Y(t\cos\theta)}{\partial \theta} &= -\int_{-\infty}^{\infty} ity\sin\theta e^{ity\cos\theta} f_Y(y)\,dy \quad \text{so } \frac{\partial \varphi_Y(t\cos\theta)}{\partial \theta}|_{\theta=0} = 0 \\
\frac{\partial \varphi_Y(-t\cos\theta)}{\partial \theta} &= \int_{-\infty}^{\infty} ity\sin\theta e^{-ity\cos\theta} f_Y(y)\,dy \quad \text{so } \frac{\partial \varphi_Y(-t\cos\theta)}{\partial \theta}|_{\theta=0} = 0
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\frac{\partial G}{\partial \theta} &= \frac{\partial}{\partial \theta}\int_0^{\infty} \frac{\varphi_X(t\sin\theta)\varphi_Y(t\cos\theta)e^{-itm_+} - \varphi_X(-t\sin\theta)\varphi_Y(-t\cos\theta)e^{itm_+}}{it}dt \\
&= \int_0^{\infty}\left[\begin{array}{c} \varphi_Y(t\cos\theta)\int_{-\infty}^{\infty} x\cos\theta e^{itx\sin\theta} f_X(x)\,dx \\ -\varphi_X(t\sin\theta)\int_{-\infty}^{\infty} y\sin\theta e^{ity\cos\theta} f_Y(y)\,dy \end{array}\right] e^{-itm_+}dt \\
&\quad + \int_0^{\infty}\left[\begin{array}{c} \varphi_Y(-t\cos\theta)\int_{-\infty}^{\infty} x\cos\theta e^{-itx\sin\theta} f_X(x)\,dx \\ -\varphi_X(-t\sin\theta)\int_{-\infty}^{\infty} y\sin\theta e^{-ity\cos\theta} f_Y(y)\,dy \end{array}\right] e^{itm_+}dt \qquad (15)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial^2 G}{\partial \theta^2} &= \int_0^{\infty}\left[\begin{array}{c} -\int_{-\infty}^{\infty} ity\sin\theta e^{ity\cos\theta} f_Y(y)\,dy \int_{-\infty}^{\infty} x\cos\theta e^{itx\sin\theta} f_X(x)\,dx \\ +\varphi_Y(t\cos\theta)\int_{-\infty}^{\infty}\left(-x\sin\theta + itx^2\cos^2\theta\right)e^{itx\sin\theta} f_X(x)\,dx \\ -\int_{-\infty}^{\infty} itx\cos\theta e^{itx\sin\theta} f_X(x)\,dx \int_{-\infty}^{\infty} y\sin\theta e^{ity\cos\theta} f_Y(y)\,dy \\ -\varphi_X(t\sin\theta)\int_{-\infty}^{\infty}\left(y\cos\theta - ity^2\sin^2\theta\right)e^{ity\cos\theta} f_Y(y)\,dy \end{array}\right] e^{-itm_+}dt \\
&\quad + \int_0^{\infty}\left[\begin{array}{c} \int_{-\infty}^{\infty} ity\sin\theta e^{-ity\cos\theta} f_Y(y)\,dy \int_{-\infty}^{\infty} x\cos\theta e^{-itx\sin\theta} f_X(x)\,dx \\ \varphi_Y(-t\cos\theta)\int_{-\infty}^{\infty}\left(-x\sin\theta - itx^2\cos^2\theta\right)e^{-itx\sin\theta} f_X(x)\,dx \\ -\int_{-\infty}^{\infty} -itx\cos\theta e^{-itx\sin\theta} f_X(x)\,dx \int_{-\infty}^{\infty} y\sin\theta e^{-ity\cos\theta} f_Y(y)\,dy \\ -\varphi_X(-t\sin\theta)\int_{-\infty}^{\infty}\left(y\cos\theta + ity^2\sin^2\theta\right)e^{-ity\cos\theta} f_Y(y)\,dy \end{array}\right] e^{itm_+}dt \qquad (16)
\end{aligned}
$$

Since $\mu_X = 0$ for the normalized distribution, we have

$$\frac{\partial G}{\partial \theta}|_{\theta=0} = 0,$$

and since $\varphi_X(0) = 1$, we have

$$
\begin{aligned}
\frac{\partial^2 G}{\partial \theta^2}|_{\theta=0} &= \int_0^\infty \left[\varphi_Y(t)\int_{-\infty}^\infty itx^2 f_X(x)\,dx - \varphi_X(0)\int_{-\infty}^\infty ye^{ity} f_Y(y)\,dy\right] e^{-itm_+}dt \\
&\quad + \int_0^\infty \left[\varphi_Y(-t)\int_{-\infty}^\infty -itx^2 f_X(x)\,dx - \varphi_X(0)\int_{-\infty}^\infty ye^{-ity} f_Y(y)\,dy\right] e^{itm_+}dt \\
&= \int_0^\infty \left[\varphi_Y(t)\,it\sigma_X^2 - \int_{-\infty}^\infty ye^{ity} f_Y(y)\,dy\right] e^{-itm_+}dt \\
&\quad + \int_0^\infty \left[-\varphi_Y(-t)\,it\sigma_X^2 - \int_{-\infty}^\infty ye^{-ity} f_Y(y)\,dy\right] e^{itm_+}dt \\
&= \int_0^\infty \left[\int_{-\infty}^\infty it\sigma_X^2\left(e^{it(y-m_+)} - e^{-it(y-m_+)}\right) f_Y(y)\,dy\right] dt \\
&\quad - \int_0^\infty \left[\int_{-\infty}^\infty y\left(e^{it(y-m_+)} + e^{-it(y-m_+)}\right) f_Y(y)\,dy\right] dt \\
&= -2\int_0^\infty \left\{\int_{-\infty}^\infty \left[t\sigma_X^2 \sin(t(y-m_+)) + y\cos(t(y-m_+))\right] f_Y(y)\,dy\right\} dt
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\frac{\partial G}{\partial m_+} &= \frac{\partial}{\partial m_+}\int_0^\infty \frac{\varphi_X(t\sin\theta)\,\varphi_Y(t\cos\theta)\,e^{-itm_+} - \varphi_X(-t\sin\theta)\,\varphi_Y(-t\cos\theta)\,e^{itm_+}}{it} dt \\
&= -\int_0^\infty \left(\varphi_X(t\sin\theta)\,\varphi_Y(t\cos\theta)\,e^{-itm_+} + \varphi_X(-t\sin\theta)\,\varphi_Y(-t\cos\theta)\,e^{itm_+}\right) dt \quad (17)
\end{aligned}
$$

Therefore, with $\varphi_X(0) = 1$, we have

$$
\begin{aligned}
\frac{\partial G}{\partial m_+}|_{\theta=0} &= -\int_0^\infty \left(\varphi_Y(t)\,e^{-itm_+} + \varphi_Y(-t)\,e^{itm_+}\right) dt \\
&= -\int_0^\infty \int_{-\infty}^\infty \left[e^{it(y-m_+)} + e^{-it(y-m_+)}\right] f_Y(y)\,dydt \\
&= -2\int_0^\infty \int_{-\infty}^\infty \cos(t(y-m_+))\, f_Y(y)\,dydt
\end{aligned}
$$

and

$$
m_+'(0) = -\frac{\frac{\partial G}{\partial \theta}|_{\theta=0}}{\frac{\partial G}{\partial m_+}|_{\theta=0}} = 0,
$$

Note that if $\theta = 0$, then $m_+(0) = m_Y$ and thus

$$
\begin{aligned}
m_+''(0) &= -\frac{\frac{\partial^2 G}{\partial \theta^2}\left(\frac{\partial G}{\partial m_+}\right)^2 + \left(\frac{\partial^2 G}{\partial m_+^2}\frac{\partial G}{\partial \theta} - 2\frac{\partial^2 G}{\partial m_+ \partial \theta}\frac{\partial G}{\partial m_+}\right)\frac{\partial G}{\partial \theta}}{\left(\frac{\partial G}{\partial m_+}\right)^3}|_{\theta=0} = -\frac{\frac{\partial^2 G}{\partial \theta^2}|_{\theta=0}}{\frac{\partial G}{\partial m_+}|_{\theta=0}} \\
&= -\frac{\int_0^\infty \left\{\int_{-\infty}^\infty \left[t\sigma_X^2 \sin(t(y-m_Y)) + y\cos(t(y-m_Y))\right] f_Y(y)\,dy\right\} dt}{\int_0^\infty \int_{-\infty}^\infty \cos(t(y-m_Y))\, f_Y(y)\,dydt}
\end{aligned}
$$

Step 2: Compute $m'_-(\theta)$ and $m''_-(\theta)$ and verify that $m'_-(0) = 0$.

As before, the distribution of the convolution $W = X\cos\theta - Y\sin\theta$ is

$$F_{X\cos\theta - Y\sin\theta}(w) = \frac{1}{2} - \frac{1}{2\pi}\int_0^\infty \frac{\varphi_X(t\cos\theta)\,\varphi_Y(-t\sin\theta)\,e^{-itw} - \varphi_X(-t\cos\theta)\,\varphi_Y(t\sin\theta)\,e^{itw}}{it}dt$$

Since $F_{X\cos\theta - Y\sin\theta}(m_-(\theta)) = 1/2$, we must have

$$\int_0^\infty \frac{\varphi_X(t\cos\theta)\,\varphi_Y(-t\sin\theta)\,e^{-itm_-} - \varphi_X(-t\cos\theta)\,\varphi_Y(t\sin\theta)\,e^{itm_-}}{it}dt = 0$$

Let us define

$$H(m_-, \theta) = \int_0^\infty \frac{\varphi_X(t\cos\theta)\,\varphi_Y(-t\sin\theta)\,e^{-itm_-} - \varphi_X(-t\cos\theta)\,\varphi_Y(t\sin\theta)\,e^{itm_-}}{it}dt$$

Then

$$m'_-(\theta) = -\frac{\partial H/\partial\theta}{\partial H/\partial m_-} \tag{18}$$

and

$$m''_-(\theta) = -\frac{\frac{\partial^2 H}{\partial\theta^2}\left(\frac{\partial H}{\partial m_-}\right)^2 + \left(\frac{\partial^2 H}{\partial m_-^2}\frac{\partial H}{\partial\theta} - 2\frac{\partial^2 H}{\partial m_-\partial\theta}\frac{\partial H}{\partial m_-}\right)\frac{\partial H}{\partial\theta}}{\left(\frac{\partial H}{\partial m_-}\right)^3} \tag{19}$$

As before, we can compute

$$\begin{aligned}
\frac{\partial H}{\partial\theta} &= \frac{\partial}{\partial\theta}\int_0^\infty \frac{\varphi_X(t\cos\theta)\,\varphi_Y(-t\sin\theta)\,e^{-itm_-} - \varphi_X(-t\cos\theta)\,\varphi_Y(t\sin\theta)\,e^{itm_-}}{it}dt \\
&= \int_0^\infty \begin{bmatrix} \varphi_Y(-t\sin\theta)\int_{-\infty}^{\infty} -x\sin\theta e^{itx\cos\theta} f_X(x)\,dx \\ +\varphi_X(t\cos\theta)\int_{-\infty}^{\infty} -y\cos\theta e^{-ity\sin\theta} f_Y(y)\,dy \end{bmatrix} e^{-itm_-}dt \\
&\quad - \int_0^\infty \begin{bmatrix} \varphi_Y(t\sin\theta)\int_{-\infty}^{\infty} x\sin\theta e^{-itx\cos\theta} f_X(x)\,dx \\ +\varphi_X(-t\cos\theta)\int_{-\infty}^{\infty} y\cos\theta e^{ity\sin\theta} f_Y(y)\,dy \end{bmatrix} e^{itm_-}dt
\end{aligned} \tag{20}$$

and

$$\begin{aligned}
\frac{\partial^2 H}{\partial\theta^2} &= \int_0^\infty \begin{bmatrix} \int_{-\infty}^{\infty} -ity\cos\theta e^{-ity\sin\theta} f_Y(y)\,dy \int_{-\infty}^{\infty} -x\sin\theta e^{itx\cos\theta} f_X(x)\,dx \\ +\varphi_Y(-t\sin\theta)\int_{-\infty}^{\infty}\left(-x\cos\theta + itx^2\sin^2\theta\right)e^{itx\cos\theta} f_X(x)\,dx \\ +\int_{-\infty}^{\infty} -itx\sin\theta e^{itx\cos\theta} f_X(x)\,dx \int_{-\infty}^{\infty} -y\cos\theta e^{-ity\sin\theta} f_Y(y)\,dy \\ +\varphi_X(t\cos\theta)\int_{-\infty}^{\infty}\left(y\sin\theta + ity^2\cos^2\theta\right)e^{-ity\sin\theta} f_Y(y)\,dy \end{bmatrix} e^{-itm_-}dt \\
&\quad - \int_0^\infty \begin{bmatrix} \int_{-\infty}^{\infty} ity\cos\theta e^{ity\sin\theta} f_Y(y)\,dy \int_{-\infty}^{\infty} x\sin\theta e^{-itx\cos\theta} f_X(x)\,dx \\ +\varphi_Y(t\sin\theta)\int_{-\infty}^{\infty}\left(x\cos\theta + itx^2\sin^2\theta\right)e^{-itx\cos\theta} f_X(x)\,dx \\ +\int_{-\infty}^{\infty} itx\sin\theta e^{-itx\cos\theta} f_X(x)\,dx \int_{-\infty}^{\infty} y\cos\theta e^{ity\sin\theta} f_Y(y)\,dy \\ +\varphi_X(-t\cos\theta)\int_{-\infty}^{\infty}\left(-y\sin\theta + ity^2\cos^2\theta\right)e^{ity\sin\theta} f_Y(y)\,dy \end{bmatrix} e^{itm_-}dt
\end{aligned} \tag{21}$$

Furthermore, we can compute

$$\begin{aligned}\frac{\partial H}{\partial m_-} &= \frac{\partial}{\partial m_-}\int_0^\infty \frac{\varphi_X(t\cos\theta)\,\varphi_Y(-t\sin\theta)\,e^{-itm_-} - \varphi_X(-t\cos\theta)\,\varphi_Y(t\sin\theta)\,e^{itm_-}}{it}dt \\ &= -\int_0^\infty \left[\varphi_X(t\cos\theta)\,\varphi_Y(-t\sin\theta)\,e^{-itm_-} + \varphi_X(-t\cos\theta)\,\varphi_Y(t\sin\theta)\,e^{itm_-}\right] dt \end{aligned} \tag{22}$$

Given that $\mu_Y = 0$, we have

$$\frac{\partial H}{\partial \theta}|_{\theta=0} = \int_0^\infty -\varphi_X(t)\,\mu_Y e^{-itm_-}dt + \int_0^\infty \varphi_X(-t)\,\mu_Y e^{itm_-}dt = 0$$

and since $\varphi_Y(0) = 1$, we have

$$\begin{aligned}\frac{\partial^2 H}{\partial \theta^2} &= \int_0^\infty \left[\varphi_Y(0)\int_{-\infty}^\infty -xe^{itx}f_X(x)\,dx + it\varphi_X(t)\int_{-\infty}^\infty y^2 f_Y(y)\,dy\right] e^{itm_-}dt \\ &\quad - \int_0^\infty \left[\varphi_Y(0)\int_{-\infty}^\infty xe^{-itx}f_X(x)\,dx + it\varphi_X(-t)\int_{-\infty}^\infty y^2 f_Y(y)\,dy\right] e^{itm_-}dt \\ &= \int_0^\infty \left[\int_{-\infty}^\infty -xe^{itx}f_X(x)\,dx + it\varphi_X(t)\,\sigma_Y^2\right] e^{-itm_-}dt \\ &\quad - \int_0^\infty \left[\int_{-\infty}^\infty xe^{-itx}f_X(x)\,dx + it\varphi_X(-t)\,\sigma_Y^2\right] e^{itm_-}dt \\ &= \int_0^\infty \left[\int_{-\infty}^\infty it\sigma_Y^2\left(e^{it(x-m_-)} - e^{-it(x-m_-)}\right) f_X(x)\,dx\right] dt \\ &\quad - \int_0^\infty \left[\int_{-\infty}^\infty x\left(e^{it(x-m_-)} + e^{-it(x-m_-)}\right) f_X(x)\,dx\right] dt \\ &= -2\int_0^\infty \left\{\int_{-\infty}^\infty \left[t\sigma_Y^2 \sin(t(x-m_-)) + x\cos(t(x-m_-))\right] f_X(x)\,dx\right\} dt\end{aligned}$$

and

$$\frac{\partial H}{\partial m_-}|_{\theta=0} = -\int_0^\infty \left[\varphi_X(t)\,e^{-itm_-} + \varphi_X(-t)\,e^{itm_-}\right] dt = -2\int_0^\infty \int_{-\infty}^\infty \cos(t(x-m_-))\,f_X(x)\,dxdt$$

As a result,

$$m_-'(0) = -\frac{\partial H/\partial\theta}{\partial H/\partial m_-}|_{\theta=0} = 0$$

Since $m_-(0) = m_X$, we have

$$\begin{aligned} m_-''(0) &= -\frac{\frac{\partial^2 H}{\partial\theta^2}\left(\frac{\partial H}{\partial m_-}\right)^2 + \left(\frac{\partial^2 H}{\partial m_-^2}\frac{\partial H}{\partial \theta} - 2\frac{\partial^2 H}{\partial m_-\partial\theta}\frac{\partial H}{\partial m_-}\right)\frac{\partial H}{\partial\theta}}{\left(\frac{\partial H}{\partial m_-}\right)^3}|_{\theta=0} = -\frac{\frac{\partial^2 H}{\partial\theta^2}|_{\theta=0}}{\frac{\partial H}{\partial m_-}|_{\theta=0}} \\ &= -\frac{\int_0^\infty \left\{\int_{-\infty}^\infty \left[t\sigma_Y^2\sin(t(x-m_X)) + x\cos(t(x-m_X))\right] f_X(x)\,dx\right\} dt}{\int_0^\infty\int_{-\infty}^\infty \cos(t(x-m_X))\,f_X(x)\,dxdt}\end{aligned}$$

Step 3: Verify the sufficient, second-order condition.

Note that the second-order derivative at $\theta = 0$ is given by

$$\begin{aligned} & m_X m_-'' (0) + m_Y m_+'' (0) \\ = & -m_X \frac{\int_0^\infty \left\{ \int_{-\infty}^\infty \left[t\sigma_Y^2 \sin\left(t\left(x - m_X\right)\right) + x\cos\left(t\left(x - m_X\right)\right) \right] f_X(x)\, dx \right\} dt}{\int_0^\infty \int_{-\infty}^\infty \cos\left(t\left(x - m_X\right)\right) f_X(x)\, dxdt} \\ & -m_Y \frac{\int_0^\infty \left\{ \int_{-\infty}^\infty \left[t\sigma_X^2 \sin\left(t\left(y - m_Y\right)\right) + y\cos\left(t\left(y - m_Y\right)\right) \right] f_Y(y)\, dy \right\} dt}{\int_0^\infty \int_{-\infty}^\infty \cos\left(t\left(y - m_Y\right)\right) f_Y(y)\, dydt} \end{aligned}$$

We want to show that

$$m_X f_X' (m_X) \geq 0, m_Y f_Y' (m_Y) \geq 0, m_X + m_Y \neq 0.$$

implies

$$m_X m_-'' (0) + m_Y m_+'' (0) < 0.$$

By the inversion formula:

$$\begin{aligned} F_X(z) &= \frac{1}{2} - \frac{1}{2\pi} \int_0^\infty \frac{\varphi_X(t)\, e^{-itz} - \varphi_X(-t)\, e^{itz}}{it} dt \\ &= \frac{1}{2} - \frac{1}{2\pi} \int_0^\infty \frac{1}{it} \int_{-\infty}^\infty \left[e^{itx} e^{-itz} - e^{-itx} e^{itz} \right] f_X(x)\, dxdt \\ &= \frac{1}{2} - \frac{1}{2\pi} \int_0^\infty \frac{1}{it} \int_{-\infty}^\infty \left[e^{it(x-z)} - e^{-it(x-z)} \right] f_X(x)\, dxdt \\ &= \frac{1}{2} - \frac{1}{2\pi} \int_0^\infty \frac{1}{it} \int_{-\infty}^\infty \left[2i \sin\left(t\left(x - z\right)\right) \right] f_X(x)\, dxdt \\ &= \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \int_{-\infty}^\infty \frac{1}{t} \sin\left(t\left(x - z\right)\right) f_X(x)\, dxdt \end{aligned}$$

Therefore,

$$\begin{aligned} f_X(z) &= -\frac{1}{\pi} \int_0^\infty \int_{-\infty}^\infty \frac{1}{t} (-t) \cos\left(t\left(x - z\right)\right) f_X(x)\, dxdt = \frac{1}{\pi} \int_0^\infty \int_{-\infty}^\infty \cos\left(t\left(x - z\right)\right) f_X(x)\, dxdt \\ f_X'(z) &= \frac{1}{\pi} \int_0^\infty \int_{-\infty}^\infty t \sin\left(t\left(x - z\right)\right) f_X(x)\, dxdt \end{aligned}$$

Hence

$$\begin{aligned} \int_0^\infty \int_{-\infty}^\infty \cos\left(t\left(x - z\right)\right) f_X(x)\, dxdt &= \pi f_X(z) \\ \int_0^\infty \int_{-\infty}^\infty t \sin\left(t\left(x - z\right)\right) f_X(x)\, dxdt &= \pi f_X'(z) \end{aligned}$$

As a result

$$
\begin{aligned}
& m_X m_-''(0) + m_Y m_+''(0) \\
= \quad & -m_X \frac{\sigma_Y^2 \int_0^\infty \int_{-\infty}^\infty t \sin(t(x-m_X)) f_X(x)\,dxdt + \int_0^\infty \int_{-\infty}^\infty x \cos(t(x-m_X)) f_X(x)\,dxdt}{\int_0^\infty \int_{-\infty}^\infty \cos(t(x-m_X)) f_X(x)\,dxdt} \\
& -m_Y \frac{\sigma_X^2 \int_0^\infty \int_{-\infty}^\infty t \sin(t(y-m_Y)) f_Y(y)\,dydt + \int_0^\infty \int_{-\infty}^\infty y \cos(t(y-m_Y)) f_Y(y)\,dydt}{\int_0^\infty \int_{-\infty}^\infty \cos(t(y-m_Y)) f_Y(y)\,dydt} \\
= \quad & -\frac{m_X}{f_X(m_X)} \left[\sigma_Y^2 f_X'(m_X) + \frac{\int_0^\infty \int_{-\infty}^\infty x \cos(t(x-m_X)) f_X(x)\,dxdt}{\pi} \right] \\
& -\frac{m_Y}{f_Y(m_Y)} \left[\sigma_X^2 f_Y'(m_Y) + \frac{\int_0^\infty \int_{-\infty}^\infty y \cos(t(y-m_Y)) f_Y(y)\,dydt}{\pi} \right]
\end{aligned}
$$

By the inversion theorem for real functions we have:

$$
\begin{aligned}
\int_0^\infty \int_{-\infty}^\infty x \cos(t(x-m_X)) f_X(x)\,dxdt &= \pi m_X f(m_X) \\
\int_0^\infty \int_{-\infty}^\infty y \cos(t(y-m_Y)) f_Y(y)\,dydt &= \pi m_Y f(m_Y)
\end{aligned}
$$

Then

$$
\begin{aligned}
& m_X m_-''(0) + m_Y m_+''(0) \\
= \quad & -\frac{m_X}{f_X(m_X)} \left[\sigma_Y^2 f_X'(m_X) + m_X f(m_X) \right] - \frac{m_Y}{f_Y(m_Y)} \left[\sigma_X^2 f_Y'(m_Y) + m_Y f(m_Y) \right] \\
= \quad & -m_X^2 \left[\sigma_Y^2 \frac{f_X'(m_X)}{m_X f_X(m_X)} + 1 \right] - m_Y^2 \left[\sigma_X^2 \frac{f_Y'(m_Y)}{m_Y f(m_Y)} + 1 \right] \\
= \quad & -\sigma_Y^2 \frac{m_X f_X'(m_X)}{f_X(m_X)} - \sigma_X^2 \frac{m_Y f_Y'(m_Y)}{f(m_Y)} - m_X^2 - m_Y^2
\end{aligned}
$$

Therefore, a sufficient condition for

$$
m_X m_-''(0) + m_Y m_+''(0) < 0
$$

is

$$
m_X f_X'(m_X) \geq 0, m_Y f_Y'(m_Y) \geq 0, \text{ and } m_X^2 + m_Y^2 \neq 0.
$$

6.3 Proof of Proposition 1

Suppose $F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \leq 1$ for all $\varepsilon > 0$. The other case is completely analogous. We first use an argument by van Zwet [1979] to claim that $m_X < \mu_X$.

Note that

$$\begin{aligned}
m_X - \mu_X &= \int_{-\infty}^{m_X} (m_X - x) f_X(x)\, dx + \int_{m_X}^{\infty} (m_X - x) f_X(x)\, dx \\
&= \int_{-\infty}^{m_X} F_X(x)\, dx - \int_{m_X}^{\infty} (1 - F_X(x))\, dx \\
&= \int_{0}^{\infty} [F_X(m_X - x) + F_X(m_X + x) - 1]\, dx
\end{aligned}$$

It follows from $m_X \neq \mu_X$ that $m_X < \mu_X$, and that $F_X(m_X - x) + F_X(m_X + x) - 1 < 0$ for some interval of x. Next, we use an argument adapted from Watson and Gordon [1986] to prove that the median function is super-additive. The super-additivity of the median function is equivalent to

$$\Pr(X + Y < m_X + m_Y) < \frac{1}{2} \tag{23}$$

Note that

$$\begin{aligned}
& \Pr(X + Y < m_X + m_Y) \\
= & \int_{m_Y}^{\infty} \int_{-\infty}^{m_X + m_Y - y} f_X(x) f_Y(y)\, dxdy + \int_{-\infty}^{m_Y} \int_{-\infty}^{m_X} f_X(x) f_Y(y)\, dxdy \\
& + \int_{m_X}^{\infty} \int_{-\infty}^{m_X + m_Y - x} f_X(x) f_Y(y)\, dxdy \\
= & \int_{m_Y}^{\infty} F_X(m_X + m_Y - y) f_Y(y)\, dy + \frac{1}{4} + \int_{m_X}^{\infty} f_X(x) F_Y(m_X + m_Y - x)\, dx \\
= & \int_{0}^{\infty} F_X(m_X - \varepsilon) f_Y(m_Y + \varepsilon)\, d\varepsilon + \int_{0}^{\infty} f_X(m_X + \varepsilon) F_Y(m_Y - \varepsilon)\, d\varepsilon + \frac{1}{4}
\end{aligned}$$

Therefore, condition (23) is equivalent to

$$4 \int_{0}^{\infty} F_X(m_X - \varepsilon) f_Y(m_Y + \varepsilon)\, d\varepsilon + 4 \int_{0}^{\infty} f_X(m_X + \varepsilon) F_Y(m_Y - \varepsilon)\, d\varepsilon < 1 \tag{24}$$

Let us define non-negative random variables X^+, X^-, Y^+, Y^- as

$$\begin{aligned}
X^+ &= X - m_X | X \geq m_X \quad \text{and} \quad X^- = m_X - X | X \leq m_X \\
Y^+ &= Y - m_Y | Y \geq m_Y \quad \text{and} \quad Y^- = m_Y - Y | Y \leq m_Y
\end{aligned}$$

Then

$$\begin{aligned}
\Pr\left(X^- > Y^+\right) &= \int_{0}^{\infty} 2F_X(m_X - \varepsilon)\, 2f_Y(m_X + \varepsilon)\, d\varepsilon \\
\Pr\left(Y^- > X^+\right) &= \int_{0}^{\infty} 2F_Y(m_X - \varepsilon)\, 2f_X(m_X + \varepsilon)\, dx
\end{aligned}$$

Therefore, condition (24) is equivalent to

$$\Pr\left(X^{-} > Y^{+}\right) + \Pr\left(Y^{-} > X^{+}\right) < 1 \tag{25}$$

A sufficient condition for (25) is

$$\Pr\left(X^{+} < \varepsilon\right) \leq \Pr\left(X^{-} < \varepsilon\right) \text{ and } \Pr\left(Y^{+} < \varepsilon\right) \leq \Pr\left(Y^{-} < \varepsilon\right) \tag{26}$$

for all $\varepsilon > 0$, and with strict inequality for some open interval of ε, because by setting $\varepsilon = Y^{+}$ and $\varepsilon = X^{+}$, respectively, we obtain

$$\Pr\left(X^{+} < Y^{+}\right) < \Pr\left(X^{-} < Y^{+}\right) \quad \text{and } \Pr\left(Y^{+} < X^{+}\right) < \Pr\left(Y^{-} < X^{+}\right)$$

and thus (25). Since X and Y are I.I.D., the sufficient condition (26) reduces to

$$\Pr\left(X^{+} < \varepsilon\right) \leq \Pr\left(X^{-} < \varepsilon\right) \text{ for all } \varepsilon > 0.$$

Equivalently,

$$\Pr\left(X - m_X < \varepsilon\right) \leq \Pr\left(m_X - X < \varepsilon\right).$$

which simplifies into the first inequality in (12). As we argued above, since $m_X \neq \mu_X$, we must have $F_X\left(m_X - \varepsilon\right) + F_X\left(m_X + \varepsilon\right) - 1 < 0$ for some open interval of ε, as desired.

6.4 Proof of Proposition 2

The first order condition (6), evaluated at $\theta = \pi/4$, is given by

$$m_{-}(\frac{\pi}{4})m_{-}'(\frac{\pi}{4}) + m_{+}(\frac{\pi}{4})m_{+}'(\frac{\pi}{4}) = 0.$$

Because X and Y are I.I.D., we must have $m_{-}(\frac{\pi}{4}) = \mu_{-}(\frac{\pi}{4}) = 0$. Therefore, it is sufficient to show

$$m_{+}'(\frac{\pi}{4}) = 0.$$

Recall from the proof of Theorem 3 that $m_{+}'(\theta)$ is given by

$$m_{+}'(\theta) = -\frac{\partial G/\partial\theta}{\partial G/\partial m_{+}}$$

where

$$G(m_{+},\theta) = \int_0^{\infty} \frac{\varphi_X\left(t\sin\theta\right)\varphi_Y\left(t\cos\theta\right)e^{-itm_{+}} - \varphi_X\left(-t\sin\theta\right)\varphi_Y\left(-t\cos\theta\right)e^{itm_{+}}}{it}dt$$

and

$$\frac{\partial G}{\partial \theta} = \int_0^{\infty} \begin{bmatrix} \varphi_Y(t\cos\theta)\int_{-\infty}^{\infty} x\cos\theta e^{itx\sin\theta} f_X(x)\,dx \\ -\varphi_X(t\sin\theta)\int_{-\infty}^{\infty} y\sin\theta e^{ity\cos\theta} f_Y(y)\,dy \end{bmatrix} e^{-itm_+}dt$$
$$+\int_0^{\infty} \begin{bmatrix} \varphi_Y(-t\cos\theta)\int_{-\infty}^{\infty} x\cos\theta e^{-itx\sin\theta} f_X(x)\,dx \\ -\varphi_X(-t\sin\theta)\int_{-\infty}^{\infty} y\sin\theta e^{-ity\cos\theta} f_Y(y)\,dy \end{bmatrix} e^{itm_+}dt$$

Since X and Y are I.I.D. and since $\cos(\pi/4) = \sin(\pi/4) = \frac{\sqrt{2}}{2}$, it is easy to verify that

$$\frac{\partial G}{\partial \theta}|_{\theta=\frac{\pi}{4}} = 0.$$

Therefore, we have $m_+'(\frac{\pi}{4}) = 0$.

7 Appendix B: More than Two Dimensions

Our main result that it is never optimal to vote on independent issues can be easily extended to higher dimensions. The idea is to apply our previous two-dimensional analysis to the first two dimensions, while keeping all other dimensions fixed. It then follows that one can improve welfare by rotating the first two dimensions. Therefore, it is never optimal to vote on independent issues.

7.1 Sub-optimality of the Zero Rotation

Consider K independent issues, denoted by X_k, $k = 1, ..., K$. We write $\mathbf{X} = (X_1, ..., X_K)^T$ and assume that all random variables X_k are normalized. Let SO_K denote the *special orthogonal group* in dimension K which consists of $K \times K$ orthogonal matrices with determinant $+1$. This group is isomorphic to the set of rotations in $\mathbb{R}^K$. A $K \times K$ orthogonal matrix $Q \in SO_K$ is a real matrix with

$$Q^T Q = QQ^T = I$$

where Q^T is the transpose of Q , and where I is the $K \times K$ identity matrix. As a result

$$Q^{-1} = Q^T$$

Each $K \times K$ special orthogonal matrix Q transforms an orthogonal system $\mathbf{X}$ into another orthogonal system while preserving the orientation in $\mathbb{R}^K$. The transformed orthogonal system $\mathbf{X}$ is denoted as $Q\mathbf{X}$. Then, the planner's objective is to choose Q in order to maximize welfare.

The zero-angle rotation is captured of course by the $K \times K$ identity matrix. In order to show that this rotation is sub-optimal for higher dimensions, consider the following special orthogonal matrix

$$Q(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & \cdots & 0 \\ \sin\theta & \cos\theta & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

It is easy to verify that $[Q(\theta)]^{-1} = [Q(\theta)]^T$, so that $Q(\theta)$ is indeed an orthogonal matrix with determinant $+1$. This matrix represents a rotation in the plane of the first two dimensions, while keeping fixed all other dimensions. Hence, for our purpose, it is sufficient to show that

$$\Delta(\theta) \equiv \mathbb{E}\,||Q(\theta)\,\mathbf{X} - \text{median}\,(Q(\theta)\,\mathbf{X})||^2 \leq \mathbb{E}\,||\mathbf{X} - \text{median}(\mathbf{X})||^2 \equiv \Delta(0)$$

for some θ close to 0 .

In particular, it is sufficient to show that $\Delta(\theta)$ has a local maximum at $\theta = 0$. But

$$Q(\theta)\,\mathbf{X} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & \cdots & 0 \\ \sin\theta & \cos\theta & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_K \end{bmatrix} = \begin{bmatrix} X_1\cos\theta - X_2\sin\theta \\ X_1\sin\theta + X_2\cos\theta \\ X_3 \\ \vdots \\ X_K \end{bmatrix}$$

Therefore,

$$\Delta(\theta) = \sum_{k=1}^{K} \sigma^2_{X_k} + \sum_{k=3}^{K} m^2_{X_k} + [\text{median}\,(X_1\cos\theta - X_2\sin\theta)]^2 + [\text{median}\,(X_1\sin\theta + X_2\cos\theta)]^2$$

and the sub-optimality of the zero-angle rotation follows directly from our two-dimensional analysis.

7.2 The Analog of the $\pi/4$ Rotation

Suppose $X_1, ..., X_K$ are I.I.D. drawn from a common distribution. What is the counterpart of $\pi/4$ rotation (or equivalently the top-down procedure) in higher dimensions? We need to look for an orthogonal matrix Q that transforms $\mathbf{X}$ into a new vector $Q\mathbf{X}$ whose one coordinate is given by the sum $X_1 + ... + X_K$ while the other coordinates consists of various differences. This is straightforward if K is even, and slightly more complicated if K is odd. For example, If $K = 4$, the orthogonal matrix Q (with determinant equal to $+1$) is given by

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} X_1 + X_2 - X_3 - X_4 \\ X_1 + X_4 - X_2 - X_3 \\ X_2 + X_4 - X_1 - X_3 \\ X_1 + X_2 + X_3 + X_4 \end{pmatrix}$$

More generally, if K is even, it is easy to see that the same condition we had before, namely the super-additivity of the median function, is again sufficient for the $\pi/4$ rotation to dominate the zero-angle rotation.

If $K = 3$, the required orthogonal matrix Q (with determinant equal to $+1$) is given

$$\begin{pmatrix} \frac{1}{6}\sqrt{6} & -\frac{1}{3}\sqrt{6} & \frac{1}{6}\sqrt{6} \\ \frac{1}{2}\sqrt{2} & 0 & -\frac{1}{2}\sqrt{2} \\ \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{6}}{6}\left(X_1 + X_3 - 2X_2\right) \\ \frac{\sqrt{2}}{2}\left(X_1 - X_3\right) \\ \frac{\sqrt{3}}{3}\left(X_1 + X_2 + X_3\right) \end{pmatrix}$$

7.3 Efficiency Bounds

As in the main text, we work here with the non-normalized random variables $X_1, ..., X_K$ representing the marginals of the distribution of ideal points. Note that, when there are K dimensions, the expected utility of choosing marginal medians under an orthogonal transformation Q is given by

$$U(Q) = -\mathbb{E}\left|\left|Q\mathbf{X} - \text{median}(Q\mathbf{X})\right|\right|^2 = -\sum_{k=1}^{K} \text{var}\left(Q_k\mathbf{X}\right) - \sum_{k=1}^{K} \left(\text{mean}\left(Q_k\mathbf{X}\right) - \text{median}(Q_k\mathbf{X})\right)^2$$

where Q_k is the k-th row of the Q matrix. The first-best expected utility is simply $-\sum_{k=1}^{K}\text{var}(Q_k\mathbf{X})$. We define the relative efficiency of transformation Q relative to the first-best as:

$$EF(Q) \equiv \frac{\sum_{k=1}^{K} \text{var}\left(Q_k\mathbf{X}\right)}{\sum_{k=1}^{K} \text{var}\left(Q_k\mathbf{X}\right) + \sum_{k=1}^{K} \left(\text{mean}\left(Q_k\mathbf{X}\right) - \text{median}(Q_k\mathbf{X})\right)^2}$$

and the maximal relative efficiency as

$$EF \equiv \max_{Q} EF(Q).$$

Again, we can apply the Hotelling-Solomons inequality to obtain that

$$EF\left(Q\right) \geq \frac{\sum_{k=1}^{K} \operatorname{var}\left(Q_k \mathbf{X}\right)}{\sum_{k=1}^{K} \operatorname{var}\left(Q_k \mathbf{X}\right) + \sum_{k=1}^{K} \operatorname{var}\left(Q_k \mathbf{X}\right)} = \frac{1}{2}$$

Analogously, we can use the Basu-DasGupta inequality to show that, for unimodal distributions, we have

$$EF > EF\left(I\right) \geq \frac{\sum_{k=1}^{K} \operatorname{var}\left(Q_k \mathbf{X}\right)}{\sum_{k=1}^{K} \operatorname{var}\left(Q_k \mathbf{X}\right) + \frac{3}{5}\sum_{k=1}^{K} \operatorname{var}\left(Q_k \mathbf{X}\right)} = \frac{5}{8}$$

Suppose now that K is even, and that $X_1, ..., X_K$ are I.I.D. with log-concave densities. Consider an orthogonal matrix $\widehat{Q}$ with

$$\frac{1}{\sqrt{K}}\widehat{Q}_{ij} = \begin{cases} \text{either } 1 \text{ or } -1 & \text{for all } j \text{ if } \quad i \neq K \\ 1 & \text{for all } j \text{ if } \quad i = K \end{cases}$$

such that for all $k \neq K$, $\widehat{Q}_k\mathbf{X}$ contains an equal number of X_k's appearing with positive and negative signs (existence???). It follows from the I.I.D. assumption that

$$\operatorname{mean}(\widehat{Q}_k\mathbf{X}) - \operatorname{median}(\widehat{Q}_k\mathbf{X}) = \begin{cases} 0 & \text{if } \quad k \neq K, \\ \frac{1}{\sqrt{K}}(\operatorname{mean}(\sum_{k=1}^{K} X_k) - \operatorname{median}(\sum_{k=1}^{K} X_k)) & \text{if } \quad k = K. \end{cases}$$

Therefore, we have

$$EF(\widehat{Q}) = \frac{\sum_{k=1}^{K} \operatorname{var}(\widehat{Q}_k\mathbf{X})}{\sum_{k=1}^{K} \operatorname{var}(\widehat{Q}_k\mathbf{X}) + \frac{1}{K}\left(\operatorname{mean}(\sum_{k=1}^{K} X_k) - \operatorname{median}(\sum_{k=1}^{K} X_k)\right)^2}$$

Given that $X_1, ..., X_K$ have log-concave densities, the convolution $Z \equiv \sum_{k=1}^{K} X_k$ also has a log-concave densities. Then the inequalities of Bobkov and Ledoux [2014] and of Ball and Böröczky [2010] together imply

$$(m_Z - \mu_Z)^2 \leq \frac{1}{f_Z^2(m_Z)} \ln^2\left(\sqrt{\frac{e}{2}}\right) \leq 12\sigma_Z^2 \ln^2\left(\sqrt{\frac{e}{2}}\right)$$

Hence,

$$EF(\widehat{Q}) \geq \frac{\sum_{k=1}^{K} \operatorname{var}(\widehat{Q}_k\mathbf{X})}{\sum_{k=1}^{K} \operatorname{var}(\widehat{Q}_k\mathbf{X}) + \frac{1}{K}12\sigma_Z^2 \ln^2\left(\sqrt{\frac{e}{2}}\right)}$$

Let $\sigma_{X_k}^2$ denote the variance of X_k. Then we have $\sigma_Z^2 = K\sigma_{X_k}^2$ and

$$\operatorname{var}(\widehat{Q}_k\mathbf{X}) = \widehat{Q}_k\widehat{Q}_k^T\sigma_{X_k}^2 = \sigma_{X_k}^2$$

since $\widehat{Q}_k\widehat{Q}_k^T = 1$ by the definition of an orthogonal matrix. Therefore, we obtain the following efficiency bound for log-concave densities:

$$EF \geq EF(\widehat{Q}) \geq \frac{K\sigma_{X_k}^2}{K\sigma_{X_k}^2 + 12\sigma_{X_k}^2 \ln^2\left(\sqrt{\frac{e}{2}}\right)} = \frac{K}{K + 12\ln^2\left(\sqrt{\frac{e}{2}}\right)}$$

For example, if $K = 4$, the bound is 93.4%. Note that this bound is increasing the number of dimensions K, and tends to 100% when K goes to infinity.

References

[1966] Bahadur, R. R. (1966), "A Note on Quantiles in Large Samples," *Annals of Mathematical Statistics* **37**(3), 577-580.

[2005] Austen-Smith D. and J.S. Banks (2005), *Positive Political Theory II: Strategy and Structure*, The University of Michigan Press, Ann Arbor, MI.

[2010] Ball, Keith M., and Karoly J. Boroczky (2010), "Stability of the Prekopa-Leindler Inequality," *Mathematika* **56**, 339-356.

[1948] Black, D. (1948), "On the Rationale of Group Decision-Making," *Journal of Political Economy* **56**(1), 23-34.

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

[1965] Barlow, R. E., and F. Proschan (1965), *Mathematical Theory of Reliability,* J. Wiley & Sons. SIAM, Philadelphia, PA.

[1997] Basu, S. and A. DasGupta (1997), "The Mean, Median, and Mode of Unimodal Distributions: A Characterization," *Theory Probab. Appl.*, **41**(2), 210–223.

[2008] Berg, Christian, and Henrik L. Pedersen (2008), "Convexity of the Median in the Gamma Distribution," *Ark. Mat.* **46**, 1–6

[1987] Bock M. E., P. Diaconis, F. W. Huffer and M. D. Perlman (1987), "Inequalities for Linear Combinations of Gamma Random Variables," *Canadian Journal of Statistics* **15** (4), 387-395.

[2014] Bobkov, S. G., and Ledoux, M. (2014), "One-dimensional Empirical Measures, Order Statistics and Kantorovich Transport Distances," working paper, University of Minesota.

[1983] Border, Kim C., and J. S. Jordan (1983), "Straightforward Elections, Unanimity and Phantom Voters," *Review of Economic Studies* **50**(1), 153-170.

[1988] Caplin, Andrew, and Barry Nalebuff (1988), "On 64%-Majority Rule," *Econometrica* **56**(4), 787-814.

[1991] Caplin, Andrew, and Barry Nalebuff (1991), "Aggregation and Social Choice," *Econometrica* **59**(1), 1-23.

[1988] Dharmadhikari,S., and Kumar, J. (1988), *Unimodality, Convexity, and Applications*, Academic Press, First Edition.

[1988] Feld, S. and Grofman, B. (1988), "Majority Rule Outcomes and the Structure of Debate in One-Issue at a Time Decision-Making," *Public Choice* **59**, 239-252.

[1987] Ferejohn, J. and Krehbiel, K. (1987), "The Budget Process and the Size of the Budget," *American Journal of Political Science* **31**(2), 296-320.

[1907] Galton, F. (1907), "Vox Populi," first published in *Nature* **75**, 450-451, 1949

[2016] Gershkov, A., Moldovanu, B. and Shi, X. (2016), "Optimal Voting Rules," *Review of Economic Studies*, forthcoming.

[1951] Gil-Pelaez, J. (1951), "Note on the Inversion Theorem," *Biometrika* **38** (3-4): 481-482.

[1948] Haldane, J.B.S. (1948), "Note on the Median of a Multivariate Distribution," *Biometrika* **35**, 414-415.

[1932] Hotelling, Harold, and Leonard M. Solomons (1932), "The Limits of a Measure of Skewness," *Annals of Mathematical Statistics* **3**(2), 141-142.

[2000] Hu, Chin-Yuan, and Gwo Dong Lin (2000), "On an Inequality for the Rayleigh Distribution," *Sankhya: The Indian Journal of Statistics* **62**(1), 36-39.

[2001] Hu, Chin-Yuan, and Gwo Dong Lin (2001), "An Inequality for the Weighted Sums of Pairwise I.I.D. Generalized Rayleigh Random Variables," *Journal of Statistical Planning and Inference* **92**, 1-5.

[1956] Ibragimov, J. A. (1956). "On the Composition of Unimodal Distributions". *Teoriya Veroyatnostey* **I,** 283-288.

[2002] Kaas, R., Dhaene, J., Vyncke, D., Goovaerts, M.J., and Denuit, M. (2002), "A Simple Geometric Proof that Comonotonc Risks Have the Convex-Largest Sum," *Astin Bulletin* **32**(1), 71-80.

[1984] Kim, K. H. and F. W. Roush (1984), "Nonmanipulation in Two Dimensions," *Mathematical Social Sciences* **8**, 29-43.

[2016] Kleiner, A. and Moldovanu, B. (2016), "Content Based Agendas and Qualified-Majorities in Sequential Voting," *American Economic Review,* forthcoming.

[1972] Kramer, G. (1972), "Sophisticated Voting over Multidimensional Choice Spaces," *Journal of Mathematical Sociology* **2**, 165-180.

[1973] Kramer, G. (1973), "On a Class of Equilibrium Conditions for Majority Rule," *Econometrica* **41**(2), 285-297.

[1966] Lehmann, E. (1966), "Some Concepts of Dependence," *Annals of Mathematical Statistics* **37**(5), 1137-1153.

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

[2006] Nelsen, R.B. (2006), *An Introduction to Copulas*, second edition, Springer, New York.

[1992] Peters, H., H. van der Stel, and T. Storcken (1992), "Pareto Optimality, Anonymity, and Strategy-Proofness in Location Problems," *International Journal of Game Theory* **21**, 221-235.

[1999] Poterba, J. M. and von Hagen, J. (eds), *Fiscal Institutions and Fiscal Perormance,* University of Chicago Press, Chiacgo.

[1973] Prekopa, A. (1973), "On Logarithmic Concave Measures and Functions," *Acta Sci. Math. (Szeged)* **34**, 335-343.

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

[1928] Szegö, G. (1928), "Über einige von S. Ramanujan gestellte Aufgaben," *Journal of the London Mathematical Society* **3**, 225-232.

[2000] Rychlik, T. (2000), "Sharp Mean-Variance Inequalities for Quantiles of Distribuition Determined by Convex and Star Orders," *Theory Probab. Appl.* **47**(2), 269-285.

[1991] Shephard, N.G. (1991), "From Characteristic Function to Distribution Function: A Simple Framework for the Theory," *Econometric Theory* **7**, 519-529.

[1979] Shepsle, K.A. (1979), "Institutional Arrangements and Equilibrium in Multidimensional Voting Models," *American Journal of Political Science* **23**, 27-60.

[1979] van Zwet, W.R. (1979), "Mean, Median, Mode II," *Statistica Neerlandica* **33** (1), 1-5.

[1986] Watson, R. and Gordon, I. (1986), "On Quantiles of Sums," *Australian Journal of Statistics* **28**(2), 192-199.

[1991] Zhou, L. (1991), "Impossibility of Strategy-Proof Mechanisms in Economies with Pure Public Goods," *Review of Economic Studies* 58, 107-120.