Theoretical Economics 14 (2019), 555–596 1555-7561/20190555

Voting on multiple issues: What to put on the ballot?

Alex Gershkov
Department of Economics, The Hebrew University of Jerusalem and School of Economics, University of Surrey

Benny Moldovanu
Department of Economics, Bonn University

Xianwen Shi
Department of Economics, University of Toronto

We study a multidimensional 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 independently and identically distributed marginals, a second-best voting mechanism attains at least 88% of the first-best efficiency. Finally, we generalize our method and some of our insights to preferences derived from distance functions based on inner products.

Keywords. Multidimensional voting, rotation, budgeting procedure, mechanism design.

JEL classification. D72, D78, D82.

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 (bottom-up), which resulted in a

Alex Gershkov: alexg@huji.ac.il
Benny Moldovanu: mold@uni-bonn.de
Xianwen Shi: xianwen.shi@utoronto.ca

A earlier version of the paper was circulated under the title "The Dimensions of Consensus." 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, Cedric Wasser, and referees for helpful comments, and various seminar and conference participants for helpful discussions. Shuangjian Zhang provided excellent research assistance. Gershkov's research is supported by a grant from Israel Science Foundation, Moldovanu's research is supported by the German Science Foundation through the Hausdorff Center for Mathematics and CRC TR-12, and Shi's research is supported by a grant from the Social Sciences and Humanities Research Council of Canada.

© 2019 The Authors. Licensed under the Creative Commons Attribution-NonCommercial License 4.0.
Available at http://econtheory.org. https://doi.org/10.3982/TE3193

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 (top-down). A large literature 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]

We analyze the problem of redefining (or bundling) the issues brought to vote in a multidimensional collective decision problem. Such methods can increase the welfare of the involved decision makers by allowing them to reach a consensus that was not possible on the original issues.

We study a multidimensional collective decision taken by simple majority voting: an example is a legislature that needs to decide on individual budgets for public goods such as, say, education and defense. Other examples are decisions on the geographical location of a desirable facility or decisions on hiring and project adoption that are based on multidimensional attributes.

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.[2]

Voters' ideal points are private information, and we study voting by simple majority on each dimension separately. As we see below, this focus yields, in combination with a decision over the dimensions that are the subject of voting, an analysis of more generality than immediately apparent.

Voting by simple majority on each dimension yields 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, the alternative that minimizes the sum of squared distances from the individual ideal points. The first-best is the coordinate-wise average (or mean) of the realized ideal points, and, thus, first-best welfare is the corresponding variance (with a minus sign).

The first-best 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.[3] Given the tension between first-best on the one hand and implementable outcomes on the other, how well does voting by simple majority perform in terms of welfare? A classical inequality due to Hotelling and Solomons (1932) implies that, for any distribution of preferences, voting by simple majority on any given issues achieves at least 50% of the first-best welfare.

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

[2]The main text deals with the two-dimensional case, while the generalization to more than two dimensions is provided in an Appendix.

[3]This observation was first made by Galton (1907), who was also the first to recommend the use of the median as a robust and nonmanipulable aggregator of opinions. His insights have been sharpened and much generalized in the literature on robust estimation.

The main insight of the present paper is that a judicious choice of the issues that are actually put to a vote (while maintaining voting by simple majority, with its desirable incentive properties) can significantly improve welfare.[4] For example, instead of voting on two separate issues, 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. More generally, we model the repackaging and bundling of issues by rotations of the orthogonal axes that define what is put to a vote. For example, suppose voters care about two separate main issues, but they actually vote on the budget of two agencies that overlap in their responsibility over these two issues. Rotations correspond then to the shifting of jurisdictions among the two agencies: they change the mix of issues under the control of each agency.

In influential work, Shepsle (1979) argued that the division of a complex decision into several different jurisdictions (*germaneness*) creates stable equilibria that would not be possible in a general, unconstrained collective decision model. His main examples are legislative committees in the U.S. congress. Viewed in light of Shepsle's theory, our goal is to endogenize the choice of jurisdictions so as to improve welfare, an issue that has not received much attention in formal studies.

A basic technical observation 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.[5] 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. The basic cause behind this phenomenon is the nonlinearity of the median function, a feature that yields a rather complex analysis.[6] To enable the use of calculus and probabilistic/statistical techniques, we focus here on the limit case where the number of voters is infinite.

Our main results are as follows.

Result 1. If the agents' ideal points in one dimension are independently distributed from the ideal points in the other dimension, then under weak conditions on the distribution of preferences, voting on the original issues is suboptimal; that is, a repackaging of the issues brought to vote via rotation (which necessarily creates some correlation among the ideal points) increases welfare. This parallels the non-optimality of separate sales in the multiproduct monopoly problem: some form of mixed bundling is always superior to separate sales (see McAfee et al. 1989).

Result 2. If the marginals of the distribution of agents' ideal points are identically distributed (not necessarily independently), we provide sufficient conditions under which the 45-degree rotation welfare is superior to no rotation. The conditions are satisfied by common distributions with independently and identically distributed (I.I.D.) marginals. We show that with I.I.D. marginals, the 45-degree rotation is always a critical point, and also provide sufficient conditions for the 45-degree rotation to be welfare maximizing.

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

[5]See Haldane (1948) or the literature on spatial voting, e.g., Feld and Grofman (1988).

[6]This is true even for common distributions of types, such as the gamma, Poisson, and log-normal. Some of our results are based on insights that go back to conjectures by Ramanujan (see Szegő 1928) and Hadamard.

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.

Result 3. We provide various lower bounds on incentive efficiency for large, nonparametric families of distributions of ideal points (such as unimodal distributions and distributions with an increasing hazard rate). 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-degree rotation of the original dimensions attains at least 88% of the first-best efficiency.

Result 4. We extend our method to the more general class of preferences induced by distance functions generated by inner-product norms. In particular, for weighted Euclidean norms, we show that voting on independent issues remains sub-optimal under the same sufficient conditions as for the Euclidean preferences.

It is possible to perform a similar analysis for goals other than efficiency, e.g., define jurisdictions that serve other purposes, such as the self-interest of an agenda setter or of a coalition of voters. Ferejohn and Krehbiel (1987) focused on controlling budgetary growth rather than efficiency, and they observed that the 1974 budget reform can be represented by a 45-degree rotation of the coordinates on which voting takes place. For that goal, we offer here precise conditions that compare the top-down and bottom-up procedures in terms of the total budget they produce, and we show that the budgeting reform can unambiguously improve welfare while having a mixed impact on the budget size.

To see how our results may fit practical voting environments, consider a legislative committee that decides on spending on several items. Each committee member has a preferred expenditure for each item. If the items are independent (i.e., the preferred expenditure level on one item is uncorrelated with the preferred level on another item), then it is not optimal to directly vote on the proposed expenditures. Instead, it may be better to vote on the budgets of two agencies that have some overlapping jurisdictions that represent a particular mix of the two issues (this is a nonzero rotation in our framework). In another example, if a committee finances regional hospitals, say, that have similar sizes and serve similar purposes, our analysis suggests that it is better to first decide the total budget for these hospitals and then divide it among hospitals. Finally, if a government, say, has to fund an activity for multiple years, it may be better first to vote on a multi-year budget and then to decide how to allocate the total budget among different years.

1.1 *Related literature*

The existence of a Condorcet winner is rare in multi-dimensional models of voting (Kramer 1973). Kramer (1972) observed, however, that voting in a variety of institutions is often sequential, issue-by-issue, and he established the existence of a sophisticated voting equilibrium if voters' preferences are continuous, convex, and separable.

The coordinate-wise median—obtained by simple-majority voting in each dimension—constitutes a basic instance of a *structure induced equilibrium* in the spirit of Shepsle (1979).[7]

Technically, our contribution builds upon and relates to several important and elegant contributions due to Moulin (1980), Border and Jordan (1983), Kim and Roush (1984), Barberà et al. (1993), and Peters et al. (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 this 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).[8] Border and Jordan (1983) removed Moulin's assumption whereby mechanisms depend only on peaks, and generalized Moulin's finding to a multidimensional setting with separable and quadratic preferences: each DIC mechanism is decomposable into a collection of one-dimensional DIC mechanisms, each described by the location of the phantom voters in the respective dimension.[9,10] Barberà et al. (1993) defined a class of multidimensional single-peaked preferences that is strictly larger than the set of all separable and quadratic preferences, and showed that, for this class, a mechanism is DIC if and only if it is a generalized marginal median.

Gershkov et al. (2017) analyzed welfare maximization in a one-dimensional setting with cardinal utilities, and derived the ex ante welfare maximizing placement of phantoms. They also showed how to avoid the phantom interpretation by implementing Moulin's mechanisms via a sequential, binary voting procedure together with a flexible qualified majority schedule.[11] Combining their result with the Border–Jordan or with the Barbera–Gul–Stacchetti decomposition results yields the welfare maximizing mechanism for the respective multidimensional settings. 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 multidimensional problem must be voted upon according to a particular institution that is sensitive to both utilities and distribution of types. Such flexibility may be difficult, if not impossible, to achieve in practice.

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

---

[7] In a multidimensional voting model with common interest, aggregate uncertainty, and two truth-motivated candidates, McMurray (2018) shows that, in equilibrium, multiple issues are consistently bundled along the 45-degree line (the major diagonal in his model).

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

[9] They also extend this finding to a setting with star-shaped and separable preferences.

[10] Most papers in the literature indeed assume separable preferences. Ahn and Oliveros (2012) is a notable exception: they prove equilibrium existence in combinatorial voting with nonseparable preferences, and provide conditions under which the Condorcet winner is implemented in the equilibrium of large elections.

[11] See also Kleiner and Moldovanu (2017) for general sufficient conditions under which sequential, binary voting procedures possess desirable properties.

The simplest multidimensional setting 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 Roush (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.[12] Peters et al. (1992) showed that, for two dimensions with an odd number of voters, voting by simple majority in each dimension (after any translation/rotation of the plane) is also Pareto efficient.[13]

Finally, it is also instructive to compare our results to those in the classical papers by Caplin and Nalebuff (1988, 1991).[14] These authors did not consider incomplete information and incentive constraints. Instead, motivated by the instability of multidimensional voting, they considered the effect of super-majority requirements on the stability of the spatial mean. For a large number of voters and 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, any coalition that prefers an alternative over the mean contains less than 64% of the voters and is thus not effective.

As mentioned above, for the log-concave case with I.I.D. marginals, our results display a mechanism that is incentive compatible for any (odd) number of voters and that achieves at least 88% of the first-best utility when this number goes to infinity. Thus, issue-by-issue simple majority voting 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.

Although our setting bears some similarity to multidimensional cheap talk, the logic of welfare gains is very different here. In those models, the multiplicity of issues helps because it improves information transmission between the sender(s) and the receiver. In a model with two senders, Battaglini (2002) shows that as long as the two senders' ideal points are linearly independent, full information revelation is possible by carefully choosing dimensions to exploit the conflict between senders. In a one-sender model, Charkraborty and Harbaugh (2007) show that the sender can credibly convey his ranking of different issues to the receiver. In our model, rotations address a very different conflict, i.e., conflict between simple majority voting and efficiency.

[12]Since both median and mean are translation equivariant, translations of the origin cannot improve welfare. It is, therefore, without loss of generality to restrict attention here to rotations.

[13]They show that a mechanism is Pareto efficient if and only if, for any realization of agents' ideal points, its allocation lies in the convex hull of the ideal points. With two or more dimensions, a generalized median with phantoms may lie outside of the convex hull, and this is true even for the standard median mechanism in three or more dimensions.

[14]These papers were also the first to use modern concentration inequalities in the economics literature.

2. The model

We consider n (odd) agents 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 $\mathbf{t}_i = (x_i, y_i)$, $i = 1, 2, \ldots, n$. The peak $\mathbf{t}_i$ is agent i's private information. Each agent i has a utility function of the form

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

where the point $\mathbf{v} \in D$ denotes the chosen alternative and where $||\cdot||$ is the standard Euclidean (l_2) norm. The peaks $\mathbf{t}_i = (x_i, y_i)$ are independently and identically distributed (I.I.D.) across agents, according to a joint distribution $F(x_i, y_i)$, with density f. Denote by μ_X (μ_Y) the expected value of x_i (y_i). Throughout the paper, we assume that $\mathbb{E}||\mathbf{t}_i||^2 < \infty$ for all $\mathbf{t}_i$.

A utilitarian planner would choose $\mathbf{v} \in D$ to maximize the average of the agents' ex ante utilities or, equivalently, minimize the expected average squared distance from the voters' peaks,

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

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

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

which we refer to as the *first-best solution*. For each fixed realization $(\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_n)$, 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)

$$-\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{t}_i - \bar{\mathbf{t}}\|^2.$$

In Section 5, we extend our analysis to preferences generated by other norms induced by inner products.

2.1 *Repackaging issues via rotations*

We consider voting by simple majority on each separate 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 "repackaging and bundling" the basic issues X and Y.

We model packaging and bundling of issues through rotations in the plane. Recall that for fixed Cartesian coordinates, rotating a point $(x, y) \in \mathbb{R}^2$ counterclockwise 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 vector (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 new random vector obtained from rotating the random vector (X, Y) by an angle of θ:

$$Z_-(\theta) = X\cos\theta - Y\sin\theta$$
$$Z_+(\theta) = X\sin\theta + Y\cos\theta.$$

Voters then vote on the new issues Z_- and Z_+ instead of the original issues X and Y.[15] By the simple majority rule, the voting outcome will be $(m_-(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n))$, where

$$m_-(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n) = \text{median}(x_1\cos\theta - y_1\sin\theta, \dots, x_n\cos\theta - y_n\sin\theta) \tag{1}$$
$$m_+(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n) = \text{median}(x_1\sin\theta + y_1\cos\theta, \dots, x_n\sin\theta + y_n\cos\theta), \tag{2}$$

are the *marginal medians* after the rotation.[16]

It is easy to verify that the mean $\bar{\mathbf{t}}$ of $\mathbf{t}_1, \dots, \mathbf{t}_n$ is *rotation equivariant*, i.e., the mean of rotated peaks is simply the rotated mean of the original peaks. In marked contrast, the marginal medians $(m_-(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n))$ are *not* rotation equivariant, i.e., rotating and taking medians is not the same as taking medians and rotating. Therefore, rotations are instruments that the planner may use to influence welfare. To illustrate, consider Figure 1 with three voters, where A, B, and C are voters' ideal points. Original coordinates are denoted by X and Y; rotated coordinates are denoted by X' and Y'. The original median is the outcome of voting along the original axes (x, y). The new median is the outcome of voting along the rotated axes (x', y'). It is clear that the mean of ideal points is rotation equivariant; the median is not.

The reason for this complex behavior is the nonlinearity of the median of random variables under convolutions, as illustrated by the following example.

[15] We abuse here notation by denoting by the same capital letters both the underlying dimensions (or issues) and the random variables that govern the distribution of peaks on those respective dimensions.

[16] Other than marginal median, there are several other multivariate generalizations of univariate median. See Small (1990) for a review of different definitions of multidimensional medians and their (lack of) equivariance properties.

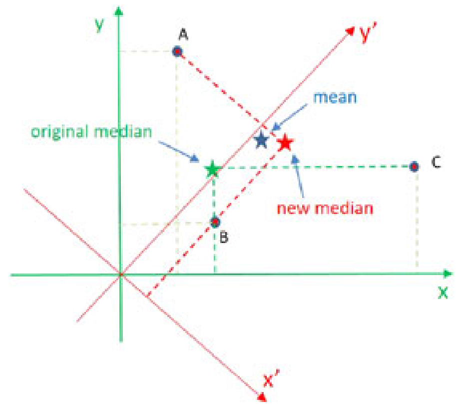

FIGURE 1. Median is not rotation equivariant.

EXAMPLE 1. Let (X, Y) denote a random vector on the plane. Suppose that X and Y are I.I.D. exponentially distributed with $f_X(x) = e^{-x}$ for all $x \geq 0$ and $f_Y(y) = e^{-y}$ for all $y \geq 0$. The means are $\mu_X = \mu_Y = 1$ and the medians are $m_X = m_Y = \ln 2$. Rotating the coordinates clockwise by $\frac{\pi}{4}$ and then projecting (X, Y) to the new coordinates, yields a new random vector $(Z_-, Z_+) = (\frac{\sqrt{2}}{2}X - \frac{\sqrt{2}}{2}Y, \frac{\sqrt{2}}{2}X + \frac{\sqrt{2}}{2}Y)$. The random variable Z_- is symmetric, so its median and mean are both equal to zero, and the mean of Z_+ equals $\frac{\sqrt{2}}{2}(\mu_X + \mu_Y) = \sqrt{2}$. In contrast, the median of Z_+ is not equal to $\frac{\sqrt{2}}{2}(m_X + m_Y)$ or, equivalently, $m_{X+Y} \neq m_X + m_Y$. To see this, note that the density of $X + Y$ is given by

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-t) f_Y(t)\, dt = \int_0^z e^{-(z-t)} e^{-t}\, dt = z e^{-z} \quad \text{for all } z \geq 0.$$

Since

$$F_{X+Y}(m_X + m_Y) = \int_0^{2\ln 2} z e^{-z}\, dz = \frac{3}{4} - \frac{1}{2}\ln 2 \approx 0.4 < F_{X+Y}(m_{X+Y}) = 1/2,$$

it follows that $m_{X+Y} > m_X + m_Y$. ◊

More generally, we could also consider an additional translation of the origin, say by a vector **w**, to obtain new orthogonal coordinates (and thus create new issues). The joint operation of rotation and translation can also be represented by a linear matrix.[17] But medians (and means) are translation equivariant and, thus, there is no extra welfare advantage from such translations. Therefore, we focus below on the family of rotations—the linear isometries with determinant +1 that fix the origin—described by the angle of rotation θ relative to standard Cartesian coordinates.

[17]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. Rotations form the subgroup that fixes the origin.

2.2 *The set of voting mechanisms*

For any rotation angle $\theta \in [0, 2\pi]$, we define the direct *marginal median mechanism* φ_θ as

$$\varphi_\theta(\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_n) = \big(m_-(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n)\big), \tag{3}$$

where $(m_-(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n))$ is the marginal median with respect to rotation θ and reported peaks $\mathbf{t}_i$ as defined in (1) and (2). Since both rotations and medians are continuous functions, $\varphi_\theta(\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_n)$ is continuous in θ and in all its other arguments.

A direct revelation mechanism $\psi(\mathbf{t}_i, \mathbf{t}_{-i})$ is dominant-strategy incentive compatible (DIC) if, for any voter i, any realizations $\mathbf{t}_i$ and $\mathbf{t}_{-i}$, and any reporting strategy profile $\hat{\mathbf{t}}_{-i}(\mathbf{t}_{-i})$ of other voters, voter i's utility $-\|\mathbf{t}_i - \psi(\mathbf{t}_i, \hat{\mathbf{t}}_{-i}(\mathbf{t}_{-i}))\|^2$ is maximized by truthfully revealing his type $\mathbf{t}_i$. It is easily seen that the direct revelation mechanism φ_θ defined in (3) is DIC. Surprisingly, as shown by Kim and Roush (1984) and Peters et al. (1992), the set of marginal median mechanisms (for all possible rotations) coincides with the entire class of anonymous, Pareto efficient, and DIC mechanisms.[18] This provides a complementary justification for our focus on simple-majority voting mechanisms.

The mechanism φ_θ can be decentralized (via an indirect mechanism) by first defining the issues (via rotations) and then voting sequentially by simple majority, one issue at a time, using a binary, sequential voting procedure with a convex agenda (such as those used by all democratic legislatures).[19] The overall outcome does not depend on the order in which the issues are put up to a vote and is the vector of marginal medians $(m_-(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n), m_+(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n))$. This forms an incidence of the structure induced equilibrium àla Shepsle (1979).

Two rotation angles, $\theta = 0$ and $\theta = \pi/4$, are of particular interest and have natural interpretations. When $\theta = 0$, voters are asked to vote on the original issues X and Y. For $\theta = \pi/4$, we have

$$m_-(\pi/4, \mathbf{t}_1, \dots, \mathbf{t}_n) = \frac{\sqrt{2}}{2}\operatorname{median}(x_1 - y_1, \dots, x_n - y_n)$$
$$m_+(\pi/4, \mathbf{t}_1, \dots, \mathbf{t}_n) = \frac{\sqrt{2}}{2}\operatorname{median}(x_1 + y_1, \dots, x_n + y_n).$$

Therefore, under the $\pi/4$ rotation, the vote is on issues $X + Y$ and $X - Y$, rather than on X and Y. Once voters have decided on $X + Y$ and $X - Y$, the planner can then

[18]A mechanism ψ is *anonymous* if, for any profile of reports $(\mathbf{t}_i, \mathbf{t}_{-i})$, $\psi(\mathbf{t}_1, \dots, \mathbf{t}_i, \dots, \mathbf{t}_n) = \psi(\mathbf{t}_{p(1)}, \dots, \mathbf{t}_{p(i)}, \dots, \mathbf{t}_{p(n)})$, where p is any permutation of the set $\{1, \dots, n\}$. A mechanism ψ is *Pareto efficient* (or Pareto optimal) 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. Note that their characterization fails in higher dimensions because anonymous, Pareto efficient, and DIC mechanisms need not exist. Hence, our analysis can be extended to higher dimensional problems, but the solution need not be ex post Pareto efficient.

[19]At each stage of a 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 et al. (2017) and Kleiner and Moldovanu (2017).

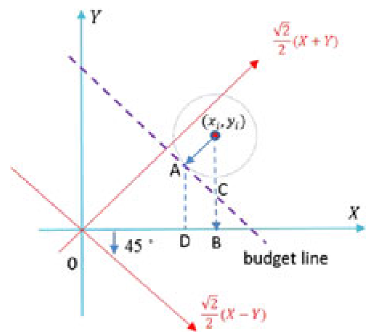

FIGURE 2. Alternative implementation of the top-down procedure.

obviously recover X and Y. The two-step voting procedure associated with the $\pi/4$ 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. In contrast, the voting procedure associated with the 0 rotation resembles the "bottom-up" budgeting procedure: agents vote on separate budgets for individual items and the total budget is gradually obtained as the sum of the individual budgets.

REMARK 1. We focus here on orthogonal coordinates. This is without loss of generality: for any equilibrium outcome obtained by voting along coordinates generated by a nonorthogonal base, there always exists an orthogonal base that yields the same voting outcome. The difference is that under a nonorthogonal base, the order in which the issues are put up to vote does matter. To illustrate, consider the following standard implementation of the $\pi/4$ rotation in practice: after the total sum $(X + Y)$ was determined, voters are asked to vote on X (or on Y) rather than on the orthogonal difference $(X - Y)$. We show that as long as $(X + Y)$ is voted upon first, any issue voted upon at the second stage that is not collinear with $X + Y$ will yield the same equilibrium outcome as under voting according to $(X - Y)$.

To see this, consider the case where voters vote first on $X + Y$ and then on X, and consider the second-stage strategy of voter i with ideal point (x_i, y_i). Then the first-stage decision imposes a budget line (the diagonal heavy dashed line in Figure 2) on which the final voting outcome must lie. Let A and B denote the points obtained by projecting (x_i, y_i) on the budget line and on the X axis, respectively, and let D denote the projection A to the X axis. Then at the second-stage voting on X, voter i's dominant strategy is to vote for point D rather than point B: whenever i is pivotal, voting D yields point A on the budget line, which is closest to his ideal point. Alternatively, A is exactly the point that i would have voted for if the second-stage vote were on the difference $X - Y$. Note that the above argument is independent of the number of voters and can be easily generalized to other nonorthogonal bases.

3. The limit case when the number of agents is large

The full probabilistic optimization problem can be rewritten as

$$(\mathcal{P}_0) \qquad \min_{\theta\in[0,2\pi]} \int_D \cdots \int_D \left(\frac{1}{n}\sum_{i=1}^{n} \left\| R(\theta)\mathbf{t}_i - \varphi_\theta(\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_n) \right\|^2\right) f(\mathbf{t}_1)\cdots f(\mathbf{t}_n)\, d\mathbf{t}_1 \cdots d\mathbf{t}_n.$$

We focus here on the solution to problem $(\mathcal{P}_0)$ when the number of agents is large. The resulting optimal mechanism will be incentive compatible, Pareto efficient, and anonymous for any (odd) number of voters. For I.I.D. random variables $\{X_i\}_{i=1}^{\infty}$ with finite mean μ_X and variance σ_X^2, we know from the central limit theorem that

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

Bahadur (1966) shows 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 $X_{(n+1)/2:n}$ is the median order statistic 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 $n \to \infty$,

$$\begin{pmatrix} m_-(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n) \\ m_+(\theta, \mathbf{t}_1, \dots, \mathbf{t}_n) \end{pmatrix} \longrightarrow \begin{pmatrix} m_-(\theta) \\ m_+(\theta) \end{pmatrix} \equiv \begin{pmatrix} \text{median}(X\cos\theta - Y\sin\theta) \\ \text{median}(X\sin\theta + Y\cos\theta) \end{pmatrix}.$$

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

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

where the two coordinates of the rotated mean are

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

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

$$(\mathcal{P}_1) \qquad \min_{\theta\in[0,2\pi]} \big(\mu_-(\theta) - m_-(\theta)\big)^2 + \big(\mu_+(\theta) - m_+(\theta)\big)^2 + \sigma_X^2 + \sigma_Y^2.$$

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 equivariance of both mean and median. Let us define the normalized random variables $\tilde{X}$ and $\tilde{Y}$ as

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

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

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

We further note that it is without loss of generality to restrict attention to rotations in the interval $[0, \pi/2]$. That is because, 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.[20] Since variances σ_X^2 and σ_Y^2 are fixed, the planner's goal under this normalization is simply to find the rotation that results in a marginal median vector with minimum norm, that is,

$$(\mathcal{P}_2) \qquad \min_{\theta \in [0, \pi/2]} \tilde{m}_-^2(\theta) + \tilde{m}_+^2(\theta).$$

To simplify notation, we drop the tilde symbol for normalized random variables where no confusion can arise.

Remark 2. We would like to comment here on the feasibility of the first-best solution.

(i) With a continuum of voters, the planner could, in principle, dictate the mean as the collective choice without seeking any input from the voters, but this would require detailed knowledge about the joint distribution of individuals' preferences. In contrast, voting by simple majority in each dimension is practical, and, indeed, often observed in reality because it is always incentive compatible and because its execution does not require any prior knowledge about the distribution. None of our theorems or propositions (e.g., Theorems 1–3 and Propositions 1–3) requires the planner to know the exact distribution: it is sufficient to know that the joint distribution belongs to a broad class.

(ii) If the number of voters is finite, the first-best solution, defined as the sample mean of the voters' ideal points, is not implementable because each agent can advantageously move the mean toward her ideal point by reporting a false peak. The individual influence on the mean is unbounded (unless the distribution of peaks is on a bounded interval). Thus, even if the number of voters is large, the possibility to tilt the mean in one's favor may still be substantial.

[20]This claim is a direct consequence of simple trigonometric identities; we omit the proof.

3.1 *Suboptimality of voting on independent issues*

In this subsection, we assume that the unrotated marginals X and Y are independent. We work on the normalized version of the planner's problem $(\mathcal{P}_2)$ and show that the 0 rotation yields a local maximum of the norm of the normalized marginal median, i.e., it leads to a local utility minimum.

Theorem 1. *Assume that X and Y are independent. The rotation with angle $\theta = 0$ is a local utility minimum if*

$$m_X f_X'(m_X) \geq 0, \qquad m_Y f_Y'(m_Y) \geq 0, \qquad m_X^2 + m_Y^2 \neq 0. \tag{4}$$

This theorem is the special case with $\beta = 1$ of Theorem 3 in Section 5.

If the random variables X and Y are unimodal, then the rotation of $\theta = 0$ is a local utility minimum whenever the median lies between the mode and the mean.[21] This alternative sufficient condition is simple and intuitive: there are elegant, general characterizations of distributions where such orders of the mode, median, and mean hold (see, for example, Dharmadhikari and Joag-Dev 1988 and Basu and DasGupta 1997).

Corollary 1. *Assume that X and Y are independent and that $m_X^2 + m_Y^2 \neq 0$. Suppose that X and Y are unimodal and satisfy*

$$M_X \leq m_X \leq \mu_X \quad \text{or} \quad \mu_X \leq m_X \leq M_X$$
$$M_Y \leq m_Y \leq \mu_Y \quad \text{or} \quad \mu_Y \leq m_Y \leq M_Y,$$

where M, m, and 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 proof of Theorem 1 proceeds as follows. The rotation $\theta = 0$ yields a local maximum of the norm of the normalized marginal median if it is a critical point

$$m_-(0)m_-'(0) + m_+(0)m_+'(0) = 0 \tag{5}$$

and if it satisfies the local second-order condition

$$m_-''(0)m_-(0) + \big(m_-'(0)\big)^2 + m_+''(0)m_+(0) + \big(m_+'(0)\big)^2 < 0. \tag{6}$$

The proof verifies that $m_-'(0) = m_+'(0) = 0$ (so condition (5) is trivially satisfied) and that condition (4) in Theorem 1 implies condition (6).

The geometric intuition of the suboptimality of voting on independent issues is illustrated in Figure 3.

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

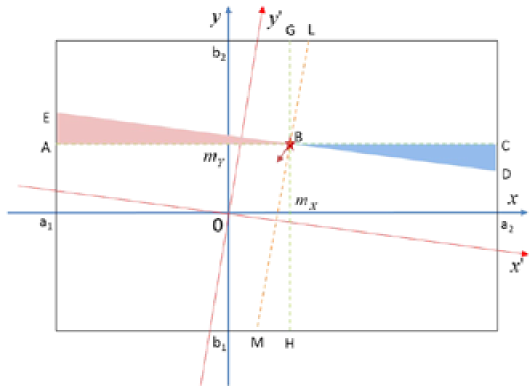

FIGURE 3. Small rotation improves welfare.

Assume that $0 = \mu_X \le m_X$ and $0 = \mu_Y \le m_Y$. We want to show that a small rotation improves welfare if $f'_X(m_X) \ge 0$ and $f'_Y(m_Y) \ge 0$. 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 shifts the new median toward the mean $(0,0)$, i.e., that the median moves toward the southwest. Consider the projection of B onto the new, rotated axes: the result obtains 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 area 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 purpose, 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 $y = m_Y + (m_X - x)\tan\theta$. Therefore, the difference between the areas ABE and BCD is

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

Since $f'_Y(m_Y) \ge 0$, F_Y is locally convex at m_Y. Therefore, for sufficiently small θ, the curve $F_Y(m_Y + (m_X - x)\tan\theta)$ with $x \in [a_1, a_2]$ lies above the tangent line $F_Y(m_Y) + f_Y(m_Y)(m_X - x)\tan\theta$; that is,

$$F_Y\bigl(m_Y + (m_X - x)\tan\theta\bigr) \ge F_Y(m_Y) + f_Y(m_Y)(m_X - x)\tan\theta.$$

As a result, for sufficiently small θ, we have

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

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

Intuitively, area ABE represents voters who have their preferred y coordinate marginally above m_Y and who, after rotation, would switch their support from alternatives above the line $y = m_Y$ to alternatives below the line $y' = m_Y$. In contrast, area BCD

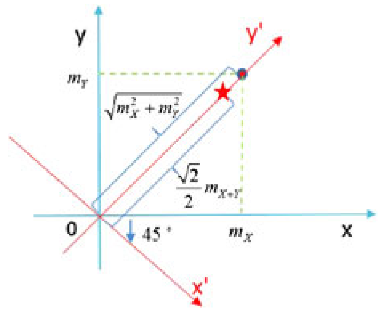

Figure 4. The $\pi/4$ rotation with symmetric marginals.

represents voters who have their y coordinate marginally below m_Y and who would switch their support from alternatives below the line $y = m_Y$ to alternatives above the line $y' = m_Y$. Since $F(y)$ is locally convex at $y = m_Y$, there are more voters in area ABE than in BCD and, thus, more than half of them will vote for alternatives below the line $y' = m_Y$; that is, the median after the rotation will be closer to the origin (the first best).

3.2 *The $\pi/4$ rotation*

In this subsection, we assume that X and Y are identically (but not necessarily independently) distributed. By symmetry,

$$m_-(\pi/4) = \text{median}\left(\frac{\sqrt{2}}{2}(X - Y)\right) = 0 = \mu_-(\pi/4)$$

and

$$m_+(\pi/4) = \text{median}\left(\frac{\sqrt{2}}{2}(X + Y)\right) = \frac{\sqrt{2}}{2}\text{median}(X + Y).$$

Hence, the $\pi/4$ rotation is a natural candidate for improving welfare. It completely eliminates the conflict that arises between efficiency and incentive compatibility along one dimension; all remaining conflict is concentrated in the other dimension, as illustrated in Figure 4 (assuming $m_X > \mu_X = 0$), where (m_X, m_Y) is the unrotated median and the star is the $\pi/4$-rotated median $m_+(\pi/4)$.

Proposition 1. *Suppose that X and Y are I.I.D. and that the density f_X satisfies the regularity condition*

$$\lim_{x\to\infty} f_X\big(\sqrt{2}m_+(\pi/4) - x\big)f_X(x)\big(2x - \sqrt{2}m_+(\pi/4)\big)^2 = 0$$
$$\lim_{x\to-\infty} f_X\big(\sqrt{2}m_+(\pi/4) - x\big)f_X(x)\big(2x - \sqrt{2}m_+(\pi/4)\big)^2 = 0.$$

Then $\theta = \pi/4$ is a critical point, i.e., it satisfies the first-order condition.

See Appendix A for proofs not provided in the main text.

The above regularity condition is satisfied if the distribution has a bounded support or a thin tail. If we could verify second-order conditions either locally or globally, then Proposition 1 could tell us whether $\theta = \pi/4$ is a local or global utility maximum. Unfortunately, the second-order conditions, evaluated at $\theta = \pi/4$, turn out to be very elusive.

Our next result offers sufficient conditions for the optimality of the $\pi/4$ rotation. It requires the following definition.

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 $\max(a, b) \geq \max\{a', b'\}$. A function $h(a, b)$ is said to be *Schur-convex* (-concave) in (a, b) if $h(a'', b'') \geq (\leq) h(a', b')$ whenever $(a'', b'') \succ (a', b')$.

Proposition 2. *Suppose that X and Y are identically distributed and that $m_X \neq \mu_X$. The $\pi/4$ rotation attains the welfare maximum if either of the following conditions holds.*

(i) We have $m_+(\theta) < \mu_+(\theta)$ for all $\theta \in [0, \frac{\pi}{4}]$, and the function

$$\Pr(X \sin\theta + Y \cos\theta \leq z)$$

is Schur-concave in $(\sin^2\theta, \cos^2\theta)$ for all $\theta \in [0, \frac{\pi}{4}]$ and all $z \in [m_X, 0]$.

(ii) We have $m_+(\theta) > \mu_+(\theta)$ for all $\theta \in [0, \frac{\pi}{4}]$, and the function

$$\Pr(X \sin\theta + Y \cos\theta \leq z)$$

is Schur-convex in $(\sin^2\theta, \cos^2\theta)$ for all $\theta \in [0, \frac{\pi}{4}]$ and $z \in [0, m_X]$.

If $\Pr(X \sin\theta + Y \cos\theta \leq z)$ is Schur-concave for all $\theta \in [0, \frac{\pi}{4}]$ and if the rotated median is always below the mean, it must hold that

$$m_X \leq m_{X \sin\theta + Y \cos\theta} \leq m_{\frac{\sqrt{2}}{2}X + \frac{\sqrt{2}}{2}Y} \leq \mu_X.$$

Hence, the distance between the mean and the rotated median $m_{X \sin\theta + Y \cos\theta}$ is smallest when $\theta = \pi/4$. The sufficient conditions in Proposition 2 involve only the model's primitives (i.e., the distributions of types) and can be, in principle, checked for any distribution.[22]

For example, we verified that $\Pr(X \sin\theta + Y \cos\theta \leq z)$ is Schur-concave if X and Y are I.I.D. exponential and, thus, the $\pi/4$ rotation is globally optimal in that case.[23] For other standard distributions such as gamma, Pareto, and Rayleigh, we used Mathematica to plot the aggregate expected welfare as a function of the rotation angle $\theta \in [0, \pi/2]$. Our simulations suggest that the $\pi/4$ rotation is optimal for these distributions, but we were unable to analytically prove it. In general, the $\pi/4$ rotation may not be optimal, as

[22]Similar Schur-concavity/convexity conditions appear in the literature: For example, if X and Y are *nonnegative* I.I.D. random variables with a log-concave density, then $\Pr(aX + bY \leq z)$ is known to be a Schur-concave function of (a^2, b^2) for all z (see Karlin and Rinott 1983). We cannot directly use this result because of the nonnegativity restriction.

[23]The verification details for the exponential distribution are available upon request.

illustrated by Example 2 below. Therefore, some restrictions on the symmetric marginals are indeed necessary for the optimality of the $\pi/4$ rotation.

Example 2. Let (X, Y) denote a random vector on the plane. Suppose that X and Y are I.I.D. according to the discrete distribution

Values of X	0	0.45	1
Probability	0.4	0.3	0.3

so that $\mu_X = \mu_Y = 0.435$ and $m_X = m_Y = 0.45$. The distribution of $X + Y$ is given by

Values of $X + Y$	0	0.45	0.90	1	1.45	2
Probability	0.16	0.24	0.09	0.24	0.18	0.09

so that $\mu_{X+Y} = 0.87$ and $m_{X+Y} = 1$. The expected utility from the 0 rotation is

$$-2(\mu_X - m_X)^2 = -2(0.435 - 0.45)^2 = -0.00045$$

and the expected utility from the $\pi/4$ rotation is

$$-\left(\frac{\sqrt{2}}{2}(\mu_{X+Y} - m_{X+Y})\right)^2 = -\left(\frac{\sqrt{2}}{2}(0.87 - 1)\right)^2 = -0.00845.$$

Therefore, the $\pi/4$ rotation is strictly dominated by the 0 rotation. Since the welfare dominance is strict, we can approximate the discrete distribution by a continuous distribution and maintain it. ♢

3.3 *When does top-down dominate bottom-up?*

We now compare the expected utility under the $\pi/4$ rotation with that under the 0 rotation. As is apparent from Figure 4, this amounts to checking whether the original coordinate-wise median vector (m_X, m_Y) is closer to the origin than the new coordinate-wise median vector $(m_{X+Y}/2, m_{X+Y}/2)$. Therefore, if $m_X < \mu_X$ and $m_X + m_Y < m_{X+Y}$ or if $m_X > \mu_X$ and $m_X + m_Y > m_{X+Y}$, then the $\pi/4$ rotation dominates the 0 rotation.

Assuming that X and Y are I.I.D., we present below a simple sufficient condition that simultaneously guarantees $m_X < (>)\mu_X$ and $m_X + m_Y < (>)m_{X+Y}$.[24] The need to control for sub-/super-additivity of medians parallels the conditions on second-highest order statistics for bundling in auctions (see Palfrey 1983).

[24] As is illustrated in Example 1, both condition $m_X < \mu_X$ and the super-additivity condition $m_X + m_Y < m_{X+Y}$ hold for the exponential distribution that is strictly concave. We show in Section A.4 of Appendix A that the super-additivity condition is satisfied for the gamma distribution (a generalization of the exponential) and the Rayleigh distribution, where the sufficient condition (7) may not be easily checked or does not hold. There we also construct, by using a copula, an example where independence is not necessary for the $\pi/4$ rotation to dominate the 0 rotation.

Proposition 3. *Suppose that X and Y are I.I.D. and that $m_X \neq \mu_X$. The expected utility at $\theta = \frac{\pi}{4}$ exceeds the expected utility at $\theta = 0$ if either*

$$F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \leq 1 \quad \textit{for all } \varepsilon > 0 \tag{7}$$

or

$$F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \geq 1 \quad \textit{for all } \varepsilon > 0. \tag{8}$$

In particular, condition (7) implies $m_X < \mu_X$ and $m_X + m_Y < m_{X+Y}$, and is satisfied if F_X is strictly concave. Condition (8) implies $m_X > \mu_X$ and $m_X + m_Y > m_{X+Y}$, and is satisfied if F_X is strictly convex.

It is worth noting that van Zwet (1979) shows that condition (7) implies $\mu_X > m_X > M_X$ and (8) implies $\mu_X < m_X < M_X$. It follows from Corollary 1 that each of the two conditions is also sufficient for the 0 rotation to be suboptimal.

Remark 3. Whenever the median function is super-additive (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 where the total budget is gradually determined.[25]

4. Bounds on relative efficiency

In this section, we provide several lower bounds on the (relative) efficiency loss of the marginal median mechanisms augmented by rotations. We keep the assumption that the number of agents is large. The various bounds are obtained under different distributional assumptions that govern the distribution of voter's ideal points, and the proofs use several classical statistical inequalities and some more recent concentration inequalities. In particular, for the log-concave case studied by Caplin and Nalebuff (1988, 1991), the lower bound is 88% of the first-best utility.

Note that each assertion in the following theorem holds for a large class of distributions and, therefore, that the results do not require exact knowledge of the particular distribution (as long as it is known that it belongs to the respective class). In particular, the optimal rotation achieves, in each case, a possibly higher relative efficiency.

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 *nonnormalized* 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, decreases as variances increase and is given by

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

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

The expected utility of rotated medians with angle θ is given by

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

Thus, the relative efficiency of the rotation with angle θ is given by

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

Two forces play a role here: 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 larger overall role.[26] 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 (e.g., both are normally distributed). In this case, we have $\mu_-(\theta)=m_-(\theta)$ and $\mu_+(\theta)=m_+(\theta)$.

A random variable X has *increasing failure rate* (IFR) if its hazard rate $f(x)/(1-F(x))$ is increasing in x.

Theorem 2. *The following relative efficiency bounds hold.*

(i) For any random variables X and Y and for any angle θ, $EF(\theta)\geq\frac{1}{2}$.

(ii) For any unimodal random variables X and Y and for any angle θ, $EF(\theta)>\frac{5}{8}$.

(iii) For any random variables X and Y that have an increasing failure rate (IFR) and that satisfy $\mu_X\leq m_X$ and $\mu_Y\leq m_Y$, and for any angle θ, $EF(\theta)>0.603$. In addition, if X and Y are I.I.D., then $EF(\frac{\pi}{4})\geq 0.753$.

(iv) For any X and Y that are identically distributed, $EF(\frac{\pi}{4})\geq\frac{2\sigma_X^2}{3\sigma_X^2+\mathrm{Cov}(X,Y)}$. Thus, when X and Y are independent, $EF(\frac{\pi}{4})\geq\frac{2}{3}$. In the polar, co-monotonic scenario, $EF(\frac{\pi}{4})=EF(0)\geq\frac{1}{2}$ and welfare cannot be improved by rotation.[27]

(v) If X and Y are I.I.D. and if each has a log-concave density, then $EF(\frac{\pi}{4})\geq 0.876$.

Proof. (i) A classical inequality due to Hotelling and Solomons (1932) says that the squared distance between the mean and median of *any* random variable is always less than the variance:

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

[26] 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 medians of convolutions are calculated.

[27] A random vector is *co-monotonic* if and only if it agrees in distribution with a random vector where all components are nondecreasing functions (or all are nonincreasing functions) of the same random variable.

Therefore,

$$\big(\mu_-(\theta) - m_-(\theta)\big)^2 \leq \sigma_-^2(\theta) = \sigma_X^2 \cos^2\theta + \sigma_Y^2 \sin^2\theta - 2\sin\theta\cos\theta\,\mathrm{Cov}(X, Y)$$
$$\big(\mu_+(\theta) - m_+(\theta)\big)^2 \leq \sigma_+^2(\theta) = \sigma_X^2 \sin^2\theta + \sigma_Y^2 \cos^2\theta + 2\sin\theta\cos\theta\,\mathrm{Cov}(X, Y).$$

We obtain the universal bound

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

(ii) For the class of unimodal distributions, 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

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

(iii) For the class of distributions with an increasing failure rate (IFR), if $\mu_X \leq m_X$, then we obtain from Rychlik (2000) that

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

and, hence, an efficiency rate of

$$EF(\theta) \geq \frac{\sigma_X^2 + \sigma_Y^2}{(\sigma_X^2 + \sigma_Y^2) + 0.656(\sigma_X^2 + \sigma_Y^2)} = \frac{1}{1 + 0.656} = 0.603.$$

If, in addition, X and Y are I.I.D., then the convolution of two such variables is again IFR (see Barlow and Proschan 1965) and we obtain

$$EF\left(\frac{\pi}{4}\right) \geq \frac{2\sigma_X^2}{2\sigma_X^2 + 0.656\sigma_X^2} = 0.753.$$

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

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

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 $(X_1, Y_1) \leq_{\mathrm{PQD}} (X_2, Y_2)$, 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.[28] We obtain that all one-dimensional variances are identical, but that $\mathrm{Cov}(X_1, Y_1) \le \mathrm{Cov}(X_2, Y_2)$. Thus, the worst case efficiency bound is higher when the variates are less positively 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 $\mathrm{Cov}(X, Y) = 0$ and where

$$EF\left(\frac{\pi}{4}\right) \ge \frac{2\sigma_X^2}{3\sigma_X^2} = \frac{2}{3}.$$

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

$$\left(\mu_+\left(\frac{\pi}{4}\right) - m_+\left(\frac{\pi}{4}\right)\right)^2 = (\mu_X - m_X)^2 \le 2\sigma_X^2$$

and we obtain

$$EF\left(\frac{\pi}{4}\right) = EF(0) \ge \frac{1}{2}.$$

This holds analogously for any other rotation.

(v) Consider now the I.I.D. case with log-concave densities.[29] Then X and Y are unimodal. Their convolution is log-concave (Prékopa 1973) and, hence, also unimodal.[30] Let $f_X = f_Y$ denote the respective log-concave densities. Bobkov and Ledoux (2016) prove that[31]

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

Ball and Böröczky (2010) prove that

$$f_X(m_X) \cdot |m_X - \mu_X| \le \ln\left(\sqrt{\frac{e}{2}}\right).$$

Combining the two inequalities above yields

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

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

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

[30] The convolution of unimodal densities need not be unimodal, but the convolution of X and Y is unimodal for any Y if and only if X is log-concave (see Ibragimov 1956).

[31] Interestingly enough, the left hand side of the inequality applies to *any* probability density on the real line.

The efficiency bound in the log-concave case then becomes

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

The above calculations also show that the improvement obtained by rotation may be significant. Just to give one example, consider the original (e.g., unrotated) distributions for which the Hotelling–Solomons bound is achieved with equality.[32] Then the 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 Appendix B, we show how the above bounds can be obtained for the case of more dimensions. For example, in the I.I.D, case, the relative efficiency tends to 1 when the number of dimensions becomes infinite.

5. Extension to other utility functions

In this section we briefly illustrate how our method can be applied to a more general class of utility functions that are based on a distance generated by an inner product. Thus, we assume that the utility of agent i with peak $\mathbf{t}_i$ from decision $\mathbf{v} \in D \subseteq \mathbb{R}^2$ is given by

$$-\Delta\left(\|\mathbf{t}_i - \mathbf{v}\|_I\right),$$

where $||\cdot||_I$ is some inner-product norm and where Δ is a strictly monotonically increasing function.

Choosing a marginal median with respect to any orthogonal coordinates yields a DIC mechanism (see Peters et al. 1993, who focus on strictly convex norms; inner-product norms are in this class).[33] Recall that two vectors are orthogonal if their inner product (that induces the distance function) is equal to zero.

For the Euclidean norm, every rotation is an isometry that fixes the origin and preserves orthogonality and orientation: it transforms a basis of orthogonal vectors into another such basis, and each oriented orthogonal basis is obtained (modulo translation) from another via a suitable rotation.

So as to proceed in an analogous fashion, we need to first identify the set of isometries: for any inner-product norm $||\cdot||_I$, this set is always an infinite multiplicative group (see Garcia-Roig 1997). Because medians and welfare measures that are based on distances are translation equivariant, it is enough, as above, to characterize the subgroup of isometries that fix the origin and that preserve orientation (i.e., their corresponding matrices have determinant $+1$). We start with the simplest case.

[32] 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.

[33] These authors also show that, as in the case of the Euclidean norm in the plane, the class of marginal medians coincides with the class of DIC, anonymous, and Pareto efficient mechanisms.

5.1 *Weighted Euclidean norm*

An agent with ideal point $\mathbf{t}_i = (x_i, y_i)$ has a weighted Euclidean preference over points $\mathbf{v} = (x, y)$ given by

$$-\beta^2(x - x_i)^2 - (y - y_i)^2,$$

with $\beta > 0$. Note that, without loss of generality, we can always normalize one of the weights to be $+1$ without changing the underlying (ordinal) preferences. Let

$$M = \begin{pmatrix} \beta^2 & 0 \\ 0 & 1 \end{pmatrix},$$

and define an inner product and its associated norm by

$$\begin{aligned} \langle (x_1, y_1), (x_2, y_2) \rangle &\equiv (x_1, y_1) M (x_2, y_2)^T \\ \|(x, y)\| &\equiv \sqrt{(x, y) M (x, y)^T} = \sqrt{\beta^2 x^2 + y^2}. \end{aligned}$$

Here the "unit circle" is an ellipse,

$$\beta^2 x^2 + y^2 = 1,$$

with axes parallel to the standard Cartesian coordinate axes. Isometries that fix the origin leave this ellipse invariant (i.e., a point on the ellipse is translated to another point on the ellipse) and can be represented by generalized "rotation" matrices of the form

$$R_\beta(\theta) = \begin{pmatrix} \cos\theta & -\dfrac{1}{\beta}\sin\theta \\ \beta\sin\theta & \cos\theta \end{pmatrix}.$$

While the mean in each coordinate is still the first-best, the welfare measure changes to incorporate the weight β. By normalizing the mean to zero, the welfare maximization problem becomes

$$\min_\theta \left[\beta^2 m_{\beta-}^2(\theta) + m_{\beta+}^2(\theta) + \beta^2\sigma_X^2 + \sigma_Y^2\right] \quad \Leftrightarrow \quad \min_\theta \left[\beta^2 m_{\beta-}^2(\theta) + m_{\beta+}^2(\theta)\right],$$

where

$$\begin{aligned} m_{\beta-}(\theta) &= \operatorname{median}\left(X\cos\theta - \frac{1}{\beta} Y\sin\theta\right) \\ m_{\beta+}(\theta) &= \operatorname{median}(\beta X\sin\theta + Y\cos\theta). \end{aligned}$$

As before, it is straightforward to verify that the minimum attained by any angle $\theta \in [\pi/2, 2\pi]$ can be attained by an angle $\theta \in [0, \pi/2]$. Hence, it is without loss of generality to restrict attention to $\theta \in [0, \pi/2]$. Instead of $\theta = \pi/4$, the rotation that yields $m_{\beta-}(\theta) = 0$ is defined here by

$$\cos\theta = \frac{1}{\beta}\sin\theta \quad \Leftrightarrow \quad \theta = \arctan\beta.$$

We now show that the insight of Theorem 1 generalizes to the present framework.

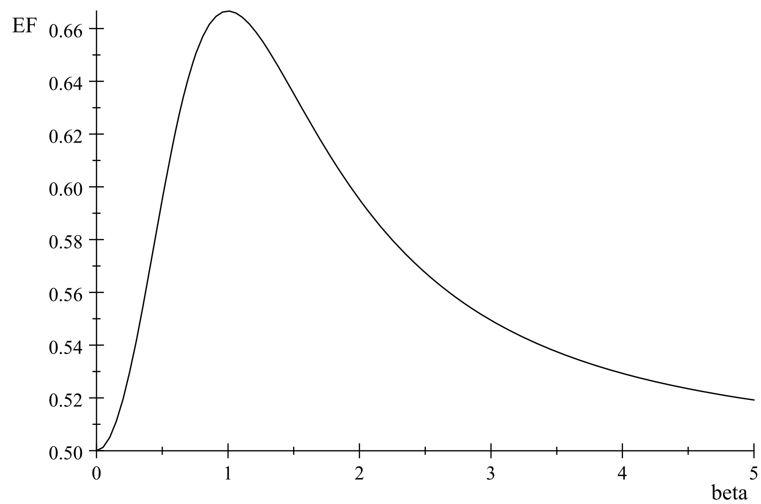

FIGURE 5. Bound on relative efficiency for I.I.D. random variables.

THEOREM 3. *Assume that X and Y are independent. The rotation with angle $\theta = 0$ is a local utility minimum if*

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

It is also straightforward to derive efficiency bounds. Here are two examples: (i) The universal bound based on the Hotelling–Solomons inequality (without any assumption on the underlying random variables) remains $\frac{1}{2}$, independently of β. (ii) If X and Y are independent, using the generalized rotation where $\theta = \arctan\beta$, we obtain

$$EF(\arctan\beta) \geq \frac{1+\beta^2}{1+2\beta^2+(1-\beta^2)\cos^2(\arctan\beta)}.$$

We depict in Figure 5 the bound as a function of β (recall that $EF(\frac{\pi}{4}) \geq \frac{2}{3}$ with $\beta = 1$). Note that the bound tends back to the universal Hotelling–Solomons bound $\frac{1}{2}$ for $\beta \to 0$ and for $\beta \to \infty$. This is intuitive since in those limit cases, one dimension becomes irrelevant and we obtain in the limit a one-dimensional voting problem where "rotations" cannot help.

5.2 *General inner-product norm*

Consider next a general norm induced by an inner product. Such a norm is generated by a symmetric, positive definite matrix Q:

$$\|(x, y)\| \equiv \sqrt{(x, y)Q(x, y)^T}.$$

The "unit circle" is now an ellipse that is possibly tilted with respect to the standard coordinates. Let A_Q be the orthogonal matrix that represents the change of variables that diagonalizes Q, and let M_Q be the obtained *diagonal* matrix.[34] Then M_Q is the matrix of a weighted Euclidean inner product, as explained above. The set of isometries that fix the origin and preserve orientation is thus given here by the composition

$$A_Q R_{M_Q}(\theta) A_Q^{-1},$$

where $R_{M_Q}(\theta)$ is the set of generalized rotations that keep invariant the untilted unit ellipse associated to the diagonal matrix M_Q, as explained in the previous subsection. Note that the unit circle (i.e., ellipse) of this norm now has axes that are parallel to the coordinate axes defined by the change of variables A_Q. In particular, the relevant "0 rotation" is the rotation corresponding to these new variables; it is suboptimal if the distribution of peaks has independent projections on these coordinates (rather than on the standard Cartesian coordinates).

6. Concluding remarks

A redefinition of issues facilitates the search for consensus among ex ante conflicting interests. We have shown that voting by simple majority on each dimension becomes a highly effective 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 induced equilibrium" can be reached in a complex multidimensional collective decision problem with incomplete information about preferences. While we have focused on welfare maximization, other goals (such as maximizing the utility of an agenda setter) can be analyzed by the same methods.

Appendix A: Omitted proofs

A.1 *Proof of Proposition 1*

If X and Y are I.I.D., then we have

$$m_-(\pi/4) = 0 \quad \text{and} \quad m_+(\pi/4) = \frac{\sqrt{2}}{2} m_{X+Y}.$$

Therefore, $\theta = \pi/4$ is a critical point if

$$0 = m_-(\pi/4)m'_-(\pi/4) + m_+(\pi/4)m'_+(\pi/4) = \frac{\sqrt{2}}{2} m_{X+Y} m'_+(\pi/4).$$

By the definition of $m_+(\theta)$,

$$\begin{aligned}
\frac{1}{2} &= F_{X\sin\theta + Y\cos\theta}\big(m_+(\theta)\big) \\
&= \int_{-\infty}^{\infty} \Pr\left(Y < \frac{m_+(\theta) - x\sin\theta}{\cos\theta}\right) f_X(x)\, dx \\
&= \int_{-\infty}^{\infty} F_Y\left(\frac{m_+(\theta) - x\sin\theta}{\cos\theta}\right) f_X(x)\, dx.
\end{aligned}$$

[34] Note that any symmetric, positive definite matrix can indeed be diagonalized, and its two eigenvalues are always real.

Since the above identity holds for all θ, we take the derivative with respect to θ and obtain

$$0=\int_{-\infty}^{\infty} f_Y\left(\frac{m_+(\theta)-x\sin\theta}{\cos\theta}\right)\left(\frac{m'_+(\theta)\cos\theta-x+m_+(\theta)\sin\theta}{\cos^2\theta}\right)f_X(x)\,dx.$$

Hence, if X and Y are I.I.D., and $\theta=\pi/4$, we have

$$0=\int_{-\infty}^{\infty} f_X\big(\sqrt{2}m_+(\pi/4)-x\big)\big(\sqrt{2}m'_+(\pi/4)-2x+\sqrt{2}m_+(\pi/4)\big)f_X(x)\,dx.$$

It follows from the convolution of the distributions for X and Y that

$$\begin{aligned}&\sqrt{2}m'_+(\pi/4)f_{X+Y}\big(\sqrt{2}m_+(\pi/4)\big)\\&\quad=\int_{-\infty}^{\infty} f_X\big(\sqrt{2}m_+(\pi/4)-x\big)\big(2x-\sqrt{2}m_+(\pi/4)\big)f_X(x)\,dx.\end{aligned}$$

Note that by a change of variable $y=\sqrt{2}m_+(\pi/4)-x$, we have

$$\begin{aligned}&\int_{-\infty}^{\infty} f'_X\big(\sqrt{2}m_+(\pi/4)-x\big)\big(2x-\sqrt{2}m_+(\pi/4)\big)^2 f_X(x)\,dx\\&\quad=\int_{-\infty}^{\infty} f'_X(y)\big(2y-\sqrt{2}m_+(\pi/4)\big)^2 f_X\big(\sqrt{2}m_+(\pi/4)-y\big)\,dy.\end{aligned}$$

Therefore,

$$\begin{aligned}0&=\int_{-\infty}^{\infty}\big[f'_X\big(\sqrt{2}m_+(\pi/4)-x\big)f_X(x)-f'_X(x)f_X\big(\sqrt{2}m_+(\pi/4)-x\big)\big]\big(2x-\sqrt{2}m_+(\pi/4)\big)^2\,dx\\&=\big[f_X\big(\sqrt{2}m_+(\pi/4)-x\big)f_X(x)\big(2x-\sqrt{2}m_+(\pi/4)\big)^2\big]_{-\infty}^{\infty}\\&\quad-\int_{-\infty}^{\infty} f_X\big(\sqrt{2}m_+(\pi/4)-x\big)f_X(x)4\big(2x-\sqrt{2}m_+(\pi/4)\big)\,dx\\&=4\sqrt{2}m'_+(\pi/4)f_{X+Y}\big(\sqrt{2}m_+(\pi/4)\big),\end{aligned}$$

where we use the assumption that

$$\begin{aligned}&\lim_{x\to\infty} f_X\big(\sqrt{2}m_+(\pi/4)-x\big)f_X(x)\big(2x-\sqrt{2}m_+(\pi/4)\big)^2\\&\quad=\lim_{x\to-\infty} f_X\big(\sqrt{2}m_+(\pi/4)-x\big)f_X(x)\big(2x-\sqrt{2}m_+(\pi/4)\big)^2=0.\end{aligned}$$

Therefore, $m'_+(\pi/4)=0$. It follows that $\sqrt{2}m_{X+Y}m'_+(\pi/4)=0$, so $\theta=\pi/4$ is indeed a critical point.

A.2 *Proof of Proposition 2*

We prove the case where $m_X<\mu_X$. The case $m_X>\mu_X$ is analogous. So as to show that $\pi/4$ is globally optimal, it is sufficient to show, for any $\beta\in[1/2,1]$,

$$m_X\le m_{\sqrt{\beta}X_1+\sqrt{1-\beta}X_2}\le m_{\frac{\sqrt{2}}{2}X_1+\frac{\sqrt{2}}{2}X_2}<0=\mu_X. \tag{9}$$

Note that for $0 \le \theta'' < \theta' \le \pi/4$, $(\sin^2\theta', \cos^2\theta') \prec (\sin^2\theta'', \cos^2\theta'')$. Hence, Schur-concavity of $\Pr(X\sin\theta + Y\cos\theta \le z)$ in $(\sin^2\theta, \cos^2\theta)$ for all $\theta \in [0, \frac{\pi}{4}]$ and all $z \in [m_X, 0]$ is equivalent to

$$\Pr(X\sin\theta + Y\cos\theta \le z) \quad \text{is weakly increasing in } \theta \in \left[0, \frac{\pi}{4}\right] \text{ for all } z \in [m_X, 0].$$

In other words,

$$h(\beta) \equiv \Pr(\sqrt{\beta}X_1 + \sqrt{1-\beta}X_2 \le z) \quad \text{is weakly increasing in } \beta \text{ for all } z \in [m_X, 0]. \tag{10}$$

Now suppose condition (10) holds. It implies that

$$\Pr(\sqrt{\beta}X_1 + \sqrt{1-\beta}X_2 \le m_X) \le \Pr(X_1 \le m_X) = \frac{1}{2} \quad \Rightarrow \quad m_X \le m_{\sqrt{\beta}X_1 + \sqrt{1-\beta}X_2}.$$

Furthermore, (10) implies that, for all $\beta \in [1/2, 1]$,

$$\Pr\left(\frac{\sqrt{2}}{2}X_1 + \frac{\sqrt{2}}{2}X_2 \le m_{\sqrt{\beta}X_1+\sqrt{1-\beta}X_2}\right) \le \Pr(\sqrt{\beta}X_1 + \sqrt{1-\beta}X_2 \le m_{\sqrt{\beta}X_1+\sqrt{1-\beta}X_2}) = \frac{1}{2},$$

which implies

$$m_{\sqrt{\beta}X_1+\sqrt{1-\beta}X_2} \le m_{\frac{\sqrt{2}}{2}X_1+\frac{\sqrt{2}}{2}X_2}.$$

Condition (9) then follows immediately and, thus, the $\pi/4$ rotation is optimal.

A.3 *Proof of Proposition 3*

Suppose that $F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \le 1$ 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

$$\begin{aligned} m_X - \mu_X &= \int_{-\infty}^{m_X} (m_X - x) f_X(x)\,dx + \int_{m_X}^{\infty} (m_X - x) f_X(x)\,dx \\ &= \int_{-\infty}^{m_X} F_X(x)\,dx - \int_{m_X}^{\infty} \big(1 - F_X(x)\big)\,dx \\ &= \int_0^{\infty} \big[F_X(m_X - x) + F_X(m_X + x) - 1\big]\,dx. \end{aligned}$$

It follows from $m_X \ne \mu_X$ that $m_X < \mu_X$. It also implies that $F_X(m_X - x) + F_X(m_X + x) - 1 < 0$ for some set of x with positive measure.

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

$$\Pr(X + Y < m_X + m_Y) < \frac{1}{2}. \tag{11}$$

Note that

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

Therefore, condition (11) is equivalent to

$$
4\int_{0}^{\infty} F_X(m_X-\varepsilon)f_Y(m_Y+\varepsilon)\,d\varepsilon+4\int_{0}^{\infty} f_X(m_X+\varepsilon)F_Y(m_Y-\varepsilon)\,d\varepsilon<1. \tag{12}
$$

Let us define nonnegative random variables X^+, X^-, Y^+, and Y^- as

$$
\begin{aligned}
X^+&=X-m_X|X\geq m_X, \qquad & X^-&=m_X-X|X\leq m_X\\
Y^+&=Y-m_Y|Y\geq m_Y, \qquad & Y^-&=m_Y-Y|Y\leq m_Y.
\end{aligned}
$$

Then

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

Therefore, condition (12) is equivalent to

$$
\Pr(X^->Y^+)+\Pr(Y^->X^+)<1. \tag{13}
$$

A sufficient condition for (13) is

$$
\Pr(X^+<\varepsilon)\leq\Pr(X^-<\varepsilon)\quad\text{and}\quad\Pr(Y^+<\varepsilon)\leq\Pr(Y^-<\varepsilon) \tag{14}
$$

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

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

and, thus, (13). Since X and Y are I.I.D., the sufficient condition (14) reduces to

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

Equivalently,

$$\Pr(X - m_X < \varepsilon) \le \Pr(m_X - X < \varepsilon),$$

which simplifies into the first inequality in (7). As we argued above, since $m_X \neq \mu_X$, we must have $F_X(m_X - \varepsilon) + F_X(m_X + \varepsilon) - 1 < 0$ for some positive measure of ε, as desired.

Finally, we show that condition (8) ((7), respectively) is satisfied if F_X is strictly convex (concave, respectively). Note that $F(X)$ is uniformly distributed, so that $E[F(X)] = 1/2$. Suppose here that F is strictly convex. The concave case can be proved analogously. By Jensen's inequality,

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

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

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

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 $F_X(m_X + \varepsilon) + F_X(m_X - \varepsilon) \ge 1$ for all $\varepsilon > 0$.

A.4 *Examples for Section 3.3*

We show here how the super-additivity condition of median is satisfied for two well known families of distributions where condition (7) is not easily checked or does not hold.[35]

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

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

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 also has a gamma distribution with parameters α and β/c. If X and Y are independent gamma random variables with parameters (α_X, β) and (α_Y, β), respectively, then $X + Y$ is also a gamma random variable 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 $\Pr(aX + bY \le t)$, $0 \le a, b \le 1$, is Schur-convex in (a, b) for all $t \le \mu_X$. Since $(1, 0) \succ (\frac{1}{2}, \frac{1}{2})$, we have $F_{\frac{1}{2}X+\frac{1}{2}Y}(t) \le F_X(t)$ for all $t \le \mu_X$. Note that $m_X < \mu_X$ for gamma distributions (Groeneveld and Meeden 1977), so we have $m_{\frac{1}{2}X+\frac{1}{2}Y} \ge m_X$ as desired.[36]

[35]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.

[36]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

$$U\left(\frac{\pi}{4}\right) = -2\sigma^2(\alpha, \beta) - (\mu_+ - m_+)^2$$

A second family is the Rayleigh distribution with cumulative distribution:

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

Suppose X and Y are I.I.D. distributed according to Rayleigh.[37] Then, according to Lemma 4 in Hu and Lin (2000), we have

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

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

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

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. We now present an example where, even though X and Y are

---

$$\begin{aligned}
&= -2\sigma^2(\alpha, \beta) - \left(\frac{\sqrt{2}\alpha}{\beta} - \frac{\sqrt{2}}{2\beta} m(2\alpha, 1)\right)^2 \\
&= -2\sigma^2(\alpha, \beta) - \frac{1}{2\beta^2}\big(2\alpha - m(2\alpha, 1)\big)^2
\end{aligned}$$

and

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

Therefore,

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

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

[37] 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.

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 Neilsen 2006 for an introduction).

Example 3. 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 the 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 the 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 $C(p,q)$:

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

With some algebra, we can derive the joint density as

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

Therefore, as in the proof of Proposition 3, we can write $\Pr(X+Y<m_X+m_Y)$ as

$$\begin{aligned}
&2\int_{m_Y}^{1}\int_0^{m_X+m_Y-y} f(x,y)\,dx\,dy+\int_0^{m_Y}\int_0^{m_X} f(x,y)\,dx\,dy\\
&\quad=2\int_{\sqrt{2}/2}^{1}\int_0^{\sqrt{2}-y}\bigl(4xy+4\delta xy\bigl(2x^2-1\bigr)\bigl(2y^2-1\bigr)\bigr)\,dx\,dy\\
&\qquad+\int_0^{\sqrt{2}/2}\int_0^{\sqrt{2}/2}\bigl(4xy+4\delta xy\bigl(2x^2-1\bigr)\bigl(2y^2-1\bigr)\bigr)\,dx\,dy\\
&\quad=\left(\frac{146}{35}-\frac{104}{35}\sqrt{2}\right)\delta-\frac{8}{3}\sqrt{2}+\frac{13}{3}\\
&\quad>0.5
\end{aligned}$$

for all $\delta\in[-1,1]$. Consequently, we have $m_{X+Y}<m_X+m_Y$. Since $F_X(x)=x^2$ is convex, $\mu_X<m_X$. It follows that the $\pi/4$ rotation dominates the 0 rotation in ex ante welfare. 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 $\mu_X>m_X$. ♢

A.5 *Proof of Theorem 3*

So as to show that $\theta=0$ is suboptimal, it is sufficient to show

$$\beta^2 m_{\beta-}(0)m'_{\beta-}(0)+m_{\beta+}(0)m'_{\beta+}(0)=0 \tag{15}$$

and

$$\beta^2 m''_{\beta-}(0)m_{\beta-}(0)+\beta^2\bigl(m'_{\beta-}(0)\bigr)^2+m''_{\beta+}(0)m_{\beta+}(0)+\bigl(m'_{\beta+}(0)\bigr)^2<0. \tag{16}$$

By definition of $m_{\beta+}(\theta)$, we note that

$$\begin{aligned}\frac{1}{2} &= F_{\beta X\sin\theta+Y\cos\theta}\big(m_{\beta+}(\theta)\big)\\ &= \int_{-\infty}^{\infty}\Pr\left(Y<\frac{m_{\beta+}(\theta)-\beta x\sin\theta}{\cos\theta}\right)f_X(x)\,dx\\ &= \int_{-\infty}^{\infty}F_Y\left(\frac{m_{\beta+}(\theta)-\beta x\sin\theta}{\cos\theta}\right)f_X(x)\,dx.\end{aligned}$$

Since it holds for all θ, we take a derivative with respect to θ to obtain

$$0=\int_{-\infty}^{\infty}f_Y\left(\frac{m_{\beta+}(\theta)-\beta x\sin\theta}{\cos\theta}\right)\left(\frac{m'_{\beta+}(\theta)\cos\theta-\beta x+m_{\beta+}(\theta)\sin\theta}{\cos^2\theta}\right)f_X(x)\,dx. \tag{17}$$

By taking the second derivative with respect to θ, we obtain

$$\begin{aligned}0 &= \int_{-\infty}^{\infty}f'_Y\left(\frac{m_{\beta+}(\theta)-\beta x\sin\theta}{\cos\theta}\right)\left(\frac{m'_{\beta+}(\theta)\cos\theta-\beta x+m_{\beta+}(\theta)\sin\theta}{\cos^2\theta}\right)^2 f_X(x)\,dx\\ &\quad+\int_{-\infty}^{\infty}\frac{f_Y\left(\dfrac{m_{+}(\theta)-\beta x\sin\theta}{\cos\theta}\right)}{\cos^4\theta}\\ &\quad\times\left(\begin{array}{c}\big[m''_{\beta+}(\theta)\cos\theta+m_{\beta+}(\theta)\cos\theta\big]\cos^2\theta\\ +2\cos\theta\sin\theta\big(m'_{\beta+}(\theta)\cos\theta-\beta x+m_{\beta+}(\theta)\sin\theta\big)\end{array}\right)f_X(x)\,dx.\end{aligned} \tag{18}$$

If $\theta=0$, then conditions (17) and (18) reduce to

$$0=\int_{-\infty}^{\infty}f_Y\big(m_{\beta+}(0)\big)\big(m'_{\beta+}(0)-\beta x\big)f_X(x)\,dx$$

and

$$\begin{aligned}0 &= \int_{-\infty}^{\infty}f'_Y\big(m_{\beta+}(0)\big)\big(m'_{\beta+}(0)-\beta x\big)^2 f_X(x)\,dx\\ &\quad+\int_{-\infty}^{\infty}f_Y\big(m_{\beta+}(0)\big)\big(m''_{\beta+}(0)+m_{\beta+}(0)\big)f_X(x)\,dx.\end{aligned}$$

Note that $m_{\beta+}(0)=m_Y$, so that we have

$$m'_{\beta+}(0)=\frac{\beta f_Y(m_Y)\displaystyle\int_{-\infty}^{\infty}xf_X(x)\,dx}{f_Y(m_Y)\displaystyle\int_{-\infty}^{\infty}f_X(x)\,dx}=\beta\mu_X=0$$

and

$$m''_{\beta+}(0)=-m_Y-\frac{f'_Y(m_Y)}{f_Y(m_Y)}\int_{-\infty}^{\infty}\beta x^2 f_X(x)\,dx.$$

Similarly, we can write

$$\frac{1}{2} = F_{X\cos\theta - \frac{1}{\beta}Y\sin\theta}\big(m_{\beta-}(\theta)\big) = \int_{-\infty}^{\infty} F_X\left(\frac{m_{\beta-}(\theta) + \frac{1}{\beta}y\sin\theta}{\cos\theta}\right) f_Y(y)\,dy.$$

Taking the derivative with respect to θ, we obtain

$$0 = \int_{-\infty}^{\infty} f_X\left(\frac{m_{\beta-}(\theta) + \frac{1}{\beta}y\sin\theta}{\cos\theta}\right)\left(\frac{m'_{\beta-}(\theta)\cos\theta + \frac{1}{\beta}y + m_{\beta-}(\theta)\sin\theta}{\cos^2\theta}\right) f_Y(y)\,dy.$$

Taking the second derivative with respect to θ, we obtain

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} f'_X\left(\frac{m_{\beta-}(\theta) + \frac{1}{\beta}y\sin\theta}{\cos\theta}\right)\left(\frac{m'_{\beta-}(\theta)\cos\theta + \frac{1}{\beta}y + m_{\beta-}(\theta)\sin\theta}{\cos^2\theta}\right)^2 f_Y(y)\,dy \\ &\quad + \int_{-\infty}^{\infty} \frac{f_X\left(\frac{m_{\beta-}(\theta) + \frac{1}{\beta}y\sin\theta}{\cos\theta}\right)}{\cos^4\theta} \\ &\quad \times \left(\begin{array}{c} \big[m''_{\beta-}(\theta)\cos\theta + m_{\beta-}(\theta)\cos\theta\big]\cos^2\theta \\ + 2\cos\theta\sin\theta\left[m'_{\beta-}(\theta)\cos\theta + \frac{1}{\beta}y + m_{\beta-}(\theta)\sin\theta\right] \end{array}\right) f_Y(y)\,dy. \end{aligned}$$

If $\theta = 0$, then the above two conditions reduce to

$$0 = \int_{-\infty}^{\infty} f_X\big(m_{\beta-}(0)\big)\left(m'_{\beta-}(0) + \frac{1}{\beta}y\right) f_Y(y)\,dy$$

and

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} f'_X\big(m_{\beta-}(0)\big)\left(m'_{\beta-}(0) + \frac{1}{\beta}y\right)^2 f_Y(y)\,dy \\ &\quad + \int_{-\infty}^{\infty} f_X\big(m_{\beta-}(0)\big)\big(m''_{\beta-}(0) + m_{\beta-}(0)\big) f_Y(y)\,dy. \end{aligned}$$

Since $m_{\beta-}(0) = m_X$, we have

$$m'_{\beta-}(0) = -\frac{1}{\beta}\frac{\int_{-\infty}^{\infty} y f_Y(y)\,dy}{\int_{-\infty}^{\infty} f_Y(y)\,dy} = -\frac{1}{\beta}\mu_Y = 0$$

and

$$m''_{\beta-}(0) = -m_X - \frac{f'_X(m_X)}{f_X(m_X)}\int_{-\infty}^{\infty}\frac{1}{\beta^2}y^2 f_Y(y)\,dy.$$

Therefore, the first-order condition (15) holds because $m'_{\beta-}(0) = m'_{\beta+}(0) = 0$. For the second-order condition (16), note that

$$\begin{aligned}
&\beta^2 m''_{\beta-}(0) m_{\beta-}(0) + \beta^2 \big(m'_{\beta-}(0)\big)^2 + m''_{\beta+}(0) m_{\beta+}(0) + \big(m'_{\beta+}(0)\big)^2 \\
&= \beta^2 m_X \left(-m_X - \frac{f'_X(m_X)}{f_X(m_X)} \int_{-\infty}^{\infty} \frac{1}{\beta^2} y^2 f_Y(y)\, dy\right) \\
&\quad + m_Y \left(-m_Y - \frac{f'_Y(m_Y)}{f_Y(m_Y)} \int_{-\infty}^{\infty} \beta^2 x^2 f_X(x)\, dx\right) \\
&= -\beta^2 m_X^2 - m_Y^2 - m_X \frac{f'_X(m_X)}{f_X(m_X)} \int_{-\infty}^{\infty} y^2 f_Y(y)\, dy \\
&\quad - m_Y \frac{f'_Y(m_Y)}{f_Y(m_Y)} \int_{-\infty}^{\infty} \beta^2 x^2 f_X(x)\, dx.
\end{aligned}$$

As a result, condition (16) is equivalent to

$$\beta^2 m_X^2 + m_Y^2 + m_X \frac{f'_X(m_X)}{f_X(m_X)} \int_{-\infty}^{\infty} y^2 f_Y(y)\, dy + m_Y \frac{f'_Y(m_Y)}{f_Y(m_Y)} \int_{-\infty}^{\infty} \beta^2 x^2 f_X(x)\, dx > 0.$$

Therefore, a sufficient condition for the sub-optimality of zero rotation is

$$m_X f'_X(m_X) \geq 0, \qquad m_Y f'_Y(m_Y) \geq 0 \quad \text{and} \quad \beta^2 m_X^2 + m_Y^2 \neq 0.$$

Appendix B: More than two dimensions

In this appendix, we sketch the generalizations of our main results (Theorems 1 and 2) to higher dimensions. Consider K independent issues, denoted by X_k, $k = 1, \dots, K$. We write $\mathbf{X} = (X_1, \dots, X_K)^T$ 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

$$Q^T Q = Q Q^T = I,$$

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 $Q\mathbf{X}$. The planner's objective is to choose Q to maximize welfare.

B.1 *The (sub)-optimality of the 0 and $\pi/4$ rotations*

Theorem 1 can be easily extended to higher dimensions by applying our previous two-dimensional analysis to rotations of the first two dimensions only (while keeping all other dimensions fixed).

Suppose now that $X_1, \ldots, X_K$ 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 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 $X_1 + \cdots + X_K$ while the other coordinates consist of various differences. 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, consider an orthogonal matrix $\widehat{Q}$ with

$$\frac{1}{\sqrt{K}} \widehat{Q}_{ij} = \begin{cases} \text{either } 1 \text{ or } -1 & \text{for all } j \text{ if } i \neq K \\ 1 & \text{for all } j \text{ if } i = K, \end{cases} \tag{19}$$

such that for all $k \neq K$, $\widehat{Q}_k\mathbf{X}$ contains an equal number of X_ks appearing with positive and negative signs. The matrix $\widehat{Q}_k$ is a *Hadamard matrix,* and the order of such a matrix must be 1, 2, or a multiple of 4. Sylvester (1867) constructed Hadamard matrices of order 2^k for every nonnegative integer k.[38] In those cases, 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 0 rotation.

B.2 *Efficiency bounds*

As in the main text, we work here with the nonnormalized random variables $X_1, \ldots, X_K$. With K dimensions, the expected utility of choosing marginal medians under an orthogonal transformation Q is given by

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

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

[38] The existence of Hadamard matrices of order $4k$ for every positive integer k is the well known *Hadamard conjecture.* It was proven for all k up to 167.

first-best as

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

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

$$EF(I) \geq \frac{\sum_{k=1}^{K} \mathrm{var}(Q_k \mathbf{X})}{\sum_{k=1}^{K} \mathrm{var}(Q_k \mathbf{X}) + \sum_{k=1}^{K} \mathrm{var}(Q_k \mathbf{X})} = \frac{1}{2}.$$

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

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

Now consider any even number K such that the Hadamard matrix exists**.** Suppose $X_1, \dots, X_K$ are I.I.D. with log-concave densities. Consider again an orthogonal matrix $\widehat{Q}$ given in (19). It follows from the I.I.D. assumption that

$$\mathrm{mean}(\widehat{Q}_k \mathbf{X}) - \mathrm{median}(\widehat{Q}_k \mathbf{X}) = \begin{cases} 0 & \text{if } k \neq K \\ \dfrac{1}{\sqrt{K}} \left(\mathrm{mean}\left(\sum_{k=1}^{K} X_k \right) - \mathrm{median}\left(\sum_{k=1}^{K} X_k \right) \right) & \text{if } k = K. \end{cases}$$

Therefore, we have

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

Given that $X_1, \dots, X_K$ 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 (2016) 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 $\widehat{Q}_k \widehat{Q}_k^T = I$ by the definition of an orthogonal matrix. Therefore, we obtain the efficiency bound for log-concave densities:

$$EF(\widehat{Q}) \geq \frac{K\sigma_{X_k}^2}{K\sigma_{X_k}^2 + 12\sigma_{X_k}^2 \ln^2\left(\sqrt{\frac{e}{2}}\right)} = \frac{1}{1 + \frac{1}{K} 12 \ln^2\left(\sqrt{\frac{e}{2}}\right)}.$$

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.[39]

Remark 4. More generally, consider any I.I.D. random variables $X_1, \ldots, X_K$ with finite means and variances, and consider K such that the Hadamard matrix exists. Then for the analog of the $\pi/4$ rotation, we obtain that

$$\begin{aligned} EF(\widehat{Q}) &= \frac{K\sigma_{X_k}^2}{K\sigma_{X_k}^2 + \frac{1}{K}\left(\operatorname{mean}\left(\sum_{k=1}^{K} X_k\right) - \operatorname{median}\left(\sum_{k=1}^{K} X_k\right)\right)^2} \\ &= \frac{\sigma_{X_k}^2}{\sigma_{X_k}^2 + \frac{1}{K^2}\left(\operatorname{mean}\left(\sum_{k=1}^{K} X_k\right) - \operatorname{median}\left(\sum_{k=1}^{K} X_k\right)\right)^2} \to 1 \quad \text{as } K \to \infty, \end{aligned}$$

where the last assertion follows from the central limit theorem.

References

Ahn, David S. and Santiago Oliveros (2012), "Combinatorial voting." *Econometrica*, 80, 89–141. [559]

Austen-Smith, David and Jeffrey S. Banks (2005), *Positive Political Theory II: Strategy and Structure*. University of Michigan Press, Ann Arbor, Michigan. [556]

[39]This asymptotic efficiency corresponds to the asymptotic efficiency of pure bundling when the number of objects goes to infinity and when marginal costs of production are zero (see Bakos and Brynjolfsson 1999).

Bahadur, R. R. (1966), "A note on quantiles in large samples." *Annals of Mathematical Statistics*, 37, 577–580. [566]

Bakos, Yannis and Erik Brynjolfsson (1999), "Bundling information goods: Pricing, profits, and efficiency." *Management Science*, 45, 1613–1630. [592]

Ball, Keith M. and Károly J. Böröczky (2010), "Stability of the Prékopa–Leindler inequality." *Mathematika*, 56, 339–356. [576, 591]

Barberà, Salvador, Faruk Gul, and Ennio Stacchetti (1993), "Generalized median voter schemes and committees." *Journal of Economic Theory*, 61, 262–289. [559]

Barlow, Richard E. and Frank Proschan (1965), *Mathematical Theory of Reliability*. John Wiley & Sons, Inc., New York, New York. [575]

Basu, Sanjib and Anirban DasGupta (1997), "The mean, median, and mode of unimodal distributions: A characterization." *Theory of Probability & Its Applications*, 41, 210–223. [568, 575]

Battaglini, Marco (2002), "Multiple referrals and multidimensional cheap talk." *Econometrica*, 70, 1379–1401. [560]

Berg, Christian and Henrik L. Pedersen (2008), "Convexity of the median in the gamma distribution." *Arkiv för Matematik*, 46, 1–6. [585]

Black, Duncan (1948), "On the rationale of group decision-making." *Journal of Political Economy*, 56, 23–34. [556]

Bobkov, Sergey and Michel Ledoux (2016), "One-dimensional empirical measures, order statistics and Kantorovich transport distances." Unpublished paper, School of Mathematics, University of Minnesota. [576, 591]

Bock, Mary Ellen, Persi Diaconis, Fred W. Huffer, and Michael D. Perlman (1987), "Inequalities for linear combinations of gamma random variables." *Canadian Journal of Statistics*, 15, 387–395. [584]

Border, Kim C. and James S. Jordan (1983), "Straightforward elections, unanimity and phantom voters." *Review of Economic Studies*, 50, 153–170. [559]

Caplin, Andrew and Barry Nalebuff (1988), "On 64%-majority rule." *Econometrica*, 56, 787–814. [560, 573]

Caplin, Andrew and Barry Nalebuff (1991), "Aggregation and social choice: A mean voter theorem." *Econometrica*, 59, 1–23. [560, 573]

Charkraborty, Archishman and Richard Harbaugh (2007), "Comparative cheap talk." *Journal of Economic Theory*, 132, 70–94. [560]

Dharmadhikari, Sudhakar and Kumar Joag-Dev (1988), *Unimodality, Convexity, and Applications*, first edition. Academic Press, San Diego, California. [568]

Feld, Scott and Bernard Grofman (1988), "Majority rule outcomes and the structure of debate in one-issue at a time decision-making." *Public Choice*, 59, 239–252. [557]

Ferejohn, John and Keith Krehbiel (1987), "The budget process and the size of the budget." *American Journal of Political Science*, 31, 296–320. [558]

Galton, Francis (1907), "Vox populi." *Nature*, 1949, 450–451. [556]

Garcia-Roig, J. L. (1997), "On the group of isometries and a system of functional equations of a real normed plane." In *Chapman & Hall/CRC Research Notes in Mathematics Series* Vol. 376 (Themistocles M. Rassias, ed.), 42–53, Addison Wesley Longman Limited, Essex, England. [577]

Gershkov, Alex, Benny Moldovanu, and Xianwen Shi (2017), "Optimal voting rules." *Review of Economic Studies*, 84, 688–717. [559, 564]

Groeneveld, Richard A. and Glen Meeden (1977), "The mode, median, and mean inequality." *The American Statistician*, 31, 120–121. [584]

Haldane, J. B. S. (1948), "Note on the median of a multivariate distribution." *Biometrika*, 35, 414–415. [557]

Hotelling, Harold and Leonard M. Solomons (1932), "The limits of a measure of skewness." *Annals of Mathematical Statistics*, 3, 141–142. [556, 574]

Hu, Chin-Yuan and Gwo Dong Lin (2000), "On an inequality for the Rayleigh distribution." *Sankhya: The Indian Journal of Statistics*, 62, 36–39. [585]

Ibragimov, Ildar Abdulovich (1956), "On the composition of unimodal distributions." *Theory of Probability & Its Applications*, 1, 255–260. [576]

Kaas, Rob, Jan Dhaene, David Vyncke, Marc J. Goovaerts, and Michel Denuit (2002), "A simple geometric proof that comonotonc risks have the convex-largest sum." *Astin Bulletin*, 32, 71–80. [576]

Karlin, Samuel and Yosef Rinott (1983), "Comparisons of measures, multivariate majorization and applications to statistics." In *Studies in Econometrics, Time Series, and Multivariate Statistics* (Leo A. Goodman, Samuel Karlin, and Takeshi Amemiya, eds.), 465–489, Academic Press, San Diego. [571]

Kim, K. H. and Fred W. Roush (1984), "Nonmanipulability in two dimensions." *Mathematical Social Sciences*, 8, 29–43. [559, 560, 564]

Kleiner, Andreas and Benny Moldovanu (2017), "Content based agendas and qualified majorities in sequential voting." *American Economic Review*, 107, 1477–1506. [559, 564]

Kramer, Gerald H. (1972), "Sophisticated voting over multidimensional choice spaces." *Journal of Mathematical Sociology*, 2, 165–180. [558]

Kramer, Gerald H. (1973), "On a class of equilibrium conditions for majority rule." *Econometrica*, 41, 285–297. [558]

Lehmann, Eric L. (1966), "Some concepts of dependence." *Annals of Mathematical Statistics*, 37, 1137–1153. [575]

McAfee, R. Preston, John McMillan, and Michael D. Whinston (1989), "Multiproduct monopoly, commodity bundling, and correlation of values." *Quarterly Journal of Economics*, 104, 371–383. [557]

McMurray, Joseph C. (2018), "Why the political world is flat: An endogenous 'left' and 'right' in multidimensional political conflict." Unpublished paper, Economics Department, Brigham Young University. [559]

Moulin, Hervé (1980), "On strategy-proofness and single-peakedness." *Public Choice*, 35, 437–455. [559]

Neilsen, Roger B. (2006), *An Introduction to Copulas*, second edition. Springer, New York. [586]

Palfrey, Thomas R. (1983), "Bundling decisions by a multiproduct monopolist with incomplete information." *Econometrica*, 51, 463–483. [572]

Peters, Hans, Hans van der Stel, and Ton Storcken (1992), "Pareto optimality, anonymity, and strategy-proofness in location problems." *International Journal of Game Theory*, 21, 221–235. [559, 560, 564]

Peters, Hans, Hans van der Stel, and Ton Storcken (1993), "Generalized median solutions, strategy-proofness and strictly convex norms." *Zeitschrift für Operations Research*, 38, 19–53. [577]

Poterba, James M., and Jürgen von Hagen, eds. (1999), *Fiscal Institutions and Fiscal Perormance*. University of Chicago Press, Chiacgo. [556]

Prékopa, András (1973), "On logarithmic concave measures and functions." *Acta Scientiarum Mathematicarum*, 34, 335–343. [576]

Rae, Douglas W. (1969), "Decision rules and individual values in constitutional choice." *American Political Science Review*, 63, 40–56. [557]

Rychlik, Tomasz (2000), "Sharp mean-variance inequalities for quantiles of distributions determined by convex and star orders." *Theory of Probability and Its Applications*, 47, 269–285. [575]

Shepsle, Kenneth A. (1979), "Institutional arrangements and equilibrium in multidimensional voting models." *American Journal of Political Science*, 23, 27–59. [557, 559, 564]

Small, Christopher G. (1990), "A survey of multidimensional medians." *International Statistical Review*, 58, 263–277. [562]

Sylvester, J. J. (1867), "Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers." *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 34, 461–475. [590]

Szegő, Giorgio (1928), "Über einige von S. Ramanujan gestellte Aufgaben." *Journal of the London Mathematical Society*, 3, 225–232. [557]

van Zwet, W. R. (1979), "Mean, median, mode II." *Statistica Neerlandica*, 33, 1–5. [573, 582]

Watson, Ray and Ian Gordon (1986), "On quantiles of sums." *Australian and New Zealand Journal of Statistics*, 28, 192–199. [582]

Co-editor Dilip Mookherjee handled this manuscript.

Manuscript received 8 February, 2018; final version accepted 18 October, 2018; available online 23 October, 2018.