

# Truthful Aggregation of Budget Proposals

RUPERT FREEMAN, University of Virginia

DAVID M. PENNOCK, Rutgers University

DOMINIK PETERS, Harvard University

JENNIFER WORTMAN VAUGHAN, Microsoft Research

We consider a participatory budgeting problem in which each voter submits a proposal for how to divide a single divisible resource (such as money or time) among several possible alternatives (such as public projects or activities) and these proposals must be aggregated into a single aggregate division. Under  $\ell_1$  preferences—for which a voter’s disutility is given by the  $\ell_1$  distance between the aggregate division and the division he or she most prefers—the social welfare-maximizing mechanism, which minimizes the average  $\ell_1$  distance between the outcome and each voter’s proposal, is incentive compatible [Goel et al. 2019]. However, it fails to satisfy a natural fairness notion of *proportionality*, placing too much weight on majority preferences. Leveraging a connection between market prices and the generalized median rules of Moulin [1980], we introduce the *independent markets* mechanism, which is both incentive compatible and proportional. We unify the social welfare-maximizing mechanism and the independent markets mechanism by defining a broad class of *moving phantom* mechanisms that includes both. We show that every moving phantom mechanism is incentive compatible. Finally, we characterize the social welfare-maximizing mechanism as the unique Pareto-optimal mechanism in this class, suggesting an inherent tradeoff between Pareto optimality and proportionality.

**Keywords:** participatory budgeting, social choice, mechanism design

## 1 INTRODUCTION

Participatory budgeting allows members of a community to collectively decide how to divide a portion of the community’s budget among a set of proposed alternatives [Cabannes 2004]. Participatory budgeting was first introduced in Brazil [Schneider and Goldfrank 2002], and it has now been used in more than 3,000 cities around the world, including New York, Boston, Chicago, San Francisco, and Toronto [Participatory Budgeting Project 2019]. It has been used to determine how to allocate the budgets of states and cities as well as housing authorities and schools.

Many participatory budgeting elections are run using a variant of  $k$ -approval voting, in which each voter chooses up to  $k$  projects to approve, and the projects with the highest number of approvals are funded, subject to budget constraints [Goel et al. 2019]. Under such a voting scheme, each proposed project is either fully funded or not funded at all. This makes sense for well-delineated projects such as renovating a school or adding an elevator to a public library. For other kinds of projects, funding decisions need not be all-or-nothing. For example, participatory budgeting could be used to decide how to divide a part of a city’s spending between its departments of health, education, infrastructure, and parks. A voter might propose a division of the tax surplus among the four departments into the fractions (30%, 40%, 20%, 10%). The city could invite each citizen to submit such a budget proposal, and they could then be aggregated by a suitable mechanism.

A first idea for aggregating the proposals would be to take the arithmetic mean. But the mean has a serious flaw as an aggregator: it’s easily manipulated. A voter preferring a (60%, 40%) division across two alternatives may vote (100%, 0%) instead in order to distort the mean calculation and move the aggregate closer to his or her true preference if the first alternative has little community support.

In this paper, we seek mechanisms that are resistant to such manipulation. In particular, we require that no voter can, by lying, move the aggregate division toward his or her preference on one alternative without moving it away from his or her preference by an equal or greater amount on other alternatives. In other words, we seek budget aggregation mechanisms that are

incentive compatible under  $\ell_1$  preferences, with each voter’s disutility for a budget division equal to the  $\ell_1$  distance between that division and the division she prefers most. One can show that such preferences emerge naturally from a model where voters have separable uniform utilities over alternatives together with a funding cap for each alternative beyond which it provides no value to the voter.

Lindner et al. [2008] and Goel et al. [2019] have shown that choosing an aggregate budget division that maximizes the utilitarian social welfare of the voters—that is, a division that minimizes the total  $\ell_1$  distance from each voter’s report—is both incentive compatible and Pareto optimal under this voter utility model. However, this utilitarian aggregate has a tendency to overweight majority preferences. For example, imagine that one hundred voters prefer (100%, 0%) while ninety-nine prefer (0%, 100%). The utilitarian aggregate is (100%, 0%) even though the mean is close to (50%, 50%). In many participatory budgeting scenarios, the latter solution is more in the spirit of consensus. For example, suppose that we vote over the allocation of education funding to schools, and imagine that each family votes for all education funding to go to their own neighborhood school. The utilitarian aggregate would earmark the entire budget to the most populous school district. However, we may prefer that funds are split in proportion to the districts’ populations. To capture this fairness constraint, we define a notion of *proportionality*, requiring that when voters are single-minded (as in this example), the fraction of the budget assigned to each alternative is equal to the proportion of voters who favor that alternative. Do there exist aggregators that are both incentive compatible and proportional?

This question is easiest to answer for two alternatives. In this case,  $\ell_1$  preferences are a special case of *single-peaked* preferences, well studied in the voting literature. The seminal results of Moulin [1980] imply that, in this setting, all incentive-compatible voting schemes correspond to inserting  $n + 1$  “phantom” proposals, where  $n$  is the number of voters, and returning the median of the  $n$  true proposals and the  $n + 1$  phantoms. We show that there exists a unique way of placing the phantoms that results in a proportional mechanism for two alternatives.

Generalizing Moulin’s phantom median mechanisms to allow for more than two alternatives is difficult. Existing proposals for such generalizations take a median in each dimension independently [Barberà and Jackson 1994; Border and Jordan 1983; Peters et al. 1992], but this strategy is doomed in our application with normalization constraints that require funding to sum up to 100%. Unlike the mean, taking a coordinate-wise median will usually fail to normalize. We address this problem by allowing the set of phantoms to continuously shift upwards, increasing the sum of the generalized medians until the aggregate becomes normalized. This idea allows us to define a very general class of *moving phantom* mechanisms. One might think that allowing the final phantom locations to depend on voters’ reports might give voters an incentive to misreport. However, we prove that every moving phantom mechanism is incentive compatible under  $\ell_1$  preferences.

Among this large family of incentive-compatible mechanisms, we find one that satisfies our proportionality requirement. This moving phantom mechanism is obtained when phantoms are placed uniformly between 0 and a value  $x \geq 0$  which increases until the coordinate-wise medians sum to one. To analyze this mechanism, we prove that the aggregate found by this mechanism can be interpreted as the clearing prices in a market system, and hence call it the *independent markets* mechanism. This reveals an unexpected connection between market prices and generalized medians that may be of broader interest. The independent markets mechanism can also be justified from a game-theoretic perspective as the unique Nash equilibrium of a natural voting game. Thus, this proportional moving phantom mechanism has two alternative interpretations:

- (1) Market interpretation: For each alternative, we set up a market in which  $x$  units of a divisible good are sold. This amount  $x$  is the same across all markets. Each voter has a value for the

good in market  $j$  that is equal to the fraction of the budget that the voter would prefer be allocated to alternative  $j$  in the budget division setting, and has \$1 to spend in each market. Increase  $x$  until the point at which the market clearing prices across these independent markets sum to \$1. At this point, the market clearing prices are exactly the aggregate division selected by the independent markets moving phantom mechanism.

- (2) Voting game: Each agent receives one credit for each alternative and may choose any amount of that credit to spend on the alternative. The outcome of the game is a normalized vector proportional to the amount spent on each alternative. Agents choose their spending in such a way as to minimize the  $\ell_1$  distance between their preferred budget division and the outcome of the game. This game has a unique Nash equilibrium outcome which is exactly the division selected by the independent markets mechanism.

By analyzing the market and Nash equilibria of these systems, we can show that our mechanism satisfies several important social choice properties, such as participation and reinforcement. We also show that our mechanism retains its proportionality guarantees even for voters that approve exactly  $k$  projects each, for some  $k \geq 1$ , and propose to split the budget equally among them.

In contrast, the independent markets mechanism fails to satisfy Pareto optimality. We show that this is unavoidable, as no proportional moving phantom mechanism is Pareto optimal. In fact, we prove that there is a *unique* moving phantom mechanism that is Pareto optimal (provided the number of alternatives is sufficiently large). In this mechanism, all phantoms start at 0 and then, one by one, transition to 1, with no two phantoms moving at the same time. This mechanism turns out to also have a phantom-free interpretation: it is equivalent to selecting the budget division that maximizes utilitarian social welfare (breaking ties in favor of divisions closer to a uniform division)—the same mechanism previously studied by Lindner et al. [2008] and Goel et al. [2019].

While the motivation of our formal model is participatory budgeting, it applies to the division of other resources such as time. For example, one might imagine using such a mechanism as a way to reach consensus among a team of conference organizers who wish to divide a day between talks, poster sessions, and social activities. Another example would be a government that needs to decide on a target energy mix (that is, how much energy should come from fossil fuels, nuclear, or renewable sources) and wishes to aggregate expert proposals. In all these applications, our class of moving phantom mechanisms can be used to make better decisions.

*Related Work.* Several recent papers study voting rules for participatory budgeting, considering both axiomatics and computational complexity, but under the assumption that projects are indivisible, so can either be fully funded or not funded at all [Aziz et al. 2018; Aziz and Shah 2020; Benade et al. 2017; Goel et al. 2019; Lu and Boutilier 2011]. The setting in which partial funding of alternatives is permitted has also been studied, but generally under a different utility model in which voters assign utility scores to the alternatives rather than having an ideal distribution [Airiau et al. 2019; Aziz et al. 2019; Bogomolnaia et al. 2005; Fain et al. 2016]. In this model, Guerdjikova and Nehring [2014] characterize a rule based on maximizing the Nash product using an axiom very similar to our proportionality property: If voters are single-minded in the sense of assigning utility 0 to all but one project, then the projects receive funding proportional to the number of supporters. Duddy [2015] proposes a rule for this setting that satisfies a stronger proportionality criterion (which gives representation guarantees to every subset of voters) and is strategyproof if voters have dichotomous preferences. This body of work includes positive results for weak versions of incentive compatibility, but impossibilities for obtaining full incentive compatibility. Garg et al. [2019] perform a Mechanical Turk study exploring preference structure in a high-dimensional continuous setting similar to ours.

Closely related to our work is an unpublished paper by Lindner et al. [2008], as well as the PhD thesis (in German language) of Lindner [2011]. Like us, they study the aggregation of points in the standard simplex. Their work focusses on the set of points that maximize utilitarian social welfare, defined in the  $\ell_1$  sense, and studies its structure. Further, they discuss a sense in which this set, considered as an irresolute voting rule, is strategyproof [Lindner 2011, Sec. 3.2.2]. Lindner [2011, Sec. 3.2.4] introduces a tie-breaking scheme to turn this rule into a resolute (single-valued) aggregation rule, and gives several interpretations of this scheme. The same tie-breaking scheme also emerges in our treatment as the unique neutral way to break ties compatible with moving phantom mechanisms. Lindner [2011, Sec. 3.4] then proposes several other aggregation rules inspired by different notions of multi-dimensional medians, some of which turn out to be strategyproof. Finally, Lindner [2011, Sec. 4] presents results of a detailed simulation study.

Also closely related is the work of Goel et al. [2019]. The primary focus of their paper is on *knapsack voting*, in which each voter submits her preferred set of projects to fully fund. However, they also consider the use of *fractional knapsack voting* in a setting in which partial funding of alternatives is permitted and voters have  $\ell_1$  preferences, which is exactly our setting. They re-discover the result of Lindner et al. [2008] that the mechanism that maximizes social welfare is incentive compatible, though they use a different tie-breaking rule (namely, lexicographic). Goel et al. do not consider other mechanisms for this setting.

More broadly, the truthful aggregation of preferences over numerical values (such as the temperature for an office) has been extensively studied. A famous result of Moulin [1980] characterizes the set of incentive-compatible voting rules under the assumption that voters have single-peaked preferences over values in  $[0, 1]$ . These voting rules are generalized median schemes. The best-known example is the standard median, in which each voter reports her ideal point in  $[0, 1]$  and the median report is selected. Other voting rules in this class insert “phantom voters” who report a fixed top choice. Barberà et al. [1993] obtained a multi-dimensional analogue of this result for  $[0, 1]^m$ , and there are further generalizations that characterize truthful rules if other constraints are imposed on the feasible set [Barberà et al. 1997]. Crucially, the constraints allowed by Barberà et al. [1997] do not include the normalization constraint that is fundamental to our setting. Several other papers [Barberà and Jackson 1994; Border and Jordan 1983; Peters et al. 1992; see Bordes et al. 2011, Section 5, for an overview] introduced multidimensional models in which one can achieve truthfulness by taking a generalized median in each coordinate, but such a strategy does not work with normalization constraints. In many of these models, all strategyproof mechanisms decompose into one-dimensional mechanisms [e.g., Border and Jordan 1983, Theorem 1]. Prior to this work, we are only aware of one multidimensional model in which truthful mechanisms that do not decompose were found: the aggregation of points in  $\mathbb{R}^d$ , where  $n = 2$  agents report an ideal point and have  $\ell_2$  preferences [Laffond 1980; Border and Jordan 1983, Example 2; Kim and Roush 1984, Example 39], though none of these mechanisms can be Pareto optimal [Kim and Roush 1984, Theorem 41].

In the computer science literature, the above-mentioned generalized median schemes have also been studied in the context of truthful facility location [Procaccia and Tennenholtz 2013]. In this context, the aim is to approximate the optimum social welfare subject to incentive compatibility.

One could apply our results to the aggregation of probabilistic beliefs. There is a large literature on *probabilistic opinion pooling* [Clemen 1989; French 1985; Genest and Zidek 1986; Intriligator 1973] which studies aggregators in this context. The main focus of that literature is to preserve stochastic and epistemic properties. To the best of our knowledge, strategic aspects have not been considered.

Finally, recently proposed rules for crowdsourcing societal tradeoffs [Conitzer et al. 2015, 2016] can also be used to aggregate budget divisions (with full support) after converting them into pairwise ratios of funding amounts, but this setting has also not been analyzed from a strategic viewpoint.

## 2 PRELIMINARIES

For  $k \in \mathbb{N}$ , we write  $[k] = \{1, \dots, k\}$ . Let  $N = [n]$  be a set of voters and  $M = [m]$  be a set of possible alternatives. Voters have structured preferences over budget *divisions*  $\mathbf{p} \in [0, 1]^m$ , with  $\sum_{j \in [m]} p_j = 1$ , where  $p_j$  is the fraction of a public resource (such as money or time) allocated to alternative  $j$ . Each voter  $i$  has a most preferred division (or *ideal point*)  $\mathbf{p}_i = (p_{i,1}, \dots, p_{i,m})$ , with their preference over other divisions induced by  $\ell_1$  distance from  $\mathbf{p}_i$ . Specifically, each voter  $i$  has a disutility for division  $\mathbf{q}$  equal to  $d(\mathbf{p}_i, \mathbf{q})$ , where  $d(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^m |x_j - y_j|$  denotes the  $\ell_1$  distance between  $\mathbf{x}$  and  $\mathbf{y}$ . Note that a voter's complete preference over all possible divisions can be deduced from their most preferred division  $\mathbf{p}_i$ .

Nehring and Puppe [2019, Sec. 3.2.2] study  $\ell_1$  preferences in a related setting, and suggest the following model where they arise naturally: Suppose voter  $i$  has an additively separable utility function  $u_{\mathbf{x}}(\mathbf{q}) = \sum_{j=1}^m \min(q_j, x_j)$ , where  $x_1, \dots, x_m \in [0, 1]$  are *funding caps* that sum to 1. According to this specification, for every alternative  $j$ , voter  $i$  derives utility linear in the amount spent on  $j$ , with no additional utility for spending over and above the cap  $x_j$ .<sup>1</sup> Note that the distribution maximizing  $u_{\mathbf{x}}$  is  $\mathbf{q} = (x_1, \dots, x_m)$ . In fact, when the funding caps are determined by voter  $i$ 's preference  $\mathbf{p}_i$ , we can check that  $i$  has  $\ell_1$  preference with ideal point  $\mathbf{p}_i$ : note that  $u_{\mathbf{p}_i}(\mathbf{q}) = 1 - \sum_{j \in \{h \in [m]: q_h > p_{i,h}\}} (p_{i,j} - q_j) = 1 - \frac{1}{2}d(\mathbf{p}_i, \mathbf{q})$ , and hence voter  $i$  prefers one distribution to another according to  $u_{\mathbf{p}_i}$  if and only if it is  $\ell_1$ -closer to  $\mathbf{p}_i$ .

A *preference profile*  $\mathbf{P} = (\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2, \dots, \hat{\mathbf{p}}_n)$  consists of a reported division  $\hat{\mathbf{p}}_i$  for each voter  $i$ . We use  $\mathbf{P}_{-i}$  to denote the reports of all voters other than  $i$ . A *budget aggregation mechanism*  $\mathcal{A}$  takes as input a preference profile  $\mathbf{P}$ , and outputs an aggregate division  $\mathcal{A}(\mathbf{P})$ . A mechanism is *continuous* if it is continuous when considered as a function  $\mathcal{A} : (\mathbb{R}^m)^n \rightarrow \mathbb{R}^m$ . We say that a mechanism is *anonymous* if its output is fixed under permutations of the voters, and *neutral* if a permutation of the alternatives in voters' inputs permutes the output in the same way.

We are interested in mechanisms that satisfy *incentive compatibility*, in the sense of dominant strategy incentive compatibility (also known as strategyproofness). Voters should not be able to change the aggregate division in their favor by misrepresenting their preference.

*Definition 2.1.* A budget aggregation mechanism  $\mathcal{A}$  satisfies incentive compatibility if, for all preference profiles  $\mathbf{P}$ , voters  $i$ , and divisions  $\mathbf{p}_i$  and  $\hat{\mathbf{p}}_i$ ,  $d(\mathcal{A}(\hat{\mathbf{p}}_i, \mathbf{P}_{-i}), \mathbf{p}_i) \geq d(\mathcal{A}(\mathbf{p}_i, \mathbf{P}_{-i}), \mathbf{p}_i)$ .

We are also interested in the basic efficiency notion of *Pareto optimality*. It should not be possible to change the aggregate so that some voter is strictly better off but no other voter is worse off.

*Definition 2.2.* A budget aggregation mechanism  $\mathcal{A}$  satisfies Pareto optimality if, for all preference profiles  $\mathbf{P}$ , and all divisions  $\mathbf{q}$ , if  $d(\mathcal{A}(\mathbf{P}), \hat{\mathbf{p}}_i) > d(\mathbf{q}, \hat{\mathbf{p}}_i)$  for some voter  $i$ , then there exists a voter  $j$  for which  $d(\mathcal{A}(\mathbf{P}), \hat{\mathbf{p}}_j) < d(\mathbf{q}, \hat{\mathbf{p}}_j)$ .

Observe that the definitions of incentive compatibility and Pareto optimality depend only on the voters' preference relations, not the exact utility model. Results pertaining to these properties therefore hold for any utility function that induces the same ordinal preferences as  $\ell_1$  utilities.

<sup>1</sup>A more expressive class of utilities would allow agents to obtain utility from different alternatives at different rates, up to the cap. Nehring and Puppe [2019] argue that this restricted model has the advantage of having low informational requirements, and that the assumption of equal rates can be justified from the principle of insufficient reason.

We also consider a fairness property that we call proportionality: Suppose each voter is single-minded, in that their ideal point is a division in which the entire budget goes to a single alternative. Then it is natural to split the resource in proportion to the number of voters supporting each alternative. For example, if 6 voters report  $(1, 0, 0)$ , 3 voters report  $(0, 1, 0)$ , and 1 voter reports  $(0, 0, 1)$ , then the aggregate should be  $(0.6, 0.3, 0.1)$ . We call this property *proportionality*.

*Definition 2.3.* A voter is *single-minded* if their preferred division is a unit vector. A budget aggregation mechanism  $\mathcal{A}$  is *proportional* if, for every preference profile  $\mathbf{P}$  consisting of only single-minded voters, and every alternative  $j$ , we have  $\mathcal{A}(\mathbf{P})_j = \sum_{i \in [m]} \hat{p}_{i,j} / n$

We note that proportionality is a fairly weak axiom, only applying to a small subset of possible profiles. However, as we will see later, it is already strong enough to be incompatible with Pareto optimality within the class of moving phantom mechanisms that we introduce in this paper.

Finally, we introduce two properties from social choice theory that make sense for mechanisms that are defined for instances with a variable number of voters. One of these is participation, which says that voters should always prefer (truthfully) participating to not contributing a report.

*Definition 2.4.* A budget aggregation mechanism  $\mathcal{A}$  satisfies *participation* if, for all  $\mathbf{p}_i$  and all preference profiles  $\mathbf{P}_{-i}$ ,  $d(\mathcal{A}(\mathbf{p}_i, \mathbf{P}_{-i}), \mathbf{p}_i) \leq d(\mathcal{A}(\mathbf{P}_{-i}), \mathbf{p}_i)$ .

The second property is reinforcement, which states that if we have two preference profiles which yield the same aggregate division, then the profile obtained by combining these profiles should yield that aggregate division as well. For  $\mathbf{P} = (\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2, \dots, \hat{\mathbf{p}}_{n_P})$  and  $\mathbf{Q} = (\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2, \dots, \hat{\mathbf{q}}_{n_Q})$ , let  $\mathbf{P} \cup \mathbf{Q} = (\hat{\mathbf{p}}_1, \dots, \hat{\mathbf{p}}_{n_P}, \hat{\mathbf{q}}_1, \dots, \hat{\mathbf{q}}_{n_Q})$ .

*Definition 2.5.* A budget aggregation mechanism  $\mathcal{A}$  satisfies reinforcement if, for all preference profiles  $\mathbf{P}$  and  $\mathbf{Q}$  with  $\mathcal{A}(\mathbf{P}) = \mathcal{A}(\mathbf{Q})$ , it holds that  $\mathcal{A}(\mathbf{P} \cup \mathbf{Q}) = \mathcal{A}(\mathbf{P}) = \mathcal{A}(\mathbf{Q})$ .

### 3 TWO ALTERNATIVES

To build intuition, we begin by considering the case in which  $m = 2$ . Due to the normalization of inputs and of the output, and with  $\ell_1$  preferences, the problem is one-dimensional: a budget division is completely determined by the fraction spent on the first of the two alternatives. This allows us to directly import Moulin's [1980] famous characterization of *generalized median* rules as the only incentive compatible mechanisms for voters with single-peaked preferences over a single-dimensional quantity.<sup>2</sup>

**THEOREM 3.1 (MOULIN 1980).** *For  $m = 2$ , an anonymous and continuous budget aggregation mechanism  $\mathcal{A}$  is incentive compatible if and only if there are  $\alpha_0 \geq \alpha_1 \geq \dots \geq \alpha_n$  in  $[0, 1]$  such that, for all profiles  $\mathbf{P}$ ,*

$$\begin{aligned}\mathcal{A}(\mathbf{P})_1 &= \text{med}(p_{1,1}, p_{2,1}, \dots, p_{n,1}, \alpha_0, \alpha_1, \dots, \alpha_n), \\ \mathcal{A}(\mathbf{P})_2 &= \text{med}(p_{1,2}, p_{2,2}, \dots, p_{n,2}, 1 - \alpha_0, 1 - \alpha_1, \dots, 1 - \alpha_n).\end{aligned}$$

The numbers  $\alpha_k$  are known as *phantoms*. Each mechanism described by Theorem 3.1 can be understood as taking the coordinate-wise median of the reported distributions, after inserting  $n + 1$  phantom voters (whose report is fixed and independent of the input profile).

One can check that  $\alpha_0, \dots, \alpha_n$  define a neutral mechanism if and only if the phantom placements are symmetric, that is if and only if  $(\alpha_0, \dots, \alpha_n) = (1 - \alpha_n, \dots, 1 - \alpha_0)$ . Note that there are  $n + 1$  phantoms but only  $n$  voters, so that the phantoms can outweigh the voters. For example, when

<sup>2</sup>Our preference model using  $\ell_1$  imposes slightly more structure than just single-peakedness, namely that voters are indifferent between points that are equidistant to their peak. However, this restriction does not enlarge the class of incentive compatible mechanisms, at least if we impose continuity [Massó and de Barreda 2011].

$\alpha_k = 1/2$  for all  $k \in \{0, \dots, n\}$  then the mechanism is just the constant mechanism returning  $(1/2, 1/2)$ . However, if we take  $\alpha_0 = 1$  and  $\alpha_n = 0$ , then these two phantoms “cancel out” and there are only  $n - 1$  phantoms left. In fact, one can check that the mechanism is Pareto optimal if and only if  $\alpha_0 = 1$  and  $\alpha_n = 0$  [Moulin 1980].

A particularly interesting example is the *uniform phantom mechanism*, obtained when placing the phantoms uniformly over the interval  $[0, 1]$ , so that  $\alpha_k = 1 - k/n$  for each  $k \in \{0, \dots, n\}$ . This placement of phantom voters appears in a paper by Caragiannis et al. [2016]. They were aiming for mechanisms whose outcome is close to the mean, and they prove that the uniform phantom mechanism yields an aggregate that is closer to the mean than that obtained from any other phantom placements, in the worst case over all profiles. The uniform phantom mechanism has other attractive properties, including being proportional in the sense of Definition 2.3.

**PROPOSITION 3.2.** *For  $m = 2$ , the uniform phantom mechanism is the unique (anonymous and continuous) budget aggregation mechanism  $\mathcal{A}$  that is both incentive compatible and proportional.*

**PROOF.** Theorem 3.1 gives us that  $\mathcal{A}$  is incentive compatible if and only if it can be written in terms of phantom medians. We therefore need only to consider the additional requirement of proportionality. The uniform phantom mechanism is proportional, because if  $\mathbf{P}$  consists of  $n - k$  voters reporting  $(1, 0)$  and  $k$  voters reporting  $(0, 1)$ , then  $\mathbf{A}(\mathbf{P})_1 = \alpha_k = (n - k)/n$ , as required.

For uniqueness, suppose  $\alpha_0, \dots, \alpha_n$  are phantom positions that induce a proportional mechanism. Let  $k \in \{0, \dots, n\}$ . We show that  $\alpha_k = 1 - k/n$ . Let  $\mathbf{P}$  be a profile consisting of only single-minded voters with  $n_1 = n - k$  voters reporting  $\hat{\mathbf{p}}_i = (1, 0)$ . Then  $\alpha_k$  is the median, and proportionality requires that  $\alpha_k = n_1/n = (n - k)/n = 1 - k/n$ .  $\square$

Another natural way to place the phantoms is one that takes the coordinate-wise median. When  $n + 1$  is even, this is achieved by placing half the phantoms at 0 and the other half at 1, outputting precisely the median of the reported values on each coordinate. When  $n + 1$  is odd, we place  $n/2$  phantoms at 0,  $n/2$  phantoms at 1, and we place a single phantom at  $1/2$  to preserve neutrality. This mechanism outputs the point between the left and right medians that is closest to  $1/2$ . The resulting mechanism returns an aggregate  $\mathbf{p}$  that minimizes the sum of distances between the reports  $(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n)$  and  $\mathbf{p}$ . We will generalize this mechanism to larger  $m$  in Section 6.

#### 4 A CLASS OF INCENTIVE COMPATIBLE MECHANISMS FOR HIGHER DIMENSIONS

For  $m = 2$ , we have a complete picture of incentive-compatible mechanisms, thanks to Moulin’s characterization. For  $m \geq 3$ , it is less clear how to construct examples of incentive-compatible mechanisms. One could try to take a generalized median in each alternative independently, but the result of such a mechanism would not respect the normalization constraint.

*Example 4.1.* Let us consider a simple numerical example. Let  $n = m = 3$ , and suppose voter reports are  $\mathbf{p}_1 = (0, 0.5, 0.5)$ ,  $\mathbf{p}_2 = (\frac{1}{3}, \frac{2}{3}, 0)$ , and  $\mathbf{p}_3 = (0.9, 0, 0.1)$ . Using the uniform phantom mechanism to aggregate the votes on each alternative independently yields  $(\frac{1}{3}, 0.5, \frac{1}{3})$ . Note that these sum to more than 1 (indeed, as we will see later, this is true in general for the uniform phantom mechanism on more than two alternatives).

A simple modification is to normalize the output of the uniform phantom mechanism by the appropriate multiplicative constant. In this example, we need to divide by  $\frac{7}{6}$ , which yields  $\mathbf{q} = (\frac{2}{7}, \frac{3}{7}, \frac{2}{7})$ . Unfortunately, this mechanism is not incentive compatible. To see this, suppose that the second voter instead reports  $\hat{\mathbf{p}}_2 = (\frac{3}{8}, \frac{5}{8}, 0)$ . The normalized uniform phantom mechanism now selects the distribution  $\mathbf{q}' = (\frac{9}{29}, \frac{12}{29}, \frac{8}{29})$ . But  $d(\mathbf{q}, \mathbf{p}_2) = \frac{4}{7}$  and  $d(\mathbf{q}', \mathbf{p}_2) = \frac{16}{29} < \frac{4}{7}$ , so voter 2 obtained a preferred outcome by manipulating.

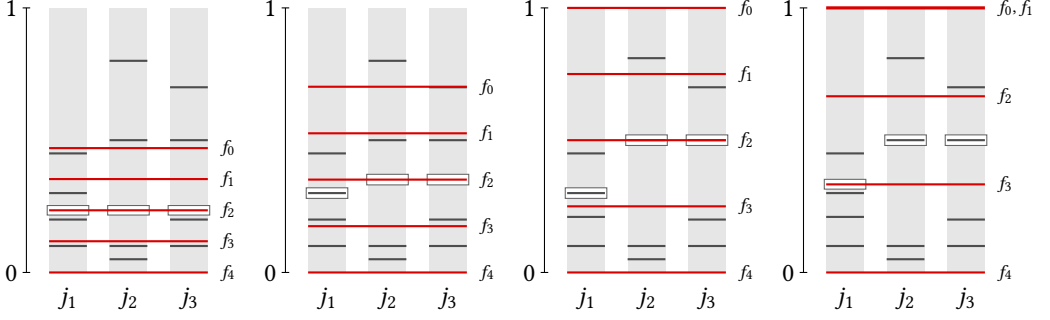


Fig. 1. A moving phantom mechanism operating on an instance with  $n = 4$  and  $m = 3$ .

However, there is a way of extending the idea of generalized medians to the higher-dimensional setting. The basic idea is that if a coordinate-wise generalized median violates the normalization constraint, then we can adjust the placement of the phantoms, increasing or decreasing the sum of the generalized medians as needed. Such a procedure might, in principle, give voters incentives to manipulate in order to affect the phantom placements. However, our class of *moving phantom mechanisms* manages to avoid this problem.

*Definition 4.2.* Let  $\mathcal{F} = \{f_k : k \in \{0, \dots, n\}\}$  be a family of functions, which we call a *phantom system*, where  $f_k : [0, 1] \rightarrow [0, 1]$  is a continuous, weakly increasing function with  $f_k(0) = 0$  and  $f_k(1) = 1$  for each  $k$ , and we have  $f_0(t) \geq f_1(t) \geq \dots \geq f_n(t)$  for all  $t \in [0, 1]$ . Then, the *moving phantom mechanism*  $\mathcal{A}^{\mathcal{F}}$  is defined so that for all profiles  $\mathbf{P}$  and all  $j \in [m]$ ,

$$\mathcal{A}^{\mathcal{F}}(\mathbf{P})_j = \text{med}(f_0(t^*), \dots, f_n(t^*), \hat{p}_{1,j}, \dots, \hat{p}_{n,j}), \quad (1)$$

where  $t^*$  is chosen so that  $t^* \in \{t : \sum_{j \in [m]} \text{med}(f_0(t), \dots, f_n(t), \hat{p}_{1,j}, \dots, \hat{p}_{n,j}) = 1\}$ . For brevity, we write  $\mathcal{F}(t) = (f_0(t), \dots, f_n(t))$  and abbreviate the median in (1) to  $\text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], j})$ .

Let us examine the definition. Each  $f_k$  represents a phantom, and the phantom system  $\mathcal{F}$  represents a “movie” in which all phantoms continuously increase from 0 to 1, with the function argument  $t$  defining an instantaneous snapshot of the phantom positions. The moving phantom mechanism  $\mathcal{A}^{\mathcal{F}}$  defined by  $\mathcal{F}$  identifies a particular snapshot in time,  $t^*$ , for which the sum of generalized medians over all coordinates is exactly 1. One can check that at least one such  $t^*$  exists, and that the output of the mechanism is independent of which of these  $t^*$  is chosen.

*PROPOSITION 4.3.* *The moving phantom mechanism  $\mathcal{A}^{\mathcal{F}}$  is well-defined for every phantom system  $\mathcal{F}$  satisfying the conditions of Definition 4.2.*

*PROOF.* First note that the function  $t \mapsto \sum_{j \in [m]} \text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], j})$  is continuous and increasing in  $t$ , because  $f_k$  is continuous and increasing, and these properties are preserved under taking the median and sum. This implies that, provided the set  $\{t : \sum_{j \in [m]} \text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], j}) = 1\}$  is non-empty, the aggregate  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})$  does not depend on the choice of  $t^*$ .

When  $t = 0$ ,  $\sum_{j \in [m]} \text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], j}) = 0$ , since all  $n + 1$  phantom entries are 0. When  $t = 1$ ,  $\sum_{j \in [m]} \text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], j}) = m > 1$ , since all  $n + 1$  phantom entries are 1. By the Intermediate Value Theorem, using continuity, there exists  $t \in [0, 1]$  with  $\sum_{j \in [m]} \text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], j}) = 1$ .  $\square$

To build intuition, we consider an example moving phantom mechanism, shown in Figure 1. There are three alternatives, each occupying a column on the horizontal axis, and four voters. Voter reports are indicated by gray horizontal line segments, with their magnitude  $\hat{p}_{i,j}$  indicated by their



vertical position. The phantom placements are indicated by the red lines and labeled  $f_0, \dots, f_4$ . For each alternative, the median of the four agent reports and the five phantoms is indicated by a rectangle.

The four snapshots shown in Figure 1 display increasing values of  $t$ . Observe that the position of each phantom (weakly) increases from left to right, as does the median on each alternative. Although the vertical axis is not labeled, for simplicity of presentation, normalization here occurs in the second image from the left. In the leftmost image, the sum of the highlighted entries is less than 1, while in the two rightmost images it is more than 1.

For simplicity, the definition of moving phantom mechanisms treats the number of voters  $n$  as fixed. To allow  $n$  to vary, it is necessary to define a family of phantom systems, one for each  $n$ . In the next two sections, we give two examples of such families, but for this section we keep the presentation simple by considering only a fixed  $n$ . On the other hand, for each fixed  $n$ , a phantom system defines a mechanism that works for any number  $m$  of projects.

Moving phantom mechanisms satisfy some important basic properties. They are all anonymous, neutral, and continuous. Here neutrality is a design choice—one could imagine defining moving phantom mechanisms for which the movement of the phantoms depends on the alternative.

We now show our main result in this section, that every moving phantom mechanism is incentive compatible. Before proving the result formally, we provide some intuition. If  $i$  changes her report from  $\mathbf{p}_i$  to  $\hat{\mathbf{p}}_i$ , the effect on the aggregate can be decomposed into two parts. First, we can think of holding the phantoms fixed at the snapshot dictated by the truthful instance, while changing  $i$ 's report to  $\hat{\mathbf{p}}_i$ . Second, we can think of repositioning the phantoms to the snapshot required to guarantee normalization of the aggregate vector after  $i$  reports  $\hat{\mathbf{p}}_i$ . To prove incentive compatibility, we show that any change that the aggregate division undergoes in the first stage can only be bad for voter  $i$ , pushing the aggregate away from  $\mathbf{p}_i$ . Change in the second stage can push the aggregate towards  $\mathbf{p}_i$ , helping voter  $i$ , but the magnitude of this change is upper bounded by the magnitude of the harmful change in the first stage.

**THEOREM 4.4.** *Every moving phantom mechanism is incentive compatible.*

**PROOF.** Let  $\mathcal{F}$  define a moving phantom mechanism  $\mathcal{A}^{\mathcal{F}}$ . Consider some report  $\hat{\mathbf{p}}_i \neq \mathbf{p}_i$ , and fix the reports of all other voters  $\mathbf{P}_{-i}$ . Let  $t^*$  determine the phantom placement for reports  $(\mathbf{p}_i, \mathbf{P}_{-i})$  and  $\hat{t}^*$  for reports  $(\hat{\mathbf{p}}_i, \mathbf{P}_{-i})$ .

Consider the effect of  $i$ 's misreport from  $\mathbf{p}_i$  to  $\hat{\mathbf{p}}_i$  while *holding the phantoms fixed at  $\mathcal{F}(t^*)$* . Then, because phantom placements are fixed on each alternative, any change that voter  $i$  can cause on alternative  $j$  by misreporting must be away from her ideal value  $p_{i,j}$ . For each  $j \in [m]$ , we have

$$\begin{aligned} \text{med}(\mathcal{F}(t^*), \hat{\mathbf{p}}_{i,j}, \mathbf{P}_{-i,j}) &\geq \text{med}(\mathcal{F}(t^*), p_{i,j}, \mathbf{P}_{-i,j}) && \text{if } p_{i,j} \leq \mathcal{A}^{\mathcal{F}}(\mathbf{p}_i, \mathbf{P}_{-i})_j < \hat{p}_{i,j}, \\ \text{med}(\mathcal{F}(t^*), \hat{\mathbf{p}}_{i,j}, \mathbf{P}_{-i,j}) &\leq \text{med}(\mathcal{F}(t^*), p_{i,j}, \mathbf{P}_{-i,j}) && \text{if } \hat{p}_{i,j} \leq \mathcal{A}^{\mathcal{F}}(\mathbf{p}_i, \mathbf{P}_{-i})_j < p_{i,j}, \\ \text{med}(\mathcal{F}(t^*), \hat{\mathbf{p}}_{i,j}, \mathbf{P}_{-i,j}) &= \text{med}(\mathcal{F}(t^*), p_{i,j}, \mathbf{P}_{-i,j}) && \text{otherwise.} \end{aligned}$$

Let  $y_j = \text{med}(\mathcal{F}(t^*), \hat{\mathbf{p}}_{i,j}, \mathbf{P}_{-i,j}) - \text{med}(\mathcal{F}(t^*), p_{i,j}, \mathbf{P}_{-i,j})$  denote the change caused on alternative  $j$  by voter  $i$ 's misreport, subject to holding the phantom placement fixed at  $\mathcal{F}(t^*)$ . By the above,

$$\sum_{j \in [m]} |p_{i,j} - \text{med}(\mathcal{F}(t^*), \hat{\mathbf{p}}_{i,j}, \mathbf{P}_{-i,j})| = \sum_{j \in [m]} |p_{i,j} - \mathcal{A}^{\mathcal{F}}(\mathbf{p}_i, \mathbf{P}_{-i})_j| + \sum_{j \in [m]} |y_j|. \quad (2)$$

Next, we consider the change that results from moving the phantoms from  $\mathcal{F}(t^*)$  to  $\mathcal{F}(\hat{t}^*)$ . Assume that  $\sum_{j \in [m]} y_j \geq 0$  (otherwise, a similar argument applies). Then  $\sum_{j \in [m]} \text{med}(\mathcal{F}(t^*), \hat{\mathbf{p}}_{i,j}, \mathbf{P}_{-i,j}) \geq 1$ , which implies that  $\hat{t}^* \leq t^*$  since the sum is monotonic in  $t$  (see the proof of Proposition 4.3).

Thus, for the aggregate division  $\mathcal{A}^{\mathcal{F}}(\hat{\mathbf{p}}_i, \mathbf{P}_{-i})$  we have that

$$\mathcal{A}^{\mathcal{F}}(\hat{\mathbf{p}}_i, \mathbf{P}_{-i})_j = \text{med}(\mathcal{F}(t^*), \hat{p}_{i,j}, \mathbf{P}_{-i,j}) \leq \text{med}(\mathcal{F}(t^*), \hat{p}_{i,j}, \mathbf{P}_{-i,j}) \quad \text{for all } j \in [m].$$

Therefore,

$$\begin{aligned} \sum_{j \in [m]} |\text{med}(\mathcal{F}(t^*), \hat{p}_{i,j}, \mathbf{P}_{-i,j}) - \mathcal{A}^{\mathcal{F}}(\hat{\mathbf{p}}_i, \mathbf{P}_{-i})_j| &= \sum_{j \in [m]} (\text{med}(\mathcal{F}(t^*), \hat{p}_{i,j}, \mathbf{P}_{-i,j}) - \mathcal{A}^{\mathcal{F}}(\hat{\mathbf{p}}_i, \mathbf{P}_{-i})_j) \\ &= \sum_{j \in [m]} (\text{med}(\mathcal{F}(t^*), \hat{p}_{i,j}, \mathbf{P}_{-i,j})) - \sum_{j \in [m]} \mathcal{A}^{\mathcal{F}}(\hat{\mathbf{p}}_i, \mathbf{P}_{-i})_j \\ &= \sum_{j \in [m]} (\text{med}(\mathcal{F}(t^*), \hat{p}_{i,j}, \mathbf{P}_{-i,j})) - 1 \\ &= \sum_{j \in [m]} (\text{med}(\mathcal{F}(t^*), \hat{p}_{i,j}, \mathbf{P}_{-i,j})) - \sum_{j \in [m]} \mathcal{A}^{\mathcal{F}}(\mathbf{p}_i, \mathbf{P}_{-i})_j \\ &= \sum_{j \in [m]} (\text{med}(\mathcal{F}(t^*), \hat{p}_{i,j}, \mathbf{P}_{-i,j}) - \mathcal{A}^{\mathcal{F}}(\mathbf{p}_i, \mathbf{P}_{-i})_j) \\ &= \sum_{j \in [m]} y_j \end{aligned} \tag{3}$$

We complete the proof by noting that

$$\begin{aligned} d(\mathbf{p}_i, \mathcal{A}^{\mathcal{F}}(\hat{\mathbf{p}}_i, \mathbf{P}_{-i})) &= \sum_{j \in [m]} |p_{i,j} - \mathcal{A}^{\mathcal{F}}(\hat{\mathbf{p}}_i, \mathbf{P}_{-i})_j| \\ &\geq \sum_{j \in [m]} |p_{i,j} - \text{med}(\mathcal{F}(t^*), \hat{p}_{i,j}, \mathbf{P}_{-i,j})| \\ &\quad - \sum_{j \in [m]} |\mathcal{A}^{\mathcal{F}}(\hat{\mathbf{p}}_i, \mathbf{P}_{-i})_j - \text{med}(\mathcal{F}(t^*), \hat{p}_{i,j}, \mathbf{P}_{-i,j})| \\ &= \sum_{j \in [m]} |p_{i,j} - \mathcal{A}^{\mathcal{F}}(\mathbf{p}_i, \mathbf{P}_{-i})_j| + \sum_{j \in [m]} |y_j| - \sum_{j \in [m]} y_j \\ &\geq \sum_{j \in [m]} |p_{i,j} - \mathcal{A}^{\mathcal{F}}(\mathbf{p}_i, \mathbf{P}_{-i})_j| = d(\mathbf{p}_i, \mathcal{A}^{\mathcal{F}}(\mathbf{p}_i, \mathbf{P}_{-i})), \end{aligned}$$

where the first and last equalities follow from the definition of the  $\ell_1$  metric, the first inequality holds by an application of the triangle inequality, the second equality follows from Equations 2 and 3, and the final inequality holds because  $|x| \geq x$  for all  $x \in \mathbb{R}$ .  $\square$

In addition to incentive compatibility, moving phantom mechanisms satisfy a natural monotonicity property that says that if some voter increases her report on alternative  $j$ , and decreases her report on all other alternatives, then the aggregate weight on alternative  $j$  should not decrease.

*Definition 4.5.* A budget aggregation mechanism  $\mathcal{A}$  satisfies *monotonicity* if, for all  $\mathbf{p}_i, \mathbf{p}'_i$  with  $p_{i,j} > p'_{i,j}$  for some  $j$  and  $p_{i,k} \leq p'_{i,k}$  for all  $k \neq j$ ,

$$\mathcal{A}(\mathbf{p}_i, \mathbf{P}_{-i})_j \geq \mathcal{A}(\mathbf{p}'_i, \mathbf{P}_{-i})_j.$$

**THEOREM 4.6.** *Every moving phantom mechanism satisfies monotonicity.*

**PROOF.** Let  $\mathbf{p}_i, \mathbf{p}'_i$  be such that  $p_{i,j} > p'_{i,j}$  for some  $j$  and  $p_{i,k} \leq p'_{i,k}$  for all  $k \neq j$ . Let  $t^*$  determine the phantom placement for reports  $(\mathbf{p}_i, \mathbf{P}_{-i})$  and  $t^{**}$  for reports  $(\mathbf{p}'_i, \mathbf{P}_{-i})$ .

Suppose that  $t'^* < t^*$ . We have

$$\mathcal{A}(\mathbf{p}'_i, \mathbf{P}_{-i})_j = \text{med}(\mathcal{F}(t'^*), p'_{i,j}, P_{-i,j}) \leq \text{med}(\mathcal{F}(t^*), p_{i,j}, P_{-i,j}) = \mathcal{A}(\mathbf{p}_i, \mathbf{P}_{-i})_j$$

where the inequality holds because  $p_{i,j} > p'_{i,j}$  and  $f_k(t^*) \geq f_k(t'^*)$  for all  $k \in \{0, \dots, n\}$ .

Next, suppose that  $t'^* > t^*$ . Then

$$\begin{aligned} \mathcal{A}(\mathbf{p}'_i, \mathbf{P}_{-i})_j &= 1 - \sum_{j' \neq j} \mathcal{A}(\mathbf{p}'_i, \mathbf{P}_{-i})_{j'} = 1 - \sum_{j' \neq j} \text{med}(\mathcal{F}(t'^*), p'_{i,j'}, P_{-i,j'}) \\ &\leq 1 - \sum_{j' \neq j} \text{med}(\mathcal{F}(t^*), p_{i,j'}, P_{-i,j'}) = \mathcal{A}(\mathbf{p}_i, \mathbf{P}_{-i})_j \end{aligned}$$

where the inequality holds as  $p_{i,j'} < p'_{i,j'}$  for all  $j' \neq j$  and  $f_k(t^*) \leq f_k(t'^*)$  for all  $k \in \{0, \dots, n\}$ .  $\square$

Let us briefly discuss how to compute the output of a moving phantom mechanism. In principle, the precise time  $t^* \in [0, 1]$  at which the output of the mechanism is normalized may have many decimal digits, and for badly-behaved  $\mathcal{F}$  it may even be irrational. For the same reason, the mechanism may produce an irrational division, so the precise computation of the output may not be possible. However, for mechanisms for which the phantom system  $\mathcal{F}$  is piecewise linear, we can show that  $t^*$  has few digits and the output is always rational, so polynomial-time computation is possible using a binary search on  $t$ . A phantom system is *piecewise linear* if we can partition time into  $r$  time intervals, such that within each interval, all the phantoms move at constant speed. All the specific moving phantom mechanisms that we consider in the remainder of the paper have a piecewise linear phantom system, and can thus be efficiently evaluated.

**THEOREM 4.7.** *Suppose that  $\mathcal{F}$  is a piecewise linear phantom system, so that there are rationals  $0 = t_1 < t_2 < \dots < t_r = 1$  such that for each  $k = 0, \dots, n$  and each  $s \in [r - 1]$ , there are rational coefficients  $a_k^s \in \mathbb{Q}$  and  $b_k^s \in \mathbb{Q}$  with*

$$f_k(t) = a_k^s \cdot t + b_k^s \quad \text{whenever } t_s \leq t \leq t_{s+1}.$$

*Let  $\mathbf{P}$  be a profile in which all reports  $p_{i,j}$  are rational. Then there exists a rational time  $t^* \in \mathbb{Q}$  for which  $\mathcal{A}^{\mathcal{F}}$  is normalized, and  $t^*$  can be specified in polynomially many bits. Hence, the outcome distribution of  $\mathcal{A}^{\mathcal{F}}$  is rational and can be computed in polynomial time.*

**PROOF.** We show that  $t^*$  and the outcome distribution  $p_1, \dots, p_m$  form the solution of a system of linear inequalities with rational coefficients. Since a system of linear inequalities with rational coefficients that has a solution always has a rational solution of polynomial size [Schrijver 1986, Thm. 10.1], the first claim of the theorem follows. For the second claim about computability, note that given a (rational) time  $t$ , we can clearly compute the outcome of  $\mathcal{A}^{\mathcal{F}}$  at time  $t$ , and we can use a binary search on  $t$  to find the time  $t^*$  when  $\mathcal{A}^{\mathcal{F}}$  is normalized. The binary search takes polynomially many steps since  $t^*$  is guaranteed to have polynomial size. (In this analysis of time complexity, the number of voters  $n$  is fixed, but the number  $m$  of alternatives is allowed to vary.)

We now construct the mentioned system of linear inequalities. Let  $t^*$  be a time when the mechanism is normalized, and choose  $s \in [r - 1]$  such that  $t_s \leq t^* \leq t_{s+1}$ . Our system of linear inequalities will have variables  $t, p_1, \dots, p_m \geq 0$ . We add inequalities (also called constraints) to the system one by one. First, we add the constraints  $\sum_{j \in [m]} p_j = 1$  and  $t_s \leq t \leq t_{s+1}$ . Next, we will add constraints that, for each alternative  $j \in [m]$ , encode  $p_j$  as the value that is the median for  $j$ .

Let  $j \in [m]$ . Suppose first that  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_j$  is a voter report, i.e.  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_j = p_{i,j}$  for some  $i \in N$ . Then we add the constraint  $p_j = p_{i,j}$ . Then, for each phantom  $k = 0, \dots, n$ , we add the constraint

$$\begin{aligned} a_k^s \cdot t + b_k^s &\leq p_j && \text{if } f_k(t^*) \leq \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j, \\ a_k^s \cdot t + b_k^s &\geq p_j && \text{if } f_k(t^*) \geq \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j, \end{aligned}$$

which implies that  $p_j$  is the median of  $p_{1,j}, \dots, p_{n,j}, f_0(t), \dots, f_n(t)$ .

Otherwise, we have that  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_j$  is not a voter report, so  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_j = f_{k'}(t^*)$  for some  $k' \in \{0, \dots, n\}$ . Then again for each phantom  $k = 0, \dots, n$ , we add the constraint

$$\begin{aligned} a_k^s \cdot t + b_k^s &\leq p_j && \text{if } f_k(t^*) \leq \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j, \\ a_k^s \cdot t + b_k^s &\geq p_j && \text{if } f_k(t^*) \geq \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j, \end{aligned}$$

which implies that  $p_j = a_{k'}^s \cdot t + b_{k'}^s$ , and that  $p_j$  is the median of  $p_{1,j}, \dots, p_{n,j}, f_0(t), \dots, f_n(t)$ .

Clearly, any solution to this system of constraints encodes an outcome of  $\mathcal{A}^{\mathcal{F}}$ , as required.  $\square$

Before we move on to particular moving phantom mechanisms, let us end this section with a tantalizing open question: Does there exist an (anonymous, neutral, continuous) incentive-compatible budget aggregation mechanism that is *not* a moving phantom mechanism? We have not been able to construct any example, and have found that some mechanisms that on first sight seem to have nothing to do with medians end up having an equivalent description as a moving phantom mechanism. For the simpler two-alternative case, we already have a characterization of all incentive-compatible mechanisms (Theorem 3.1). This class can equivalently be described in terms of moving phantoms, and so the answer to our question for  $m = 2$  is *no*.

**THEOREM 4.8.** *For  $m = 2$ , moving phantom mechanisms are the only budget aggregation mechanisms that satisfy anonymity, neutrality, continuity, and incentive compatibility.*

**PROOF.** Certainly all moving phantom mechanisms satisfy these properties. For the other direction, we know from Theorem 3.1 that any mechanism  $\mathcal{A}$  satisfying these properties can be described as a generalized median with phantoms  $\alpha_0, \dots, \alpha_n$  satisfying, due to neutrality,  $\{\alpha_0, \dots, \alpha_n\} = \{1 - \alpha_0, \dots, 1 - \alpha_n\}$ . We show that  $\mathcal{A}$  is equivalent to a moving phantom mechanism. Define  $\mathcal{A}^{\mathcal{F}}$  using a phantom system  $\mathcal{F}$  for which there exists a  $t^* \in [0, 1]$  with  $f_k(t^*) = \alpha_k$  for every  $k \in \{0, \dots, n\}$ . Then, for every preference profile  $\mathbf{P}$ , we have that  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_1 = \text{med}(\mathcal{F}(t^*), \mathbf{P}_{i \in [n], j}) = (\alpha_0, \dots, \alpha_n, \mathbf{P}_{i \in [n], j})$ , and  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_2 = 1 - \mathcal{A}^{\mathcal{F}}(\mathbf{P})_1$ , matching the output of  $\mathcal{A}$ .  $\square$

## 5 THE INDEPENDENT MARKETS MECHANISM

We have seen that uniform phantoms is uniquely proportional for  $m = 2$ . By a similar argument to the proof of Proposition 3.2, it is easy to see that any family of functions  $\mathcal{F}$  that generates uniform phantoms at some snapshot will be proportional, and will reduce to the uniform phantom mechanism for  $m = 2$ . However, this leaves a large class of moving phantom mechanisms to choose from. In this section, we identify a particular moving phantom mechanism that generalizes the uniform phantom mechanism for arbitrary  $m$ . Its outcome can be interpreted as a market equilibrium, which will be useful for analyzing the mechanism.

*Definition 5.1.* The independent markets mechanism ( $\mathcal{A}^{IM}$ ) is the moving phantom mechanism defined by the phantom system  $f_k(t) = \min\{t(n - k), 1\}$  for each  $k \in \{0, \dots, n\}$ .

To visualize the phantom placement, observe that for any  $t \leq 1/n$ , phantoms are being placed at  $0, t, 2t, \dots, nt$ . Once  $t$  reaches  $1/n$ , phantoms continue to grow in the same manner, but the higher

phantoms get capped at 1.<sup>3</sup> This is actually the mechanism that we displayed in Figure 1. Note that, when  $t = 1/n$ , the phantom placement is uniform on  $[0, 1]$  (as is the case in the third panel of Figure 1); thus,  $\mathcal{A}^{IM}$  reduces to the uniform phantom mechanism for  $m = 2$ .

*Example 5.2.* Consider the profile  $\mathbf{P}$  from Example 4.1 with voter reports  $\mathbf{p}_1 = (0, 0.5, 0.5)$ ,  $\mathbf{p}_2 = (\frac{1}{3}, \frac{2}{3}, 0)$ , and  $\mathbf{p}_3 = (0.9, 0, 0.1)$ . Consider the placement of the  $n + 1 = 4$  phantoms when  $t = \frac{2}{9}$ . They are placed at  $f_0(t) = \frac{6}{9}$ ,  $f_1(t) = \frac{4}{9}$ ,  $f_2(t) = \frac{2}{9}$ ,  $f_3(t) = 0$ . The generalized medians are

$$\begin{aligned} \text{med}\{f_0(t), f_1(t), f_2(t), f_3(t), p_{1,1}, p_{2,1}, p_{3,1}\} &= \text{med}\{\frac{6}{9}, \frac{4}{9}, \frac{2}{9}, 0, 0, \frac{1}{3}, 0.9\} = \frac{1}{3}, \\ \text{med}\{f_0(t), f_1(t), f_2(t), f_3(t), p_{1,2}, p_{2,2}, p_{3,2}\} &= \text{med}\{\frac{6}{9}, \frac{4}{9}, \frac{2}{9}, 0, 0.5, \frac{2}{3}, 0\} = \frac{4}{9}, \\ \text{med}\{f_0(t), f_1(t), f_2(t), f_3(t), p_{1,3}, p_{2,3}, p_{3,3}\} &= \text{med}\{\frac{6}{9}, \frac{4}{9}, \frac{2}{9}, 0, 0.5, 0, 0.1\} = \frac{2}{9}. \end{aligned}$$

Because  $\frac{1}{3} + \frac{4}{9} + \frac{2}{9} = 1$ , these are normalized, and hence we have  $t^* = \frac{2}{9}$ . Thus  $\mathcal{A}^{IM}(\mathbf{P}) = (\frac{1}{3}, \frac{4}{9}, \frac{2}{9})$ .

### 5.1 Market Interpretation

Why do we call this mechanism the independent markets mechanism? To explain this, we first establish a connection between the market clearing price in a simple single-good market and the median of some familiar-looking numbers.

Suppose we are selling a single divisible good, of which a total amount of  $x \in [0, \infty)$  is available. Each of  $n$  voters has a budget of 1, and a value  $v_i \in [0, \infty)$  per unit of the good. We assume voters have quasilinear utilities  $u_i(x_i) = x_i \cdot (v_i - \pi)$ , where  $\pi \geq 0$  is the price per unit of the good. For a price  $\pi$ , the demand of voter  $i$ ,  $D_i(\pi) \in [0, \infty]$  must satisfy

$$\begin{aligned} D_i(0) &= \infty \\ D_i(\pi) &= 1/\pi \quad \text{for all } 0 < \pi < v_i \\ D_i(v_i) &\in [0, 1/v_i] \\ D_i(\pi) &= 0 \quad \text{for all } \pi > v_i \end{aligned}$$

Thus, each voter demands as much of the good as their budget of 1 allows at price  $\pi$ , as long as the price per unit is lower than their value per unit. When the price is equal to the voter's value, we allow the demand to be anywhere between 0 and  $\frac{1}{\pi}$ , since the voter is indifferent between purchasing any amount of the good at this price, subject to her budget. The *market clearing price*  $c$  is the price at which the supply of the good ( $x$ ) equals the total demand. Formally, we say that  $c$  is a market clearing price if there exist demands  $D_i(c)$  such that

$$\sum_{i \in [n]} D_i(c) = x. \tag{4}$$

It turns out that a unique market clearing price  $c$  always exists and is equal to the median of the  $n$  voter values  $v_i$  and of  $n + 1$  "phantom values" which are uniformly distributed on the interval  $[0, n/x]$ . To the best of our knowledge, this connection has not previously been appreciated in the literature.

**LEMMA 5.3.** *In the market defined above, a unique market clearing price  $c$  always exists and equals*

$$\text{med}(0, 1/x, \dots, (n-1)/x, n/x, v_1, \dots, v_n).$$

**PROOF.** We distinguish the cases that the median is a phantom  $a/x$  or a voter report  $v_i$ .

Suppose that the median is  $a/x$  for some  $a \in \{0, \dots, n\}$ . Then we can partition the (phantom and voter report) entries, with the exception of the phantom at  $a/x$ , into sets  $A$  and  $B$  with  $|A| = |B| = n$ ,

<sup>3</sup>As written,  $f_n(1) = 0$ , but Definition 4.2 requires  $f_k(1) = 1$  for all  $k$ . This detail does not matter here, since normalization is always achieved without moving phantom  $n$ , but one could write  $f_n$  in a different form to satisfy Definition 4.2 without it changing the behavior of the mechanism.

where  $A$  consists only of entries less than or equal to  $a/x$ , and  $B$  consists only of entries greater than or equal to  $a/x$ . Note that  $B$  contains  $n - a$  phantoms, and thus  $n - (n - a) = a$  voter reports.

We show that  $\pi = a/x$  is a market clearing price. At this price, we can set the demand of each voter  $i \in N$  whose report is in  $B$  to  $D_i(\pi) = 1/\pi = x/a$ , and we can set the demand of each voter whose report is not in  $B$  to 0. Since  $B$  contains  $a$  voter reports, the total demand of all voters is  $a \cdot x/a = x$ , and hence  $\pi = a/x$  is a market clearing price. To see that it is the unique clearing price, note that at any price  $\pi < a/x$ , each voter whose report is in  $B$  has demand  $D_i(\pi) = 1/\pi > x/a$ , so the total demand of all voters in  $B$  is greater than  $x$ . At any price  $\pi > a/x$ , each voter whose report is in  $A$  has demand  $D_i(\pi) = 0$ , while each of the  $a$  voters whose report is in  $B$  has demand at most  $D_i(\pi) = 1/\pi < x/a$ , so the total demand of all voters is less than  $x$ .

Next, suppose that the median  $y$  is not a phantom entry, and say that we have  $a/x < y < (a+1)/x$  for some  $a \leq n - 1$  (note that the median cannot be greater than  $n/x$ , because it cannot be higher than the largest phantom value). Then we can partition the (phantom and voter report) entries, not including a single voter with  $v_i = y$  (one such voter must exist because the median coincides with some entry, and no phantom entry lies at  $y$ ), into sets  $A$  and  $B$  each of size  $n$ , where  $A$  consists only of entries less than or equal to  $y$ , and  $B$  consists only of entries greater than or equal to  $y$ . Again,  $B$  contains  $n - a$  phantom reports, and thus  $a$  voter reports.

We show that  $\pi = y$  is a market clearing price. For each voter  $i$  with  $v_i = y$  whose report is in  $B$ , let  $D_i(v_i) = 1/v_i = 1/y$ . For each voter  $i$  with  $v_i = y$  whose report is in  $A$ , let  $D_i(\pi) = 0$ . Finally, for the single voter  $i$  with  $v_i = y$  whose report is neither in  $A$  nor in  $B$ , set  $D_i(v_i) = x - a/y$ . Note that  $D_i(v_i) \in [0, 1/v_i]$  because  $a/x < y$  implies  $x - a/y > 0$  and  $y < (a+1)/x$  implies  $x - a/y < 1/y$ . Therefore, these settings are consistent with the requirements of a demand function. At price  $\pi = y$ , each of the  $a$  voters whose report is in  $B$  has demand  $D_i(y) = 1/y$ , the voter whose report is in neither  $A$  nor  $B$  has demand  $D_i(y) = x - a/y$ , and all other voters have demand 0, so the total demand is exactly  $x$ . Hence,  $\pi = y$  is a market clearing price. To see that it is the unique clearing price, note that at all prices  $\pi < y$ , each voter  $i$  whose report is in  $B$ , as well as the voter whose report is in neither  $A$  nor  $B$  has demand  $1/\pi > 1/y$ . The total demand is thus greater than  $(a+1) \cdot 1/y > x$ . Moreover, at all prices  $\pi > y$ , all voters  $i$  whose report is not in  $B$  have demand  $D_i(\pi) = 0$ , and all voters  $i$  whose report is in  $B$  have demand  $D_i(\pi) \leq 1/\pi < 1/y$ . The total demand is thus less than  $a \cdot 1/y < x$ . Therefore,  $y$  is the unique clearing price.  $\square$

The “market” connection to independent markets is now clear: For each alternative  $j$ , we set up a market in which we sell an amount  $x$  of a good; this amount is the same across markets. Voter  $i \in [n]$  has value  $\hat{p}_{i,j}$  for the good sold in market  $j$ , and has a budget of 1 in each market. The markets are “independent” because, while each voter is engaged in every market, the budget of 1 for each market can only be used to buy the good sold in that market. Using Lemma 5.3, we can derive the market clearing prices in each of these markets. If we write  $t = 1/x$ , then these prices correspond exactly to the output of  $\mathcal{A}^{IM}$  with the phantoms as placed at time  $t$ . Changing the phantom placement by varying  $t$  to normalize the output is equivalent to varying the amount  $x$  of the good sold in each market until the clearing prices across markets sum to 1. While we prevent phantoms from moving above 1 in the definition of independent markets—complying with Definition 4.2—the exact positions of these phantoms do not affect the clearing price since all reports are at most 1.

Returning to Example 5.2, we can verify the outcome using the market interpretation, by setting the quantity of goods to be sold in each market to  $x^* = 1/t^* = 4.5$ . In the market corresponding to alternative 1, the market clears at price  $\pi_1 = \frac{1}{3}$ , at which price voter 3 demands  $1/\pi_1 = 3$  goods, voter 1 demands nothing, and the demand of voter 2 can be set to  $D_2(\frac{1}{3}) = 1.5$ , which lies between 0

and  $1/v_i = 3$ . Supply therefore equals demand. It can be checked that the market prices also match the independent markets outcome for alternatives 2 and 3.

An interesting property of independent markets is that normalization can always be achieved with  $x^* \geq n$ , that is, a supply of at least  $n$  units of each good.

**PROPOSITION 5.4.** *Suppose we run independent markets with supply  $x \leq n$ . Then the clearing prices sum to at least 1.*

**PROOF.** Let  $\pi_1, \dots, \pi_m$  be the clearing prices for supply  $x$ . We may assume that for each agent  $i \in [n]$ , there exists a project  $j$  for which  $\pi_j < \hat{p}_{i,j}$ , since otherwise we have  $\sum_{j \in [m]} \pi_j \geq \sum_{j \in [m]} \hat{p}_{i,j} = 1$  and we are done. Thus, for each agent there exists at least one market where the agent spends all available funds of 1 (by definition of demand). Since each agent spends at least 1, total spending across all markets is at least  $n$ . Now, the cost of purchasing all available goods is  $x \cdot \sum_{j \in [m]} \pi_j$ . Since the markets clear, this is the overall amount of money that the agents spend. Hence  $x \cdot \sum_{j \in [m]} \pi_j \geq n$ . Since  $x \leq n$ , we obtain  $\sum_{j \in [m]} \pi_j \geq 1$ , as desired.  $\square$

In the moving phantom interpretation, Proposition 5.4 means that  $t^* \leq \frac{1}{n}$  will always suffice for achieving normalization, so that the phantom movement will never need to continue past the uniform phantom placement  $\mathcal{F}^{IM}(\frac{1}{n}) = (1, \frac{n-1}{n}, \dots, \frac{1}{n}, 0)$ . In turn, this shows that if we used the uniform phantom mechanism on each project independently (as discussed in Example 4.1) the outcome will always sum to at least 1. The independent markets mechanism always returns a distribution  $\mathbf{p}$  where each coordinate is at most the result of taking uniform phantoms in each project independently.

The market system we have described yields an incentive-compatible aggregator, since it corresponds to a moving phantom mechanism. There are other market-based aggregation mechanisms described in the literature, most famously the parimutuel consensus mechanism of Eisenberg and Gale [1959]. That mechanism differs from ours in that voters have only a single budget of 1 which they can use in all of the markets. (The supply of goods can be fixed at  $x = n$ , which guarantees that prices are normalized, because total spending is fixed.) For the case  $m = 2$ , it does not matter whether markets are independent or not, and our mechanism is equivalent to the one of Eisenberg and Gale [1959]. It follows that the parimutuel consensus mechanism is incentive compatible for  $m = 2$  (in our  $\ell_1$  sense). However, for  $m \geq 3$ , the mechanism is manipulable,<sup>4</sup> and hence cannot be represented as a moving phantom mechanism. We point the reader to the work of Garg et al. [2018] for a detailed overview of other settings in which market mechanisms have been used in the context of public decision making.

## 5.2 Voting Game Interpretation

We have seen descriptions of  $\mathcal{A}^{IM}$  as a moving phantom mechanism and as clearing prices of a market system. We next give a game-theoretic description: the independent markets mechanism can be seen as the unique Nash equilibrium outcome of a voting game inspired by range voting [Smith 2000]. The game works as follows. Each voter receives one “credit” per alternative, and chooses how much of that credit to place on the alternative. That is, each voter chooses a vector  $\mathbf{s}_i \in [0, 1]^m$ . The outcome of the game is the division  $\mathbf{p}$  where  $p_j$  is proportional to  $\sum_{i \in [n]} s_{i,j}$ , the number of credits spent on alternative  $j$ . Voters choose their spending so as to obtain an outcome that is as close to their ideal point  $\hat{\mathbf{p}}_i$  as possible, according to  $\ell_1$  distance.

<sup>4</sup>Let  $\mathbf{p}_1 = (0, 0.5, 0.5)$ ,  $\mathbf{p}_2 = (0.5, 0.5, 0)$ . Parimutuel consensus yields prices  $(1/3, 1/3, 1/3)$ , at distance  $2/3$  from  $\mathbf{p}_1$ . If voter 1 instead reports  $\hat{\mathbf{p}}_1 = (0, 0, 1)$ , the price vector is  $(0.25, 0.25, 0.5)$ , at distance  $0.5$  from  $\mathbf{p}_1$ . This is an example of an agent manipulating prices in a competitive market, which has been studied in more general settings [e.g., Roberts and Postlewaite 1976].

**THEOREM 5.5.** *The voting game defined above has a unique outcome  $\mathbf{p}$  that can be obtained in Nash equilibrium, and it is equal to the outcome of the independent markets mechanism.*

The idea of the proof is to set  $s_{i,j}$  equal to the amount that agent  $i$  spends in the market for alternative  $j$ , under the market setup that we described earlier. We show that when every agent casts “vote”  $s_i$ , the system is at (its unique) equilibrium, and that the spending is proportional to the independent markets outcome.

**PROOF.** Let  $\mathbf{s}_{-i}$  denote the spending profile of all voters other than  $i$ , and let  $\mathbf{q}$  denote the aggregate division in which the weight on an alternative is proportional to the number of credits spent on that alternative. Suppose that  $\mathbf{s}_i$  is a best response for voter  $i$ . We first show that for every alternative  $j$ ,  $q_j < p_{i,j}$  implies  $s_{i,j} = 1$  and  $q_j > p_{i,j}$  implies  $s_{i,j} = 0$  (we only prove the former, but the latter follows from a symmetric argument). That is, voters will always prefer to increase their spending on alternatives that they consider “undervalued”, and decrease their spending on “overvalued” alternatives.

To prove this, assume for a contradiction that  $q_j < p_{i,j}$  and  $s_{i,j} < 1$ . Since  $\mathbf{q}$  and  $\mathbf{p}_i$  are both normalized, there must exist some alternative  $j'$  for which  $q_{j'} > p_{i,j'}$ . By marginally increasing  $s_{i,j}$  while holding  $s_{i,\hat{j}}$  constant for all  $\hat{j} \neq j$ , voter  $i$  can move the aggregate division from  $\mathbf{q}$  to  $\tilde{\mathbf{q}}$ , where  $\tilde{q}_j = q_j + \varepsilon < p_{i,j}$ , and  $\tilde{q}_{\hat{j}} = q_{\hat{j}} - \varepsilon_j$  for all  $\hat{j} \neq j$ , with  $\sum_{\hat{j} \neq j} \varepsilon_j = \varepsilon$  and  $\tilde{q}_{j'} > p_{i,j'}$ . This implies that

$$\begin{aligned} d(\mathbf{q}, \mathbf{p}_i) - d(\tilde{\mathbf{q}}, \mathbf{p}_i) &= |q_j - p_{i,j}| - |\tilde{q}_j - p_{i,j}| + |q_{j'} - p_{i,j'}| - |\tilde{q}_{j'} - p_{i,j'}| + \sum_{\hat{j} \neq j, j'} (|q_{\hat{j}} - p_{i,\hat{j}}| - |\tilde{q}_{\hat{j}} - p_{i,\hat{j}}|) \\ &\geq \varepsilon + \varepsilon_{j'} - \sum_{\hat{j} \neq j, j'} \varepsilon_j = 2\varepsilon_{j'} > 0 \end{aligned}$$

and hence  $d(\mathbf{q}, \mathbf{p}_i) > d(\tilde{\mathbf{q}}, \mathbf{p}_i)$ , contradicting that  $\mathbf{s}_i$  is a best response for  $i$ .

Next, we show that there exists an equilibrium of the voting game that produces the same outcome as independent markets. To this end, consider the market interpretation. For every alternative, some amount  $x^* = 1/t^*$  of a divisible good is sold to voters with a budget of 1 each. The amount of money that each voter spends in each market defines some spending profile  $\mathbf{s}$ .<sup>5</sup> Importantly, whenever  $p_{i,j} > \mathcal{A}^{IM}(\mathbf{P})_j$ , voter  $i$  spends their full budget on alternative  $j$  (i.e.  $s_{i,j} = 1$ ), and whenever  $p_{i,j} < \mathcal{A}^{IM}(\mathbf{P})_j$ , voter  $i$  spends nothing on alternative  $j$  (i.e.  $s_{i,j} = 0$ ). Further, the total amount spent by all voters on alternative  $j$  is  $x^* \cdot \mathcal{A}^{IM}(\mathbf{P})_j$  – the amount of goods sold multiplied by the price per unit – which is proportional to the aggregate division  $\mathcal{A}^{IM}(\mathbf{P})$ . Therefore, the induced spending profile does produce the same aggregate division as independent markets when aggregated under the rules of the voting game. We now show that this spending profile is a Nash equilibrium.

To do so, consider the spending vector  $\mathbf{s}_i$  of some voter  $i$ , and suppose they have a better response  $\tilde{\mathbf{s}}_i$ . Denote the aggregate division when  $i$  spends  $\mathbf{s}_i$  by  $\mathbf{q}$  (this division is  $\mathcal{A}^{IM}(\mathbf{P})$ , but we use  $\mathbf{q}$  for short), and the aggregate division when  $i$  spends  $\tilde{\mathbf{s}}_i$  by  $\tilde{\mathbf{q}}$ . Since  $d(\tilde{\mathbf{q}}, \mathbf{p}) < d(\mathbf{q}, \mathbf{p})$ , there must exist some alternative  $j$  with  $q_j < p_{i,j}$  and  $q_j < \tilde{q}_j$ , or with  $q_j > p_{i,j}$  and  $q_j > \tilde{q}_j$ . Suppose without loss of generality that the former case holds (if not, a very similar argument applies). Because  $q_j < p_{i,j}$ , we have  $s_{i,j} = 1$  by definition of  $\mathbf{s}$ . Therefore, the only way for

$$q_j = \frac{s_{i,j} + \sum_{i' \neq i} s_{i',j}}{\sum_{j'} s_{i,j'} + \sum_{i' \neq i} \sum_{j'} s_{i',j'}} < \frac{\tilde{s}_{i,j} + \sum_{i' \neq i} s_{i',j}}{\sum_{j'} \tilde{s}_{i,j'} + \sum_{i' \neq i} \sum_{j'} s_{i',j'}} = \tilde{q}_j$$

<sup>5</sup>This spending profile is not unique, since for voters with  $p_{i,j}$  equal to the clearing price for alternative  $j$ , there is some flexibility as to which voters pay for and get assigned goods, and which do not.



is if we have  $\sum_{j'} s_{i,j'} > \sum_{j'} \tilde{s}_{i,j'}$ . But, because  $s_{i,j'} = 0$  for all  $j'$  with  $q_{j'} > p_{i,j'}$ , it must be the case that  $q_{j'} \leq p_{i,j'}$  for all alternatives with  $s_{i,j'} > \tilde{s}_{i,j'}$ . Further, any alternative with  $\tilde{q}_{j'} < q_{j'}$  must have  $\tilde{s}_{i,j'} < s_{i,j'}$ ; this follows from the fact that  $\sum_{j'} s_{i,j'} > \sum_{j'} \tilde{s}_{i,j'}$ . Putting the previous two observations together, we have that  $q_{j'} \leq p_{i,j'}$  for all alternatives with  $\tilde{q}_{j'} < q_{j'}$ . Therefore

$$\begin{aligned} d(\mathbf{q}, \mathbf{p}_i) - d(\tilde{\mathbf{q}}, \mathbf{p}_i) &= \sum_{\tilde{q}_{j'} < q_{j'}} (|q_{j'} - p_{i,j'}| - |\tilde{q}_{j'} - p_{i,j'}|) + \sum_{\tilde{q}_{j'} \geq q_{j'}} (|q_{j'} - p_{i,j'}| - |\tilde{q}_{j'} - p_{i,j'}|) \\ &= \sum_{\tilde{q}_{j'} < q_{j'}} (\tilde{q}_{j'} - q_{j'}) + \sum_{\tilde{q}_{j'} \geq q_{j'}} (|q_{j'} - p_{i,j'}| - |\tilde{q}_{j'} - p_{i,j'}|) \\ &\leq \sum_{\tilde{q}_{j'} < q_{j'}} (\tilde{q}_{j'} - q_{j'}) + \sum_{\tilde{q}_{j'} \geq q_{j'}} (\tilde{q}_{j'} - q_{j'}) = 0 \end{aligned}$$

which contradicts that  $\tilde{\mathbf{s}}_i$  is a better response than  $\mathbf{s}_i$  for voter  $i$ .

Next we show that the voting game has a *unique* equilibrium aggregate division.<sup>6</sup> We know that the independent markets aggregate  $\mathbf{q}$  is an equilibrium with spending profile  $\mathbf{s}$ . For contradiction, suppose that there is some other equilibrium aggregate  $\tilde{\mathbf{q}}$  with spending profile  $\tilde{\mathbf{s}}$ . Then there is an alternative  $j$  for which  $\tilde{q}_j > q_j$  and an alternative  $j'$  for which  $\tilde{q}_{j'} < q_{j'}$ . Thus, there are weakly fewer voters with  $p_{i,j} \geq \tilde{q}_j$  than with  $p_{i,j} > q_j$ . From our discussion of best responses in the first part of the proof, we know that only voters with  $p_{i,j} \geq \tilde{q}_j$  can have  $\tilde{s}_{i,j} > 0$ , and that all voters with  $p_{i,j} > q_j$  have  $s_{i,j} = 1$ . It follows that  $\sum_i s_{i,j} \geq \sum_i \tilde{s}_{i,j}$ . But, because  $\tilde{q}_j > q_j$ , global spending across all alternatives must be lower under  $\tilde{\mathbf{s}}$  than under  $\mathbf{s}$ , that is,  $\sum_i \sum_j s_{i,j} > \sum_i \sum_j \tilde{s}_{i,j}$ . By an identical argument,  $\sum_i s_{i,j'} \leq \sum_i \tilde{s}_{i,j'}$ , implying that  $\sum_i \sum_j s_{i,j} < \sum_i \sum_j \tilde{s}_{i,j}$ , contradicting the previous sentence. Hence, the equilibrium aggregate division  $\mathbf{q} = \mathcal{A}^{IM}(\mathbf{P})$  is unique.  $\square$

To illustrate the voting game interpretation, consider again Example 5.2. Define spending vectors  $\mathbf{s}_1 = (0, 1, 1)$ ,  $\mathbf{s}_2 = (0.5, 1, 0)$ ,  $\mathbf{s}_3 = (1, 0, 0)$ . These vectors sum to  $(1.5, 2, 1)$ , which is proportional to  $\mathcal{A}^{IM}(\mathbf{P}) = (\frac{1}{3}, \frac{4}{9}, \frac{2}{9})$ . It can be checked that no voter wishes to unilaterally change their spending vector. For example, if voter 1 increases the amount she spends on alternative 1, she will increase the first coordinate of the outcome while decreasing the second and third coordinates, all of which increase the distance between the outcome and her preferred budget.

### 5.3 Properties of Independent Markets

Using the different interpretations of  $\mathcal{A}^{IM}$ , we will now establish that this mechanism satisfies several appealing axioms from social choice theory. As we mentioned at the beginning of Section 5, since  $\mathcal{A}^{IM}$  includes the uniform phantom placement, it satisfies proportionality (Definition 2.3).<sup>7</sup> In fact, we can prove that independent markets satisfies proportionality for the more general class of *k-approval profiles*.

*Definition 5.6.* Let  $k \geq 1$ . A budget division  $\mathbf{p}_i$  is *k-uniform* if it is uniform over  $k$  entries. We say that voter  $i$  approves a project  $j$  if  $p_{i,j} = \frac{1}{k}$ . A mechanism satisfies *k-proportionality* if for every profile consisting only of  $k$ -uniform votes, for each project  $j$ , the outcome spending on  $j$  is proportional to the number of voters who approve  $j$ .

<sup>6</sup>In contrast, the exact spending profile is clearly not unique. For instance, in an instance with only a single voter, that voter can enforce their belief exactly as long as they spend on each outcome in proportion to their belief, with no restriction on the magnitude of their spending.

<sup>7</sup>This condition is not necessary; a necessary and sufficient condition is that for all  $i_1, i_2, \dots, i_m$  with  $\sum_{j=1}^m i_j = n$ , there exists  $t$  for which  $f_{n-i_j}(t) = i_j/n$  for each  $j \in [m]$ .

For each  $k \geq 1$ , the phantom system of the independent markets mechanism places phantoms uniformly between 0 and  $\frac{1}{k}$  at some point. Our proof shows that any moving phantom mechanism with this property satisfies  $k$ -proportionality.

**THEOREM 5.7.**  $\mathcal{A}^{IM}$  satisfies  $k$ -proportionality for all  $k \leq m$ . In fact,  $\mathcal{A}^{\mathcal{F}}$  is  $k$ -proportional,  $2 \leq k \leq m - 2$ , if and only if there exists  $t \in [0, 1]$  with  $\mathcal{F}(t) = (f_0(t), \frac{n-1}{n} \cdot \frac{1}{k}, \dots, \frac{1}{n} \cdot \frac{1}{k}, 0)$ , where  $f_0(t) \geq \frac{1}{k}$ .

**PROOF.** Let  $k \leq m$ . Suppose that  $n_j$  voters approve alternative  $j$ . Thus  $\sum_{j=1}^m n_j = nk$ , since each voter approves  $k$  alternatives. When  $\mathcal{F}(t) = (f_0(t), \frac{n-1}{n} \cdot \frac{1}{k}, \dots, \frac{1}{n} \cdot \frac{1}{k}, 0)$  and  $f_0(t) \geq \frac{1}{k}$ , we have

$$\text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], j}) = \frac{n_j}{n} \cdot \frac{1}{k}.$$

Note that  $\sum_{j=1}^m \text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], j}) = 1$ , so the output is normalized for phantom placement  $\mathcal{F}(t)$ . Since the output on each alternative is proportional to the number of voters approving that alternative,  $\mathcal{A}^{\mathcal{F}}$  is  $k$ -proportional.

Conversely, suppose there is no such  $t$ . Then there must exist phantoms  $f_i$  and  $f_h$ ,  $i \in \{1, \dots, n\}$  and  $h \in \{0, \dots, n-1\}$ , and  $t' \in [0, 1]$  for which  $f_i(t') > \frac{n-i}{n} \cdot \frac{1}{k}$  and  $f_h(t') < \frac{n-h}{n} \cdot \frac{1}{k}$ . Assume without loss of generality that  $i > h$ . Consider a profile in which voters  $1, \dots, n-i$  approve alternatives  $1, \dots, k$ , voters  $n-i+1, \dots, n-h$  approve alternatives  $2, \dots, k+1$ , and voters  $n-h+1, \dots, n$  approve alternatives  $3, \dots, k+2$ . This profile is well-defined since each voter approves  $2 \leq k \leq m-2$  alternatives. Further,  $n-i$  voters approve alternative 1 and  $n-h$  voters approve alternative 2. Therefore, we have  $\text{med}(\mathcal{F}(t'), \mathbf{P}_{i \in [n], 1}) = \min(f_i(t'), \frac{1}{k}) > \frac{n-i}{n} \cdot \frac{1}{k}$  and  $\text{med}(\mathcal{F}(t'), \mathbf{P}_{i \in [n], 2}) = f_h(t') < \frac{n-h}{n} \cdot \frac{1}{k}$ . Since  $\text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], j})$  is increasing in  $t$  for every alternative  $j$ , there does not exist a  $t$  for which  $\text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], 1}) = \frac{n-i}{n} \cdot \frac{1}{k}$  and  $\text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], 2}) = \frac{n-h}{n} \cdot \frac{1}{k}$ , as required by  $k$ -proportionality.  $\square$

An equivalent definition of  $k$ -proportionality says that on profiles consisting only of  $k$ -uniform votes, the mechanism should equal the mean of the reported distributions. This suggests a further generalization of this proportionality notion: on profiles where each voter  $i$  submits a  $k_i$ -uniform distribution (where the  $k_i$  may vary), the mechanism should equal the mean. However, it is easy to see that this property is incompatible with incentive compatibility: Let  $n = m = 3$ , and consider the two profiles  $\mathbf{P} = ((1, 0, 0), (0.5, 0.5, 0), (0.5, 0, 0.5))$  and  $\mathbf{P}' = ((1, 0, 0), (0, 1, 0), (0.5, 0, 0.5))$ . If mechanism  $\mathcal{A}$  were to satisfy the most general proportionality requirement, this would imply  $\mathcal{A}(\mathbf{P}) = (\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$  and  $\mathcal{A}(\mathbf{P}') = (\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$ . Thus,  $\mathcal{A}$  is not incentive compatible, because in profile  $\mathbf{P}$ , voter 2 has an incentive to misreport to obtain profile  $\mathbf{P}'$ . Note that the outcome at  $\mathbf{P}'$  is closer to voter 2's ideal point than the outcome at  $\mathbf{P}$  for every project. Thus, this general proportionality criterion is incompatible even for much weaker notions of incentive compatibility.

Independent markets satisfies several other desirable properties. By using our equivalent descriptions of this mechanism, it is possible to see this easily. For example, using the voting game interpretation, it becomes clear that independent markets satisfies participation.

**THEOREM 5.8.**  $\mathcal{A}^{IM}$  satisfies participation.

**PROOF.** Under  $\mathcal{A}^{IM}$ , adding a new voter who agrees with the aggregate division does not change the aggregate:  $\mathcal{A}^{IM}(\mathcal{A}^{IM}(\mathbf{P}_{-i}), \mathbf{P}_{-i}) = \mathcal{A}^{IM}(\mathbf{P}_{-i})$ . This can be seen by noting that for the new voters, it is an equilibrium spending strategy to spend nothing, thus not changing the spending profile. Then by incentive compatibility,  $d(\mathcal{A}^{IM}(\mathbf{p}_i, \mathbf{P}_{-i}), \mathbf{p}_i) \leq d(\mathcal{A}^{IM}(\mathcal{A}^{IM}(\mathbf{P}_{-i}), \mathbf{P}_{-i}), \mathbf{p}_i) = d(\mathcal{A}^{IM}(\mathbf{P}_{-i}), \mathbf{p}_i)$ .  $\square$

Using the market interpretation, we can verify that independent markets satisfies reinforcement.

**THEOREM 5.9.**  $\mathcal{A}^{IM}$  satisfies reinforcement.

PROOF. Let  $\mathbf{P}$  and  $\mathbf{R}$  be profiles with  $\mathcal{A}^{IM}(\mathbf{P}) = \mathcal{A}^{IM}(\mathbf{R}) = \mathbf{q}$ . We utilize the market interpretation of independent markets. Suppose that for profile  $\mathbf{P}$ , market prices are normalized when  $x_P^*$  goods are sold and for profile  $\mathbf{R}$ , market prices are normalized when  $x_R^*$  goods are sold. Now consider the combined profile  $\mathbf{P} \cup \mathbf{R}$  when  $x_P^* + x_R^*$  goods are sold. For every alternative  $j$ , and every price  $\pi \in [0, 1]$ , the total demand is equal to the total demand in profile  $P$  at price  $\pi$  plus the total demand in profile  $R$  at price  $\pi$ , since a voter's demand depends only on their valuation and the price  $\pi$ . Likewise, the total supply is, by definition, the sum of the supplies in each individual instance. Therefore, the market clearing prices in the combined instance when  $x_P^* + x_R^*$  goods are sold are equal to the (normalized) aggregate vector  $\mathbf{q}$ .  $\square$

## 6 PARETO OPTIMALITY AND SOCIAL WELFARE

We now study moving phantom mechanisms from an efficiency perspective, focusing on Pareto optimality. We will discover that there is only a single moving phantom mechanism that is Pareto optimal. This mechanism can be interpreted as maximizing utilitarian welfare. On the other hand, the independent markets mechanism is not Pareto optimal. If voter 1 reports  $(0.8, 0.2, 0)$  and voter 2 reports  $(0.8, 0, 0.2)$ , then independent markets chooses  $(0.6, 0.2, 0.2)$ , which is dominated by  $(0.8, 0.1, 0.1)$ . On this example, independent markets even fails to be *range respecting*, which requires that  $\min_{i \in [n]} \hat{p}_{i,j} \leq \mathcal{A}(\mathbf{P})_j \leq \max_{i \in [n]} \hat{p}_{i,j}$  for all  $j \in [m]$ .

One can show that a moving phantom mechanism  $\mathcal{A}^{\mathcal{F}}$  is range respecting if  $f_0(t) = 1$  and  $f_n(t) = 0$  for all  $t \in [0, 1]$  except for an initial period where phantom 0 moves from 0 to 1 while all other phantoms remain at 0, and a period at the end where phantom  $n$  moves from 0 to 1 while all other phantoms are at 1.<sup>8</sup> Using this condition, it is easy to construct moving phantom mechanisms that are both proportional and range respecting. For instance, one can modify  $\mathcal{A}^{IM}$  so that the highest phantom is always at 1. This makes the mechanism range respecting while preserving  $k$ -proportionality. Alternatively, we can take the phantom system

$$f_k(t) = \max\{0, 1 - (1 - t)(n - k)\} \quad \text{for each } k \in \{0, \dots, n\},$$

that is equivalent to placing phantoms uniformly between  $y$  and 1 (rather than between 0 and  $x$ , as independent markets does).<sup>9</sup> Using Proposition 5.4, it follows that  $y \leq 0$ , and so the resulting mechanism is range respecting. This mechanism is proportional, but it fails  $k$ -proportionality.

While many moving phantom mechanisms are range respecting, it is much more difficult to find a mechanism in this class which is Pareto optimal. Usually, it is possible to construct a profile in which the mechanism outputs a vector  $\mathbf{p}$  all of whose entries are phantom reports, and then a Pareto improvement can be obtained by perturbing this vector in the directions where the majority of voter reports lie. Such a perturbation is not possible if the phantoms lie at 0 or 1, which turns out to be the only escape. As we prove below, no mechanism  $\mathcal{A}^{\mathcal{F}}$  can be Pareto optimal if there is any time point  $t$  when two phantoms are both strictly between 0 and 1.

This condition is extremely restrictive, and a moment's thought reveals that there is only one phantom system which avoids having two interior phantoms: All phantoms start at 0, and then, one by one, one of the phantoms is moved to 1. At each  $t$ , at most one phantom lies strictly between 0 and 1 while travelling. We call this phantom system  $\mathcal{F}^*$ . It can be formalized as

$$f_k(t) = \begin{cases} 0 & 0 \leq t \leq \frac{k}{n+1}, \\ t(n+1) - k & \frac{k}{n+1} < t < \frac{k+1}{n+1}, \\ 1 & \frac{k+1}{n+1} \leq t \leq 1. \end{cases}$$

<sup>8</sup>This mirrors a result in Section 3; if the outer two phantoms are at 0 and 1, the  $n - 1$  remaining phantoms cannot outweigh the  $n$  voter reports.

<sup>9</sup>We thank an anonymous reviewer for suggesting this mechanism.

Below we will show that  $\mathcal{A}^{\mathcal{F}^*}$  precisely corresponds to the budget aggregation mechanism that maximizes voter welfare, breaking ties in favor of the maximum entropy division. It will immediately follow that  $\mathcal{A}^{\mathcal{F}^*}$  is indeed Pareto optimal. Combined with Theorem 6.1 below, which shows that all other moving phantom mechanisms are Pareto inefficient, this implies that the welfare-maximizing mechanism is the unique Pareto-optimal moving phantom mechanism.

### 6.1 Characterizing Pareto Optimality

The proof of Theorem 6.1 shows, by induction, that each phantom needs to move all the way to 1 before the next phantom can leave its position at 0. In case this does not happen, based on the approximate phantom positions, we construct a profile where the mechanism is Pareto inefficient. These constructions are of two kinds: an easier case when the interior phantoms are low (lying below  $\frac{1}{n(n-1)}$ ), and a more involved case when one of the phantoms has moved higher. In both cases, our constructions utilize two types of alternatives. More voters report “high” probabilities on alternatives of the first type than on alternatives of the second type. The constructions work so that if two phantoms simultaneously take values between 0 and 1, then the mechanism outputs middling values on all alternatives. Social welfare can be improved by increasing the output on alternatives of the first type, and decreasing the output on alternatives of the second type. By incorporating enough symmetry between voters, we guarantee that social welfare gains are shared equally, so obtain a Pareto improvement.

**THEOREM 6.1.** *A moving phantom mechanism  $\mathcal{A}^{\mathcal{F}}$  cannot be Pareto optimal for any  $m \geq n^2$  unless  $\mathcal{A}^{\mathcal{F}} = \mathcal{A}^{\mathcal{F}^*}$ .*

**PROOF.** Consider a Pareto-optimal moving phantom mechanism  $\mathcal{A}^{\mathcal{F}}$ . We begin by showing that if there exists a  $t$  with  $f_0(t) < 1$  and  $f_1(t) > 0$  then  $\mathcal{A}^{\mathcal{F}}$  can be equivalently expressed as a moving phantom mechanism that does not have such a  $t$ . It will be helpful to order the entries  $\{p_{i,j}\}$  from largest to smallest on every alternative  $j$ . We denote the relabeled entries  $\bar{p}_{1,j} \geq \dots \geq \bar{p}_{n,j}$ . Suppose that  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_j = \text{med}(f_0(t^*), \dots, f_n(t^*), \hat{p}_{1,j}, \dots, \hat{p}_{n,j}) < \bar{p}_{n,j}$  for some  $j$ . Then  $\mathcal{A}^{\mathcal{F}}$  is not Pareto optimal, because increasing  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_j$  and decreasing  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_{j'}$  for any coordinate with  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_{j'} > \bar{p}_{n,j'}$  is a Pareto improvement (such a coordinate must exist, because  $\sum_j \bar{p}_{n,j} \leq 1$ ). Therefore, for all preference profiles  $\mathbf{P}$ ,  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_j = \text{med}(f_0(t^*), \dots, f_n(t^*), \hat{p}_{1,j}, \dots, \hat{p}_{n,j}) \geq \bar{p}_{n,j}$ . This implies that  $f_0(t^*) \geq \bar{p}_{n,j}$  and so the exact position of  $f_0(t^*)$  has no effect on the mechanism. It would be equivalent to move phantom  $f_0$  to position 1 before moving phantom  $f_1$ .

A very similar argument can be used to show that there cannot exist a  $t$  for which  $f_{n-1}(t) < 1$  and  $f_n(t) > 0$ . For  $n = 1$  and  $n = 2$ , this already pins down a unique Pareto-optimal moving phantom mechanism, so we assume  $n \geq 3$  and focus on the intermediate phantoms for the rest of the proof. Suppose that there exists some index  $1 \leq k \leq n - 2$  for which  $f_k(t) < 1$  and  $f_{k+1}(t) > 0$  for some  $t$ . If no such  $k$  exists, then  $\mathcal{F} = \mathcal{F}^*$  and we are done. The remainder of the proof proceeds in two parts.

*Part 1.* We show that for any  $t$  for which  $f_k(t) < \frac{1}{n(n-1)}$ , it must be the case that  $f_{k+1}(t) = 0$ . Suppose otherwise for contradiction, that there exists a  $t'$  for which  $f_k(t') < \frac{1}{n(n-1)}$  but  $f_{k+1}(t') > 0$ . We define an instance with  $m = n^2$  alternatives, which are divided into two kinds: we write  $X = \{1, \dots, n(n-1)\}$  and  $Y = \{n(n-1) + 1, \dots, n^2\}$ . Thus,  $|X| = n(n-1)$  and  $|Y| = n$ . For each voter

$i \in [n]$ , we also introduce the sets

$$\begin{aligned} X_i &= \{((i-1)(n-1) + 1, \dots, (i-1)(n-1) + n^2 - n - kn + k) \bmod n(n-1)\} \subseteq X, \\ Y_i &= \{n(n-1) + (i \bmod n), n(n-1) + (i+1 \bmod n), \dots, \\ &\quad \dots, n(n-1) + (i+n-k-2 \bmod n)\} \subseteq Y. \end{aligned}$$

We have  $|X_i| = n^2 - n - kn + k = (n-1)(n-k)$  and  $|Y_i| = n - k - 1$ . Voter  $i \in [n]$  reports

$$\hat{p}_{i,j} = \begin{cases} \frac{1}{n^2 - kn - 1} & j \in X_i \cup Y_i, \\ 0 & \text{otherwise.} \end{cases}$$

Note that  $\sum_{j=1}^m \hat{p}_{i,j} = 1$  because  $|X_i \cup Y_i| = n^2 - kn - 1$ . Further, since each voter makes  $|X_i| = (n-1)(n-k)$  non-zero reports among alternatives in  $X$ , and the constructed votes treat all of these alternatives symmetrically, each alternative receives  $n-k$  non-zero reports. Similarly, each alternative in  $Y$  receives  $n-k-1$  non-zero reports. Therefore,  $\text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], j}) = \min(f_k(t), \frac{1}{n^2 - kn - 1})$  for  $j \in X$  and  $\text{med}(\mathcal{F}(t), \mathbf{P}_{i \in [n], j}) = \min(f_{k+1}(t), \frac{1}{n^2 - kn - 1})$  for  $j \in Y$ .

By assumption,  $f_k(t^*) < \frac{1}{n(n-1)} < \frac{1}{n^2 - kn - 1}$  and  $f_{k+1}(t^*) > 0$ . It will therefore necessarily be the case that

$$\begin{aligned} \mathcal{A}^f(\mathbf{P})_j &= f_k(t^*) < \frac{1}{n(n-1)} \quad \text{for all } j \in X, \text{ and} \\ \mathcal{A}^f(\mathbf{P})_j &= f_{k+1}(t^*) > 0 \quad \text{for all } j \in Y. \end{aligned}$$

To see this, suppose that  $f_k(t^*) \geq \frac{1}{n(n-1)}$ , which implies that  $t^* > t'$ . In turn, this implies that  $\mathcal{A}^f(\mathbf{P})_j = \min(f_k(t^*), \frac{1}{n^2 - kn - 1}) \geq \frac{1}{n(n-1)}$  for  $j \in X$  and that  $\mathcal{A}^f(\mathbf{P})_j \geq \text{med}(\mathcal{F}(t'), \mathbf{P}_{i \in [n], j}) = \min(f_{k+1}(t'), \frac{1}{n^2 - kn - 1}) > 0$  for  $j \in Y$ , violating normality of  $\mathcal{A}^f(\mathbf{P})$ . Conversely, if  $f_{k+1}(t^*) = 0$ , then  $t^* < t'$ , which implies that  $\mathcal{A}^f(\mathbf{P})_j = \min(f_k(t^*), \frac{1}{n^2 - kn - 1}) \leq \min(f_k(t'), \frac{1}{n^2 - kn - 1}) < \frac{1}{n(n-1)}$  for  $j \in X$ , again violating normality.

But this is not Pareto optimal. Consider, for some small enough  $\varepsilon > 0$ , increasing  $\mathcal{A}^f(\mathbf{P})_j$  by  $\varepsilon$  on alternatives  $j \in X$ , and decreasing  $\mathcal{A}^f(\mathbf{P})_j$  by  $\varepsilon(n-1)$  on alternatives  $j \in Y$ . For every voter  $i$ , there are  $|X_i| = (n-k)(n-1)$  alternatives on which the aggregate moves  $\varepsilon$  closer to  $i$ 's report,  $|X| - |X_i| = k(n-1)$  alternatives for which the aggregate moves  $\varepsilon$  farther from  $i$ 's report,  $|Y_i| = n - k - 1$  alternatives on which the aggregate moves  $\varepsilon(n-1)$  farther from  $i$ 's report, and  $|Y| - |Y_i| = k+1$  alternatives for which the aggregate moves  $\varepsilon(n-1)$  closer to  $i$ 's report. Summing these up, the change moves the aggregate

$$((n-k)(n-1) - k(n-1) - (n-1)(n-k-1) + (n-1)(k+1))\varepsilon = (2n-2)\varepsilon > 0$$

closer to  $\hat{\mathbf{p}}_i$  in  $\ell_1$  distance. Since the new aggregate is closer to every voter, it is a Pareto improvement.

*Part 2.* We now show that if  $f_k(t) < 1$  then it must be the case that  $f_{k+1}(t) = 0$ . For contradiction suppose otherwise. Let  $\bar{t} = \max\{t : f_{k+1}(t) = 0\}$  be the final snapshot at which  $f_{k+1}(t) = 0$ . By assumption,  $f_k(\bar{t}) < 1$ . We define an instance similar to that above, defining  $X, Y, X_i, Y_i$  as before. For each  $i \in [n]$ , we further partition  $X_i$  into two pieces, letting

$$\begin{aligned} X_i^1 &= \{((i-1)(n-1) + 1, \dots, (i-1)(n-1) + n-1) \bmod n(n-1)\} \subseteq X_i, \\ X_i^2 &= X_i \setminus X_i^1 = \{((i-1)(n-1) + n, \dots, (i-1)(n-1) + n^2 - n - kn + k) \bmod n(n-1)\}. \end{aligned}$$

Note that  $|X_i^1| = n-1$  and  $|X_i^2| = (n-k-1)(n-1)$ .

Let  $\delta > 0$  (we will determine the exact value of  $\delta$  later). We now specify the voter reports on  $X$ , handling the reports on alternatives in  $Y$  later. For every voter  $i$  and each  $j \in X$ , we let

$$\hat{p}_{i,j} = \begin{cases} \frac{1-f_k(\bar{t})-\delta}{n(n-1)-1} & j \in X_i^1 \setminus \{1\}, \\ \frac{1-f_k(\bar{t})}{n(n-1)-1} & j \in X_i^2 \setminus \{1\}, \\ f_k(\bar{t}) + \delta & j \in (X_i^1 \cup X_i^2) \cap \{1\}, \\ 0 & j \in X \setminus (X_i^1 \cup X_i^2). \end{cases}$$

Because we know that  $f_k(\bar{t}) \geq \frac{1}{n(n-1)}$  from Part 1 of the proof, we have that

$$\frac{1-f_k(\bar{t})-\delta}{n(n-1)-1} < \frac{1-f_k(\bar{t})}{n(n-1)-1} \leq \frac{1-\frac{1}{n(n-1)}}{n(n-1)-1} = \frac{1}{n(n-1)} \leq f_k(\bar{t}) < f_k(\bar{t}) + \delta, \quad (5)$$

which in particular implies that we have distributed more probability mass on  $X$  for voters  $i$  with  $1 \in X_i^1 \cup X_i^2$  than for other voters. To set  $\delta$ , we need to choose some value that guarantees  $\sum_{j \in X} \hat{p}_{i,j} < 1$  for all  $i$ . By (5), it is sufficient to set  $\delta$  so that

$$f_k(\bar{t}) + \delta + ((n-k)(n-1) - 1) \frac{1-f_k(\bar{t})}{n(n-1)-1} < 1. \quad (6)$$

To see that such a value of  $\delta$  exists, note that the left-hand side of (6) is continuous in  $\delta$  and takes a value strictly less than 1 when  $\delta = 0$ :

$$\begin{aligned} f_k(\bar{t}) + ((n-k)(n-1) - 1) \frac{1-f_k(\bar{t})}{n(n-1)-1} &< f_k(\bar{t}) + (n(n-1) - 1) \frac{1-f_k(\bar{t})}{n(n-1)-1} \\ &= f_k(\bar{t}) + 1 - f_k(\bar{t}) = 1. \end{aligned}$$

We now specify voter reports on the alternatives in  $Y$ : For each  $i$ , we distribute the remaining positive mass  $1 - \sum_{j \in X} p_{i,j}$  evenly among  $j \in Y_i$ , and set  $\hat{p}_{i,j} = 0$  for  $j \in Y \setminus Y_i$ .

As before, each alternative  $j \in X$  receives  $n-k$  non-zero reports. For  $j = 1$ , all of these reports are  $\hat{p}_{i,1} = f_k(\bar{t}) + \delta$ . For  $j \geq 2$ , there exist  $n-k-1$  voters with  $\hat{p}_{i,j} = \frac{1-f_k(\bar{t})}{n(n-1)-1}$  and a single voter with  $\hat{p}_{i,j} = \frac{1-f_k(\bar{t})-\delta}{n(n-1)-1}$  (since the sets  $X_i^1$  are pairwise disjoint). For every alternative  $j \in Y$ , there are  $n-k-1$  non-zero reports. It follows that

$$\begin{aligned} \text{med}(\mathcal{F}(\bar{t}), \mathbf{P}_{i \in [n], 1}) &= f_k(\bar{t}), \\ \text{med}(\mathcal{F}(\bar{t}), \mathbf{P}_{i \in [n], j}) &= \frac{1-f_k(\bar{t})-\delta}{n(n-1)-1} \quad \text{for all } j \in X \setminus \{1\}, \\ \text{med}(\mathcal{F}(\bar{t}), \mathbf{P}_{i \in [n], j}) &= f_{k+1}(\bar{t}) = 0 \quad \text{for all } j \in Y. \end{aligned}$$

The sum of these generalized medians is  $1-\delta$ . Therefore  $t$  needs to increase to achieve normalization; that is,  $t^* > \bar{t}$ . By the definition of  $\bar{t}$ , for any  $t > \bar{t}$ , we have that  $f_{k+1}(t) > 0$ , and therefore

$$\mathcal{A}^{\mathcal{F}}(\mathbf{P})_j = \text{med}(\mathcal{F}(t^*), \mathbf{P}_{i \in [n], j}) > 0 \quad \text{for all } j \in Y.$$

From this, we can deduce that  $f_k(t^*) < f_k(\bar{t}) + \delta$ , since otherwise we have

$$\begin{aligned} \mathcal{A}^{\mathcal{F}}(\mathbf{P})_1 &= \text{med}(\mathcal{F}(t^*), \mathbf{P}_{i \in [n], 1}) \geq f_k(\bar{t}) + \delta, \\ \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j &\geq \text{med}(\mathcal{F}(\bar{t}), \mathbf{P}_{i \in [n], j}) = \frac{1-f_k(\bar{t})-\delta}{n(n-1)-1} \quad \text{for all } j \in X \setminus \{1\}, \\ \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j &> 0 \quad \text{for all } j \in Y, \end{aligned}$$

which would yield  $\sum_{j=1}^m \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j > 1$ . Similarly, we must have  $f_{k+1}(t^*) < \frac{1-f_k(\bar{t})}{n(n-1)-1}$ , since otherwise

$$\begin{aligned} \mathcal{A}^{\mathcal{F}}(\mathbf{P})_1 &\geq \text{med}(\mathcal{F}(\bar{t}), \mathbf{P}_{i \in [n], 1}) = f_k(\bar{t}) \\ \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j &= \text{med}(\mathcal{F}(t^*), \mathbf{P}_{i \in [n], j}) = \frac{1-f_k(\bar{t})}{n(n-1)-1} \quad \text{for all } j \in X \setminus \{1\}, \\ \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j &> 0 \quad \text{for all } j \in Y, \end{aligned}$$

again yielding  $\sum_{j=1}^m \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j > 1$ .

To summarize, we are guaranteed that

$$\begin{aligned} f_k(\bar{t}) &\leq \mathcal{A}^{\mathcal{F}}(\mathbf{P})_1 < f_k(\bar{t}) + \delta \\ \frac{1-f_k(\bar{t})-\delta}{n(n-1)-1} &\leq \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j < \frac{1-f_k(\bar{t})}{n(n-1)-1} \quad \text{for all } j \in X \setminus \{1\}, \\ 0 &< \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j \quad \text{for all } j \in Y. \end{aligned}$$

Now we can define a Pareto improvement to  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})$  of the same form as previously. For some small enough  $\varepsilon$ , increase the aggregate by  $\varepsilon$  on alternatives  $j \in X$ , and decrease the aggregate by  $\varepsilon(n-1)$  on alternatives  $j \in Y$ .

For each voter  $i$  with  $\hat{p}_{i,1} = f_k(\bar{t}) + \delta$ , the new aggregate is  $\varepsilon$  better than  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})$  on alternative 1,  $\varepsilon$  worse on alternatives  $j \in X_i^1 \setminus \{1\}$ , with  $\hat{p}_{i,j} = \frac{1-f_k(\bar{t})-\delta}{n(n-1)-1}$ ,  $\varepsilon$  better on alternatives  $j \in X_i^2$ , with  $\hat{p}_{i,j} = \frac{1-f_k(\bar{t})}{n(n-1)-1}$ , and  $\varepsilon$  worse on alternatives  $j \in X \setminus X_i$ , with  $\hat{p}_{i,j} = 0$ . On alternatives in  $Y$ , there are  $k+1$  alternatives on which the aggregate moves  $\varepsilon(n-1)$  closer to voter  $i$ 's report (on alternatives for which voter  $i$  reports  $p_{i,j} = 0$ ), and  $n-k-1$  alternatives for which the aggregate may or may not have moved  $\varepsilon(n-1)$  farther from voter  $i$ 's report. Summing these up, we get that the aggregate has moved at least

$$(1 - (n-2) + (n-k-1)(n-1) - k(n-1) - (n-1)(n-k-1) + (n-1)(k+1))\varepsilon = 2\varepsilon > 0$$

closer to  $\mathbf{p}_1$  in  $\ell_1$  distance.

For each voter  $i$  with  $\hat{p}_{i,1} = 0$ , the new aggregate is  $\varepsilon$  worse than  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})$  on alternatives  $j \in X_i^1$ , with  $p_{i,j} = \frac{1-f_k(\bar{t})-\delta}{n(n-1)-1}$ , and  $\varepsilon$  better on alternatives  $j \in X_i^2$ , with  $p_{i,j} = \frac{1-f_k(\bar{t})}{n(n-1)-1}$ , and  $\varepsilon$  worse on alternatives  $j \in X \setminus X_i$ , with  $p_{i,j} = 0$ . On alternatives in  $Y$ , there are  $k+1$  alternatives on which the aggregate moves  $\varepsilon(n-1)$  closer to voter  $i$ 's report (on alternatives for which voter  $i$  reports  $p_{i,j} = 0$ ), and  $n-k-1$  alternatives for which the aggregate may or may not have moved  $\varepsilon(n-1)$  farther from voter  $i$ 's report. Summing these up, we get that the net change of  $\ell_1$  distance from  $\mathbf{p}_i$  from moving the aggregate is

$$(-1 - (n-2) + (n-k-1)(n-1) - k(n-1) - (n-1)(n-k-1) + (n-1)(k+1))\varepsilon = 0.$$

Thus, the aggregate has moved strictly closer to some voters and weakly closer to all voters, and hence the change forms a Pareto improvement, contradicting our assumption that  $\mathcal{A}^{\mathcal{F}}$  is Pareto optimal.

Finally, our construction uses  $m = n^2$  alternatives, but we can extend it to larger  $m$  by adding dummy alternatives that no voter puts any weight on.  $\square$

While the proof of this result requires a large enough  $m$ , note that moving phantom mechanisms apply to problems with an arbitrary number of projects, so no mechanism other than  $\mathcal{A}^{\mathcal{F}}$  can provide Pareto optimality for all cases.

Above, we constructed a mechanism that is range respecting and proportional, though it is not Pareto optimal. Using our characterization, it is immediate that no Pareto-optimal moving phantom mechanism can be proportional: On profiles consisting of single-minded voters,  $\mathcal{A}^{\mathcal{F}}$  selects a division that is also single-minded, following the plurality. Hence, it is not proportional.

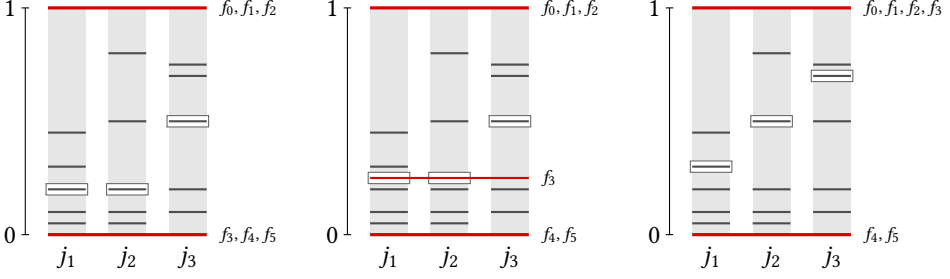


Fig. 2. Snapshots of the phantom system  $\mathcal{F}^*$  with  $t < t^*$  (left),  $t = t^*$  (center), and  $t > t^*$  (right) on an instance with  $n = 5$ ,  $m = 3$ .

**COROLLARY 6.2.** *For  $m \geq n^2$ , no moving phantom mechanism is proportional and Pareto optimal.*

## 6.2 Maximizing Utilitarian Social Welfare

Having narrowed down the space of Pareto-optimal moving phantom mechanisms to at most one mechanism, let us examine the behavior of  $\mathcal{A}^{\mathcal{F}^*}$  with the assistance of Figure 2, which takes the same form as Figure 1. Recall from the proof of Theorem 6.1 that  $\bar{p}_{1,j} \geq \dots \geq \bar{p}_{n,j}$  denote the entries  $\{p_{i,j}\}$ , relabeled and ordered from largest to smallest. At the snapshot of  $\mathcal{F}^*$  for which  $f_0(t) = \dots = f_k(t) = 1$  and  $f_{k+1}(t) = \dots = f_n(t) = 0$ , the generalized median selects the order statistic  $\bar{p}_{n-k,j}$  for all  $j$ . We see this in Figure 2 where, in the left image,  $k = 2$  and the generalized median is the  $n - k = 3$ rd highest report on each alternative, and in the right image  $k = 3$  and the  $n - k = 2$ nd highest reports are chosen.

We can think of  $\mathcal{F}^*$  as partitioning the phantom “movie” into periods defined by which phantom is moving. Initially, all phantoms are at 0, and the generalized medians are 0 for each  $j \in [m]$ . Then phantom  $f_0$  moves to 1, and the generalized medians are  $\bar{p}_{n,j}$ . As phantom  $f_k$  moves from 0 to 1, the generalized medians progress from  $\bar{p}_{n-k+1,j}$  to  $\bar{p}_{n-k,j}$ , until all phantoms reach 1 and the generalized medians are all 1. By (a discrete analogue of) the intermediate value theorem, there must exist some value  $I$  for which  $\sum_{j \in [m]} \bar{p}_{I+1,j} \leq 1$  and  $\sum_{j \in [m]} \bar{p}_{I,j} \geq 1$ , and this transition is made during the period in which phantom  $n - I$  is moving. In Figure 2, we have  $I = 2$  because the sum of the third-highest entries is less than 1 (see the left image), while the sum of the second-highest entries is more than 1 (the right image).

Normalization therefore occurs during the movement of phantom  $f_{n-I}$ , and the final value  $\mathcal{A}^{\mathcal{F}^*}(\mathbf{P})_j$  lies in the interval  $[\bar{p}_{I+1,j}, \bar{p}_{I,j}]$ . If  $f_{n-I}(t^*)$  lies in this interval, then  $\mathcal{A}^{\mathcal{F}^*}(\mathbf{P})_j = f_{n-I}(t^*)$ , otherwise  $\mathcal{A}^{\mathcal{F}^*}(\mathbf{P})_j$  is equal to the endpoint of the interval closest to  $f_{n-I}(t^*)$ . This is depicted in the center image of Figure 2, where  $f_3(t^*)$  lies between the second and third-highest reports on the first two alternatives, but below the third-highest report on the third alternative.

Finding the exact value of  $f_{n-I}(t^*)$ , and therefore the output  $\mathcal{A}^{\mathcal{F}^*}(\mathbf{P})$ , can be thought of as finding the “most equal” division, subject to interval constraints on each alternative. This problem has been studied before, and the (unique) value of  $f_{n-I}(t^*)$  can be found in  $O(m \log m)$  time by the Divvy algorithm of Gulati et al. [2012].

Given a profile  $\mathbf{P}$ , the *social cost* of an outcome  $\mathbf{p}$  is  $\sum_{i \in [n]} d(\hat{\mathbf{p}}_i, \mathbf{p})$ , and the (utilitarian) social welfare of  $\mathbf{p}$  is the negation of the social cost. In general, there may be multiple divisions that maximize social welfare. For example, if  $m = 2$ , one voter reports  $(1, 0)$  and another reports  $(0, 1)$ , then all divisions have the same social cost of 2. As it turns out, any division that satisfies the upper and lower bound constraints of  $\bar{p}_{I,j}$  and  $\bar{p}_{I+1,j}$  maximizes social welfare. This was derived



independently in a different context by Nehring and Puppe [2019], and related results are obtained by Lindner [2011].

LEMMA 6.3. *A division  $\mathbf{q}$  maximizes social welfare if and only if  $\bar{p}_{I+1,j} \leq q_j \leq \bar{p}_{I,j}$  for all  $j$ .*

PROOF. Let  $\mathbf{q}$  be a division with  $\bar{p}_{I+1,j} \leq q_j \leq \bar{p}_{I,j}$  for all  $j$ , with normalization of  $\mathbf{q}$  implying that  $\sum_{j \in [m]} \varepsilon_j = 1 - \sum_{j \in [m]} \bar{p}_{I+1,j}$ . Then the social cost of  $\mathbf{q}$  is

$$\begin{aligned} \sum_{j \in [m]} \sum_{i \in [n]} |\bar{p}_{i,j} - q_j| &= \sum_{j \in [m]} \left( \sum_{i \in [n]} |\bar{p}_{i,j} - \bar{p}_{I+1,j}| + \sum_{i \geq I+1} \varepsilon_j - \sum_{i \leq I} \varepsilon_j \right) \\ &= \sum_{j \in [m]} \sum_{i \in [n]} |\bar{p}_{i,j} - \bar{p}_{I+1,j}| + (n - 2I) \left( 1 - \sum_{j \in [m]} \bar{p}_{I+1,j} \right) \end{aligned}$$

Because this expression does not depend on  $\varepsilon_j$ , all such divisions  $\mathbf{q}$  have the same social cost.

We now show that this distance is minimal. Let  $\mathbf{q}$  be a division that does not satisfy  $\bar{p}_{I+1,j} \leq q_j \leq \bar{p}_{I,j}$  for some  $j$ . Suppose  $q_j > \bar{p}_{I,j}$  (the case where  $q_j < \bar{p}_{I+1,j}$  can be handled similarly). By the definition of  $I$ , there must exist some alternative  $j'$  for which  $q_{j'} < \bar{p}_{I,j'}$ . Now, consider the division  $\mathbf{q}'$  defined by  $q'_j = q_j - \varepsilon > \bar{p}_{I,j}$  and  $q'_{j'} = q_{j'} + \varepsilon < \bar{p}_{I,j'}$ , with  $\mathbf{q}'$  and  $\mathbf{q}$  equal on all other coordinates. Compare  $\mathbf{q}$  and  $\mathbf{q}'$  in terms of  $\ell_1$  distance from the reports. They are indistinguishable on all alternatives other than  $j$  and  $j'$ . On alternative  $j$ ,  $\mathbf{q}'$  is  $\varepsilon$  closer than  $\mathbf{q}$  to all entries  $\bar{p}_{i,j}$  with  $i \geq I$ , and at most  $\varepsilon$  farther from all other entries. On alternative  $j'$ ,  $\mathbf{q}'$  is  $\varepsilon$  closer than  $\mathbf{q}$  to all entries  $\bar{p}_{i,j'}$  with  $i \leq I$ , and at most  $\varepsilon$  farther from all other entries. Therefore, of the  $2n$  entries on alternatives  $j$  and  $j'$ ,  $\mathbf{q}'$  is  $\varepsilon$  closer than  $\mathbf{q}$  to at least  $n + 1$  of them, and no more than  $\varepsilon$  farther than  $\mathbf{q}$  from the other  $n - 1$ . Therefore,  $\mathbf{q}$  does not maximize social welfare.  $\square$

As a corollary of Lemma 6.3, we immediately obtain that  $\mathcal{A}^{\mathcal{F}}$  maximizes social welfare, and is therefore Pareto optimal. Social welfare-maximizing mechanisms have been considered before; all that is needed is a suitable tiebreaking procedure to select a single division from the set of maximizers. Goel et al. [2019] suggest breaking ties by selecting the lexicographically largest welfare maximizer, but this is not neutral. Lindner [2011] proposes a different way to break ties, which is neutral: select the welfare maximizer  $\mathbf{p}$  that is closest to the uniform distribution according to  $\ell_2$  distance  $\sum_{j \in [m]} (p_j - \frac{1}{m})^2$ . As we now show, it turns out that  $\mathcal{A}^{\mathcal{F}}$  implements this tiebreaking method.

THEOREM 6.4. *For every profile  $\mathbf{P}$ ,  $\mathcal{A}^{\mathcal{F}}$  selects the division that minimizes  $\ell_2$  distance to the uniform division, among those that maximize social welfare.*

PROOF. From the earlier discussion, we know that  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_j = \text{med}\{\bar{p}_{I+1,j}, \bar{p}_{I,j}, f_{n-I}(t^*)\} \in [\bar{p}_{I+1,j}, \bar{p}_{I,j}]$  for all  $j \in [m]$ . Therefore, it maximizes social welfare.

It remains to show that, subject to these constraints,  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})$  minimizes  $\ell_2$  distance to the uniform division. Consider any other division  $\mathbf{q} \neq \mathcal{A}^{\mathcal{F}}(\mathbf{P})$  with  $\bar{p}_{I+1,j} \leq q_j \leq \bar{p}_{I,j}$ . Then there must exist an alternative  $j$  for which  $\bar{p}_{I+1,j} \leq \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j < q_j \leq \bar{p}_{I,j}$ . Further, because  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_j < \bar{p}_{I,j}$ , it must be the case that  $f_{n-I}(t^*) \leq \mathcal{A}^{\mathcal{F}}(\mathbf{P})_j = \text{med}\{\bar{p}_{I+1,j}, \bar{p}_{I,j}, f_{n-I}(t^*)\}$ . There must also be an alternative  $j'$  for which  $\bar{p}_{I+1,j'} \leq q_{j'} < \mathcal{A}^{\mathcal{F}}(\mathbf{P})_{j'} \leq \bar{p}_{I,j'}$ , with  $\mathcal{A}^{\mathcal{F}}(\mathbf{P})_{j'} \leq f_{n-I}(t^*)$ .

Putting these together, we have that  $q_j < f_{n-I}(t^*) < q_{j'}$ . We also know that  $\bar{p}_{I+1,j} < q_j$  and  $q_{j'} < \bar{p}_{I,j'}$ . Therefore, adjusting  $q_j$  to  $q_j - \varepsilon$  and  $q_{j'}$  to  $q_{j'} + \varepsilon$ , for  $\varepsilon$  small enough that none of the above strict inequalities are violated, both (1) reduces  $\ell_2$  distance to uniform, and (2) respects social-welfare maximization. The former is immediately clear when  $q_{j'} < \frac{1}{m}$  and  $q_j > \frac{1}{m}$ , and

follows from strict convexity of the function  $x \mapsto x^2$  when either  $\frac{1}{m} \leq q_j$  or  $q_j < \frac{1}{m}$ . Therefore,  $\mathbf{q}$  is not the unique  $\ell_2$  distance-minimizing division among social welfare maximizers, so  $\mathcal{A}^{\mathcal{F}^s}$  is.  $\square$

We note that the tiebreaking method of Lindner [2011] is equivalent to choosing the welfare-maximizing division with largest Shannon entropy  $-\sum_{j \in [m]} p_j \log p_j$ . The proof of Theorem 6.2 goes through almost identically, replacing the convexity of the function  $x \mapsto x^2$  by the convexity of  $x \mapsto x \log x$ .

Like independent markets,  $\mathcal{A}^{\mathcal{F}^s}$  satisfies participation. The argument proceeds along exactly the same lines. For an individual voter, participating and reporting the existing aggregate has no effect, while reporting truthfully is at least as beneficial, by incentive compatibility.

**THEOREM 6.5.**  *$\mathcal{A}^{\mathcal{F}^s}$  satisfies participation.*

The utilitarian mechanism also satisfies reinforcement, which follows from the additivity of social welfare.

**THEOREM 6.6.**  *$\mathcal{A}^{\mathcal{F}^s}$  satisfies reinforcement.*

**PROOF.** Let  $\mathbf{P}$  and  $\mathbf{R}$  be profiles with  $\mathcal{A}^{\mathcal{F}^s}(\mathbf{P}) = \mathcal{A}^{\mathcal{F}^s}(\mathbf{R}) = \mathbf{q}$ . Let  $\mathbf{q}'$  be some division with higher entropy than  $\mathbf{q}$ . Then it must be the case that  $d(\mathbf{q}', \mathbf{P}) > d(\mathbf{q}, \mathbf{P})$  and  $d(\mathbf{q}', \mathbf{R}) > d(\mathbf{q}, \mathbf{R})$ . Since the  $\ell_1$  distance is additive,  $d(\mathbf{q}', \mathbf{P} \cup \mathbf{R}) > d(\mathbf{q}, \mathbf{P} \cup \mathbf{R})$ .

Now let  $\mathbf{q}'$  be some division with  $d(\mathbf{q}', \mathbf{P} \cup \mathbf{R}) < d(\mathbf{q}, \mathbf{P} \cup \mathbf{R})$ . Then  $d(\mathbf{q}', \mathbf{P}) + d(\mathbf{q}', \mathbf{R}) < d(\mathbf{q}, \mathbf{P}) + d(\mathbf{q}, \mathbf{R})$ . But this is impossible, since  $\mathbf{q}$  maximizes social welfare for profiles  $\mathbf{P}$  and  $\mathbf{R}$ .

Therefore, any division  $\mathbf{q}' \neq \mathbf{q}$  either has strictly lower social welfare on profile  $\mathbf{P} \cup \mathbf{R}$ , or equal social welfare but lower entropy. So  $\mathcal{A}^{\mathcal{F}^s}(\mathbf{P} \cup \mathbf{R}) = \mathbf{q}$ .  $\square$

## 7 CONCLUSION

We considered the problem of aggregating budget proposals for participatory budgeting. Inspired by the generalized median mechanisms of Moulin [1980], we introduced the broad class of moving phantom mechanisms and proved that all mechanisms in this class are incentive compatible under  $\ell_1$  voter preferences. We analyzed two moving phantom mechanisms in detail: one that maximizes social welfare while violating the natural fairness notion of proportionality, and another that satisfies proportionality while violating Pareto optimality.

In addition to the properties discussed in the main body of the paper, some other contrasts can be drawn between these two mechanisms. For example, suppose that  $n$  is large, and  $n/2$  voters report  $(1, 0)$  while the other  $n/2$  report  $(0, 1)$ . Then, under  $\mathcal{A}^{\mathcal{F}^s}$ , a single voter can completely dictate the outcome. Under independent markets, we can show that the ability of a single voter to affect the outcome vanishes as  $n \rightarrow \infty$ .

In many participatory budgeting applications, projects will have fixed costs, and so it does not make sense to direct only a small amount of funding to such a project. One can enrich our model by allowing for funding lower bounds, so that a project receives either no funding or at least its funding lower bound. However, in this model, one can show that there does not exist an anonymous, range respecting, and incentive-compatible mechanism [Peters 2019, Section 4.7].

It would be interesting to further analyze the utilitarian mechanism  $\mathcal{A}^{\mathcal{F}^s}$ . For example, can it be implemented using a voting game similar to the one we found for independent markets? Of course, there are many other moving phantom mechanisms that we have not considered. A natural question to investigate is what other properties can be achieved by other phantom systems. And finally, as we mentioned earlier, when there are only two outcomes, we know that all (anonymous, neutral, and continuous) incentive compatible budget aggregation mechanisms can be represented

as moving phantom mechanisms. It remains an open question whether this continues to hold when the number of outcomes  $m$  is more than 2.

Some other open questions are whether there exist incentive-compatible aggregation mechanisms when we define preferences with respect to the  $\ell_2$  or the  $\ell_\infty$  distance. Also, while generalized medians in the one-dimensional single-peaked setting are even *group* strategyproof, this is not in general true for moving phantom mechanisms, and it is unclear whether sensible group strategyproof mechanisms exist for budget aggregation with  $\ell_1$  preferences.

## ACKNOWLEDGEMENTS

We thank anonymous reviewers at JET and at the ACM EC 2019 conference for helpful suggestions that improved the presentation. We thank Klaus Nehring and Clemens Puppe for useful discussions, and Alexandra Osipov for pointing out a mistake in an earlier version. Compared to the working paper, Theorem 4.7, Proposition 5.4, and Theorem 5.7 are new. Dominik Peters was supported by EPSRC and by ERC under grant 639945 (ACCORD).

## REFERENCES

- S. Airiau, H. Aziz, I. Caragiannis, J. Kruger, J. Lang, and D. Peters. 2019. Portioning using ordinal preferences: fairness and efficiency. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*. 11–17.
- H. Aziz, A. Bogomolnaia, and H. Moulin. 2019. Fair mixing: the case of dichotomous preferences. In *Proceedings of the 20th ACM Conference on Economics and Computation (EC)*. 753–781.
- H. Aziz, B. E. Lee, and N. Talmon. 2018. Proportionally representative participatory budgeting: Axioms and algorithms. In *Proceedings of the 17th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. 23–31.
- H. Aziz and N. Shah. 2020. Participatory Budgeting: Models and Approaches. In *Pathways Between Social Science and Computational Social Science: Theories, Methods, and Interpretations*, Tamás Rudas and Gábor Péli (Eds.). Springer.
- S. Barberà, F. Gul, and E. Stacchetti. 1993. Generalized Median Voter Schemes and Committees. *Journal of Economic Theory* 61, 2 (1993), 262–289.
- S. Barberà and M. Jackson. 1994. A characterization of strategy-proof social choice functions for economies with pure public goods. *Social Choice and Welfare* 11, 3 (1994), 241–252.
- S. Barberà, J. Masso, and A. Neme. 1997. Voting under constraints. *Journal of Economic Theory* 76, 2 (1997), 298–321.
- G. Benade, A. D. Procaccia, S. Nath, and N. Shah. 2017. Preference Elicitation for Participatory Budgeting. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*. 376–382.
- A. Bogomolnaia, H. Moulin, and R. Stong. 2005. Collective Choice Under Dichotomous Preferences. *Journal of Economic Theory* 122, 2 (2005), 165–184.
- K. C. Border and J. S. Jordan. 1983. Straightforward elections, unanimity and phantom voters. *Review of Economic Studies* 50, 1 (1983), 153–170.
- G. Bordes, G. Laffond, and M. Le Breton. 2011. Euclidean preferences, option sets and strategyproofness. *SERIEs* 2, 4 (2011), 469–483.
- Y. Cabannes. 2004. Participatory Budgeting: A Significant Contribution to Participatory Democracy. *Environment and Urbanization* 16, 1 (2004), 27–46.
- I. Caragiannis, A. D. Procaccia, and N. Shah. 2016. Truthful Univariate Estimators. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)*. 127–135.
- R. T. Clemen. 1989. Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting* 5 (1989), 559–583.
- V. Conitzer, M. Brill, and R. Freeman. 2015. Crowdsourcing Societal Tradeoffs. In *Proceedings of the 14th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. 1213–1217.
- V. Conitzer, R. Freeman, M. Brill, and Y. Li. 2016. Rules for Choosing Societal Tradeoffs. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*. 460–467.
- C. Duddy. 2015. Fair sharing under dichotomous preferences. *Mathematical Social Sciences* 73 (2015), 1–5.
- E. Eisenberg and D. Gale. 1959. Consensus of Subjective Probabilities: The Pari-Mutuel Method. *The Annals of Mathematical Statistics* 30, 1 (1959), 165–168.
- B. Fain, A. Goel, and K. Munagala. 2016. The Core of the Participatory Budgeting Problem. In *Proceedings of the 12th Conference on Web and Internet Economics (WINE)*. 384–399.
- S. French. 1985. Group consensus probability distributions: A critical survey. *Bayesian Statistics* 2 (1985), 183–202.
- N. Garg, A. Goel, and B. Plaut. 2018. Markets for Public Decision-making. arXiv:1807.10836. (2018).

- N. Garg, V. Kamble, A. Goel, D. Marn, and K. Munagala. 2019. Iterative Local Voting for Collective Decision-making in Continuous Spaces. *Journal of Artificial Intelligence Research* 64 (2019), 315–355.
- C. Genest and J.V. Zidek. 1986. Combining Probability Distributions: A Critique and an Annotated Bibliography. *Statist. Sci.* 1, 1 (1986), 114–148.
- A. Goel, A. K. Krishnaswamy, S. Sakshuwong, and T. Aitamurto. 2019. Knapsack voting for participatory budgeting. *ACM Transactions on Economics and Computation (TEAC)* 7, 2 (2019).
- A. Guerdjikova and K. Nehring. 2014. Weighing Experts, Weighing Sources: The Diversity Value. (2014). Working Paper.
- A. Gulati, G. Shanmuganathan, X. Zhang, and P.J. Varman. 2012. Demand Based Hierarchical QoS Using Storage Resource Pools.. In *USENIX Annual Technical Conference*. 1–13.
- M.D. Intriligator. 1973. A probabilistic model of social choice. *The Review of Economic Studies* 40, 4 (1973), 553–560.
- K. H. Kim and F. W. Roush. 1984. Nonmanipulability in two dimensions. *Mathematical Social Sciences* 8, 1 (1984), 29–43.
- G. Laffond. 1980. *Révélation des Préférences et Utilités Unimodales*. Ph.D. Dissertation. Laboratoire d'économétrie du CNAM, Paris.
- T. Lindner. 2011. *Zur Manipulierbarkeit der Allokation öffentlicher Güter: Theoretische Analyse und Simulationsergebnisse*. KIT Scientific Publishing.
- Tobias Lindner, Klaus Nehring, and Clemens Puppe. 2008. Allocating Public Goods via the Midpoint Rule. In *Proceedings of the 9th International Meeting of the Society for Social Choice and Welfare*.
- T. Lu and C. Boutilier. 2011. Budgeted Social Choice: From Consensus to Personalized Decision Making. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*. 280–286.
- J. Massó and I. M. de Barreda. 2011. On strategy-proofness and symmetric single-peakedness. *Games and Economic Behavior* 72, 2 (2011), 467–484.
- H. Moulin. 1980. On strategy-proofness and single-peakedness. *Public Choice* 35 (1980), 437–455.
- K. Nehring and C. Puppe. 2019. *Resource allocation by frugal majority rule*. Working Paper Series in Economics 131. Karlsruhe Institute of Technology (KIT).
- Participatory Budgeting Project. 2019. Participatory Budgeting Project. <https://www.participatorybudgeting.org/what-is-pb/>. (2019). Accessed: 2019-02-12.
- D. Peters. 2019. *Fair Division of the Commons*. Ph.D. Dissertation. DPhil thesis, University of Oxford.
- H. Peters, H. van der Stel, and T. Storcken. 1992. Pareto optimality, anonymity, and strategy-proofness in location problems. *International Journal of Game Theory* 21, 3 (1992), 221–235.
- A. D. Procaccia and M. Tennenholtz. 2013. Approximate mechanism design without money. *ACM Transactions on Economics and Computation* 1, 4 (2013), article 18.
- D. J. Roberts and A. Postlewaite. 1976. The incentives for price-taking behavior in large exchange economies. *Econometrica* 44, 1 (1976), 115–127.
- A. Schneider and B. Goldfrank. 2002. Budgets and ballots in Brazil: Participatory budgeting from the city to the state. (2002). IDS Working Paper 149.
- A. Schrijver. 1986. *Theory of Linear and Integer Programming*. John Wiley & Sons.
- W. D. Smith. 2000. Range Voting. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.32.9150>. (2000).