Proportional Participatory Budgeting with Additive Utilities

#10089

Abstract

We study voting rules for participatory budgeting, where a group of voters collectively decides which projects should be funded using a common budget. We allow the projects to have arbitrary costs, and the voters to have arbitrary additive valuations over the projects. We formulate two axioms that guarantee proportional representation to groups of voters with common interests. To the best of our knowledge, all known rules for participatory budgeting do not satisfy either of the two axioms; in addition we show that the most prominent proportional rule for committee elections, Proportional Approval Voting, cannot be adapted to arbitrary costs nor to additive valuations so that it would satisfy our axioms of proportionality. We construct a simple and attractive voting rule that satisfies one of our axioms (for arbitrary costs and arbitrary additive valuations), and that can be evaluated in polynomial time. We prove that our other stronger axiom is also satisfiable, though by a computationally more expensive and less natural voting rule.

1 Introduction

Consider an abstract model where there is a group of agents who have preferences over a set of options. Each option has a cost, and the goal is to select a subset of options whose total cost does not exceed a predefined budget. This model provides a formal framework for a number of real-life scenarios. Perhaps the most natural example is Participatory Budgeting (PB). Through a voting system, PB allows residents of a city (the agents) to decide which projects (the options) will be funded by the government. In recent years, PB has been started in many cities around the world [Cabannes, 2004, Aziz and Shah, 2020], and in some cases is used to decide a significant fraction of the city budget. For example, in Paris, PB has been run every year since 2014, and since 2016 the total amount of funding for PB in Paris has been more than 100 million euros annually. Besides PB, the formal model captures the problem of electing a representative committee for a group of voters, say a faculty board or a parliament [Faliszewski et al., 2017, Lackner and Skowron, 2020], but also appears useful in situations that do not involve humans. For example, our framework describes the problem of selecting validators in consensus protocols, such as the blockchain [Cevallos and Stewart, 2020], the problem of selecting web pages that should be displayed in response to user queries, where the selected set of web pages should be useful for different types of user profiles [Skowron et al., 2017] or the problem of locating public facilities [Skowron et al., 2016, Byrka et al., 2018]. Algorithms for PB can even be used for improving the quality of genetic algorithms [Faliszewski et al., 2016]. While there are numerous applications of the model that we consider, for concreteness we will use terminology referring to participatory budgeting.

In this paper, we focus on designing algorithms (which we call aggregation rules) for selecting projects. We are interested in rules that are *fair* in the sense that each agent has roughly equal influence on the outcome.This implies that the selected subset of projects must proportionally represent the views of the voters, and so every group of voters with similar preferences should have an appropriate portion of the available budget allocated to fulfilling those preferences. Our results contribute to the study of fairness in algorithmic decision making. Notably, while many works on this topic aim to be fair to

Submitted to 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Do not distribute.

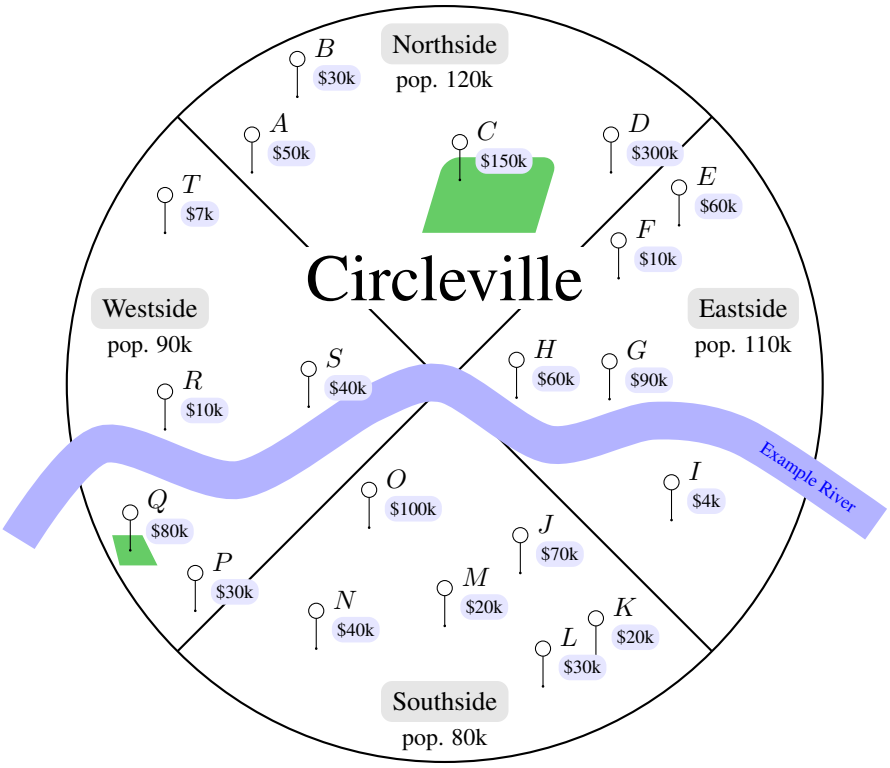


Figure 1: Map of Circleville, showing the locations and costs of the PB project proposals.

agents described by feature vectors (that for example contain demographic information), we only use agents' preferences. Fair algorithms will need to identify cohesive groups of agents on their own.

To understand the constraints that fairness and proportionality place on the decision procedure, let us start by discussing the way cities implement PB today. To count the votes, most cities use a variant of a simple protocol: Each voter is allowed to vote for a certain number of project proposals. Then, the projects with the highest number of votes are funded, until the budget limit is reached. While simple and intuitive, this is a bad voting rule. To see this, consider Circleville, a fictional city divided into four districts. A map of the city is shown in Figure 1. The districts all have similar sizes, but Northside has the largest population. Suppose \$400k have been allocated to PB, and suppose that all the project proposals are of a local character (such as school renovations), so we can assume that residents only vote for projects that concern their own district. For example, every Northside resident will cast votes for projects A, B, C, and D, but no one else votes for these. Because Northside is the most populous district, the Northside projects will all receive the highest number of votes, and the voting rule described will spend the entire budget on Northside projects. The 280k residents of the other districts are left empty-handed.

To circumvent this obvious issue, many cities have opted to hold separate elections for each district. The budget is divided in advance between the districts (e.g., in proportion to their number of residents), each project is assigned to a district, and voters only vote in their local election. While this avoids the issue of spending the entire budget in Northside, this fix introduces many other problems. For example, projects on the boundary of two districts (such as A and P) need to be assigned to one of them. Residents of the other district may be in favor of the boundary project, but cannot vote for it. Thus boundary projects are less likely to be funded, even if they would be more valuable overall. Similarly, projects without a specific location that benefit the entire city cannot be handled. Also, interest groups that are not geographic in nature will be underserved; for instance, parents across the city might favor construction of a large playground (project C), but with separate district elections, parents cannot form a voting block. Similarly, bike riders across the city cannot express their joint interest in the construction of a bike trail along Example River (projects R, S, H, and G).

To solve these problems, it seems desirable to hold a single city-wide election, but use a voting system that ensures that money is spent proportionally. The voting system should automatically and endogenously identify groups of voters who share common interests, and make sure that those groups are appropriately represented. This aim has been identified by several researchers [Aziz et al., 2018b], but no convincing proposal for a proportional voting rule has emerged so far. Indeed, no good formalization of "proportionality" for the PB context has been identified in the literature, except for the concept of the *core*. However, the core is a very demanding requirement, and there are situations where it fails to exist [Fain et al., 2018].

In this paper, we formalize proportionality for participatory budgeting as an axiom called *extended justified representation* (EJR). The axiom requires that no group of voters with common interests is underserved. We construct a simple and attractive voting rule that satisfies EJR for approval preferences, and that satisfies EJR up to one project for general additive valuations. We then discuss a potential strengthening of EJR—which we call *fully justified representation* (FJR)—and show that this strengthening is still satisfiable, albeit by a different voting rule.

Both our proportionality axiom and our voting rule are generalizations of concepts that have been introduced in the literature on multi-winner voting [Faliszewski et al., 2017]. That literature can be seen as handling a special case of PB, where all projects cost the same amount of money. This is often called the *unit cost assumption*. It turns out that the unit cost assumption substantially simplifies the problem. Further, much of the relevant literature studies rules that work with *approval ballots*, where voters are allowed to approve or disapprove each project. In our paper, we allow any additive valuations (not just 0/1), which is more expressive. Indeed, the proportionality axioms and voting rules that we introduce all work for general additive valuations. This is notable, since allowing additive valuations introduces significant conceptual difficulty. We provide formal arguments explaining that most prominent multi-winner voting rules do to not naturally extend to additive valuations (or at least not gracefully). Even though our main rule generalizes a known multi-winner voting rule, this generalization required developing new conceptual ideas and new proof techniques (our proofs are significantly different from those used in the literature on multi-winner voting).

2 Preliminaries

For each $t \in \mathbb{N}$, write $[t] = \{1, 2, \dots, t\}$. An *election* is a tuple $(N, C, \mathrm{cost}, \{u_i\}_{i \in N})$, where:

- $N = [n]$ and $C = \{c_1, \dots, c_m\}$ are the sets of *voters* and *candidates* (or *projects*).
- $\mathrm{cost}\colon C \to \mathbb{Q}_+$ is a function that for each $c \in C$ assigns the *cost* that needs to be paid if c is selected. For each $T \subseteq C$, we write $\mathrm{cost}(T) = \sum_{c \in T} \mathrm{cost}(c)$ for the total cost of T.
- For each voter $i \in N$, the function $u_i\colon C \to [0, 1]$ defines i's additive *utility function*. If a set $T \subseteq C$ of candidates is implemented, i's overall utility is $u_i(T) = \sum_{c \in T} u_i(c)$. For a subset $S \subseteq N$ of voters, we further write $u_S(T) = \sum_{i \in S} \sum_{c \in T} u_i(c)$ for the total utility enjoyed by S if T is implemented. We assume that $u_N(c) > 0$ for each $c \in C$, that is, every candidate is assigned positive utility by at least one voter.

The voters have a fixed common *budget* which we normalize to 1. A subset of candidates $W \subseteq C$ is *feasible* if $\mathrm{cost}(W) \leq 1$. Our goal is to choose a feasible subset of candidates, which we call an *outcome*, based on voters' utilities. An *aggregation rule* (or, in short, a *rule*) is a function that for each election returns a set of feasible outcomes, called the *winning outcomes*.[1]

There are two interesting special cases of our model:

Committee elections. In this case, there exists $k \in \mathbb{N}$ (the committee size) such that each candidate costs $1/k$. Then W is an outcome if and only if $|W| \leq k$. In this special case we also refer to outcomes as *committees*, and we say that the election satisfies the *unit cost assumption*.

Approval-based elections. In this case, for each $i \in N$ and $c \in C$ it holds that $u_i(c) \in \{0, 1\}$. The *approval set* of voter i is $A(i) := \{c \in C\colon u_i(c) = 1\}$, and we say that i *approves* candidate c if $c \in A(i)$. If $c \in A(i) \cap W$, we say that c is a *representative* of i.

Often we combine of these special cases, and study approval-based committee elections.

[1] Typically, a rule will be selecting a single winning outcome, but ties are possible. For the results of this paper it does not matter how these ties are broken.

3 Rule X

Recently, Peters and Skowron [2020] introduced an aggregation rule called *Rule X* for approval-based committee elections. In that setting, they showed that Rule X satisfies a combination of appealing proportionality properties. Here, we extend it to the more general model of participatory budgeting, that is, to the model with arbitrary costs and utilities.

Definition 1 (Rule X)**.** Each voter is initially given an equal fraction of the budget, i.e., each voter is given $1/n$ dollars. We start with an empty outcome $W = \emptyset$ and sequentially add candidates to W. To add a candidate c to W, we need the voters to pay for c. Write $p_i(c)$ for the amount that voter i pays for c; we will need that $\sum_{i \in N} p_i(c) = \mathrm{cost}(c)$. We write $p_i(W) = \sum_{c \in W} p_i(c) \leq \frac{1}{n}$ for the total amount i has paid so far. For $\rho \geq 0$, we say that a candidate $c \notin W$ is *ρ-affordable* if

$$\sum_{i \in N} \min\left(\tfrac{1}{n} - p_i(W), u_i(c) \cdot \rho\right) = \mathrm{cost}(c).$$

If no candidate is ρ-affordable for any ρ, Rule X terminates and returns W. Otherwise it selects a candidate $c \notin W$ that is ρ-affordable for a minimum ρ. Individual payments are given by

$$p_i(c) = \min\left(\tfrac{1}{n} - p_i(W), u_i(c) \cdot \rho\right) \qquad \square$$

Intuitively, when Rule X adds a candidate c, it asks voters to pay an amount proportional to their utility $u_i(c)$ for c; in particular, the cost per unit of utility is ρ. If a voter does not have enough money, the rule asks the voter to pay all the money the voter has left, which is $\frac{1}{n} - p_i(W)$. Throughout the execution of Rule X, the value of ρ increases. Thus, candidates are added in decreasing order of utility per dollar that the voters get from the candidates. In comparison to the work of Peters and Skowron [2020], the new elements in our definition of the rule are (1) the formula according to which the costs of selected projects are divided among the voters, and (2) the algorithm specifying in which order the candidates are selected; these are critical to ensure that the rule is proportional.

3.1 Extended Justified Representation (EJR)

The first notion of proportionality that we examine is Extended Justified Representation (EJR). This axiom was first proposed for approval-based committee elections [Aziz et al., 2017]. Even for the special case of approval-based committee elections, only few rules are known to satisfy EJR [Aziz et al., 2017, 2018a, Peters and Skowron, 2020], but Rule X is one of them. In this section, we introduce a generalization of EJR to the PB model and show that our generalization of Rule X continues to satisfy EJR.

We first recall the definition of EJR for approval-based committee elections. Intuitively, this axiom ensures that every large enough group of voters whose approval sets have a large enough intersection must obtain a fair number of representatives. For example, if a group of voters forms an α-fraction of the whole population and if this group agrees on sufficiently many candidates, then it should be allowed to decide about an α-fraction of the elected candidates. Formally, this is achieved by excluding the possibility that each member of the group approves less than $\lfloor \alpha k \rfloor$ elected candidates.

Definition 2 (Extended Justified Representation for approval-based committee elections)**.** We say that a group of voters S is *ℓ-cohesive* for $\ell \in \mathbb{N}$ if $|S| \geq \ell/k \cdot n$ and $|\bigcap_{i \in S} A(i)| \geq \ell$. A rule $\mathcal{R}$ satisfies *extended justified representation* if for each election instance E and each ℓ-cohesive group S of voters there exists a voter $i \in S$ such that $|A(i) \cap \mathcal{R}(E)| \geq \ell$.

At first sight it is unintuitive that we only require that at least one voter obtain ℓ representatives. However, the strengthening of EJR that requires each member of S to obtain ℓ representatives is impossible even on very small instances [Aziz et al., 2017]. Still, even with only the at-least-one guarantee, EJR has plenty of bite [Aziz et al., 2018a, Skowron, 2018, Peters and Skowron, 2020].

The generalization of this axiom to the PB model is not straightforward and to the best of our knowledge none has been proposed in the literature.[2] To warm up, let's first relax the unit cost assumption, but stay in the approval-based setting. Then EJR should state the following.

[2] Aziz et al. [2018b] generalize the weaker axiom of Proportional Justified Representation (PJR) [Sánchez-Fernández et al., 2017] beyond unit costs, but they operate in a non-standard utility model where voters care more about more expensive projects.

Definition 3 (Extended Justified Representation for approval-based elections)**.** We say that a group of voters S is T-cohesive for $T \subseteq C$ if $|S| \geq \mathrm{cost}(T) \cdot n$ and $T \subseteq \bigcap_{i \in S} A(i)$. A rule $\mathcal{R}$ satisfies *extended justified representation* if for each election instance E and each T-cohesive group S of voters there exists a voter $i \in S$ such that $|A(i) \cap \mathcal{R}(E)| \geq |T|$.

Thus, cohesiveness now requires that the group S can identify a collection of projects T that they all approve and that is affordable with their fraction of the budget ($|S| \geq \mathrm{cost}(T) \cdot n$). Note that voters $i \in S$ obtain utility $u_i(T) = |T|$ from T; EJR requires that at least one member of S must attain this utility in the election outcome.

To further generalize EJR beyond approvals is more difficult, because the notion of a candidate who is approved by all members of S does not have an analogue. Instead, we quantify cohesion by calculating the minimum utility that any member of S assigns to each project in T.

Definition 4 (Extended Justified Representation)**.** A group of voters S is (α, T)-*cohesive*, where $\alpha \colon C \to [0; 1]$ and $T \subseteq C$, if $|S| \geq \mathrm{cost}(T) \cdot n$ and if $u_i(c) \geq \alpha(c)$ for all $i \in S$ and $c \in T$. A rule $\mathcal{R}$ satisfies *extended justified representation* if for each election instance E and each (α, T)-cohesive group of voters S there exists a voter $i \in S$ such that $u_i(\mathcal{R}(E)) \geq \sum_{c \in T} \alpha(c)$.

Again, an (α, T)-cohesive group of voters S can propose the projects in T, since they are affordable with S's share of the budget. The values $(\alpha(c))_{c \in T}$ denote how much the coalition S agrees about the desirability of the projects in T. In particular, we have $u_i(T) \geq \sum_{c \in T} \alpha(c)$ for each $i \in S$. Consequently, Definition 4 prohibits any outcome in which every voter in S gets utility strictly lower than $\sum_{c \in T} \alpha(c)$; hence there must exists $i \in S$ such that $u_i(\mathcal{R}(E)) \geq \sum_{c \in T} \alpha(c)$.

We do not know if Rule X satisfies EJR in the general PB model. However, we can show that it satisfies a mild relaxation, which requires EJR to hold "up to one project".

Definition 5 (Extended Justified Representation Up To 1 Project)**.** A rule $\mathcal{R}$ satisfies extended justified representation *up to one project* if for each election instance E and each (α, T)-cohesive group of voters S there exists a voter $i \in S$ such that either $u_i(\mathcal{R}(E)) \geq \sum_{c \in T} \alpha(c)$ or for some $a \in C$ it holds that $u_i(\mathcal{R}(E) \cup \{a\}) > \sum_{c \in T} \alpha(c)$.

It is worth noting that in the approval-based model, Definitions 4 and 5 are actually equivalent, because the "up to one project" option never applies: Consider an (α, T)-cohesive group of voters S. Since voters' utilities are 0/1, we may assume that for each $c \in T$ we have $\alpha(c) = 1$: if $\alpha(c) > 0$ this is clear; otherwise we can remove c from T without losing cohesiveness. Thus, the cohesiveness condition is equivalent to the condition that every voter approves every candidate in T. Finally, note that in the approval model, due to the strict inequality, both conditions $u_i(\mathcal{R}(E)) \geq \sum_{c \in T} \alpha(c)$ and $\exists_{a \in C}. u_i(\mathcal{R}(E) \cup \{a\}) > \sum_{c \in T} \alpha(c)$ boil down to $|A(i) \cap \mathcal{R}(E)| \geq \sum_{c \in T} \alpha(c) = |T|$.

Our main result is that Rule X satisfies EJR up to one project, and hence it satisfies EJR in the approval-based model.

Theorem 1. *Rule X satisfies EJR up to one project in the participatory budgeting model.*

Proof. For the sake of contradiction, assume that there exists an election E, a subset of voters $S \subseteq N$ and a set of candidates $T \subseteq C$ such that: (i) $\mathrm{cost}(T) \leq {}^{|S|}/_{n}$, (ii) $u_i(c) \geq \alpha(c) > 0$ (candidates with $\alpha(c) = 0$ can be skipped) for each $i \in S$ and $c \in T$, and (iii) $u_i(\mathcal{R}(E) \cup \{a\}) \leq \sum_{c \in T} \alpha(c)$ for each $i \in S$ and $a \in T$.

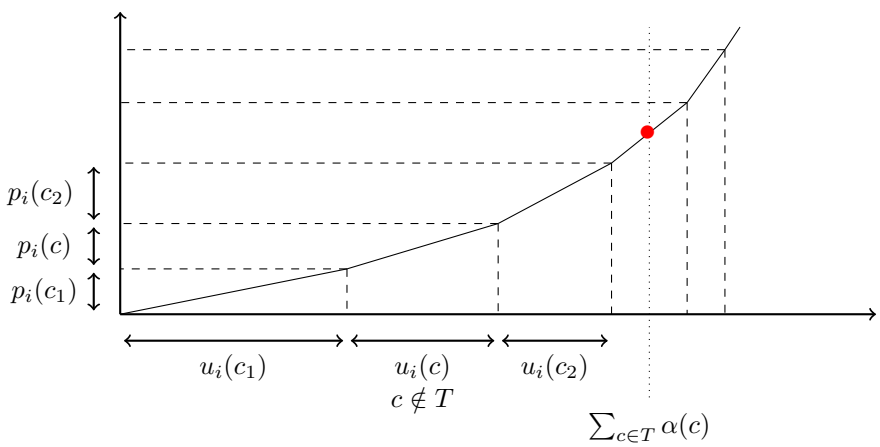
Assume for a while that the voters from S have unrestricted initial budgets, and let us analyze how Rule X would proceed in such a case. For simplicity of presentation, without loss of generality, let us rename the candidates in T so that $T = \{c_1, \ldots, c_t\}$ and so that for $1 \leq i < j \leq t$ candidate c_i is picked by Rule X before candidate c_j.

Whenever a candidate $c \in T$ is selected, the voters pay for this candidate. Voter i pays $p_i(c)$ dollars for c, and in return, she gets utility $u_i(c)$. Thus, the price-per-utility she pays equals $\rho_i(c) = {}^{p_i(c)}/_{u_i(c)}$. Rule X works in a way that all voters from S who pay for c obtain the same price-per-utility ratio, i.e., for all $i, j \in S$ and $c \in C$ we have that $\rho_i(c) = \rho_j(c)$. Further, this price-per-utility equals at most ${}^{\mathrm{cost}(c)}/_{u_S(c)}$, independently of whether any voters from $N \setminus S$ pay for c or not (if no voters from

$N \setminus S$ pays for c, then the price-per-utility equals exactly $\mathrm{cost}(c)/u_S(c)$):

$$\rho_i(c) = \frac{p_i(c)}{u_i(c)} = \frac{p_i(c) \cdot \frac{\sum_{j \in S} u_j(c)}{\sum_{j \in S} u_j(c)}}{u_i(c)} = \frac{1}{\sum_{j \in S} u_j(c)} \cdot \sum_{j \in S} \frac{p_i(c)}{u_i(c)} \cdot u_j(c) = \frac{\sum_{j \in S} p_j(c)}{\sum_{j \in S} u_j(c)} \leq \frac{\mathrm{cost}(c)}{u_S(c)}.$$

Since $u_i(c) \geq \alpha(c)$ for each $i \in S$ and $c \in T$, the price-per-utility for $c \in T$ equals at most $\mathrm{cost}(c)/|S|\alpha(c)$. Now, consider the voter who in the first possible iteration uses more than $1/n$ dollars[3]. For this voter, call her i, let us consider the function f defined as follows. For each value x, the function f returns the price that i needs to pay to achieve the utility x. We make this function continuous, by assuming that the candidates are divisible. That is, if the voter pays p for her first paid candidate c with utility $u_i(c)$, then $f(u_i(c)/2) = p/2$, $f(u_i(c)/3) = p/3$, and so on. The key observation is that the function f is convex. This is because Rule X selects the candidates in increasing order of price-per-utility. This function is depicted below.



We are interested in the value $f(\sum_{c \in T} \alpha(c))$. This value would be maximized if the fragments of the function with the lowest slope were the shortest. However, we know that the part of the function that corresponds $\rho_i(c)$ must be of length at least equal to $u_i(c) \geq \alpha(c)$. Thus:

$$f\left(\sum_{c \in T} \alpha(c)\right) \leq \sum_{c \in T} \alpha(c) \cdot \rho_i(c) \leq \sum_{c \in T} \alpha(c) \cdot \frac{\mathrm{cost}(c)}{|S|\alpha(c)} = \sum_{c \in T} \frac{\mathrm{cost}(c)}{|S|} = \frac{\mathrm{cost}(T)}{|S|} \leq \frac{1}{n}.$$

Now, consider the first moment when i uses more than $1/n$ dollars. Until this time moment, Rule X behaves exactly in the same way as if the voters from S had their initial budgets set to $1/n$ (this follows from how we chose i). Further, we know that in this moment, if we chose a candidate that would be chosen if the voters had unrestricted budgets, then the utility of voter i would be greater than $\sum_{c \in T} \alpha(c)$. This gives a contradiction, and completes the proof. □

Theorem 1 establishes Rule X as a prime candidate for voting under a budget constraint, showing that it satisfies a demanding fairness property. This makes it the first known rule that can give such strong proportionality guarantees in the model with additive utilities and arbitrary costs. In the literature on the special case of approval-based committee elections, another rule has received much attention: Proportional Approval Voting (PAV). This rule satisfies EJR when assuming unit costs. But we can show by an example that PAV badly violates EJR when we go beyond unit costs. Let us briefly recall the definition of this rule.

Definition 6 (Proportional Approval Voting (PAV))**.** Consider an approval-based election. PAV selects a feasible outcome that maximizes $\sum_{i \in N} \mathrm{H}(|A(i) \cap W|)$, where $\mathrm{H}(r)$ is the r-th harmonic number, i.e., $\mathrm{H}(r) = \sum_{j=1}^{r} 1/j$.

[3]The only case when there is no such voter is when every candidate $c \in C$ with $u_S(c) > 0$ has already been elected before. But then the utility of i from the elected outcome is clearly at least $\sum_{c \in T} u_i(c) \geq \sum_{c \in T} \alpha(c)$.

Example 1 below constructs an instance on which PAV does not satisfy EJR. In fact, this example shows that PAV does not even approximate EJR up to a constant factor.

Example 1. Fix a constant $r \in \mathbb{N}$ ($r \geq 2$), and consider the following approval-based profile:

$$\begin{array}{ll} r^2 - 1 \text{ voters:} & \{a_1, a_2, \ldots, a_r\}, \\ 1 \text{ voter:} & \{b_1, b_2, \ldots, b_r\}. \end{array}$$

The candidates $a_1, a_2, \ldots, a_r$ cost $1/r$ dollars each; the candidates $b_1, b_2, \ldots, b_r$ cost $1/r^3$ dollars each. EJR requires that the one voter who approves candidates $b_1, \ldots, b_r$ must approve at least r candidates in the outcome. However, PAV selects $\{a_1, a_2, \ldots, a_r\}$, leaving the voter with nothing. □

In fact, in Appendix A, we prove that no rule that, like PAV, globally maximizes an objective function over voter utilities can satisfy proportionality. Further, in Appendix C we argue that another rule for proportional approval-based committee elections, Phragmén's rule, does not extend to the PB setting.

3.2 Approximating the Core

An important proportionality property that has been proposed for PB is the *core* [Aziz et al., 2017, Fain et al., 2018]. It strengthens EJR by removing the cohesiveness requirement on groups of voters.

Definition 7 (The Core)**.** For a given election instance $(N, C, \text{cost}, \{u_i\}_{i \in N})$, we say that an outcome W is in the *core*, if for every $S \subseteq N$ and $T \subseteq C$ with $|S| \geq \text{cost}(T) \cdot n$ there exists $i \in S$ such that $u_i(W) \geq u_i(T)$. We say that an election rule $\mathcal{R}$ satisfies the *core property* if for each election instance E the winning outcome $\mathcal{R}(E)$ is in the core.

While this is a clean and appealing notion, unfortunately, there are elections where no outcome is in the core, even with unit costs.

Example 2. [4] We have 6 voters and 6 candidates with unit costs, and $k = 3$. Utilities satisfy

$$\begin{array}{lll} u_1(c_1) > u_1(c_2) > 0, & u_2(c_2) > u_2(c_3) > 0, & u_3(c_3) > u_3(c_1) > 0; \\ u_4(c_4) > u_4(c_5) > 0, & u_5(c_5) > u_5(c_6) > 0, & u_6(c_6) > u_6(c_4) > 0, \end{array}$$

and all unspecified utilities are equal to 0. Let $W \subseteq C$ be any feasible outcome, so $|W| \leq 3$. Then either $|W \cap \{c_1, c_2, c_3\}| \leq 1$ or $|W \cap \{c_4, c_5, c_6\}| \leq 1$. Without loss of generality assume the former, and again without loss assume that $c_2 \notin W$ and $c_3 \notin W$. Then $S = \{v_2, v_3\}$ and $T = \{c_3\}$ block W and show it is not in the core, since $2 = |S| \geq \text{cost}(T) * n = \frac{1}{3} \cdot 6 = 2$ and both $u_2(c_3) > u_2(c_1) \geq u_2(W)$ and $u_3(c_3) > u_3(c_1) \geq u_3(W)$. □

Notably, this example is not approval-based. It is unknown whether the core is always non-empty for approval-based elections (with or without the unit cost assumption).

In the committee context, Peters and Skowron [2020] showed that Rule X returns an outcome that never violates the core too badly; formally, Rule X provides a multiplicative approximation to the core.[5] We can generalize this result to the general PB setting: Rule X continues to provide a good multiplicative approximation to the core property.

Definition 8. For $\alpha \geq 1$, we say that an outcome is in the *α-core* if for every $S \subseteq N$ and $T \subseteq C$ with $\sum_{c \in T} \text{cost}(c) \leq |S|/n$ there exists $i \in S$ and $c \in T$ such that $u_i(\mathcal{R}(E) \cup \{c\}) \geq \frac{u_i(T)}{\alpha}$.

Theorem 2. *Given an election E, by $u_{\max}$ we denote the highest utility a voter can get from a feasible outcome. Analogously, by $u_{\min}$ we denote the smallest, yet still positive utility a voter can get from a feasible outcome:*

$$u_{\max} = \max_{i \in N} \max_{\text{cost}(W) \leq 1} u_i(W) \qquad \textit{and} \qquad u_{\min} = \min_{i \in N} \min_{u_i(W) > 0} u_i(W).$$

Rule X satisfies the α-core property for $\alpha = 4 \log(2 \cdot u_{\max}/u_{\min})$.

All proofs omitted from the main text appear in Appendix D. The bound of $\alpha = 4 \log(2 \cdot u_{\max}/u_{\min})$ is asymptotically tight, and the hard instance can be constructed even for the approval-based committee-election model (there, $u_{\max}/u_{\min} \leq k$).

[4] This example is adapted from Fain et al. [2018, Appendix C].

[5] Our approximation notion is different from the one proposed by Fain et al. [2018] and the one proposed by Cheng et al. [2019] and Jiang et al. [2020].

Let us finally mention that in the context of approval-based committee elections EJR can be viewed as a weakening of the core. The core prohibits situations where a group of voters S can propose a sufficiently small set of candidates T such that each voter from S prefers T to the committee at hand. EJR additionally specifies that the voters can deviate only if they agree on the proposed set of candidates T. Intuitively, while the core prohibits any rational deviation, EJR prohibits only deviations towards outcomes that are unanimously approved, thus, in a way the deviations that the voters can most easily agree on. A similar relation holds between the two properties defined for additive utilities. This is yet another explanation why EJR is an important and well-justified axiom.

3.3 Other properties of Rule X

In Appendices B and C we discuss two other important properties—priceability and exhaustiveness. Let us briefly discuss the latter one. Exhaustiveness requires that a voting rule spends its entire budget. Of course, due to the discrete model, we cannot guarantee that the rule will spend exactly 1 dollar (i.e., the entire budget); however, we can require that no additional project is affordable.

Definition 9 (Exhaustiveness, Aziz et al., 2018b)**.** An election rule $\mathcal{R}$ is *exhaustive* if for each election instance E and each non-selected candidate $c \notin \mathcal{R}(E)$ it holds that $\mathrm{cost}(\mathcal{R}(E) \cup \{c\}) > 1$.

Notably, Rule X fails to be exhaustive. Yet, in Appendix C we discuss a few modifications which make the rule exhaustive. Further in Appendix H we experimentally compare these modifications.

4 Greedy Cohesive Rule

In Section 3 we discussed the EJR axiom for the PB model, and saw that it is implemented by Rule X. We will now propose a strengthening of EJR, and show a rule that satisfies the new strong property. Interestingly, even in the approval-based committee-election model our new property is substantially stronger than EJR, and hence this new rule provides the strongest known proportionality guarantees. On the other hand, compared to Rule X, it is computationally expensive and arguably less natural.

4.1 Full Justified Representation (FJR)

Our new proportionality axiom strengthens EJR by weakening its requirement that groups must be cohesive. Thus, the new axiom guarantees representation to groups that are only partially cohesive.

Definition 10 (Full Justified Representation (FJR))**.** We say that a group of voters S is *weakly* (β, T)*-cohesive* for $\beta \in \mathbb{R}$ and $T \subseteq C$, if $|S| \geq \mathrm{cost}(T) \cdot n$ and $u_i(T) \geq \beta$ for every voter $i \in S$. A rule $\mathcal{R}$ satisfies *full justified representation (FJR)* if for each election instance E and each weakly (β, T)-cohesive group of voters S there exists a voter $i \in S$ such that $u_i(\mathcal{R}(E)) \geq \beta$.

In the approval-based committee-election model, FJR boils down to the following requirement: Let S be a group of voters, and suppose that each member of S approves at least β candidates from some set $T \subseteq C$ with $|T| \leq \ell$, and let $|S| \geq {}^{\ell}/_{k} \cdot n$. Then at least one voter from S must have at least β representatives in the committee. It is clear that in the special case of $\beta = \ell$, we obtain Definition 2, hence FJR implies EJR. The same implication holds in the general PB model.

Proposition 1. *FJR implies EJR in the general PB model.*

It is easy to see that FJR is implied by the core property (cf. Definition 7). It is related to, but stronger than, some other relaxations of the core discussed by Peters and Skowron [2020, Section 5.2].

Since no known aggregation rule satisfies FJR, one might wonder whether FJR existence can be guaranteed. It turns out that it can: we present a rule satisfying this strong notion of proportionality.

Definition 11 (Greedy Cohesive Rule (GCR))**.** The *Greedy Cohesive Rule* (GCR) is defined sequentially as follows: we start with an empty outcome $W = \emptyset$. At each step, we search for a value $\beta > 0$, a set of voters $S \subseteq N$, and a set of candidates $T \subseteq C$ such that S is weakly (β, T)-cohesive. If such values of β, S, and T do not exist, then we stop and return W. Otherwise, we pick values of β, S, and T that maximize β, breaking ties in favor of smaller $\mathrm{cost}(T)$.[6] We add all the candidates from T to W, remove all voters in S from the election and repeat the search.

[6] This way of breaking ties is important for our analysis of the priceability of GCR in Appendix F.1. The proof of Theorem 3 does not use it, so GCR satisfies FJR for any way of breaking ties.

Let us first check that the Greedy Cohesive Rule always selects a feasible outcome (i.e., that does not exceed the budget limit). Indeed, whenever the algorithm adds some set T to W, then by definition of weakly cohesive groups, we have $|S| \geq \mathrm{cost}(T) \cdot n$, and hence it removes at least $\mathrm{cost}(T) \cdot n$ voters after this step. Thus, if GCR selects an outcome with total cost $\mathrm{cost}(W)$, then it must have removed at least $\mathrm{cost}(W) \cdot n$ voters during its execution. Hence $\mathrm{cost}(W) \leq 1$.

Theorem 3. *Greedy Cohesive Rule satisfies FJR.*

In Appendix F we discuss other important properties as well as some drawbacks of GCR.

5 Conclusion

In this paper, we have formulated two axioms, EJR and FJR, that capture the idea of proportionality in the participatory budgeting (PB) model. We have argued that none of the prominent committee election rules extend to the PB model so that it would satisfy even much weaker forms of proportionality. We have designed a simple and natural rule for the PB model, Rule X. It satisfies EJR and other proportionality-related properties, and it is computable in polynomial time. The stronger of our two properties, FJR, is also satisfiable, albeit by a different and arguably less natural voting rule. It is an interesting open question whether there exists a natural voting rule that satisfies FJR and shares other desirable properties of Rule X.

There are numerous results which are not included in the main text. Specifically, in Appendices B and C we discuss other properties of Rule X, and in Appendix F we provide a more detailed analysis of GCR. In Appendix G we explain that our rules can be directly applied to a model where voters have ordinal preferences, that is, when they express their preferences by ranking the candidates from the most to the least preferred one. Notably, in the ordinal model Rule X satisfies the axiom of proportionality for solid coalitions, which is perhaps the strongest known axiom of proportionality for ordinal preferences.

In Appendix H, we present experiments comparing different variants of Rule X, Phragmén's rule, and PAV. These experiments are based on real data from participatory budgeting elections carried out in several major Polish cities (we looked at nine different elections). In our first experiment, we compared three different strategies of making Rule X exhaustive. We observed substantial differences between different variants of Rule X. We conclude that Rule X gives a lot of flexibility to a mechanism designer, because it often selects outcomes that do not spend all of the budget, while still satisfying strong fairness requirements like EJR. Depending on the specific objectives, a mechanism designer can choose to complete this outcome using different strategies.

In our second experiment we compared Rule X, Phragmén's rule, and PAV. We observed that for approval-based utilities the results returned by Rule X and Phragmén's rule are comparably good, both in terms of the total utility obtained by the voters and in terms of proportionality. On the other hand, if we take a model with more fine-grained utilities, the difference between the two rules becomes apparent. This difference is unsurprising since Phragmén's rule does not take into account the more fine-grained information on utilities, but operates only on approval ballots. On the other hand, our results suggest that there is indeed a considerable advantage of using rules (like Rule X) that take into account the full information contained in cardinal additive utilities. We conclude that Rule X performs as well as Phragmén's rule for approval ballots and outperforms it when more fine-grained information on voters' utilities are available. Somewhat surprisingly, we show that the sequential variant of PAV produces highly disproportional outcomes compared to Phragmén's rule and Rule X. Altogether, our experiments confirm our theoretical results and suggest that Rule X outperforms the other two rules in terms of proportionality and/or efficiency.

As we discussed in the introduction, many cities run PB by dividing the overall budget between districts and running separate elections in each district. In particular, this is true in the Polish cities that provide our experimental election data. We claimed in the introduction that this practice of separate elections leads to inferior outcomes. We designed a final experiment to study this question. Our results show a visible advantage of using global rules such as Rule X over separate district elections. For example, Rule X always produces outcomes with a more equal distribution of voter utility, and in most cases also provides a higher total utility in comparison to the rules that are in actual use in the elections we examined.

References

H. Aziz and B. Lee. The expanding approvals rule: improving proportional representation and monotonicity. *Social Choice and Welfare*, 54(1):1–45, 2020.

H. Aziz and N. Shah. Participatory budgeting: Models and approaches. In Tamás Rudas and Gábor Péli, editors, *Pathways Between Social Science and Computational Social Science: Theories, Methods, and Interpretations*. Springer, 2020.

H. Aziz, M. Brill, V. Conitzer, E. Elkind, R. Freeman, and T. Walsh. Justified representation in approval-based committee voting. *Social Choice and Welfare*, 48(2):461–485, 2017.

H. Aziz, E. Elkind, S. Huang, M. Lackner, L. Sánchez-Fernández, and P. Skowron. On the complexity of extended and proportional justified representation. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI-2018)*, pages 902–909, 2018a.

H. Aziz, B. Lee, and N. Talmon. Proportionally representative participatory budgeting: Axioms and algorithms. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2017)*, pages 23–31, 2018b.

G. Benade, S. Nath, A. D. Procaccia, and N. Shah. Preference elicitation for participatory budgeting. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI-2017)*, pages 376–382, 2017.

J. Byrka, P. Skowron, and K. Sornat. Proportional approval voting, harmonic k-median, and negative association. In *In proceedings of 45th International Colloquium on Automata, Languages, and Programming (ICALP-2018)*, pages 26:1–26:14, 2018.

Y. Cabannes. Participatory budgeting: a significant contribution to participatory democracy. *Environment and Urbanization*, 16(1):27–46, 2004.

Alfonso Cevallos and Alistair Stewart. A verifiably secure and proportional committee election rule. *arXiv preprint arXiv:2004.12990*, 2020.

Y. Cheng, Z. Jiang, K. Munagala, and K. Wang. Group fairness in committee selection. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 263–279, 2019.

B. Fain, K. Munagala, and N. Shah. Fair allocation of indivisible public goods. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 575–592, 2018. Extended version arXiv:1805.03164.

P. Faliszewski, J. Sawicki, R. Schaefer, and M. Smolka. Multiwinner voting in genetic algorithms for solving illposed global optimization problems. In *Proceedings of the 19th International Conference on the Applications of Evolutionary Computation*, pages 409–424, 2016.

P. Faliszewski, P. Skowron, A. Slinko, and N. Talmon. Multiwinner voting: A new challenge for social choice theory. In U. Endriss, editor, *Trends in Computational Social Choice*, pages 27–47. AI Access, 2017.

P. Faliszewski, P. Skowron, S. Szufa, and N. Talmon. Proportional representation in elections: STV vs PAV. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2019)*, pages 1946–1948, 2019.

Z. Jiang, K. Munagala, and K. Wang. Approximately stable committee selection. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 463–472, 2020.

M. Lackner and P. Skowron. Consistent approval-based multi-winner rules. In *Proceedings of the 19th ACM Conference on Economics and Computation (EC-2018)*, pages 47–48, 2018.

M. Lackner and P. Skowron. Approval-based committee voting: Axioms, algorithms, and applications. Technical Report arXiv:2007.01795 [cs.GT], arXiv.org, 2020.

D. Peters and P. Skowron. Proportionality and the limits of welfarism. In *Proceedings of the 2020 ACM Conference on Economics and Computation*, pages 793–794, 2020. Extended version arXiv:1911.11747.

D. Peters, G. Pierczynski, N. Shah, and P. Skowron. Market-based explanations of collective decisions. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI-2021)*, 2021.

L. Sánchez-Fernández, E. Elkind, M. Lackner, N. Fernández, J. A. Fisteus, P. Basanta Val, and P. Skowron. Proportional justified representation. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI-2017)*, pages 670–676, 2017.

P. Skowron. Proportionality degree of multiwinner rules. Technical Report arXiv:1810.08799 [cs.GT], arXiv.org, 2018.

P. Skowron, P. Faliszewski, and J. Lang. Finding a collective set of items: From proportional multirepresentation to group recommendation. *Artificial Intelligence*, 241:191–216, 2016.

P. Skowron, M. Lackner, M. Brill, D. Peters, and E. Elkind. Proportional rankings. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI-2017)*, pages 409–415, 2017.

Dariusz Stolicki, Stanisław Szufa, and Nimrod Talmon. Pabulib: A participatory budgeting library. *arXiv preprint arXiv:2012.06539*, 2020.

N. Talmon and P. Faliszewski. A framework for approval-based budgeting methods. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI-2019)*, pages 2181–2188, 2019.

T. N. Thiele. Om flerfoldsvalg. In *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger*, pages 415–441. 1895.

N. Tideman and D. Richardson. Better voting methods through technology: The refinement-manageability trade-off in the Single Transferable Vote. *Public Choice*, 103(1–2):13–34, 2000.

D. Woodall. Properties of preferential election rules. *Voting Matters*, 3:Paper 4, 1994.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes]
   (c) Did you discuss any potential negative societal impacts of your work? [N/A] Our work is theoretical in spirit and has no potential negative societal impacts.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes] See Section 2.
   (b) Did you include complete proofs of all theoretical results? [Yes] Either in the main text, or in Appendix D.
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] The code is provided in the 'code' catalog in the supplementary material. The link to the data and the detailed description of the experiments are provided in Appendix H.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [N/A] Our experiments do not require training.
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A] The usage of randomization in our experiments is very limited and did not affect the results. We mention it in Appendix H.

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [N/A] We say that the algorithms we compute are polynomial, consequently their execution does not have any special technical requirements and can be reproduced even on a single standard computer.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes] As we write in Appendix H, we use the data from the Pabulib library.
   (b) Did you mention the license of the assets? [Yes] See Appendix H.
   (c) Did you include any new assets either in the supplemental material or as a URL? [No]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A] The data we use is publicly available.
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [Yes] We mention in Appendix H that it is not the case.
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]