Grzegorz Pierczyński

Supervised by dr hab. Piotr Skowron

PROPORTIONAL PARTICIPATORY BUDGETING

University of Warsaw
Faculty of Mathematics, Informatics and Mechanics
Warsaw, 2023

Abstract

The goal of the dissertation is a formal analysis of the problem of selecting winning projects in *participatory budgeting* (PB), a problem in the field of computational social choice with multiple applications within and beyond computer science (for example, in the political domain, blockchain protocols, search engines or genetic algorithms). It can be formulated as follows: we have a set of projects and certain amount of available funds. Each project is associated with some cost. There is a set of voters who gain utilities from the projects (the form of which may differ depending on the model). The problem is to choose the subset of projects whose joint total cost does not exceed the available budget and is fair with respect to the voters' preferences. In this dissertation we will focus on the problem of group fairness, called *proportionality*. Here we assume that each group of voters should get a representation proportional to its size ("every $x\%$ of the voters should decide about spending $x\%$ of the budget"). The meaning of this phrase is intuitive in simple cases when we can split voters into separate groups supporting identical projects. However, in many applications of the participatory budgeting problem it is not the case: the groups of voters supporting certain projects may overlap and their preferences are not fully cohesive. In such cases, a formal analysis is required to compare different outcomes to one another.

Before our research this problem has been studied only under some simplifying assumption (like equal costs of the projects or specific forms of voters' preferences) and the obtained results still appeared to be complex. In particular, it turns out that the idea of proportionality can be formalized in many substantially different ways. The well-established approach in the computational social choice is then to try to formalize this idea via multiple axioms specifying desired properties of outcomes.

We contribute to the studies of this model in the following way: in the first part of the dissertation we propose a new algorithm, called Method of Equal Shares, together with an extensive theoretical and experimental study of its properties. It is notable that our results have lead to changing the statutes of participatory budgeting, switching from the simple majoritarian rules to Method of Equal Shares, in two Polish communities (Wieliczka and Świecie) and in one Swiss city (Aarau). It shows the practical impact of our results.

The second part of the dissertation is focused on designing strict mathematical definitions (axioms) capturing the idea of proportionality in the PB setting, analyzing the logical connections between various axioms as well as their satisfiability. We focus here on the strongest possible guarantees, in particular, stronger than the ones provided by Method of Equal Shares. However, we show that without additional assumptions they are not satisfiable in general by polynomial-time computable algorithms.

Our results have been published in the form of four separate papers:

(1) Dominik Peters, Grzegorz Pierczyński, Nisarg Shah, and Piotr Skowron. Market-based explanations of collective decisions. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI-2021)*, 2021a

(2) Dominik Peters, Grzegorz Pierczyński, and Piotr Skowron. Proportional participatory budgeting with additive utilities. *Advances in Neural Information Processing Systems (NeurIPS-2021)*, 34:12726–12737, 2021b

(3) Grzegorz Pierczyński and Piotr Skowron. Core-stable committees under restricted domains. In *Proceedings of the 18th Conference on Web and Internet Economics (WINE-2022)*, pages 311–329, 2022

(4) Piotr Faliszewski, Jarosław Flis, Dominik Peters, Grzegorz Pierczyński, Piotr Skowron, Dariusz Stolicki, Stanisław Szufa, and Nimrod Talmon. Participatory budgeting: Data, tools and analysis. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI-2023)*, pages 2667–2674, 2023

Keywords: computational social choice, voting, participatory budgeting, committee elections, Method of Equal Shares, proportionality, fairness, core, priceability

Streszczenie

Celem ninejszej rozprawy doktorskiej jest formalna analiza problemu wyboru zwycięskich projektów w budżecie partycypacyjnym (ang. *participatory budgeting*, PB), problemu z zakresu obliczeniowej teorii wyboru społecznego, mającego wiele zastosowań zarówno w dziedzinie informatyki, jak i poza nią (np. w głosowaniach politycznych, protokołach blockchain, wyszukiwarkach internetowych czy algorytmach genetycznych). Problem ten można sformułować w następujący sposób: mamy zbiór projektów oraz pewną ilość dostępnych środków. Realizacja każdego projektu wiąże się z pewnym kosztem. Oprócz tego mamy zbiór wyborców, różniących się preferencjami na temat projektów (forma wyrazu tych preferencji może się różnić w zależności od modelu). Problem polega na wyborze podzbioru projektów, których łączny koszt nie przekracza dostępnego budżetu i który sprawiedliwie (proporcjonalnie) odzwierciedla preferencje wyborców. Zakładamy tutaj, że każda grupa wyborców powinna mieć reprezentację proporcjonalną do swojego rozmiaru ("każda grupa $x\%$ wyborców powinna decydować o przeznaczeniu $x\%$ dostępnych środków"). Zasada ta jest intuicyjnie zrozumiała w prostych przypadkach, gdy możemy podzielić wyborców na rozłączne grupy, głosujące na projekty o kosztach proporcjonalnych do swoich rozmiarów. Jednak w wielu zastosowaniach problemu budżetu partycypacyjnego nie jest to możliwe: grupy wyborców wspierających określone projekty mogą się pokrywać, a ich preferencje nie być w pełni spójne. W takich sytuacjach, do oceny proporcjonalności wybranego zbioru projektów i porównywania różnych wyników ze sobą, wymagana jest formalna analiza.

Problem ten był dotychczas badany głównie w uproszczonej wersji, przy dodatkowych założeniach (takich jak jednakowe koszty projektów albo uproszczone preferencje wyborców), co i tak prowadziło do złożonych rezultatów. W szczególności, okazuje się wówczas że ideę proporcjonalności można sformalizować na wiele istotnie róznych sposobów. Stąd też powszechnie przyjętym podejściem w obliczeniowej teorii wyboru społecznego jest podejście aksjomatyczne, określające pożądane właściwości zwycięskich podzbiorów w poszczególnych sytuacjach.

Nasz wkład w badania nad tym problemem jest następujący: w pierwszej części rozprawy proponujemy nowy algorytm, Metodę Równych Udziałów, wraz z głęboką teoretyczną i eksperymentalną analizą jej własności. Warto wspomnieć, że nasze wyniki doprowadziły do zmiany statutów budżetu partycypacyjnego i używania Metody Równych Udziałów w praktyce, w dwóch polskich gminach (Wieliczka i Świecie) i jednym szwajcarskim mieście (Aarau). Dowodzi to istotnego znaczenia praktycznego wyników opublikowanych w niniejszej rozprawie.

Druga część rozprawy skupia się na projektowaniu aksjomatów formalizujących ideę proporcjonalności w modelu PB oraz analizie zależności logicznych i spełnialności tych aksjomatów. Skupiamy się tu na możliwie najsilniejszych gwarancjach sprawiedliwości, w szczególności silniejszych od tych gwarantowanych przez Metodę Równych Udziałów. Jednak nasze wyniki wskazują na to, że w ogólnym modelu, bez dodatkowych założeń, nie są one możliwe do spełnienia w czasie wielomianowym.

Nasze wyniki zostały opublikowane w formie czterech osobnych artykułów naukowych:

(1) Dominik Peters, Grzegorz Pierczyński, Nisarg Shah, and Piotr Skowron. Market-based explanations of collective decisions. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI-2021)*, 2021a

(2) Dominik Peters, Grzegorz Pierczyński, and Piotr Skowron. Proportional participatory budgeting with additive utilities. *Advances in Neural Information Processing Systems (NeurIPS-2021)*, 34:12726–12737, 2021b

(3) Grzegorz Pierczyński and Piotr Skowron. Core-stable committees under restricted domains. In *Proceedings of the 18th Conference on Web and Internet Economics (WINE-2022)*, pages 311–329, 2022

(4) Piotr Faliszewski, Jarosław Flis, Dominik Peters, Grzegorz Pierczyński, Piotr Skowron, Dariusz Stolicki, Stanisław Szufa, and Nimrod Talmon. Participatory budgeting: Data, tools and analysis. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI-2023)*, pages 2667–2674, 2023

Słowa kluczowe: obliczeniowa teoria wyboru społecznego, głosowanie, budget partycypacyjny, wybory komitetowe, Metoda Równych Udziałów, proporcjonalność, sprawiedliwość, core, priceability

Acknowledgments (Podziękowania)

I would like to express my heartfelt gratitude to the individuals who played pivotal roles in my PhD journey.

First and foremost, I am deeply grateful to my supervisor, dr hab. Piotr Skowron, who encouraged me to start my academic career. His kindness, infectious passion for research, boundless enthusiasm, and commitment to my growth as a scientist have been absolutely extraordinary. Piotr, it is a privilege to have you as my mentor and friend.

I extend my thanks to all the co-authors of my research papers included in the dissertation. Collaborating with each of you has been an enriching experience, and I am grateful for the contributions you made to our joint work.

Dziękuję z całego serca mojej rodzinie, szczególnie rodzicom i siostrze. Wasza miłość, cierpliwość i zrozumienie są dla mnie bezcennym wsparciem pozwalającym przetrwać nawet najtrudniejsze chwile.

Szczególne podziękowania kieruję również do Magdaleny Karczewskiej, której wsparcie i niezachwiana wiara w moje możliwości były dla mnie źródłem siły i motywacji.

To all my friends, and all those who have contributed, directly or indirectly, to my academic and personal growth, I extend my sincere appreciation.

* * *

My research has been supported by the Foundation for Polish Science (FNP) and the National Science Centre (NCN) with the following grants:

- FNP HOMING grant no. Homing/2017-4/40 (*Normative Comparison of Multiwinner Election Rules*),
- NCN OPUS grant no. 2018/29/B/ST6/00174 (*Conflicts with Multiple Battlefields and Discrete Resources*),
- NCN OPUS grant no. 2019/35/B/ST6/02215 (*Multiwinner Election Rules: Beyond Scoring Protocols*),
- NCN PRELUDIUM grant no. 2022/45/N/ST6/00271 (*Proportional Participatory Budgeting*),
- NCN OPUS grant no. 2018/31/B/ST6/03201 (*Group Centrality Measures: Axioms, Algorithms and Applications*).

Contents

Chapter 1

Introduction

A growing list of cities nowadays uses Participatory Budgeting (PB) to decide how to spend a part of their budgets [De Vries et al., 2022, Wampler et al., 2021]. Through voting, PB allows the residents of a city to decide which from the considered projects will be funded by the city. The advantages of such initiatives are numerous—thanks to it, the citizens are encouraged to get involved in the problems of their local community and to propose socially valuable initiatives.

The overall process is quite complex. First, groups of citizens register projects. Each project is described in detail, and its cost is estimated. Typically, project proposals are verified by a special commission—projects not meeting certain formal criteria might be rejected. Second, citizens vote for the projects, by casting paper or online ballots. There are several possible types of ballots—for example, in Warsaw citizens choose up to a few projects they approve (an example of an approval ballot is presented in Figure 1). The third and final step of the procedure is running a voting rule—an algorithm, which takes as input the voters' ballots and returns the subset of winning projects. The total cost of the winning outcome should not exceed the budget constraint. In this dissertation, we focus on this last stage—we want to design algorithms, which elect outcomes that *proportionally* represent the views of the voters. What does it mean in practice and why is it so important?

To answer these questions, let us take a look at the most commonly used voting rule, hereinafter called *Utilitarian Greedy*. It works as follows: initially, the elected outcome is empty. Then we iterate over projects starting from the ones with the highest total number of approvals and add them to the outcome until the budget is exhausted. If adding a project would exceed the budget, it is skipped. While simple and intuitive, it is an example of an algorithm which may elect outcomes not reflecting the actual views of the society. To see this, consider the following simple example:

Example 1.1. Suppose that all the projects have the same cost and the budget limit allows to fund any 10 of them. Each project belongs to one of three categories. We have 10 *red* projects approved by 40% of voters, 10 *blue* projects approved by another 30% of voters and 10 *green* ones approved by yet another 30% of voters. For example, each color might represent a district of the city where the corresponding project will be conducted. Each group of voters approves projects from only a single category.

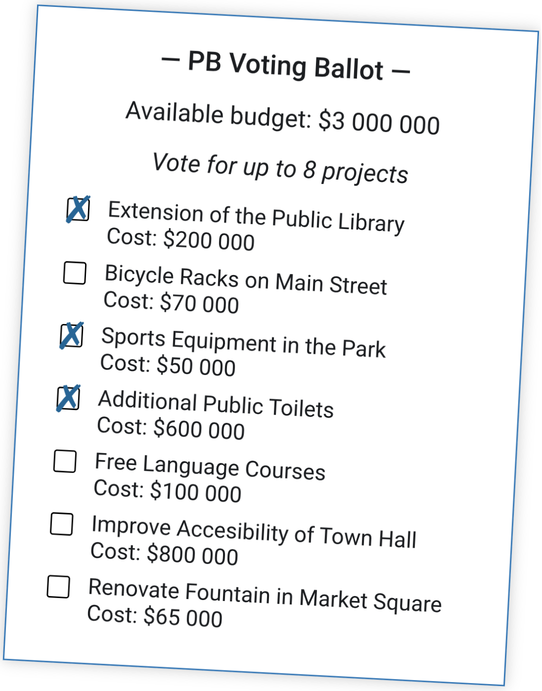


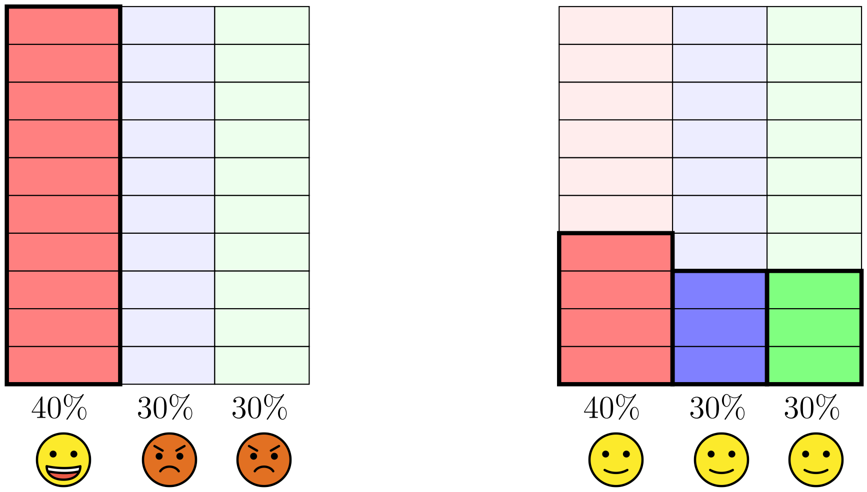
Figure 1: An example of an approval ballot.

In this case, the Utilitarian Greedy rule selects only red projects (Figure 2a) and 60% of the voters are left empty-handed. ⌟

We can intuitively see that the outcome elected above is severely unfair. We would rather expect selecting only 4 red projects (costing 40% of the overall budget), together with 3 blue and 3 green ones (Figure 2b). This example illustrates the fact that under Utilitarian Greedy *the winner takes all*: the majority (or rather the largest minority) of voters may decide about the full budget, leaving the remaining part of the society with nothing. If a city implemented this rule directly, it could lead to, for example, the domination of large districts over smaller ones.

Note that under Utilitarian Greedy, project submitters have an incentive to follow a very simple strategy: in Example 1.1 submitters of blue and green projects could merge them into 10 artificial blue-green projects costing twice more each. Then such projects could receive 60% of votes and dominate the election. This process is indeed observable in practice: for example, in Wieliczka in 2021,[1] the whole election was completely dominated by the rivalry of two projects, with names: "Improving the level of security in the towns: Mietniów, Pawlikowice, Chorągwica, Grajów, Dobranowice, Jankówka, Raciborsko, Lednica Górna, Podstolice, Gorzków, Janowice" (southern towns of the municipality) and "Improving life conditions and the level of security in towns: Brzegi, Byszyce, Czarnochowice, Grabie, Kokotów, Mała Wieś, Strumiany, Sułków, Śledziejowice, Węgrzce Wielkie, Zabawa" (northern towns of the munici-

[1] All the data concerning real-life elections presented in the dissertation are publicly available within our participatory budgeting library: https://pabulib.org



(a) Outcome elected by Utilitarian Greedy

(b) The proportional outcome

Figure 2: Illustration of Example 1.1. Each project is depicted as a red, blue or green box, depending on the project's category. The percentages mean the quantities of disjoint groups of voters approving the same projects (the ones placed above them). Projects marked with more intense colors are elected. Emojis depict (informally) levels of satisfaction of each group. For example, in Figure 2a only the projects supported by the first 40% of voters and so these voters have high satisfaction. The remaining voters are left empty-handed and so they are highly dissatisfied.

pality). In this situation, no smaller and more specific projects had even a chance of winning.

To avoid this problem, city councils modify Utilitarian Greedy in the following way: they divide voters and projects into districts upfront. Then the budget is split among districts proportionally to their population and the rule is run in each district separately. Besides, some projects are marked as "citywide" and some fixed fraction of the budget is allocated to them.

There are numerous problems with this approach. First of all, it does not actually solve the problem of proportionality—it only takes into account territorial groups of interests, while in a city we can have also non-territorial ones. For example, we can easily imagine that the colors of projects from Example 1.1 do not correspond to different districts, but rather to different topics, like building green areas, bicycle infrastructure or playgrounds—and the voters do not actually care whether the selected projects are close to them or not.

Even if we care only about territorial groups of interests, the problems remain—first, the borders of the districts are also defined arbitrarily. Voters living close to the border of their district may gain satisfaction from projects located on the other side of the border, yet they cannot vote for them. Second, a voter can work or spend her free time in a different district than the one she lives—in this case, her assignment to only one district is artificial as well. For example, in Gdańsk in 2020, where voters could vote for projects from various districts, over 15% of voters did so. Third, and most importantly, the problem of proportionality *within* each district still remains untouched.

In fact, even under the most mild and favorable assumption—that we take into account only districtwise groups of interests—the current solution is still not satisfactory. Note that it is arbitrary whether a project should be classified as citywide or districtwise, and how much money should be spent on citywide projects. It may result in electing inefficient outcomes, which is indeed easily observable in practice. For example, in Warsaw in 2021 a project "A pavement on the Modlińska Street" costing 630,000 PLN was funded in Białołęka district, receiving 1,932 votes. On the other hand, a cheaper project "Trees and shrubs on the Modlińska Street", costing only 430,000 PLN and concerning the very same street, was registered as a citywide project. Since the competition between citywide projects was much more fierce than between districtwise ones, the latter project was not selected despite receiving 12,463 votes in total (including 4,365 from the voters in Białołęka district). As a result, instead of electing a cheaper and more popular projects, a more expensive and less popular one was chosen. Similar examples can be found in participatory budgeting elections from various years and cities.

As we can see, there is an actual need for proportional voting rules—the ones that would endogenously identify coherent groups of voters (without any arbitrary decisions) and satisfy them proportionally to their size. However, the idea of proportionality is still very vague. The meaning of this term is clear only in simple examples like Example 1.1, which appear extremely rarely in practice. For example, what would be the proportional outcome, if we modified this example by adding 2 *yellow* projects, approved by the supporters of blue and green projects, and 2 *violet* projects, approved by the supporters of blue and red projects (see Figure 3)?

Now it is much harder to indisputably call only one outcome proportional. We probably feel that the original outcome (4 red, 3 blue and 3 green projects, depicted in Figure 3a) is still an acceptable choice and approving additional projects should at the very least not affect the voters' guarantees negatively. At the same time, we might allow for replacing some red/blue projects with violet ones, and some blue/green projects with yellow ones. But can we do it in any way? For example, is the outcome containing 2 violet, 2 yellow, 2 red, 1 green and 3 blue projects (Figure 3b), clearly favoring the middle 30% of the voters, proportional? Or shall we prefer a replacement that is more "balanced" with respect to the voters' satisfaction (like the one in Figure 3c)?

This example shows that in more complex scenarios, the question whether an outcome is proportional is no longer straightforward and needs to be analyzed formally. The most typical approach is to say that every outcome satisfying certain formal conditions (*axioms*) is proportional. However, axioms need to be designed with care. Consider for example, the following definition, formalizing the aforementioned "feeling".

Definition Attempt 1.2. We say that an outcome W is proportional, if for every group S of $x\%$ of voters, approving jointly some projects worth at least $x\%$ of the budget, every voter from S approves some elected projects from W worth at least $x\%$ of the budget.

According to this definition, in the modified Example 1.1 every outcome depicted in Figure 3 is proportional. At the first sight, Definition Attempt 1.2 appears then to be quite intuitive and reasonable, maybe even a bit weak—yet there exist elections where it is impossible to be

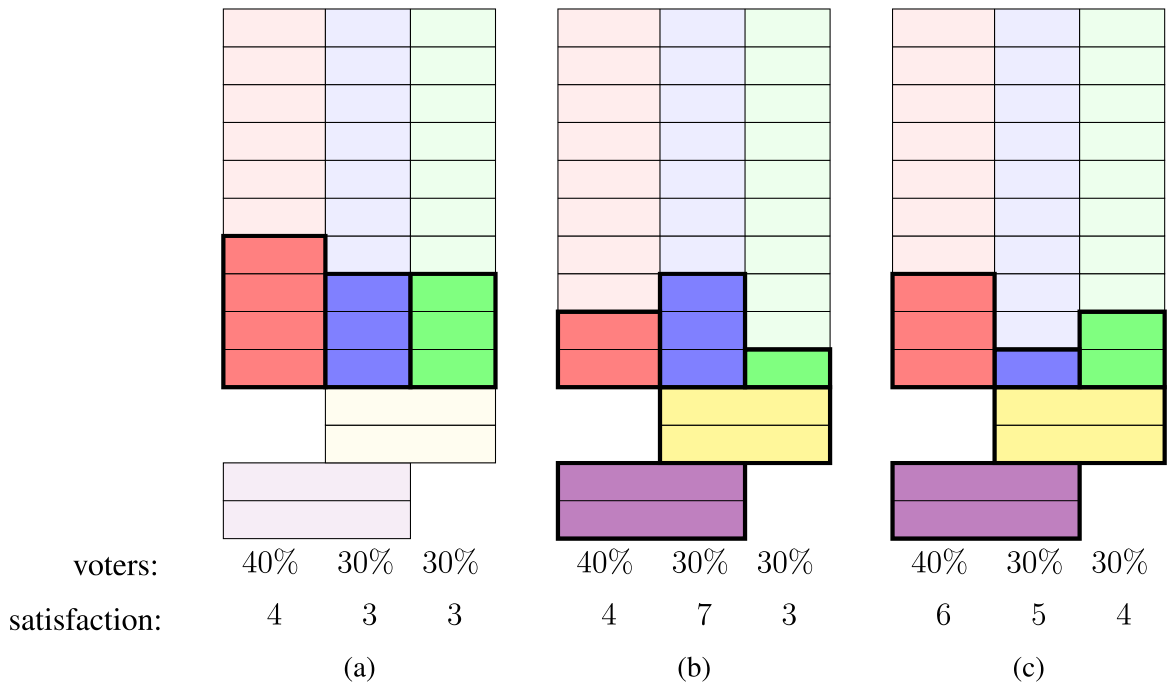


Figure 3: A modified Example 1.1 with three examples of possibly proportional outcomes. Numbers below the groups of voters denote their satisfaction from the outcome (numbers of approved elected projects).

satisfied.

Example 1.3 ([Aziz et al., 2018]). Consider an election with four projects a, b, c, d, and 12 voters, with the following approval sets:

$$
\begin{array}{llll}
1\colon \{a, d\} & 4\colon \{a, b\} & 7\colon \{b, c\} & 10\colon \{c, d\} \\
2\colon \{a\} & 5\colon \{b\} & 8\colon \{c\} & 11\colon \{d\} \\
3\colon \{a\} & 6\colon \{b\} & 9\colon \{c\} & 12\colon \{d\}.
\end{array}
$$

Suppose that each project costs 100,000 dollars and the budget is 300,000 dolars. Hence, the budget allows to fund any three projects. Now, according to Definition Attempt 1.2, the group of one third of the voters $\{1, 2, 3, 4\}$ deserves one project, as it has a commonly approved project (a) costing one third of the budget. Since voters 2 and 3 approve only a, definition Definition Attempt 1.2 actually requires that a is elected. Now observe that we can repeat this reasoning also for groups $\{4, 5, 6, 7\}$, $\{7, 8, 9, 10\}$, and $\{10, 11, 12, 1\}$ and projects b, c, and d respectively. Thus, all four projects have to be in a winning committee, which would exceed the budget. ⌟

As we can see, defining proportionality axioms is not straightforward when the groups of voters may overlap. Note that in Example 1.3 still the costs of all the projects are equal and the voters vote via approval ballots—considering elections with unequal costs or more complex voters' ballots makes the problem naturally even more difficult. Another challenge is that, since real-life elections usually contain large numbers of voters and projects, we prefer definitions that are satisfiable by polynomial-time computable voting rules.

1.1 Our contribution

The dissertation is based on the following research papers:

(1) Dominik Peters, Grzegorz Pierczyński, and Piotr Skowron. Proportional participatory budgeting with additive utilities. *Advances in Neural Information Processing Systems (NeurIPS-2021)*, 34:12726–12737, 2021b

(2) Piotr Faliszewski, Jarosław Flis, Dominik Peters, Grzegorz Pierczyński, Piotr Skowron, Dariusz Stolicki, Stanisław Szufa, and Nimrod Talmon. Participatory budgeting: Data, tools and analysis. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI-2023)*, pages 2667–2674, 2023

(3) Grzegorz Pierczyński and Piotr Skowron. Core-stable committees under restricted domains. In *Proceedings of the 18th Conference on Web and Internet Economics (WINE-2022)*, pages 311–329, 2022

(4) Dominik Peters, Grzegorz Pierczyński, Nisarg Shah, and Piotr Skowron. Market-based explanations of collective decisions. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI-2021)*, 2021a

All the publications have several authors. I participated in every technical contribution presented in the dissertation.

The results published within the first two papers form the first part of the dissertation. The remaining ones are included in the second part. Now let us briefly summarize our contribution within both of them.

Method of Equal Shares for Participatory Budgeting

Our main contribution is the following: we designed a polynomial-time computable voting rule, called Method of Equal Shares (or Equal Shares, in short). It is based on the idea that voters should *pay* for the projects they support with the fictional *money*. Intuitively, this fictional money represents their voting power. The formal description of this algorithm together with illustrative examples is included in Chapter 3. In the case of Example 1.1, Method of Equal Shares elects the outcomes depicted in Figure 2b and for its modified version, the outcome depicted in Figure 3c. Thus, the outcomes returned in these cases by the rule are proportional according to our intuition.

To prove formally that Equal Shares elects arguably proportional outcomes in any case, we extend the literature with new axioms of proportionality and show that these axioms are satisfied by our algorithm (see Chapter 4). The first axiom is *Extended Justified Representation up to one project* (EJR-1), which is similar in spirit to Definition Attempt 1.2, yet slightly weaker—and therefore, satisfiable. Intuitively, both definitions can be formulated as "every group of voters whose members have cohesive preferences, shall be represented proportionally to its size". The differences between them follow from different ways of formalizing the vague phrases *cohesive preferences* and *a group shall be represented*. Our next axiom, *priceability*, extends the notion introduced by Peters and Skowron [2020]. Similarly to Equal Shares, it is based on the idea of "paying" for projects by the voters and specifies conditions under which these payments form a fair market. Method of Equal Shares satisfies this axiom, which intuitively ensures that all voters have a similar amount of influence on the elected outcome. Additionally, we consider the very strong and well-established notion of *core-stability* [Aziz et al., 2017a], derived from cooperative game theory. It is not satisfiable for unconstrained voters' preferences [Fain et al., 2018], yet we show that Equal Shares provides its good approximation.

Further, in Chapter 5 we present an experimental analysis, showing which outcomes Equal Shares would elect in practice, if it was used in real-life PB elections. First, we collected data from over 800 hundred elections (mainly from Poland) which are publicly available within our participatory budgeting library, called *Pabulib* (`https://pabulib.org`). Second, we compared the quality of outcomes elected by Method of Equal Shares to the ones elected by Utilitarian Greedy, using various metrics. These metrics measure both efficiency of an outcome (how much satisfaction it provides in total to the voters) and its fairness (how equally the satisfaction is distributed among the voters). For example, one of our metrics of efficiency measures the average number of selected projects that are approved by a voter.

Note that there might appear a trade-off between efficiency and proportionality—in Example 1.1, the average satisfaction of the voters from the outcome depicted in Figure 2a is 4 $(40\% \cdot 10 + 60\% \cdot 0)$, while for the outcome depicted in Figure 2b it is only 3.4 $(40\% \cdot 4 + 30\% \cdot 3 + 30\% \cdot 3)$. Moreover, we can see that Utilitarian Greedy is an efficiency-driven voting rule. One could therefore expect that Equal Shares is better than Utilitarian Greedy in terms of fairness, yet worse in terms of efficiency. Surprisingly, our experiments show that it is not the case—outcomes elected by Method of Equal Shares have comparable efficiency to the ones elected by Utilitarian Greedy, at the same time being much more fair. Hence, proportionality does not require sacrificing much efficiency in practice.

Finally, we show that Equal Shares is more robust to changing the type of the ballot used in the election. The code used for the experiments is publicly available for researchers conducting similar experiments in the future, within the Python package *Pabutools* (`https://pypi.org/project/pabutools`).

Voters' preferences over projects can be provided in various forms—in the definition of Equal Shares and its aforementioned analysis we assumed that voters' preferences are cardinal, which means that the utility of each voter from each project can be expressed as a number, for example from 0 to 10. However, our rule is very flexible and can be also adapted to the very different setting of *ordinal* preferences. Here, each voter provides only a ranking of projects, from the most preferred one to the least preferred one, and compares the outcomes lexicographically. In Chapter 6 we present this alternative variant of Equal Shares and prove that it satisfies

Inclusion Proportionality for Solid Coalitions [Aziz and Lee, 2021]. To our knowledge, it is the strongest satisfiable proportionality axiom for ordinal preferences.

Beyond Equal Shares: Stronger Notions of Proportionality

In the second part of the dissertation, we consider several other notions of proportionality, that are stronger than those satisfied by Method of Equal Shares. We start with the significant strengthening of the EJR-1 axiom—*Full Justified Representation* (FJR), presented in Chapter 7. It is notable that, according to our results, FJR is the strongest notion of proportionality which is known to be always satisfiable—unfortunately, the problem of finding outcomes satisfying it is NP-hard. Further, we present an exponential algorithm satisfying FJR, called *Greedy Cohesive Rule* (GCR) and study its properties. We prove that, on one hand, together with a specific completion GCR satisfies priceability, which suggests that it might have better theoretical properties than Equal Shares. On the other hand, apart from the running time (which makes it impossible to use in large-scale elections), it also behaves less naturally than Equal Shares on certain specific classes of elections. Besides, unlike Equal Shares, GCR loses its proportionality properties when adapted to ordinal preferences.

In Chapter 8 we consider the axiom of the core-stability, mentioned in the previous part. Determining whether it is satisfiable for approval-based preferences is a famous long-standing open problem in computational social choice [Aziz et al., 2017a], while it is known not to be satisfiable for other types of preferences [Fain et al., 2018, Aziz et al., 2017b]. Moreover, verifying whether a given outcome is in the core is CoNP-complete even for approval-based preferences [Brill et al., 2022]. Since this axiom is so demanding, there are multiple papers considering various approximated [Jiang et al., 2020, Peters and Skowron, 2020] or randomized [Cheng et al., 2019] variants of the core. We study this problem from a different perspective—instead of weakening the axiom, we check whether it is satisfiable under the additional assumptions about voters' preferences. We prove that core-stability is satisfiable in polynomial time for some specific well-known classes of elections, where the voters' preferences are somehow *structured* (that is, there are no *weird* votes). For example, the classes we consider cover and generalize the structure in which both voters and candidates can be placed in one-dimensional Euclidean space and all the voters prefer closer projects to the further ones. Our algorithm works for both approval-based and ordinal voters' preferences. However, it is custom-engineered to work only for elections where the preferences of 100% of voters are well-structured. This fact blocks the possibility to use this algorithm in practice, since it is naturally not the case for real-life elections. Hence, an interesting open question for future research is whether there exist voting rules satisfying the strongest proportionality notions in general and at the same time core-stability under restricted domains. Our results show that it is not the case for the all most important voting rules considered in the literature.

Finally, in Chapter 9 we aim to design the strongest possible axioms of proportionality—even stronger than core-stability—for the case of approval-based voters' utilities. We designed two such axioms, *stable priceability* and *balanced stable priceability*. Stable priceability has an advantage in being strictly stronger than core-stability for all elections, while balanced stable priceability is incomparable to both notions in general, but in many specific cases seems to be the most intuitive and restrictive axiom. For example, in the modified Example 1.1, core-stability still accepts all three considered outcomes. Meanwhile, both our new axioms rule out the outcome depicted in Figure 3a and balanced stable priceability additionally rules out the one depicted in Figure 3b. It is notable that stable pricebility, in contrast to core-stability, is verifiable in polynomial time.

Both axioms are similar in spirit to priceability—they view the election as a market, in which voters pay for the projects with their voting power—yet they are much stronger, requiring that the payments in general satisfy a specific market equilibrium. In particular, we show that stable priceability can be viewed as *Lindahl equilibrium* [Foley, 1970] for the model with public indivisible goods.

In general, both of them are not satisfiable—however, since stable priceability can be formulated as an integer linear program and we desinged a polynomial heuristic algorithm for finding balanced stable priceable outcomes, we were able to check that both axioms are usually possible to be satisfied in more than 90% of randomly generated elections from various distributions.

Summarizing this part of the dissertation, as we can see, no currently known polynomial-time computable algorithm can provide stronger proportionality guarantees for unrestricted voters' preferences than Equal Shares. Therefore, it still remains our main proposition for large-scale real-life PB elections.

1.2 Impact of Our Results on the Real World

The results presented in the dissertation lead to implementing Method of Equal Shares in practice in 2023, in two Polish municipalities (Wieliczka and Świecie) and in one Swiss city (Aarau). Since in Wieliczka and Aarau the voting process has already finished, let us present the overall results, showing that the choice of a voting rule actually does matter in practice.

Equal Shares in Wieliczka

In 2023, Wieliczka organized the "Green Million" PB election. Within this event, voters could decide about spending 1,000,000 Polish zlotys on ecological projects. In total, 64 projects were submitted (and positively verified) and 6586 city residents participated in the election. Voters voted via approval ballots and there was no restriction on the number of projects one could approve.

The full list of submitted and elected projects is available for example, under the following link: `https://equalshares.net/resources/zielony-milion`. Under this link, we can also find the comparison of the results elected by Equal Shares to the ones that would have been elected if Utilitarian Greedy rule was used. Let us now discuss the benefits from using our voting rule.

Fewer excluded voters. Using Equal Shares, 18% of voters did not receive any of the projects they voted for. If Utilitarian Greedy was used, this percentage would grow up to 28% . This means that our method allowed to significantly limit the number of voters who had no real influence on the outcome of the elections. Moreover, as we can see in Figure 4, most of the voters who did not receive any projects voted for only one project—among voters who voted for at least 3 projects, this percentage was only 3% for Equal Shares, compared to 9% under Utilitarian Greedy.

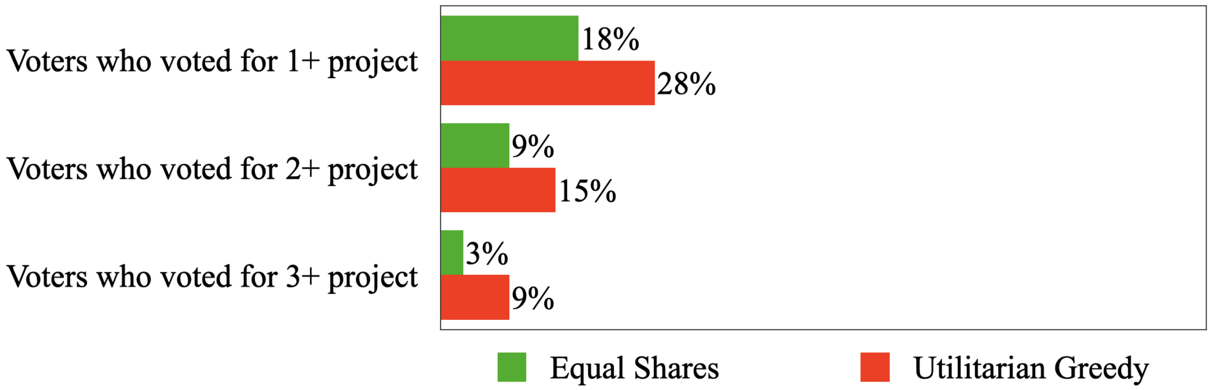


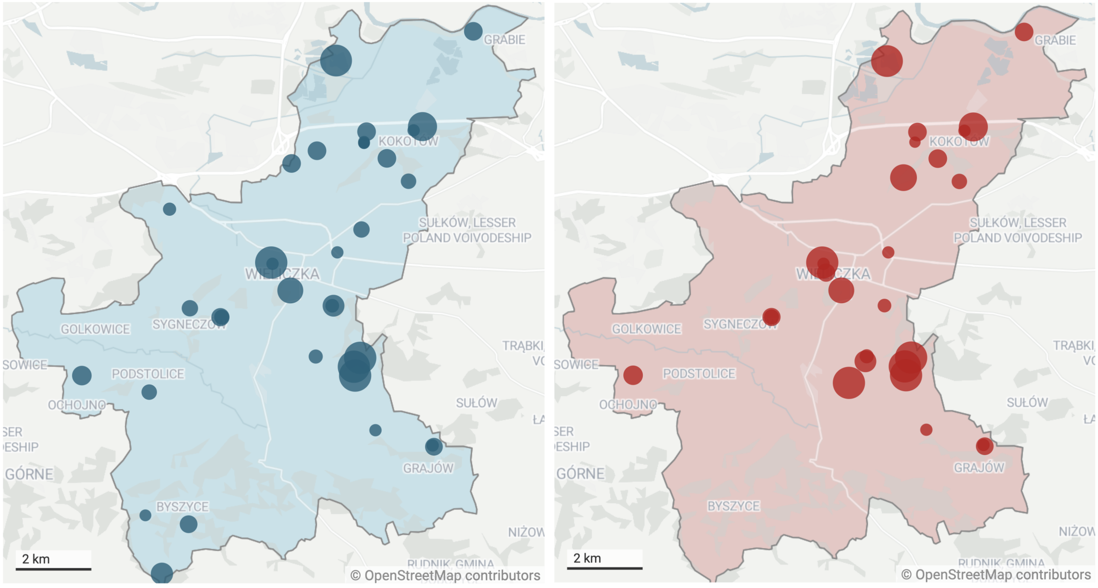
Figure 4: Percentage of voters in Wieliczka who did not receive any project that they voted for.

Higher efficiency. Utilitarian Greedy would select 4 projects that Method of Equal Shares does not select. These 4 projects together received 1453 votes. With the money saved, Method of Equal Shares selects 11 other projects, which together received 2414 votes. On average, a voter approves 1.61 projects in the outcome selected by Method of Equal Shares, while for Utilitarian Greedy, it would be 1.47. Hence, our algorithm translates about 1,000 votes more into the final result and is better by 10% in terms of average satisfaction.

No regional bias. Based on GPS coordinates, in Figure 5 we marked projects on the map of Wieliczka that were selected using Method of Equal Shares and which would be selected using Utilitarian Greedy. Thanks to Method of Equal Shares, we see a more even distribution of projects; in particular, we see that the residents of southern and western part of the municipality were not excluded from the decision-making process. In the case of Method of Equal Shares, for each resident of the municipality, their distance from the nearest selected project is no more than 3.4 km; in the case of Utilitarian Greedy, this value would be as much as 6.1 km. This improvement is also depicted in Figure 6, where for each point on the map we show the distance to the nearest elected project.

Equal Shares in Aarau

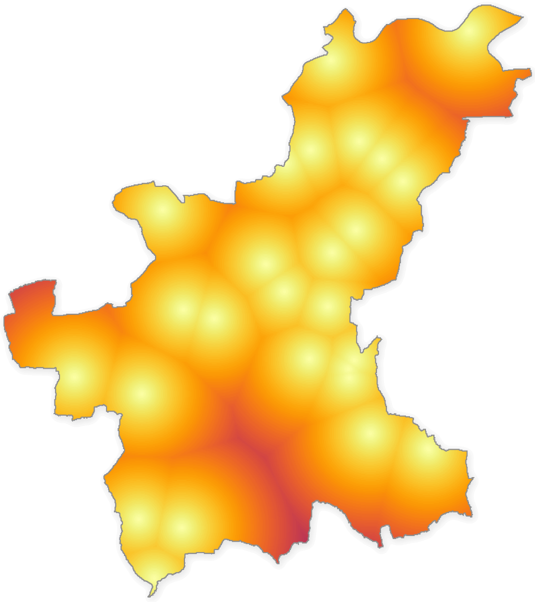
Contrary to Wieliczka, we do not have direct access to the voting data from Aarau, hence we cannot present a detailed analysis in this case. However, at the official website of Aarau



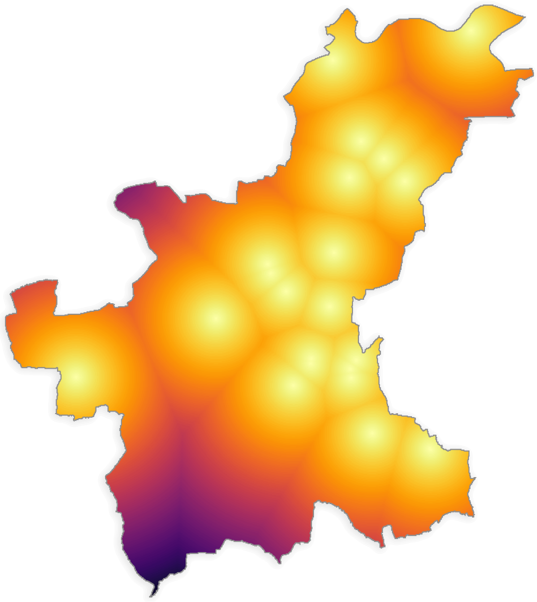
(a) Equal Shares (b) Utilitarian Greedy

Figure 5: Map of elected projects in Wieliczka. The sizes of the circles marked on the map correspond to the costs of the selected projects.

(https://www.stadtidee.aarau.ch/abstimmung.html/2114) we can see two diagrams comparing the actual results elected by Method of Equal Shares to the hypothetical scenario if Utilitarian Greedy was used. The first one is presented in a slightly modified way in Figure 7. It shows that under Utilitarian Greedy, the whole budget would have been spent only on citywide projects and the projects from the Centrum district (7 projects in total). On the other hand, Equal Shares elected 17 projects, representing districts more proportionally to their population. The same effect is visible in Figure 8 (an analogous map to Figure 5), showing that thanks to using Equal Shares, projects from different parts of the city were elected instead of only those from the central one.



(a) Equal Shares



(b) Utilitarian Greedy

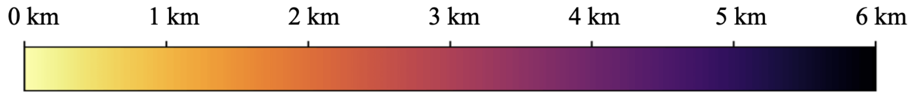
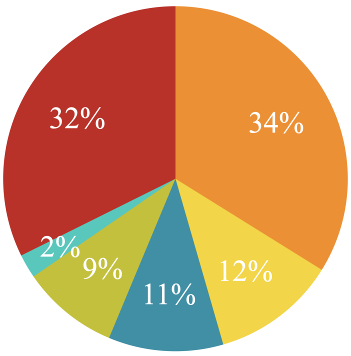
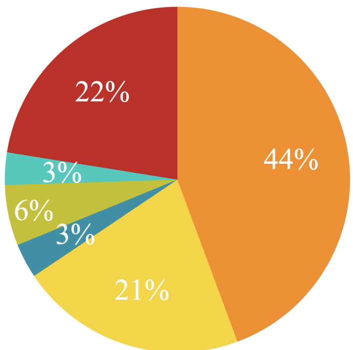


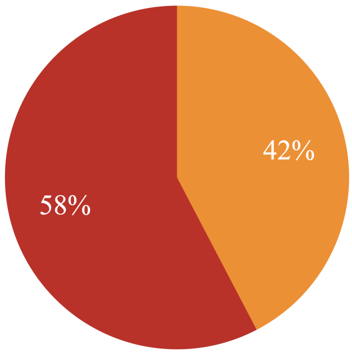
Figure 6: Heatmap presenting distances to the closest project in Wieliczka. Darker points indicate that there is no selected project near a given location.



(a) Percentage of votes



(b) Equal Shares



(c) Utilitarian Greedy

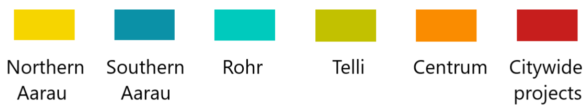


Figure 7: Distribution of the budget among districts in Aarau.

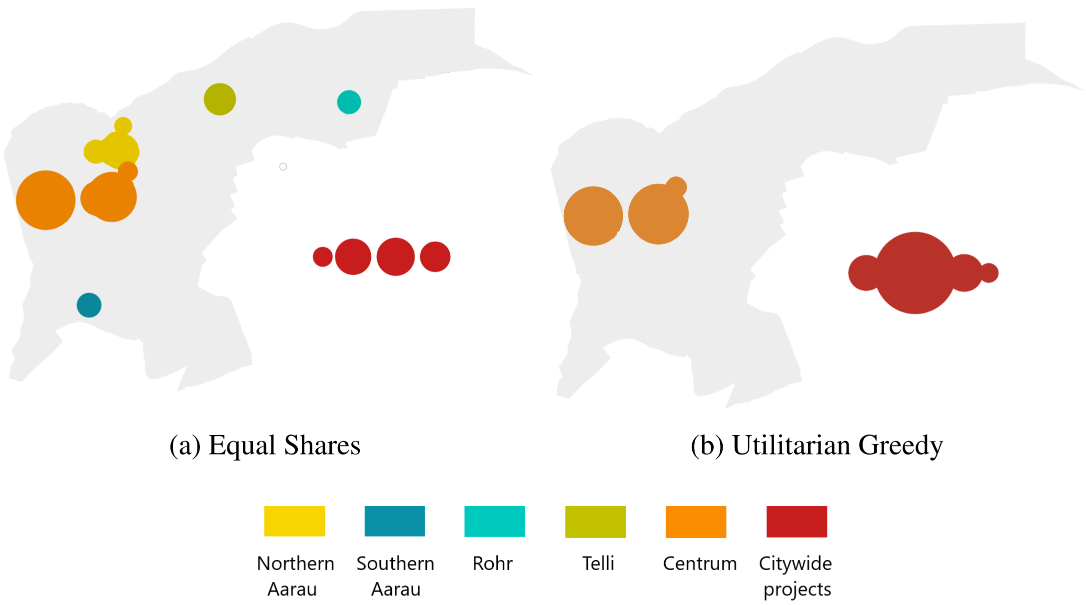


Figure 8: Map of elected projects in Aarau. The size of the circles marked on the map corresponds to the costs of the selected projects.

1.3 Other Possible Applications of our Research

Improving the organization of real-life PB elections in cities is probably the most natural real-life application of our research. However, the mathematical framework developed within the dissertation has also other possible applications, from and beyond the field of computer science.

It is especially true for the special case of the PB model where all the projects have the same costs. This case is often called in the literature the *committee election* model, since in this case participatory budgeting election boils down to electing up to a certain number of projects (as we could earlier see in Example 1.1 and Example 1.3). This problem can be therefore interpreted as electing a *committee* of fixed size from the set of possible *candidates*. Naturally, our theory developed for the general PB model applies to the committee election model as well.

Below, we present a few applications that match the model of committee elections.

Elections of Representative Bodies

The committee election model can be directly applied to model various elections of representative bodies, for example, municipal councils, or supervisory boards in companies and non-profit organizations. Note that in these cases it is usually desirable that a committee proportionally represents different groups of voters.

In the case of national parliamentary elections, some countries require that the elections should be proportional. However, here in the vast majority of cases it is ensured by grouping candidates into parties, which makes the problem theoretically simpler. The notable counterexamples are Ireland, Australia and Malta [Bowler and Grofman, 2000], which use the proportional voting rule called Single Transferable Vote [Tideman, 1995]. The same rule is used for local political elections in Australia, Ireland, New Zealand, Canada, United Kingdom, and the United States of America. However, this rule requires using ranking ballots, that are typically considered to be cognitively harder than approval ones [Zwicker, 2016]. Moreover, the rule itself has some important axiomatic flaws.[2] The theory presented within the dissertation may therefore lead to using better voting rules in such scenarios.

Blockchain Protocols

Every blockchain architecture needs to provide a mechanism for obtaining consensus in its peer-to-peer network. The first proposed idea was Proof-of-Work (PoW)—in order to add a new block, an agent should show the validation that certain amount of a specific computational effort has been expended. However, this approach has been recently criticized since it requires a high consumption of energy.

Therefore, a lot of newest blockchain projects replace the PoW protocol with Proof-of-Stake (PoS), in which validators participate in block production with a frequency proportional to their token holdings, as opposed to their computational power. One of the variants of PoS is Nominated Proof-of-Stake (NPoS), introduced by the designers of Polkadot network in 2020. In this variant, only certain k users are validators and they are elected each day (*era*) in approval elections by all the users in the network [Cevallos and Stewart, 2021, Burdges et al., 2020].

For aggregating votes, the designers of Polkadot used a polynomial proportional algorithm, called the Phragmén's rule. It is a well-established voting rule for the committee election model, first proposed in XIX century [Phragmén, 1894]. Phragmén's rule is very similar in spirit to Method of Equal Shares, since both rules are based on the concept of *buying* candidates by voters. In the dissertation we build a theory that allows to design and analyze such and similar voting rules.

Search Engines

Imagine that the word "tree" is typed in a search engine. What results should be displayed for an anonymous search—the ones about plants, graph theory, algorithms and data structures, or maybe STL library? Should the engine concentrate on news, tutorials, videos, encyclopedic information or pictures (and within the last category—photos, gifs or icons)?

Usually an engine needs to select a specific number (let it be 10) of the most relevant results. If we have already obtained profiles of potential users of our engine and their popularity, we know that for example, 60% of users typing "tree" would be interested in results about plants, and 35% in the ones about data structures. In such case, probably the "winner-takes-all" solution (display 10 results about plants) would be undesirable—typically, an engine would

[2] For example, under STV there might happen a paradoxical situation in which improving the position of a candidate in some voters' rankings decreases her chances of winning. Formally, it means that STV is not monotonous [Woodall, 1997].

display ca. 6 results about plants and 3-4 about data structures. Naturally, the groups may also overlap (for example, some users may be interested both in graph theory and in the implementation of trees in C++). Note that this problem can be solved by using algorithms for the committee election model, where profiles of the users correspond to the groups of voters (the popularity of the profile is the size of corresponding group) and potential results are candidates [Dwork et al., 2001, Skowron et al., 2017].

Facility Location

Consider a situation where a city council has to decide about the location of k new public facilities (for example, schools) in a city. The goal is to minimize the citizens' travel time to the nearest school. Such situations can be modelled as committee elections: possible locations of the schools are candidates, citizens are voters and the utility of each citizen from each location is inversely proportional to the travel time she needs to reach it. Indeed, there are multiple papers applying different concepts from committee elections in this problem [Faliszewski et al., 2017, Anshelevich and Zhu, 2021, Feldman et al., 2016]. Note that if we assume that each school has some fixed capacity, then the ideal choice of locations should be proportional with respect to distribution of the density of population: the more populous an area is, the more schools it requires.

Genetic Algorithms

Recently, committee elections have also found application in improving selection procedures for genetic algorithms [Faliszewski et al., 2016]. In this context, in every epoch the population of individuals is both the set of voters and candidates; individuals need to "elect" some fixed number of them to survive. The (ordinal) utilities are constructed with respect to (1) the specified fitness function, and (2) selfish desires of the individuals (first, each individual wants to survive itself; second, it wants other individuals with similar genes to survive).

After testing several committee election rules, the authors have found that proportional ones behave best (in particular, they avoid selecting individuals from large clusters only). However, they have tested only few methods and this approach is still preliminary—the theory presented in the dissertation may help to achieve even better results in this area.

Chapter 2

The Model

A participatory budgeting (PB) election is a tuple (N, C, b), where $N = \{1, 2, \dots, n\}$ is the set of *voters* and $C = \{c_1, c_2, \dots, c_m\}$ is the set of *candidate projects* (*projects*, in short). Each project c is associated with its cost, denoted as $\mathrm{cost}(c) \in \mathbb{N}$. We assume that the cost function is additive, that is the cost of the set of projects is the sum of costs of the projects in the set:

$$\mathrm{cost}(T) = \sum_{c \in T} \mathrm{cost}(c) \qquad \text{for each } T \subseteq C.$$

The parameter $b \in \mathbb{N}$ is the *budget constraint*: we say that a subset of projects (outcome) $T \subseteq C$ is *feasible* if $\mathrm{cost}(T) \leq b$. The goal is to elect a feasible outcome that, in some sense, best reflects the voters' preferences.

2.1 Preferences

The voters have preferences over the candidates. In this dissertation we consider two types of preferences:

Cardinal Preferences. Here, we assume that for each voter $i \in N$ we are given a utility function $u_i \colon C \to \mathbb{N}$ that quantifies the voter's level of appreciation towards different projects. We say that a voter i *supports* (or is a *supporter* of) a project c if $u_i(c) > 0$.

Ordinal Preferences. Here, the projects are not compared quantitatively. Instead, each voter $i \in N$ has a strict ranking (a linear order) $\succ_i$ over projects. A voter $i \in N$ prefers project c over c' if $c \succ_i c'$.

Since our goal is to select a subset of projects, we need a way to compare different sets from the perspective of a voter. For that we extend the voters' preferences over individual projects to their preferences over the sets as follows. Whenever we speak of cardinal preferences, we will assume that the corresponding utility functions are additive, that is:

$$u_i(T) = \sum_{c \in T} u_i(c), \text{ for each } i \in N \text{ and } T \subseteq C.$$

The ordinal preferences, on the other hand, are extended in a *lexicographical* way. Specifically, a voter $i \in N$ prefers an outcome W_1 to an outcome W_2 if and only if there exists a project $c \in W_1 \setminus W_2$ such that $c \succ_i c'$ for each $c' \in W_2 \setminus W_1$.

Note that it is also common in the literature to assume that the preferences over the sets are given directly [Benoit and Kornhauser, 1991, Lacy and Niou, 2000, Lang and Xia, 2016]. This, however, implies that the preferences might have an exponential size in the number of projects and such preferences are typically hard to elicit. For that reason in this dissertation we only consider the voters' preferences over the individual projects, which are then extended to the preferences over the sets, as we mentioned before.

2.2 Ballots

The voters express their preferences by casting ballots. Perhaps the most common examples of ballots are:

Approval ballots. Each voter simply indicates which projects she supports. In other words, a voter submits a subset of projects she approves of (see Figure 1). Such ballots are used in participatory budgeting elections, for example, in Warsaw, Wrocław and Zabrze.

Score ballots. Here, a voter assigns points to each project. A higher number of points assigned to a project means that the voter more strongly supports the respective project (this type of ballots is used in Częstochowa and Gdańsk).

Ranking ballots. Here, a voter provides a ranking (a linear order) over projects. Such ballots are used, for example, in Kraków.

The types of ballots listed above are a bit idealized compared to their practical analogues. For approval ballots there is usually a restriction that one can approve at most only x projects ($x = 15$ in Warsaw, $x = 1$ in Zabrze and Wrocław). For score ballots, there is typically a restriction that the total number of points assigned to projects should not exceed some x ($x = 10$ in Częstochowa, $x = 20$ in Gdańsk). Finally, ranking ballots are typically truncated to only x top projects ($x = 3$ in Kraków).

2.3 Relation Between Preferences and Ballots

Typically, an election designer does not have a direct access to the voters' preferences; we can use only the ballots cast by the voters. However, in this dissertation, we assume there is no distortion caused by the imperfect preference elicitation [Procaccia and Rosenschein, 2006, Boutilier et al., 2015, Anshelevich et al., 2018, 2021] and we do not consider strategic aspects of voting [Gibbard, 1973, Conitzer and Walsh, 2016]. In other words, we assume that the ballots perfectly reflect the voters' preferences. In further chapters we will then mostly assume that the voters' preferences are a part of the model; the only part where we make a clear distinction between preferences and ballots is Chapter 5.

Below we precisely describe how the preferences can be inferred from the ballots.

2.3.1 Approval and Score Ballots

When a voter $i \in N$ casts an approval or a score ballot, she directly assigns some kind of score to each project $c \in C$ (approval ballots can be viewed in this context as a special case of score ballots, where each project is given either 0 or 1 point). The most straightforward interpretation of this score is to assume that it directly corresponds to $u_i(c)$, the cardinal utility gained from c by voter i. In this case we speak of *direct* score-utility mapping.

This approach is undoubtedly very intuitive, yet it is not the only possible one. Let us once again consider the ballot depicted in Figure 1. There, the voter approves two projects, "Sports Equipment in the Park", costing $50,000, and "Additional Public Toilets", costing $600,000. Under the direct score-utility mapping we would assume that the voter would be equally happy from funding any of these two projects, even though the latter costs 12 times more than the former. An alternative approach is to assume that expensive projects generally carry more value to the voters, independently from the differences in score. In such case, we speak of the *costwise* score-utility mapping. Formally, $u_i(c)$ is defined as a multiplication of $\mathrm{cost}(c)$ and the number of points assigned to c by i. This mapping has also an intuitive interpretation. For example, for approval ballots the voter's utility from an outcome W corresponds to the total amount of money designated to the projects that the voters approves.

When analyzing voters' preferences, we will be particularly interested in the ones obtained from approval ballots. Hence, we define two classes of *approval-based* utilities and *approval-based cost* utilities that are special cases of cardinal utilities. In the first class we assume that $u_i(c) \in \{0, 1\}$ for each $i \in N$, $c \in C$ (which corresponds to the direct score-utility mapping), while in the second class we assume that $u_i(c) \in \{0, \mathrm{cost}(c)\}$ for each $i \in N$, $c \in C$ (which corresponds to the costwise score-utility mapping).

2.3.2 Ranking Ballots

Given ranking ballots, we can directly infer the voters ordinal preferences over the projects. However, this is not the only possible interpretation. In practice, cities use ranking ballots rather to elicit cardinal utilities, by assuming that placing a project at a specific position in a ranking corresponds to assigning a specific number of points to her. For example, Kraków uses the so called Borda score [Black, 1976] which assumes that, given a ranking, the scores of projects ranked from the bottom to the top position are consecutive natural numbers (1, 2, 3, $\dots$). In such case, the above discussion about the direct and the costwise score-utility mappings applies also to ranking ballots.

2.4 Voting Rules

Algorithms which compute feasible outcomes, given PB election are called *voting rules* (or *rules* in short). Each such rule takes an election as input and returns a set of winning outcomes (we are mostly interested in electing a single winning outcome, but ties are allowed).

In the dissertation we analyze voting rules mostly from the *axiomatic* perspective. Axioms in computational social choice are formal conditions specifying how a voting rule should behave in certain situations. In particular, proportionality axioms for the PB setting attempt to formally define the requirements needed to claim that a rule is proportional. They are typically defined only with respect to outcomes (like our Definition Attempt 1.2 from Chapter 1), specifying which outcomes are proportional for a given election. However, in such case, they can be naturally extended to voting rules—we say that a voting rule satisfies a proportionality axiom $\mathcal{A}$, if for each election it selects only outcomes satisfying $\mathcal{A}$. We say that a proportionality axiom $\mathcal{A}$ is *satisfiable*, if for every election there exists an outcome satisfying $\mathcal{A}$.

2.5 The Committee Election Model

An interesting special case of our model is the *committee election model*: here $\mathrm{cost}(c) = 1$ for all $c \in C$ (the *unit costs assumption*). Note that then feasibility boils down to choosing an outcome of size at most b. As discussed in Chapter 1, the most common application of this model are various elections of representative bodies, such as parliaments or supervisory boards—hence, in this model we will typically refer to the candidate projects as *candidates* and outcomes as *committees*. Note that in the committee election model there is no difference between direct and costwise score-utility mapping.

Since the unit cost assumption is often studied together with the assumption that voters' utilities are approval-based [Lackner and Skowron, 2022], we will use the term *approval-based committee election model* for the model with both assumptions.

Part I

Method of Equal Shares for Participatory Budgeting

Chapter 3

The Description of Method

In this chapter we introduce Method of Equal Shares (for brevity, we often refer to the rule as simply *Equal Shares*). Our rule is an extension of the algorithm designed by Peters and Skowron [2020] for approval-based committee election model, initially known as Rule X. We generalized this idea to the model with arbitrary costs and utilities. Before presenting its formal definition, let us start with the general intuition standing behind the algorithm.

A natural approach to ensure that selected projects proportionally correspond to the views of different voters is to directly map each selected project c to a subset of voters whose views are meant to be "represented" by c. Our idea is the following: we interpret an election as a market in which voters have some fictional money, corresponding to their voting power. A voter pays for a project only if she supports it (and hence, the project, if selected, would be compatible with her views). The projects are purchased (elected) sequentially. We adopt here the following principles:

1. Voters pay only for elected projects and the cost of every such project needs to be covered. The rule stops if no project can be afforded by her supporters.

2. Each voter has equal initial endowment. Because the price of each project is equal to its cost, and the total amount of spent money should not exceed b, the most natural value of the voter's initial endowment is b/n.

3. The rule prefers (*caeteris paribus*):

   (a) cheaper projects to more expensive ones,

   (b) projects with higher support to the ones from which each voter gains strictly lower utility.

4. The price of each elected project is distributed among her supporters proportionally to their utilities—or at least, as proportional as possible, without violating the previous principles. That is, if from some project voter 1 gains utility 1, voter 2 gains utility 3, and voter 3 gains utility 6—then voter 1 should ideally pay 10% of the project's cost, voter 2 should pay 30% and voter 3 should pay 60%. However, if for example, voter 3 does

not have enough remaining money, she is allowed to pay less. In general, we say that payment distribution should *minimize the maximal "payment-per-utility" ratio of all the voters gaining nonzero utility from a project*—for an ideally proportional distribution this ratio is equal for all the voters, as we can see in Figure 9.

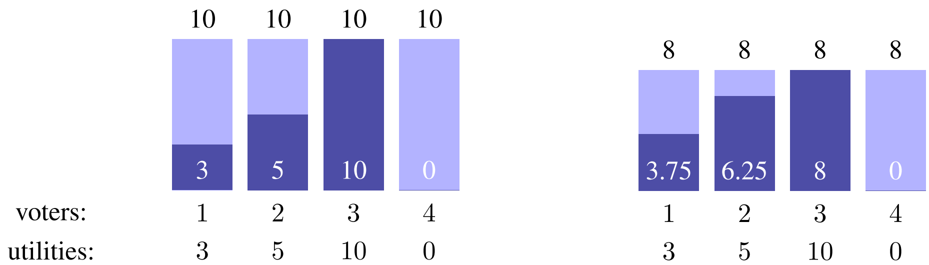


Figure 9: Consider a project c worth 18 dollars and 4 voters. The utilities of the voters from c are 3, 5, 10 and 0 respectively. Since 4 is a non-supporter, she does not participate in paying for c. In the scenario depicted in the left-hand side picture, each voter has 10 dollars left. Then, if the payments are distributed proportionally to the utilities, the payment-per-utility ratio of each supporter is 1. Note that under any other distribution, this ratio for some voter would be greater. In the right-hand side scenario, each voter has 8 dollars left. Now the proportional payment distribution is impossible, since 3 does not have enough money. Therefore we split the payment as follows: 3 pays all her savings (8 dollars). The remaining cost of c (10 dollars) needs to be covered by 1 and 2. From the previous observation, we know that the maximal payment-per-utility will be minimized, if they pay proportionally for this rest (so that voter 1 pays 3.75 dollars and voter 2 pays 6.25 dollars). Indeed, then the payment-per-utility ratio for 1 and 2 will be $^{3.75}/_{3} = ^{6.25}/_{5} = 1.25$ and for 3 it will be $^{8}/_{10} = 0.8$—hence, the maximal payment-per-utility ratio will be 1.25.

Let us denote this *maximal payment-per-utility ratio* of some project c by ρ. Note that then every voter $i \in N$ pays for c either her all savings or $u_i(c) \cdot \rho$ dollars.

Now we can observe that if in each round, we will choose a project minimizing ρ, both conditions mentioned in the third principle are achieved. Indeed, both decreasing the cost of a project and increasing voters' satisfaction from it would result in decreasing ρ.

Having this intuition, we are now ready to present the formal definition of Equal Shares:

Definition 3.1 (Method of Equal Shares)**.** Each voter is initially given an equal fraction of the budget, that is, each voter $i \in N$ is given $b_i \leftarrow {}^{b}/_{n}$ dollars. For $\rho \geq 0$, we say that a project $c \notin W$ is *ρ-affordable* if

$$\sum_{i \in N} p_i(c) = \mathrm{cost}(c), \qquad \text{where } p_i(c) = \min\left(b_i, u_i(c) \cdot \rho\right).$$

The algorithm starts with an empty committee W. Then it greedily chooses a project c that is ρ-affordable for minimal ρ and updates the voters' individual budgets: $b_i \leftarrow b_i - p_i(c)$ for each $i \in N$ (the value of $p_i(c)$ is the i's payment for an elected project c, for which she is charged). If no project is ρ-affordable for any ρ, Method of Equal Shares stops and returns W.

3.1 Technical Details

Algorithm 1 shows an implementation of Definition 3.1 as a pseudocode.

Algorithm 1: Implementation of Method of Equal Shares

```
W ← ∅.
For each voter i ∈ N, b_i ← b/|N|
while true do
    for c ∈ C \ W do
        if \sum_{i∈N:u_i(c)>0} b_i < cost(c) then
            ρ(c) ← ∞ (project c is not affordable)
        else
            Let 1, . . . , t be a list of all voters i ∈ N with u_i(c) > 0, ordered so
             that b_1/u_1(c) ≤ · · · ≤ b[t]/u_t(c) .
            for s = 1, . . . , t do
                ρ(c) ← (cost(c) − (b_1 + · · · + b[s − 1]))/(u_s(c) + · · · + u_t(c))
                if ρ(c) · u_s ≤ b[s] then
                    break (we have found the ρ-value)
    if min_{c∈C\W} ρ(c) = ∞ then
        return W
    c ← argmin_{c∈C\W} ρ(c) (break ties arbitrarily)
    W ← W ∪ {c}
    for i ∈ N such that u_i(c) > 0 do
        b_i ← b_i − min{b_i, ρ(c) · u_i(c)}
```

The only non-obvious part of the computation is how to determine the value of ρ for each project—we will use here the intuition described in the caption of Figure 9. Suppose that the algorithm has selected the set W thus far. To calculate the value of ρ for a project $c \in C \setminus W$, note first that only voters with positive utility for c (its "supporters") will pay for it. Label the supporters of c as $1, \dots, t$. Some of the supporters will spend all their remaining money for c, and others only part of it. Like in the pseudocode of the algorithm, write $b[i] = {}^{b}\!/_{n} - p_i(W)$ for the amount of money leftover for i at the current point in time. Rewriting the defining equation of ρ, we have that c is ρ-affordable if and only if

$$\sum_{i \in N} u_i(c) \cdot \min\left(b[i]/u_i(c), \rho\right) = \mathrm{cost}(c).$$

Note that the supporters who spend all their remaining money will have the minimum attained in the first coordinate; hence these voters have the lowest values of $b_i/u_i(c)$. Algorithm 1 sorts the supporters of c by this value, and then iterates through all $s \in [t]$ and checks whether the cost of c can be covered with voters $1, \dots s-1$ spending all their remaining money and $s, \dots, t$

spending $\rho \cdot u_i(c)$ for some ρ. If that is possible, then we can choose ρ such that

$$(b[1] + \cdots + b[s-1]) + \sum_{j=s}^{t} \rho \cdot u_j(c) = \text{cost}(c) \iff \rho = \frac{\text{cost}(c) - (b[1] + \cdots + b[s-1])}{u_s(c) + \cdots + u_t(c)}.$$

Note that it is only possible if ρ satisfies $\rho \cdot u_s \leq b[s]$. If this last condition fails, the algorithm needs to continue its iteration and try the next value for s. If the condition is satisfied, we have found the correct value of ρ.

This algorithm can easily be implemented to run in time $O(m^2 n \log n)$, assuming we can do arithmetic in constant time. To achieve this time bound, we need to store the current values of the partial sums $b[1] + \cdots + b[s]$ and $u_{s+1}(c) + \cdots + u_t(c)$, so that in each iteration of the inner for-loop, we can update these sums in constant time.

3.2 Examples

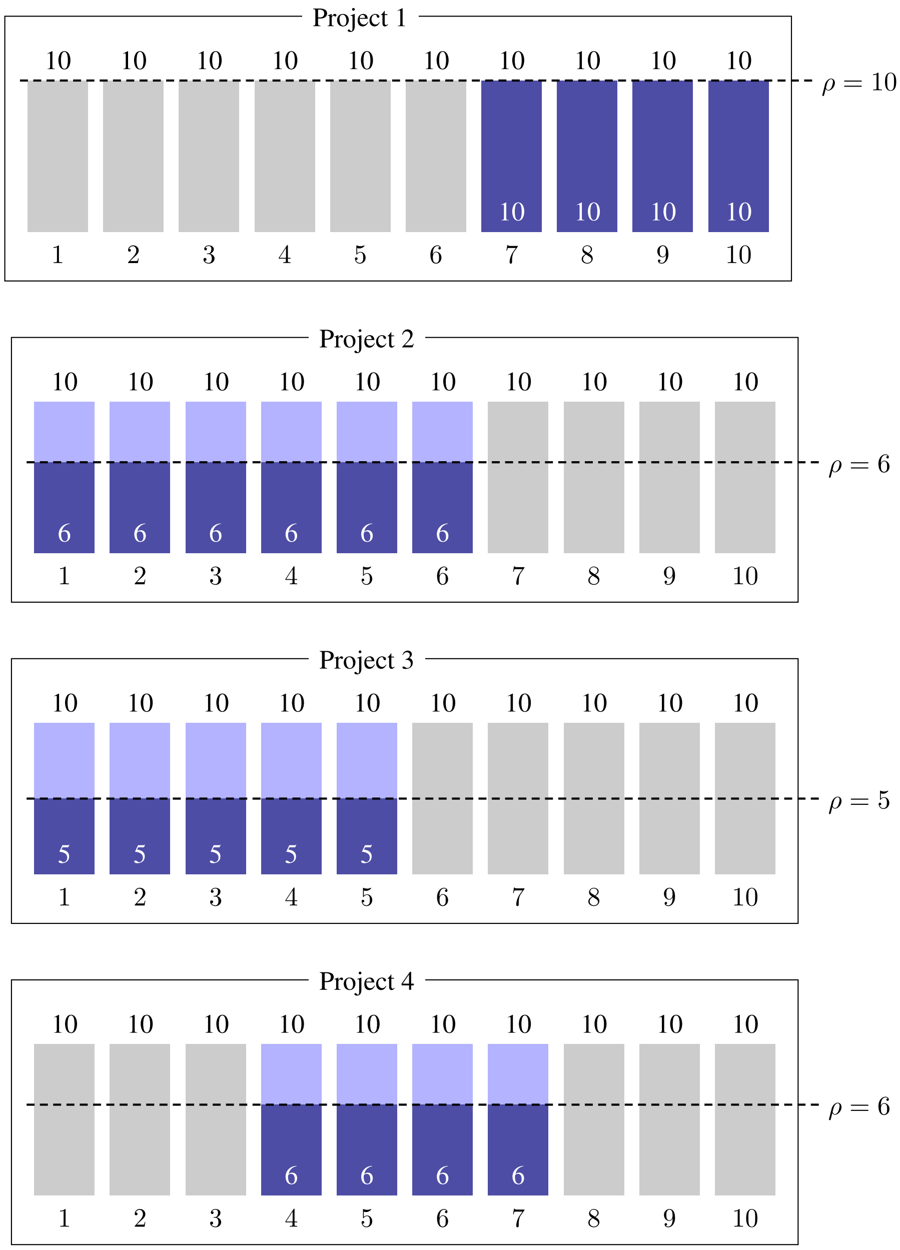
Let us now present three examples illustrating how the rule works. As a warm-up, let us first consider how Equal Shares would work in special cases where voters have approval-based utilities or approval-based cost utilities. Then we will present an example for arbitrary cardinal utilities.

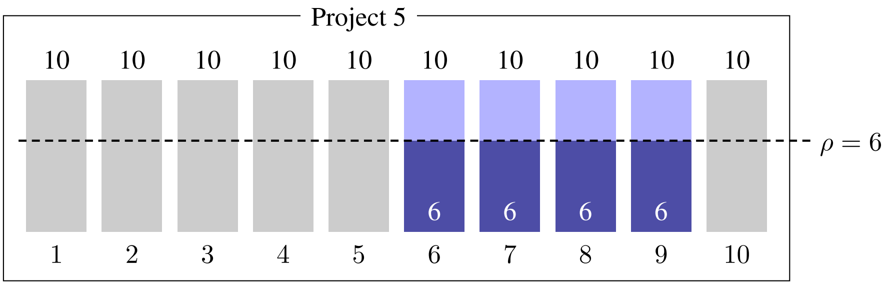
3.2.1 Special Case: Approval-Based Utilities

In the case of approval-based utilities, the definition of the algorithm becomes more straightforward. At each step, a project $c \notin W$ is ρ-affordable if and only if the cost of c can be covered by the voters approving c in such a way that the maximum payment of any voter is ρ. Voters who have less than ρ dollars left, spend all their money, and the other voters pay exactly ρ. This way, the cost is shared as equally as possible among the supporters of a project. See the following example for an illustration of the determination of ρ. Suppose we have 5 projects, 10 voters $N = \{1, \ldots, 10\}$ with approval-based utilities, and a budget of $b = \$100$. The project costs and utilities are shown in the table below.

	cost	1	2	3	4	5	6	7	8	9	10
Project 1	\$40	0	0	0	0	0	0	1	1	1	1
Project 2	\$36	1	1	1	1	1	1	0	0	0	0
Project 3	\$25	1	1	1	1	1	0	0	0	0	0
Project 4	\$24	0	0	0	1	1	1	1	0	0	0
Project 5	\$24	0	0	0	0	0	1	1	1	1	0

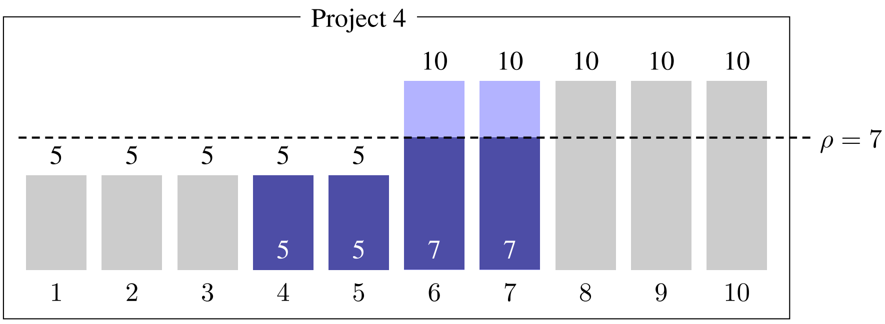
Sharing the available budget equally among the voters, everyone starts with \$10. We can now calculate, for each project, the value ρ such that the project is ρ-affordable. As discussed above, because we have approval-based utilities, this just entails spreading the cost as equally as possible among the project supporters. The values of ρ for different projects and the payment distribution is depicted below, in a way similar to Figure 9. In particular, columns represent the remaining money of the voters—gray for the non-supporters, violet for the supporters. The amount of money that would have been paid if a specific project was selected is marked with a darker shade of violet.





Equal Shares will choose to implement Project 3, since it has the lowest value of ρ.

Then voters 1 through 5 each pay \$5 for it. We can note that Project 2, costing \$36, is now not affordable anymore: its supporters have \$35 left. However, Projects 1, 4 and 5 are still affordable. For Projects 1 and 5 the values of ρ remain unchanged, since their supporters did not pay for Project 3. For Project 4, however, it is not the case:



Note that even though Projects 4 and 5 have the same cost and the same number of supporters, we can see that they now induce different ρ-values. The reason for this difference is that voters 4 and 5 supporting Project 4 have already been satisfied by Project 3. Hence, electing Project 4 would overrepresent these two voters at the expense of the supporters of Project 5.

Thus, Equal Shares next selects Project 5, and voters 6 through 9 each pay \$6 for it. We can check that at this point, no project is affordable. Thus, Equal Shares stops and selects the winning outcome $W = \{\text{Project } 3, \text{Project } 5\}$. Note that $\text{cost}(W) = \$49$ while $b = \$100$, and thus a large part of the budget was not spent by Equal Shares. We will discuss this issue in Section 3.3.

3.2.2 Special Case: Approval-Based Cost Utilities

In the case of approval-based cost utilities, the definition of Method of Equal Shares is quite similar to the one for approval-based utilities. Here, since all the supporters of each project $c \in C$ gain the same utilities from it, we again aim to spread the voters' payments for c as equally as possible. However, now the value of ρ corresponds no longer to the maximal payment of a voter, yet rather to *the maximal payment of a voter divided by a project's cost*. Intuitively, it favors more expensive projects, assuming they provide more satisfaction to the voters.

In fact, the process of determining which project should be selected in a given round can be even more simplified in this setting. Fix any round and a project $c \in C$. Suppose that we have already divided the payments for the cost of c as equally as possible. Then every voter $i \in N$ pays for c either b_i or some $p(c)$ dollars (equal to $\mathrm{cost}(c) \cdot \rho$). Let us denote by X the set of voters i such that $p(c) \leq b_i$. We have that:

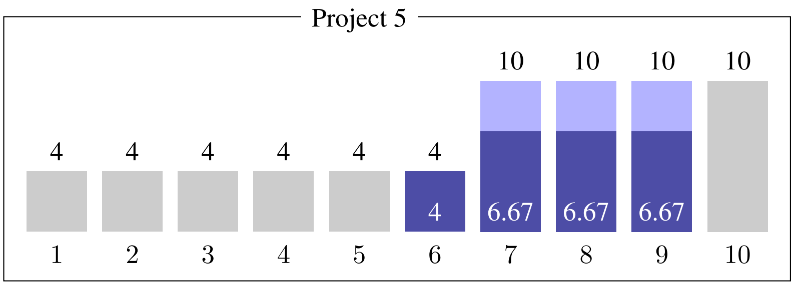
$$
\begin{aligned}
\mathrm{cost}(c) &= \sum_{i \in X} p(c) + \sum_{i \in N \setminus X} b_i, \\
\mathrm{cost}(c) &= |X| \cdot p(c) + \sum_{i \in N \setminus X} b_i, \\
\frac{\mathrm{cost}(c)}{p(c)} &= |X| + \sum_{i \in N \setminus X} \frac{b_i}{p(c)}, \\
\frac{1}{\rho} &= |X| + \sum_{i \in N \setminus X} \frac{b_i}{p(c)}.
\end{aligned}
$$

The last equality can be interpreted as follows: instead of choosing a project c minimizing ρ, we can choose the one maximizing the number of supporters. However, if a voter does not have enough money to pay the equal share $p(c)$, her vote is counted fractionally, as ${}^{b_i}/_{p(c)}$—for example, if a voter pays only 70% of the amount someone else needs to pay, her vote is counted with weight 0.7. The presented formulation is usually seen as more intuitive for the general audience than the original one. In particular, it has been used to define Method of Equal Shares in the PB statutes of Wieliczka and Świecie.

An interesting corollary from the presented observation is that in the first round the affordable project with the highest number of votes is always selected (since in the first round all the values of b_i are equal, a project that is affordable is paid equally by all the voters), which was not the case for approval-based utilities. Let us consider once again the previous example, yet assuming now approval-based cost utilities.

	cost	1	2	3	4	5	6	7	8	9	10
Project 1	\$40	0	0	0	0	0	0	40	40	40	40
Project 2	\$36	36	36	36	36	36	36	0	0	0	0
Project 3	\$25	25	25	25	25	25	0	0	0	0	0
Project 4	\$24	0	0	0	24	24	24	24	0	0	0
Project 5	\$24	0	0	0	0	0	24	24	24	24	0

All the projects are affordable and the distribution of payments in the first round is the same as in the previous example. Therefore, as we noted before, we choose the project with the highest number of approvals, namely Project 2. After that, Project 3 and Project 4 are no longer affordable. Let us consider now Project 1 and Project 5. Both projects gained 4 approvals. Since no supporter of Project 1 paid for Project 2, the payment distribution is still equal and Project 1 gains 4 full votes in the second round. For Project 5 the situation is different:



As we can see, Project 5 gained 3 "full" votes and 1 "fractional" one, counted as 0.6 vote (4 divided by $^{20}/_{3}$). Hence, Project 1 is better and it is elected. After that, no project is affordable (voters 7 to 10 run out of money) and the rule stops, returning the outcome $\{\text{Project } 1, \text{Project } 2\}$. Comparing this result to the one under approval-based utilities, note that we have elected two strictly more expensive projects and have spent a significantly greater amount of the budget ($\$76$ instead of $\$49$). However, we still could have selected one more project, either Project 4 or Project 5, to spend the whole available budget of $\$100$. Let us once again leave this discussion for later (see Section 3.3) and, to ensure the definition is fully clear, present the final example for general cardinal utilities.

3.2.3 General Case

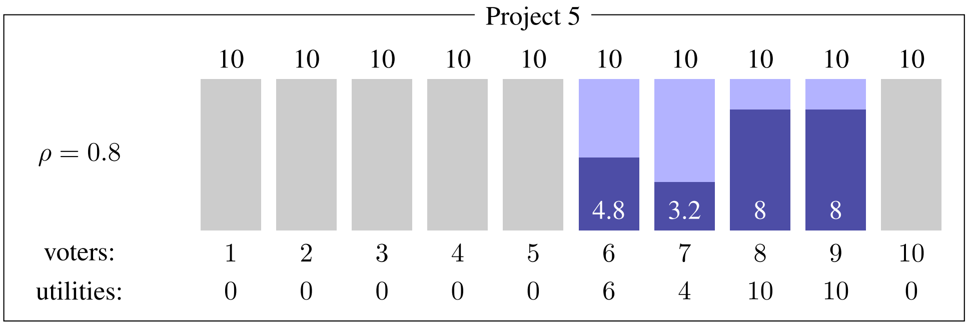
Let us once again consider an analogous example to the previous ones, yet with some arbitrary cardinal utilities.

	cost	1	2	3	4	5	6	7	8	9	10
Project 1	\$40	0	0	0	0	0	0	10	10	10	10
Project 2	\$36	3	4	5	1	2	3	0	0	0	0
Project 3	\$25	5	4	4	1	1	0	0	0	0	0
Project 4	\$24	0	0	0	4	1	3	4	0	0	0
Project 5	\$24	0	0	0	0	0	6	4	10	10	0

The budget is still $\$100$, so again every voter starts with $\$10$. Now to speed up the calculations, let us first calculate the *best-case* ρ for every project—the minimal possible value of ρ under which each voter would pay their proportional share of the cost of the project (possibly not attainable). For each project $c \in C$, it is equal to $\mathrm{cost}(c)$ divided by $\sum_{i \in N} u_i(c)$. We obtain:

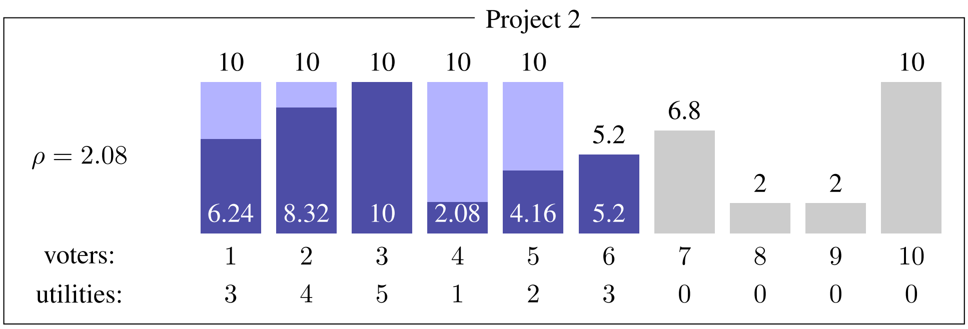
	best-case ρ
Project 1	1
Project 2	2
Project 3	2.27
Project 4	2
Project 5	0.8

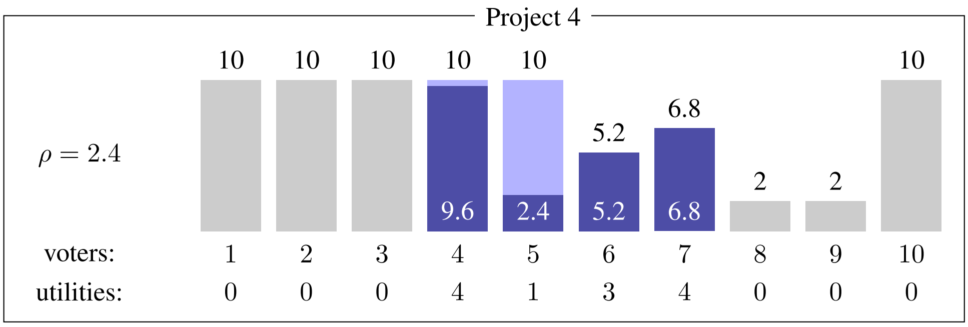
Now let us start from computing the actual ρ for Project 5, according to the procedure described in Algorithm 1. It is simple, since it is possible to split the payments proportionally to the utilities:



Note that now we can immediately say that Project 5 is elected in the first round—we do not need to compute the values of ρ for remaining projects, since they would be greater or equal to their values of "best-case ρ", which are higher than the ρ of Project 5.

In the second round, let us consider the next projects with the lowest "best-case ρ". The first one is Project 1, which is yet not affordable after electing Project 5. The next possible choices are Project 4 or Project 2. Now we can see that for both of them the values of "best-case" and "actual" ρ differ:





Since for Project 2 the value of ρ is smaller both from the ρ of Project 4 and from "best-case ρ" of Project 3, we can again say immediately that Project 2 is elected. Note that now the value of "best-case ρ" of Project 4 can be updated to 2.4 (since in further rounds its ρ can only get worse). Finally, we observe that no project is affordable in the next round and the algorithm stops, returning the outcome $W = \{\text{Project } 2, \text{Project } 5\}$.

Our publicly available implementation of Equal Shares in the *Pabutools* Python package actually includes the presented optimization, keeping in memory the values of "best-case ρ" and sorting the projects by them in each round. Even though the worst-case time complexity of the optimized algorithm is the same as in the standard implementation presented in Algorithm 1 (assuming $m \leq n$), for large real-life elections it is faster by more than 95% on average.

We can intuitively see that Equal Shares behaves in a proportional way and balances the influence of the voters on the elected outcome—the number of unrepresented voters is at most one in every case. Let us now consider how to modify it so that it would utilize a greater fraction of the initial budget.

3.3 Exhaustive Variants of Equal Shares

We have already noticed that in all the examples in Section 3.2, Equal Shares spent intuitively too little fraction of the budget. Of course, when the costs of the projects are unequal, we can never guarantee that a rule will spend exactly the entire budget. However, we can require that no project can be added to the elected outcome without violating the budget constraint.

Definition 3.2 (Exhaustiveness, Aziz and Lee, 2021)**.** An election rule $\mathcal{R}$ is *exhaustive* if for each election E and each non-selected candidate $c \notin \mathcal{R}(E)$ it holds that:

$$\mathrm{cost}(\mathcal{R}(E) \cup \{c\}) > b.$$

Under the unit cost assumption, the above definition boils down to the requirement that we should elect a committee of size *exactly* b. Note that no outcome returned by Equal Shares in the examples presented in Section 3.2 was exhaustive.

In some contexts, violating this property may actually be a desirable feature of Equal Shares, especially if unspent budget can be used in other productive ways (such as in the next year's PB election). Arguably, when an outcome is non-exhaustive under this rule, no remaining project has sufficient support to justify its expense; on that view, no further projects should be funded. In other situations, unspent budget may not be reusable, such as when the budget comes from a grant where unspent money needs to be returned (and the relevant decision makers do not obtain value from the grant-maker's alternative activities), or when the "budget" is time (for example, when we use PB to plan activities for a day-long company retreat). In such situation, one might prefer an exhaustive rule.

Let us now present a few strategies to "complete" Method of Equal Shares.

Completion by Utilitarian Greedy The simplest way to complete the outcome elected by Equal Shares is to use another rule that is known to be exhaustive. The natural candidate is the Utilitarian Greedy protocol intuitively described in Chapter 1 for approval ballots. In fact, this is practically the only rule used in practice for Participatory Budgeting.[1] For score or ranking ballots, the protocol picks in each round the project maximizing the total score. When defining this procedure formally, one could wonder, whether the above description assumes direct or costwise score-utility mapping. In our opinion, the second option is much more reasonable; then the definition for the general case is the following:

Definition 3.3 (Utilitarian Greedy (UG))**.** We start with an empty outcome $W = \emptyset$, and repeatedly select a project c maximizing the ratio: $\left(\sum_{i \in N} u_i(c)\right) / \mathrm{cost}(c)$. If $\mathrm{cost}(W) + \mathrm{cost}(c) \leq b$ then we add project c to W; otherwise, we remove the project from consideration and repeat, until no more projects remain.

The above definition has a very clear goal: it aims at maximizing the total utility of the voters, $\sum_{i \in N} u_i(W)$. Indeed, the Utilitarian Greedy rule is *optimal up to one project* for this objective [Dantzig, 1957], that is, for each outcome W returned by UG there exists $p \notin W$ s.t.:

$$\sum_{i \in N} u_i(W \cup \{p\}) \geq \max_{W' \colon \mathrm{cost}(W') \leq b} \sum_{i \in N} u_i(W').$$

An advantage of completing Equal Shares with Utilitarian Greedy is that it is intuitively efficiency-driven—it spends the remaining budget aiming for maximizing the highest total utility.

Completion by varying the budget In this variant, we evaluate the method with a budget $b' \geq b$ that is higher than the actually available budget b. We increase the value of b' gradually (each time by some fixed amount ε), and after each increase we recompute the outcome from scratch using Equal Shares. We stop when we reach an exhaustive outcome or when the next increase of b' would cause us to exceed the original budget b. The selected outcome is then typically exhaustive, but formally there is no guarantee that the budget is exhausted, as it can be seen in the following example.

[1] The only notable exception is Paris, using the rule called Majority Judgment [Balinski and Laraki, 2010]. We skip this rule in our analysis, since it focuses neither on maximizing utilitarian welfare, nor on providing proportionality guarantees. It has been proved that in some cases it returns highly unintuitive outcomes [Laslier, 2019].

Example 3.4. Suppose that $b = 1$, we have 3 candidates and 3 voters. The first 2 voters approve $\{c_1\}$, and the third one approves $\{c_2, c_3\}$. We have $\mathrm{cost}(c_1) = 1$ and $\mathrm{cost}(c_2) = \mathrm{cost}(c_3) = 1/3$. The only exhaustive outcomes are $\{c_1\}$ and $\{c_2, c_3\}$. However, neither of them is elected by Equal Shares completed by varying the budget—indeed, to buy both c_2 and c_3, the third voter needs to control at least $2/3$ dollars. Then the first two voters control $4/3$ dollars and can buy candidate c_1, a contradiction. On the other hand, to buy c_1, the first two voters need to control at least 1 dollar. Then, the third voter controls at least $1/2$ dollars and buys c_2 or c_3, a contradiction. ⌟

However, our experiments in Chapter 5 suggest that Equal Shares with varying the budget spends nearly the whole budget on real-life elections, especially on the larger ones (about $98\% \cdot b$), and therefore is "nearly" exhaustive. In order to spend the remaining small fraction of the budget, we can complete the outcome with the use of another rule as in the previous variant.

One could ask why we increase the budget gradually, instead of using binary search to find the optimal value of b' faster. We may intuitively feel that giving more money to the voters should result in electing a more expensive outcome. However, it is not the case, even for the committee election model, as presented in Example 3.5.

Example 3.5. Consider the following committee election with $n = 14$ voters and $m = 15$ candidates.

$$\begin{aligned}
&4 \text{ voters}: \{a_1, ..., a_8, b_1\}\\
&4 \text{ voters}: \{a_1, ..., a_8, b_2\}\\
&1 \text{ voter}: \ \{c_1, c_2, c_3, b_1, b_2\}\\
&2 \text{ voters}: \{c_1, c_2, c_3\}\\
&1 \text{ voter}: \ \{d_1, d_2, d_3, b_1, b_2\}\\
&2 \text{ voters}: \{d_1, d_2, d_3\}
\end{aligned}$$

If we set the budget to $b' = 14$ (so that each voter is given 1 dollar), then Equal Shares chooses $a_1, \ldots, a_8, c_1, c_2, c_3, d_1, d_2, d_3$, that is, it selects 14 candidates. However, if we increase the budget to $b' = 49/3$ (so that each voter is given $7/6$ dollar), then the rule first chooses $a_1, \ldots, a_8$, and next it selects b_1 and b_2. For b_1, b_2 each voter pays $1/6$. Thus, each voter who approve both b_1 and b_2 is left with $5/6$ dollar. Then we can see that the supporters of c- and d- candidates have less than 3 dollars in total. Thus, only two c- and d-candidates can be selected. It means that at most 13 candidates can be selected, which shows the lack of monotonicity. ⌟

An advantage of varying the budget is that a larger part of the budget is spent in a way such that the voting power is equally shared. Besides, this variant appeared to the best one according to our experiments on real-life data described in Chapter 5. Hence, it has been recommended by us to use in practice and implemented in Wieliczka, Świecie and Aarau.

Completion by perturbation Since Equal Shares works for general additive valuations, there is another way for us to make it exhaustive. Recall that Equal Shares fails to be exhaustive in situations where the remaining projects' supporters do not have sufficient funds left. However, in elections where $u_i(c) > 0$ for all $i \in N$ and $c \in C$, every voter supports every candidate, and thus this problem never occurs. In fact, Equal Shares is exhaustive when run on profiles of this type.

Proposition 3.6. *Consider an election E such that $u_i(c) > 0$ for each $i \in N$ and $c \in C$. The outcome returned by Method of Equal Shares for E is exhaustive.*

Proof. For the sake of contradiction assume that an outcome W returned by Equal Shares for an election E is not exhaustive. Then, there exists a candidate $c \notin W$ such that $\mathrm{cost}(W \cup \{c\}) \leq b$. The voters paid in total $\mathrm{cost}(W)$ for W; their total initial budget was b, thus after W is selected they all have at least $\mathrm{cost}(c)$ unspent money. However, this means that at the end of the execution of Equal Shares there exists a possibly very large value of ρ such that

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

Consequently, c (or some other candidate) would be selected by Equal Shares, a contradiction. □

The most straightforward way to use this idea is to override each $u_i(p) = 0$ with $u_i(p) = \varepsilon$ for some very small value of ε. However, in the *Pabutools* package we have included an implementation which in fact does not depend on the ε: in the first step, we run the standard Equal Shares algorithm for the original election. Then, in the second step, for each unelected project c we assume that all the voters with positive score over c spend all their remaining money to cover as much of the cost of c as possible. The remaining part of the cost is covered by the voters $i \in N$ such that $u_i(c) = 0$ so that the maximal payment of such a voter, β, is minimized. We select the project minimizing β and update voters' individual budgets. Intuitively, this implementation corresponds to setting ε to some *infinitely low* value, so that tweaking the utilities has no effect until no projects are affordable by their supporters.

An advantage of this variant is that it is probably the most natural one; it uses purely the internal mechanism of Equal Shares to elect the whole exhaustive outcome.

In Chapter 5 we will present the comparison of all the variants, based on the real-life data analysis. However, independently from which variant we choose, the most important axiomatic properties of Equal Shares are actually the same—therefore in the next chapter we will consider only the basic variant of Equal Shares, presented in Definition 3.1.

Chapter 4

Axiomatic Properties of Equal Shares

After reading the definition of Method of Equal Shares and seeing how it works on the examples from Section 3.2, we probably already have an intuition that this rule in some sense represents groups of voters proportionally to their size. Indeed, consider a simple election from Example 1.1. Now, under Method of Equal Shares, the first 40% of voters control 40% of the budget, while the next two groups of 30% of voters control 30% of the budget, each. Hence, since the groups do not overlap, the voters will eventually buy 4 red projects, 3 blue and 3 green ones.

Observation 4.1. Note that the above argumentation is valid independently from the way how exactly particular projects are chosen in each round. To claim that Equal Shares is proportional in such simple cases it is enough to note it constructs a system of payments satisfying the following simple conditions: (1) every project needs to have its cost covered in order to be selected, (2) all the voters have equal initial endowments, specifically equal to at least b/n, (3) the voters spend money only on the selected projects they support, and (4) the rule stops only if no project is affordable.

We further explore the above observation in Section 4.3, combining the aforementioned conditions into the *priceability* axiom [Peters and Skowron, 2020]. However, let us first focus on perhaps the most important axiomatic property of Equal Shares—Extended Justified Representation (EJR) [Aziz et al., 2017a]. This axioms indicates that the rule is proportional also in the most general PB setting, that is the setting with arbitrary costs of the projects and arbitrary cardinal utilities. Here, the exact way the rule is designed would be crucial—specifically, the fact that in each round we select a project and the corresponding voters' payments so that the value of ρ is minimal.

4.1 Extended Justified Representation (EJR)

Extended Justified Representation (EJR) was first proposed for approval-based committee election model [Aziz et al., 2017a]. Even for approval-based utilities, only few rules are known to satisfy EJR and Method of Equal Shares is one of them [Peters and Skowron, 2020]. In this section, we introduce a generalization of EJR to the PB model with cardinal utilities and show that Equal Shares still satisfies EJR.

We first present the original definition of EJR by Aziz et al. [2017a]. The intuition behind this axiom is similar to the already presented one in Definition Attempt 1.2—as a warm-up, let us recall this definition in a slightly adapted way for the approval-based committee election model:

Definition Attempt 4.2. We say that an outcome W is proportional, if for every group of voters S, approving jointly at least some $\ell \geq {}^{|S|}/_{n} \cdot b$ candidates, every voter from S approves ℓ candidates in W.

As we already noticed in Example 1.3, this definition is not always satisfiable. The idea of Extended Justified Representation is therefore to keep the premise of this condition (we call groups S satisfying this premise *ℓ-cohesive groups*), yet weaken the conclusion—now we require that only *one* member of S is satisfied, not all of them.

Definition 4.3 (Extended Justified Representation for the approval-based committee election model [Aziz et al., 2017a])**.** A group of voters S is *ℓ-cohesive* for $\ell \in \mathbb{N}$ if

$$|S| \geq \frac{\ell}{b} \cdot n \quad \text{and} \quad |\bigcap_{i \in S} A_i| \geq \ell.$$

We say that an outcome W satisfies Extended Justified Representation (EJR), if for every ℓ-cohesive group of voters S, at least one voter from S approves ℓ candidates in W.

After this weakening, outcomes satisfying EJR not only always exist [Aziz et al., 2017a], but also are possible to be found by polynomial-time computable algorithms [Aziz et al., 2018, Peters and Skowron, 2020].

At first sight, one could worry that with only the at-least-one guarantee, EJR is a weak property. However, it is not the case—in particular, it implies that the members of a cohesive group have high utility on average [Aziz et al., 2018, Skowron, 2021]. The intuition here is the following: consider an ℓ-cohesive group S. EJR requires that there exists $i \in S$ approving ℓ candidates. However, the group $S \setminus \{i\}$ is still at least $(\ell - 1)$-cohesive. Hence, there also exists a voter $j \in S$, $j \neq i$, approving $\ell - 1$ canididates. Continuing this reasoning, we obtain that voters in S have an average utility of ${}^{\ell-1}/_{2}$.

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. To warm up, let us first relax the unit cost assumption, but keep the assumption that the voters have approval-based utilities, that is, they gain from each selected approved project utility 1, regardless of the project's cost.

Example 4.4. Consider an election with approval-based utilities. Let $n = 1000$ and $b = 100$. Consider five disjoint groups of voters, S_1, S_2, S_3, S_4 and S_5, each of which consists of 30 voters with the same approval sets. Hence, looking only at sizes of groups, intuitively each of them is large enough to have the right to decide about spending 3 dollars (the fraction of the budget proportional to their sizes). Moreover, we assume that each group approves the projects worth at least 3 dollars (which was the condition stated in Definition Attempt 1.2). However, looking closer at the approval sets, we can see they are very different:

- Group S_1 approves three projects a_1, a_2, a_3, each of which costs 1 dollar,
- Group S_2 approves two projects, b_1 costing 2 dollars and b_2 costing 1 dollar,
- Group S_3 approves two projects, c_1 and c_2, costing 2 dollars each, and a project c_3 costing 1 dollar,
- Group S_4 approves one project d, costing 3 dollars,
- Group S_5 approves one project e, costing 4 dollars.

It is therefore crucial to notice that each group has a right to spending 3 dollars only if they can do it without exceeding this treshold. Now we probably intuitively feel that:

- Group S_1 clearly deserves all the 3 projects they approve, as they would do in the committee election model,
- So does group S_2—the total cost of the set $\{b_1, b_2\}$ they approve does not exceed 3.
- Group S_3 does not deserve all the projects they approve, because their total cost of 5 exceeds the proportional share of the budget. However, there exists a subset, for example, $\{c_1, c_3\}$ approved by S_3 which has cost 3, and approving additional projects should not affect their guarantees negatively. Hence, S_3 is entitled to the set $\{c_1, c_3\}$.
- Group S_4 deserves project $\{d\}$,
- Group S_5 does not deserve anything—the only way to give them any satisfaction is to select e, which would exceed their proportional share of the budget. ⌟

As we can see, if the projects have different costs then instead of speaking about ℓ-cohesive groups (groups that are large enough to be entitled to some ℓ candidates they jointly approve) it is more convenient to speak about T-*cohesive* groups for some $T \subseteq C$—groups that are large enough to be entitled to some subset of projects T they jointly approve. Naturally, the fact that a group is T-cohesive does not mean that *exactly* set T should be funded—otherwise, the problem of satisfiability would remain (recall Example 1.3). We would rather say that, analogously to the committee election model, EJR should require that at least one voter from a T-cohesive group should be satisfied from the elected outcome *as if T was elected*—hence, should get at least the utility of $|T|$.

Definition 4.5 (Extended Justified Representation for elections with approval-based utilities)**.** We say that a group of voters S is T-cohesive for $T \subseteq C$ if

$$\frac{|S|}{n} \geq \frac{\mathrm{cost}(T)}{b} \quad \text{and} \quad T \subseteq \bigcap_{i \in S} A_i.$$

An outcome W satisfies *Extended Justified Representation* if for each $T \subseteq C$, and each T-cohesive group S of voters there exists a voter $i \in S$ such that $|A_i \cap W| \geq |T|$.

The definition for the approval-based cost utilities would be the same, with the only change that in the last inequality we should have $|A_i \cap \mathcal{R}(E)| \geq \mathrm{cost}(T)$. Intuitively, when we are entitled to some set T of approved projects, then under approval-based utilities we would be equally happy from any other $|T|$ approved projects as we would be from T. Under the approval-based cost utilities it is no longer the case and instead we require a set of projects worth $\mathrm{cost}(T)$ from which we would gain the same utility as from T.

As we can see, while in the premise of the EJR definition (the definition of a cohesive group) we are interested in projects' costs, in the conclusion we only care about voters' utilities from them—the fact that in the committee election model we had the same value of ℓ in both parts, turned out to be merely coincidential.

If we try to generalize EJR beyond approval-based utilities, the definition of T-cohesiveness becomes meaningless, because the notion of set T approved by all voters in S does not have an analogue. However, if we keep the intuition that from every project $c \in T$ every voter in S gains an identical utility, we can actually resign from this condition. Let us consider the following (not-yet-final) definition:

Definition Attempt 4.6 (First attempt to define EJR for cardinal utilities)**.** We say that a group of voters S is T*-cohesive* for $T \subseteq C$ if

$$\frac{|S|}{n} \geq \frac{\mathrm{cost}(T)}{b} \qquad \text{and} \qquad u_i(c) = u_j(c) \text{ for each } c \in T, i, j \in S.$$

An outcome W satisfies *Extended Justified Representation* if for each $T \subseteq C$, and each T-cohesive group S of voters there exists a voter $i \in S$ such that $u_i(W) \geq u_i(T)$.

Note that this definition in the approval-based (cost) model, boils down to the previous ones. We do not even need to add the condition that voters from S gain nonzero utilities from projects in T, since adding such projects to T never helps group S to improve their satisfaction.

However, the above definition is still somewhat weak. Suppose, for example, that a group of three voters $S = \{1, 2, 3\}$ is large enough to be entitled to a project c. However, they gains utilities of 10, 5 and 7 from c respectively. According to the above definition, they do not deserve anything. On the other hand, if they all gained utility 5 from c, they would form a $\{c\}$-cohesive group entitled to satisfaction of 5.

Now we should have the proper intuition to formulate the final definition of EJR: we want to strengthen Definition Attempt 4.6 so that increasing a voter's utility from a project would never reduce her proportionality guarantees. In the above example, if all the voters from S agree that c is worth *at least* 5, at least one of them should gain utility 5 from the elected outcome. This leads us to the final definition of EJR:

Definition 4.7 (Extended Justified Representation)**.** A group of voters S is (α, T)*-cohesive*, where $\alpha \in \mathbb{N}$ and $T \subseteq C$, if

$$\frac{|S|}{n} \geq \frac{\mathrm{cost}(T)}{b} \qquad \text{and} \qquad \alpha \leq \sum_{c \in T} \min_{i \in S} u_i(c).$$

An outcome W satisfies *Extended Justified Representation* if for each $\alpha \in \mathbb{N}$, $T \subseteq C$, and each (α, T)-cohesive group of voters S there exists a voter $i \in S$ such that $u_i(W) \geq \alpha$.

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 value α denotes how much the coalition S agrees about the desirability of the projects in T. Consequently, Definition 4.7 prohibits any outcome in which every voter in S gets utility strictly lower than α.

Now EJR is a much more demanding property in the model with general cardinal utilities than its approval-based counterpart. Consider the special case where there is only one voter, $N = \{1\}$. Then any outcome W satisfying EJR must solve the *knapsack* problem, that is, it must maximize $\sum_{c \in W} u_1(c)$ subject to the budget constraint, since otherwise an optimum knapsack T would witness an EJR violation. Because the knapsack problem is weakly NP-hard, this presents a difficulty for a rule to satisfy EJR. Knapsack is also weakly NP-hard when assuming that the value of each item equals its weight (this is the *subset sum* problem), so satisfying EJR is weakly NP-hard even for approval-based cost utilities.

Proposition 4.8. *Unless P = NP, no aggregation rule that can be computed in strongly polynomial time can satisfy EJR in the general PB model.*

Equal Shares can be computed in strongly polynomial time, and indeed it fails EJR when the utilities are not approval-based, even with keeping the unit cost assumption.

Example 4.9 (Equal Shares fails EJR.)**.** Suppose that the budget is 2 dollars. There are two voters 1, 2, and three projects: c_1, c_2, c_3. The utilities are the following: $u_1(c_1) = u_2(c_1) = 2$, $u_1(c_2) = u_2(c_3) = 3$, and $u_1(c_3) = u_2(c_2) = 0$. Then c_1 is $1/4$-affordable and c_2 and c_3 are both $1/3$-affordable. So c_1 is elected (with both voters paying equal amounts). Now nothing is affordable, and thus $W = \{c_1\}$. But then $S = \{1\}$ and $T = \{c_2\}$ witness the EJR violation. ⌟

In the second part of the dissertation (specifically, in Chapter 7) we will prove that for every election, an EJR outcome does exist, but we do not know an efficiently computable method that finds such an outcome. However, we can show that Equal Shares selects an outcome that satisfies a mild relaxation of EJR, which requires that EJR holds "up to one project".

Definition 4.10 (Extended Justified Representation up to one project (EJR-1))**.** A rule $\mathcal{R}$ satisfies Extended Justified Representation *up to one project* if for each election E and each (α, T)-cohesive group of voters S there exists a voter $i^* \in S$ such that either $u_{i^*}(\mathcal{R}(E)) \geq \alpha$ or for some $c^* \in T$ it holds that $u_{i^*}(\mathcal{R}(E) \cup \{c^*\}) > \alpha$.

It is worth emphasizing that in the approval-based utilities model, Definitions 4.7 and 4.10 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 score and approval-based, we may assume that $\alpha = |T|$: indeed, for each $c \in T$, if $\min_{i \in S} u_i(c) > 0$ then $\min_{i \in S} u_i(c) = \max_{i \in S} u_i(c) = 1$; otherwise, as we have mnetioned earlier, we can remove c from T without losing cohesiveness. Finally, note that in the approval-based model, due to the strict

	Approval-based utilities	Approval-based cost utilities	Cardinal utilities
Unit costs	EJR	EJR	EJR-1
General costs	EJR	EJR-1 †	EJR-1 †

Table 4.1: Equal Shares and Extended Justified Representation (see Theorem 4.11 and Example 4.9).
†: Unless P = NP, no strongly polynomial time method (such as Equal Shares) can satisfy EJR.

inequality, both conditions $u_{i^*}(\mathcal{R}(E)) \geq \alpha$ and $\exists_{c^* \in T}. u_{i^*}(\mathcal{R}(E) \cup \{c^*\}) > \alpha$ boil down to $|A_{i^*} \cap \mathcal{R}(E)| \geq \alpha = |T|$.

Our main result is that Method of Equal Shares satisfies EJR up to one project in the general PB model. By the previous observation, it hence satisfies EJR in the approval-based utilities model (even when not imposing unit costs), see Table 4.1.

Theorem 4.11. *Method of Equal Shares satisfies EJR up to one project in the participatory budgeting model.*

Proof. Let $S \subseteq N$ be a non-empty group of voters, and let $T \subseteq C$ be a proposal with $\mathrm{cost}(T)/b \leq |S|/n$. For each $c \in C$, write $\alpha_c = \min_{i \in S} u_i(c)$ We assume that $\alpha_c > 0$ for all $c \in T$ (otherwise we can delete c from T). If W is the output of Equal Shares, we will show that there exists a voter $i^* \in S$ such that either $u_{i^*}(W) \geq \alpha$, or there is a project $c^* \in T$ such that $u_{i^*}(W \cup \{c^*\}) > \alpha$.

In this proof, we will consider three runs of Equal Shares in different variations:

(A) Equal Shares run on the original election (thus, outputting W).

(B) Equal Shares run so that voters in S are not bound by their budget constraint when paying for projects in T. To make this formal, in the definition of Equal Shares, we redefine the notion of ρ-affordability so that $c \in T$ is ρ-affordable if

$$\underbrace{\sum_{i \in S} \rho \cdot u_i(c)}_{\text{no budget limit}} + \underbrace{\sum_{i \in N \setminus S} \min\{b_i, \rho \cdot u_i(c)\}}_{\text{with budget limit}} = \mathrm{cost}(c), \tag{4.1}$$

and $c \in C \setminus T$ is ρ-affordable if

$$\sum_{i \in S} \min\{\underbrace{\max\{b_i, 0\}}_{b_i \text{ may be } < 0}, \rho \cdot u_i(c)\} + \sum_{i \in N \setminus S} \min\{b_i, \rho \cdot u_i(c)\} = \mathrm{cost}(c),$$

with payments defined as these equations suggest (namely, i's payment is the value of the ith term of the sum).

(C) Equal Shares run on a smaller election where only projects in T and only voters in S exist, and each voter has an unlimited budget $b_i = \infty$. In addition, we set $u_i(c) = \alpha_c$ for all $i \in S$ and $c \in T$.

Note that in variations (B) and (C), all projects in T will be elected (eventually) as $\alpha_c > 0$ for all $c \in T$, and the voters in S have unrestricted budgets when buying projects in T.

If at the end of the execution of (B) all voters in S have spent strictly less than b/n, then (B) has selected all projects in T and no voter has overshot their budget. Thus (A) also elects all of T, so $u_i(W) \geq u_i(T) \geq \alpha$ for all $i \in S$, and we are done. Otherwise, let $i^* \in S$ be the first voter in S who during the execution of (B) spends at least b/n. Suppose this happens just after (B) adds project c^* to the outcome. Write $W_{(B)}$ for the set of projects selected by (B) up to but excluding c^*. Note that (A) has also selected all projects in $W_{(B)}$, because until that point the two rules behave identically.

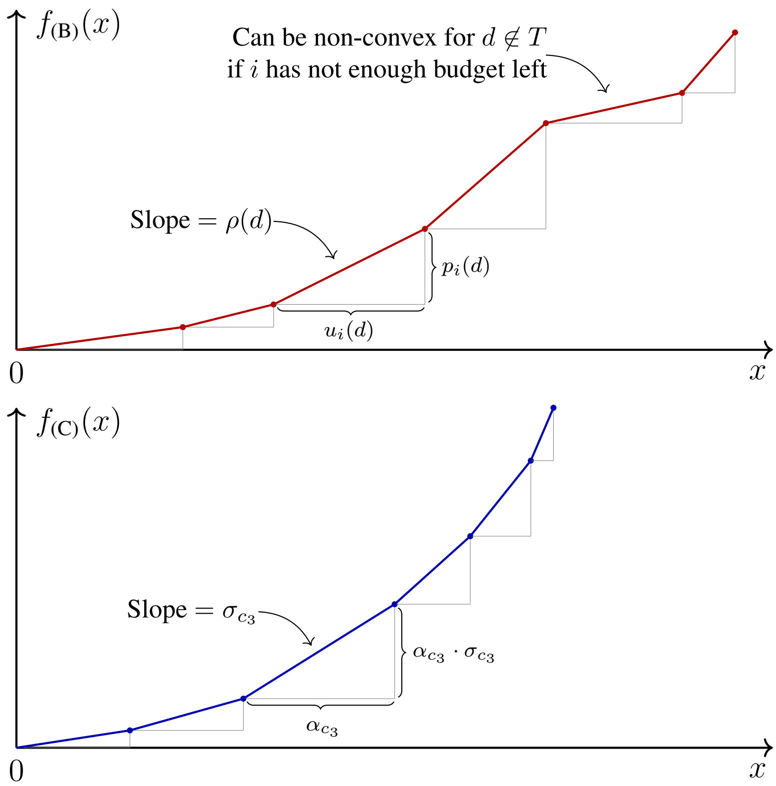


Figure 10: Illustration of functions $f_{(B)}(x)$ and $f_{(C)}(x)$.

Next, we will lower bound the utility that i^* receives under (B) by the time i^* has spent at least b/n. To do so, we define a function $f_{(B)}$ so that for a number x, $f_{(B)}(x)$ is the amount of money that i^* had to spend during the execution of (B) until i^* receives utility x. We make this into a continuous, piecewise linear function, so that to get a β-fraction of the utility of a project one needs to spend a β-fraction of the total spending for that project. See Figure 10. In other words, $f_{(B)}$ consists of a sequence of line segments where the segment corresponding to $d \in C$

has length $u_i(d)$ and height $p_i(d)$ (the amount that i^* paid for d). Note that the segment has a slope of usually $\rho(d)$, but it can be lower than $\rho(d)$ in case $d \notin T$ and i^*'s budget was not enough to pay the full amount $\rho(d) \cdot u_{i^*}(d)$ for d.

We can define a function $f_{\text{(C)}}$ in exactly the same way with respect to the execution of (C), based on the same voter i^*. The function $f_{\text{(C)}}$ is easy to understand. For each $c \in T$, let us write $\sigma_c = \text{cost}(c)/(|S|\alpha_c)$, and let us label $T = \{c_1, \dots, c_r\}$ such that $\sigma_{c_1} \leq \dots \leq \sigma_{c_r}$. Note that under (C) each not-yet-selected $c \in T$ is σ_c-affordable, because all voters have unlimited budgets and

$$\sum_{i \in S} \sigma_c \cdot \alpha_c = |S| \cdot \sigma_c \cdot \alpha_c = \text{cost}(c).$$

It follows that $f_{\text{(C)}}$ consists of a sequence of line segments of length α_c and slope σ_c, one for each $c \in T$. These line segments come in increasing order of σ_c, that is,in the order $c_1, \dots, c_r$, because Equal Shares always selects the ρ-affordable project with lowest ρ.

We claim that

$$f_{\text{(C)}}(x) \geq f_{\text{(B)}}(x) \quad \text{for all } x \in [0, \alpha]. \tag{4.2}$$

(Intuitively, this says that under (C) the money of i is used less efficiently for i than under (B).) The inequality certainly holds at $x = 0$ because both functions take the value 0. To establish (4.2) for other x, we will show that the slope of $f_{\text{(C)}}$ is always at least as high as the slope of $f_{\text{(B)}}$ (except of course when x is a point joining two line segments, where the slope is not defined, but this only applies to finitely many points).

Let us first note the following useful fact:

Under (B), at each step, any not-yet-selected $c \in T$ is ρ-affordable for some $\rho \leq \sigma_c$. (4.3)

Informally, the fact holds because there are extra voters in (B) compared to (C), and the voters in S have weakly higher utility for c in (B). Formally, looking at the definition (4.1) of affordability in (B), fact (4.3) follows because

$$\sum_{i \in S} \sigma_c \cdot u_i(c) + \sum_{i \in N \setminus S} \min\{b_i, \sigma_c \cdot u_i(c)\} \geq \sum_{i \in S} \sigma_c \cdot u_i(c) \geq \sum_{i \in S} \sigma_c \cdot \alpha_c = \text{cost}(c),$$

where the last step holds because c is σ_c-affordable during the execution of (C).

Now, let $x' \in [0, \alpha]$ be any point that is not a boundary point (for either $f_{\text{(B)}}$ or $f_{\text{(C)}}$). Say that x' lies in the interior of the line segment corresponding to $d \in C$ of $f_{\text{(B)}}$ (call this Segment 1) and in the interior of the line segment corresponding to $c_s \in T$ of $f_{\text{(C)}}$ (call this Segment 2). See Figure 11 for an illustration. Consider the time point t when (B) chose to add d to its outcome (but before it actually added d). At time t, i^*'s utility under (B) was equal to the x-coordinate of the left endpoint of Segment 1, and thus less than x'. Further, at time t, it cannot be the case that (B) has already selected all of the candidates $c_1, \dots, c_s$, because then i^*'s utility would be at least $\alpha_{c_1} + \dots + \alpha_{c_s}$ which is the x-coordinate of the right endpoint of Segment 2 and thus more than x'. Hence there is some c_p, $p \in [s]$, such that (B) has not selected c_p before time t. By fact (4.3), c_p is ρ'-affordable in (B) at time t for some $\rho' \leq \sigma_{c_p}$. Since (B) always selects a candidate that is ρ-affordable for the smallest ρ, it must be the case that d is ρ-affordable for

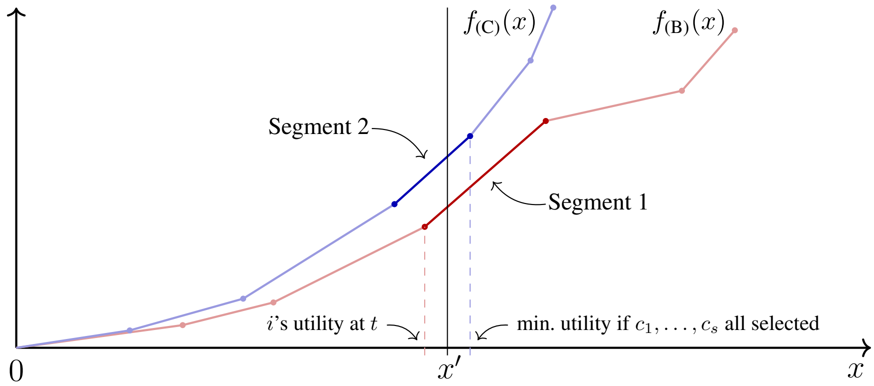


Figure 11: Illustration of the proof of claim (4.2).

some $\rho \leq \rho'$. Thus, the slope of Segment 1 is at most ρ and hence at most σ_{c_p}. On the other hand, Segment 2 has slope σ_{c_s}. Note that $\sigma_{c_p} \leq \sigma_{c_s}$ because $p \leq s$. Thus Segment 1 has slope weakly lower than Segment 2. Since this is true for all x' (not on boundary points), our claim (4.2) follows.

Next, note that

$$f_{(C)}(\alpha) = \sum_{c \in T} \sigma_c \cdot \alpha_c = \sum_{c \in T} \frac{\text{cost}(c)}{|S|} = \frac{\text{cost}(T)}{|S|} \leq \frac{b}{n}.$$

Using (4.2), it follows that

$$f_{(B)}(\alpha) \leq \frac{b}{n}. \tag{4.4}$$

Finally, consider the point in time just after (B) adds c^* to its output. Voter i^* has spent at least b/n at this point. There are two cases:

(i) i^* has spent exactly b/n. In this case c^* is also selected by (A) because the rules behave identically until this point. Now (4.4) implies that $u_{i^*}(W_{(B)} \cup \{c^*\}) \geq \alpha$. Hence $u_{i^*}(W) \geq \alpha$.

(ii) i^* has spent strictly more than b/n. In this case, by definition of (B), we have $c^* \in T$. Now (4.4) implies that $u_{i^*}(W_{(B)} \cup \{c^*\}) > \alpha$. Because $W_{(B)} \subseteq W$, this implies $u_{i^*}(W \cup \{c^*\}) > \alpha$.

In both cases, we conclude that W satisfies EJR up to one project. $\square$

To emphasize the importance of Theorem 4.11 and show how demanding EJR-1 is, let us consider another well-established voting rule, Proportional Approval Voting (PAV).

Definition 4.12 (Proportional Approval Voting (PAV))**.** For approval-based voters' preferences, *Proportional Approval Voting (PAV)* selects a feasible outcome maximizing $\sum_{i \in N} \mathrm{H}(|A_i \cap W|)$, where $\mathrm{H}(r)$ is the rth harmonic number, namely:

$$\mathrm{H}(r) = \sum_{j=1}^{r} \frac{1}{j}.$$

Apart from Equal Shares, it is the only natural voting rule known to satisfy EJR in the committee election model with approval-based preferences. Besides, in this model it is known to provide the highest possible guarantees regarding the average satisfaction of the voters [Skowron, 2021]. However, it completely loses its proportional properties when the costs of the projects are unequal. Example 4.13 shows that, PAV does not even satisfy EJR up to r projects, for any $r \geq 0$.

Example 4.13 (PAV fails EJR)**.** Let $r \in \mathbb{N}$ ($r \geq 2$), $b = r^3$, 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 r^2 dollars each; the candidates $b_1, b_2, \ldots, b_r$ cost 1 dollar 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. ⌟

4.2 Approximation of the Core

Looking how challenging it was to extend the idea of cohesive groups beyond approvals, one could wonder what would happen if we resigned from this requirement at all—and provide the same guarantees as in Definition Attempt 4.6 to every group of voters that is large enough to deserve a certain set of projects. Then we obtain an important notion of proportionality, adapted from cooperative game theory, called the *core* [Aziz et al., 2017a, Fain et al., 2018].

Definition 4.14 (The Core)**.** For an election $E = (N, C, b)$, an outcome W is *core-stable* if for every $S \subseteq N$ and $T \subseteq C$ with

$$\frac{|S|}{n} \geq \frac{\mathrm{cost}(T)}{b}$$

there exists $i \in S$ such that $u_i(W) \geq u_i(T)$. The set of all core-stable outcomes is called *the core* of E.

Core-stability is a stronger guarantee than EJR. It allows any group S to present an arbitrary "counter-proposal" T that they can afford, and guarantees that at least one member $i \in S$ would prefer to stick with the core-stable outcome W, so $u_i(W) \geq u_i(T)$. EJR only guarantees that $u_i(W) \geq \sum_{c \in T} \min_{j \in S} u_j(c)$. Thus, EJR only respects counter-proposals T if they have broad agreement within the coalition S. This is arguably a reasonable restriction, since such coalitions can more easily coordinate to "complain" against the selected W. Still, it is interesting whether we can provide the strong core guarantee.

Unfortunately, there are elections where the core is empty, even with unit costs.

Example 4.15 (Core might be empty). [1] We have 6 voters and 6 candidates with unit costs, and $b = 3$. Utilities satisfy

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

and all other 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 assume that $c_2 \notin W$ and $c_3 \notin W$. Then $S = \{2, 3\}$ and $T = \{c_3\}$ block W, since $1/3 = |S|/n \geq \mathrm{cost}(T)/b = 1/3$ 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, in this example we do not have approval-based utilities. It is unknown whether the core-stability is always satisfiable under this additional assumption (even if the costs are unequal). The further investigation of this problem is presented in Chapter 8.

In this section, let us generalize the result of Peters and Skowron [2020]: Equal Shares provides a multiplicative approximation of the core. It is actually not the best possible result—for general cardinal utilities, a constant factor 9.27-approximation always exists and a constant factor 67.37-approximation is polynomial-time computable (both results are by Munagala et al. [2022])—yet it shows that Equal Shares returns an outcome that never violates core-stability too badly.

Definition 4.16. For an election (N, C, b) and $\alpha \geq 1$, an outcome W is *α-core-stable* if for every $S \subseteq N$ and $T \subseteq C$ with

$$\frac{|S|}{n} \geq \frac{\mathrm{cost}(T)}{b}$$

there exists $i^* \in S$ and $c^* \in T$ such that $u_{i^*}(\mathcal{R}(E) \cup \{c^*\}) \geq u_{i^*}(T)/\alpha$.

The above definition means that the α-core provides guarantees the same guarantees as the core for α times larger groups—for example, in the committee election model, under 2-core each group needs to contain at least $2 \cdot n/b$ voters instead of n/b to have a right to one candidate.

Theorem 4.17. *Given an election E, let $u_{\max}$ be the highest utility a voter can get from a feasible outcome. Let $u_{\min}$ we denote the smallest, yet positive utility a voter can get from a feasible outcome:*

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

Then the outcome selected by Equal Shares is always α-core-stable for $\alpha = 4 \log(2 \cdot u_{\max}/u_{\min})$.

Proof. For notational simplicity, without loss of generality we assume $b = 1$ (by scaling the costs of all projects appropriately). Towards a contradiction, assume there exist an election E, a winning outcome $W \in \mathcal{R}(E)$, a subset of voters $S \subseteq N$, and a subset of candidates $T \subseteq C$

[1] This example is adapted from [Fain et al., 2018, Appendix C] so as to satisfy the unit cost assumption.

with $\sum_{c \in T} \mathrm{cost}(c) \leq {|S|}/{n}$ such that for every $i \in S$ and $c \in T$ it holds that $u_i(W \cup \{c\}) < {u_i(T)}/{\alpha}$.

Now consider a fixed subset $S' \subseteq S$, and let

$$\Delta(S') = \sum_{i \in S'} \big(u_i(T) - u_i(W)\big).$$

Similarly as in the proof of Theorem 4.11, assume for the moment that the voters from S' can spend more than their budgets when paying for candidates in T. Let us analyze how Equal Shares would proceed in this case. By the pigeonhole principle it follows that in each step of the rule, there exists a not-yet-elected candidate $c \in T \setminus W$ such that

$$\frac{u_{S'}(c)}{\mathrm{cost}(c)} \geq \frac{\Delta(S')}{\mathrm{cost}(T)}.$$

Indeed, if for each $c \in T \setminus W$ we had ${u_{S'}(c)}/{\mathrm{cost}(c)} < {\Delta(S')}/{\mathrm{cost}(T)}$, then

$$\Delta(S') \leq \sum_{c \in T \setminus W} u_{S'}(c) < \sum_{c \in T \setminus W} \mathrm{cost}(c) \cdot \frac{\Delta(S')}{\mathrm{cost}(T)} \leq \Delta(S'),$$

a contradiction.

Thus, in each step of the rule there exists a not-yet-elected candidate $c \in T \setminus W$ that is ρ-affordable for some $\rho \leq {\mathrm{cost}(T)}/{\Delta(S')}$. Thus, in each step, Equal Shares selects some candidate $c \in C \setminus W$ that is ρ-affordable for some $\rho \leq {\mathrm{cost}(T)}/{\Delta(S')}$, because Equal Shares always minimizes ρ. Hence, the cost-per-utility that voters pay for the selected candidates is at most ${\mathrm{cost}(T)}/{\Delta(S')}$. Now, consider the first moment when some voter in S', say i, uses more than the voter's initial budget ${1}/{n}$. Until this time moment, Equal Shares behaves exactly in the same way as if the voters from S' had their initial budgets set to ${1}/{n}$. Further, we know that in this moment, if we chose a candidate $c \in T$ that would be chosen if the voters had unrestricted budgets, then the voter i would pay more than ${1}/{n}$ in total, and thus, would get utility more than ${\Delta(S')}/{\mathrm{cost}(T) \cdot n}$. Since we assumed $u_i(W \cup \{c\}) < {u_i(T)}/{\alpha}$, we get that

$$\frac{u_i(T)}{\alpha} > u_i(W) + u_i(c) > \frac{1}{n} \cdot \frac{\Delta(S')}{\mathrm{cost}(T)}.$$

Since $\alpha \geq 2$, and so $u_i(T) - u_i(W) \geq {u_i(T)}/{2}$, we get that

$$u_i(T) - u_i(W) \geq \frac{u_i(T)}{2} > \frac{\alpha \Delta(S')}{2n \cdot \mathrm{cost}(T)}.$$

Let $S" = S' \setminus \{i\}$. Clearly, we have

$$\Delta(S") = \Delta(S') - (u_i(T) - u_i(W)) \leq \Delta(S') \left(1 - \frac{\alpha}{2n \cdot \mathrm{cost}(T)}\right).$$

The above reasoning holds for each $S' \subseteq S$. Thus, we start with $S' = S$ and apply it recursively, in each iteration decreasing the size of S' by 1. After $|S|/2$ iterations we are left with a subset S_e such that

$$\Delta(S_e) \leq \Delta(S) \left(1 - \frac{\alpha}{2n \cdot \mathrm{cost}(T)}\right)^{\frac{|S|}{2}} \leq \Delta(S) \left(1 - \frac{\alpha}{2n \cdot \mathrm{cost}(T)}\right)^{\frac{\mathrm{cost}(T)n}{2}} < \Delta(S) \left(\frac{1}{e}\right)^{\frac{\alpha}{4}}.$$

Now, observe that $\Delta(S_e) \geq |S|/2 \cdot u_{\min}$ (for each $i \in S$ it must hold that $u_i(T) - u_i(W) > 0$) and that $\Delta(S) \leq |S| \cdot u_{\max}$. Thus, we get that

$$\frac{|S|}{2} u_{\min} \cdot e^{\frac{\alpha}{4}} < |S| \cdot u_{\max},$$

which is equivalent to $e^{\alpha/4} < 2 \cdot u_{\max}/u_{\min}$ and, further, to $\alpha < 4\log(2 \cdot u_{\max}/u_{\min})$. This gives a contradiction and completes the proof. □

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 committee-election model with approval-based utilities (there, $u_{\max}/u_{\min} \leq b$). The precise construction is given by Peters and Skowron [2020].

4.3 Priceability of Equal Shares

At the beginning of this chapter, we have noticed (within Observation 4.1) that the basic proportionality of Equal Shares follows just from the fact that it simulates a market satisfying some few basic fairness conditions. The possibility to justify an outcome via such a mechanism can be seen as a yet another proportionality notion [Peters and Skowron, 2020].

Let us first introduce some notation. Fix an election $E = (N, C, b)$. We say that a tuple $p = (\mathrm{end}, C_p, \{p_i\}_{i \in N})$ is a *price system*, where $\mathrm{end} \geq b/n$ is the *initial endowment*, $C_p \subseteq C$ is the *supported outcome*, and $\{p_i\}_{i \in N}$ is the collection of voters' *payment functions* ($p_i \colon C_p \to \mathbb{R}_{\geq 0}$ for each $i \in N$), if the following conditions are satisfied:

1. Each voter pays only for projects she supports:
$$u_i(c) = 0 \implies p_i(c) = 0 \qquad \text{for each } i \in N \text{ and } c \in C_p.$$

2. Each voter has the same initial endowment of end:
$$\sum_{c \in C_p} p_i(c) \leq \mathrm{end} \qquad \text{for each } i \in N.$$

3. The cost of each project from C_p is fully paid:
$$\sum_{i \in N} p_i(c) = \mathrm{cost}(c) \qquad \text{for each } c \in C_p.$$

Now the *priceability* axiom can be formulated as follows:

Definition 4.18 (Priceability)**.** An outcome W is said to be *priceable* if there exists a price system supporting W, that is, $p = (\mathrm{end}, C_p, \{p_i\}_{i \in N})$ where $W = C_p$, satisfying the following condition:

(P) For each project $c \notin C_p$, the unspent budget of her supporters is at most $\mathrm{cost}(c)$:

$$\sum_{i \in N : u_i(c) > 0} \left(\mathrm{end} - \sum_{c' \in C_p} p_i(c') \right) \leq \mathrm{cost}(c) \quad \text{for each } c \notin C_p.$$

Intuitively, condition (P) ensures that the amount of unspent money is small enough so that no unsupported candidate can be additionally purchased.[2] We separate here this condition from the 3 aforementioned ones (in contrary to Peters and Skowron [2020]), since in Chapter 9 we will consider price systems not satisfying this conditon.

Hence, the priceability axiom requires that each outcome is possible to be justified by a fair market. Priceability does not place any restrictions on how the rule splits the project's cost among supporters. The concept also allows initial voters' endowments higher than $^b/_n$; an outcome is priceable if there exists *some* budget limit for which it is priceable.

It is known that Equal Shares is priceable in the committee election model with approval-based utilities [Peters and Skowron, 2020], and in the general PB model this property is also preserved—indeed, the rule explicitly constructs a price system satisfying the above conditions.

One could wonder why this axiom is important—for example, what are the advantages of priceability over EJR or core-stability. Let us present then an example showing that in some cases priceability rules out some clearly unfair outcomes that do not violate the core (and hence, EJR).

Example 4.19. Fix an integer b as the budget constraint. Consider the following committee election with $n = b^2$ voters divided into $b + 1$ groups G_0, G_1, …, G_b. Group G_0 consists of b voters approving some b candidates. The remaining groups contain $b - 1$ voters each. Each of these groups approve one unique candidate. This election together with two sample outcomes (a green one and a blue one) are depicted in Figure 12.

For this election, the subset $W_1 = \{c_1, \ldots, c_b\}$ (the candidates marked green) is core-stable (and hence, also satisfies EJR). In fact W_1 would be uniquely selected by the following natural rule: "among all committees satisfying EJR, select the one which garnered most approvals".

This outcome gives zero satisfaction to a majority of the voters. Naturally, it is not possible to give a nonzero utility to all the groups of voters (since there are $b + 1$ of them), yet one can argue that, for example, the blue committee $W_2 = \{c_1, c_{b+1}, \ldots, c_{2b-1}\}$, is a much fairer choice. This committee is priceable, as witnessed by any price system supporting W_2 with $\mathrm{end} = {}^1/_{b-1}$.

[2]One could argue that then the inequality in the condition should be strict. The role of weak inequality here is to allow tie-breaking—otherwise, priceability would be unsatisfiable for simple symmetrical elections (for example $b = 1$, $C = \{c_1, c_2\}$, 50% of voters approving only c_1 and the other 50% approving only c_2).

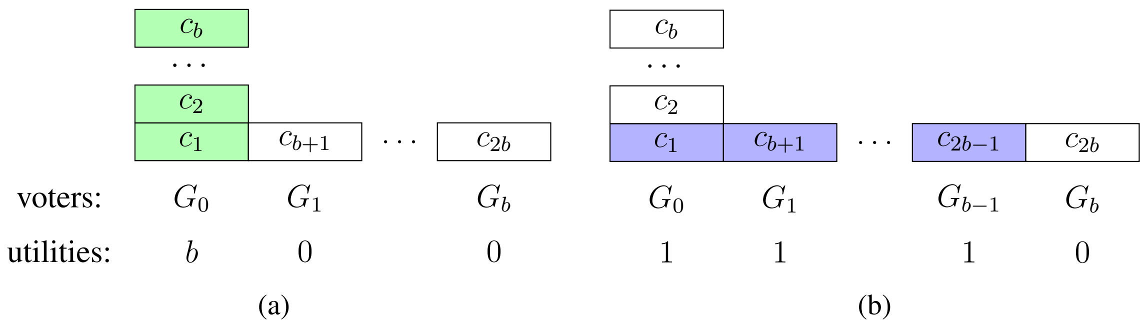


Figure 12: An illustration of Example 4.19. To interpret it, recall Figure 2 and Figure 3—here again we assume that candidates correspond to the boxes and the voters approve the candidates placed above them.

Then condition (P) is satisfied (note that group G_b has exactly 1 unspent dollar in total, yet to violate (P) they would need to have strictly more than 1 dollar). On the other hand, to purchase any committee containing at least two candidates approved by G_0 (for example, W_1), this group should have at least 2 dollars in total. Then the initial endowment should be equal to at least $2/b$ and every neglected group of voters would have enough unspent money to violate (P). ⌟

Intuitively, the problem with core-stability here is that the guarantees provided by this axiom are based on the fixed quota n/b—hence, groups containing $n/b - 1$ voters are not entitled to anything. Priceability is much more flexible here, since instead of a fixed quota it uses the variable end.

In the approval-based committee election model there always exist outcomes that are both priceable and exhaustive (recall Definition 3.2) [Peters and Skowron, 2020]. In the general PB model, these two properties are mutually exclusive—it can be seen in Example 3.4. Therefore, in particular, no exhaustive variant of Equal Shares presented in Section 3.3 is priceable. However, the "completion by varying the budget" is priceable unless it is further completed by another rule—it might be therefore seen as a good compromise between the priceability condition and the desire to spend a larger part of the budget than the basic variant of Equal Shares.

To conclude this chapter, we can see that Method of Equal Shares has very strong fairness properties. It provides very strong proportionality guarantees to groups of voters with cohesive preferences (EJR-1), approximates the strongest known notions of proportionality (the core) and balances the influence of the voters on the elected outcome (priceability). In our opinion, all these factors combined establish this rule as a prime candidate for voting in the model with additive utilities and arbitrary costs. In the next chapter we will analyze whether these good theoretical properties translate into good behavior in practice.

Chapter 5

Evaluation of Equal Shares on the Real-Life Data

In this chapter we take a step towards understanding how various voting rules for participatory budgeting operate in practice. We do this by collecting and analyzing data from over 800 PB elections. The data collected by us has been published as the open library, called *Pabulib*. In Appendix A we define the *.pb* format, which we recommend for representing PB elections, and which is used in our library.

The aim of *Pabulib* is to gather participatory budgeting data from as many cities and as many countries as possible, but currently most of the elections come from several large cities in Poland[1] (in particular, from Warsaw, with a population of 1.7 million people; from Kraków, Wrocław, and Gdańsk, with populations between 500 000 and 1 million; and from Częstochowa, Zabrze, and Katowice, with populations between 150 000 and 300 000).

Having this data, we have performed a comparative analysis of the Utilitarian Greedy rule (recall Definition 3.3), which is currently used for all the elections in *Pabulib* to the three exhaustive variants of Method of Equal Shares presented in Section 3.3. Specifically, we considered:

- The completion by Utilitarian Greedy (for brevity, we will call it further ES-U),
- The completion by increasing gradually voters' endowments (by 1% in each step), additionally completed by Utilitarian Greedy if needed (further called ES-Inc),
- The completion by the perturbation of utilities (further called ES-Eps), yet implemented without their actual perturbation as described in Section 3.3.

Additionally, we assumed that each of these four voting rules can be run in four variants. First of all, recall the discussion over the relation between ballots and voters' preferences, presented in Section 2.3. All the cities that we analyze (even those using ranking ballots) map the ballots to cardinal utilities. Our analysis includes the comparison whether it is better for a rule to

[1]Poland is a good source of PB elections because the law requires every major city to spend at least 0.5% of its annual budget through PB. In 2022, over 43% of Polish cities with populations above 5,000 organized PB elections, spending 630.5 million PLN in total [Martela et al., 2023].

use the direct or the costwise score-utility mapping. For simplicity, we will refer to a rule $\mathcal{R}$ using direct/costwise score-utility mapping as a *direct-mapping/costwise-mapping variant of* $\mathcal{R}$, respectively.

Next, there are two different ways of applying each rule to a specific election in a city.

- The *districtwise* (D) variant which corresponds to how the cities currently organize their elections: a separate election is run in each city district; typically there is also one additional election involving the same set of voters but concerning citywide projects (these are projects that are potentially interesting to voters from multiple districts). The outcome for a given city and year is obtained by adding together the outcomes of these smaller elections.

- The *citywide* (C) variant, a natural alternative to the current solution. Here we put all the projects from different districts and categories in the same pool, and we kept the original voters' ballots. Thus, for a fixed city and year we have a single election.

The combination of direct-mapping/costwise-mapping and districtwise/citywide variants yields four possible variants of each rule. All the studied cities currently use the costwise-mapping districtwise variant of Utilitarian Greedy.

Our results lead to several observations that should be taken into account by election designers.

1. We see that costwise-mapping variants of the rules select fewer but larger projects compared to their direct-mapping counterparts.

2. The proportional methods such as the all variants of Equal Shares result in much fairer solutions, where the different voters' opinions are better represented; the difference is substantial.

3. Running citywide elections instead of districtwise ones result in a significant improvement in terms of the average utility as well as the number of voters whose opinions are taken into account.

4. Among the considered completions of Equal Shares, ES-Inc seems to give the best results with respect to fairness and comparable results to ES-UG with respect to efficiency, while ES-Eps is the worst with respect to both criteria.

5. The additional restrictions placed for ballots (like the ones mentioned in Section 2.2) affect the process of preparing project proposals. For example, restricting approval ballots to voting for only one project discourage submitting small and medium projects.

6. Equal Shares is more robust to changing the ballot types than the Utilitarian Greedy rule.

To conclude the above observations, we suggest that the cities should use the costwise and citywide variant of the ES-Inc rule. The question which type of ballots is the best is typically seen as a trade-off between the simplicity and the expressiveness of a voting process (approval ballots are the simplest yet the least expressive; score ballots are the most expressive yet the most demanding; ranking ballots are in the middle with respect to both categories). However, the recent lab experiments by Benade et al. [2023] suggest that the majority of people do not see the additional expressiveness provided by score ballots as an advantage; what is more, they often subjectively claim that approval ballots are even *more* (!) expressive for them. Nevertheless, our recommended rule is very robust to changing the type of ballots (on average, under score and approval ballots approximately 80% of an outcome remains unchanged), which suggests that the choice of a ballot type is a less important problem than the choice of a voting rule. However, based on the fifth aforementioned observation, we discourage the cities from adding too restrictive additional conditions on the ballots.

5.1 Basic Metrics of Fairness and Efficiency

In this section we describe several basic metrics for comparing outcomes returned by different voting rules and present the results of our analysis. We want to measure both efficiency (the total social welfare generated by a rule) and fairness (the level of egalitarian spread of the social welfare among voters) of voting rules.

Average utility. It is probably the most natural metric of efficiency. We define it for each outcome W as ${}^{1}/{}_{n} \sum_{i \in N} u_i(W)$.

Dominance margin. It is a metric for comparative analysis of two rules, R_1 and R_2, defined as the fraction of voters who enjoy a strictly higher utility from the outcome of R_1 than from the one of R_2. A related metric is the **improvement margin**, defined for given rules R_1 and R_2 as the dominance margin of R_1 over R_2 minus the dominance margin of R_2 over R_1. Note that such a way of comparing outcomes can be viewed as a metric balancing fairness and efficiency (a rule may be better according to the dominance/improvement margin either because it generates more social welfare, or because it spreads it more equally).

Exclusion ratio. This metric is defined as the fraction of voters who support none of the selected projects (intuitively, they were excluded by the voting process).

Power inequality. This fairness metric is based on the concept recently introduced by Lackner et al. [2021] which, informally speaking, measures the amount of spending that different voters had an influence on [Maly et al., 2023]. The measure assumes that the supporters of a selected project contributed to the decision on spending money on this project proportionally to their utilities. Consequently, given an outcome W, voter i's share is:

$$\text{share}_i(W) = \sum_{p \in W} \frac{u_i(p)}{\sum_{j \in N} u_j(p)} \cdot \text{cost}(p).$$

Note that the shares of the voters sum up to the total cost of the selected projects. In an ideally fair solution we would like all these shares to be equal (that is, equal to b/n). Hence, power inequality is defined as a normalized ℓ_1-distance between these two distribution:

$$\frac{1}{n} \cdot \sum_{i \in N} \frac{|\mathrm{share}_i(W) - b/n|}{b/n}.$$

We used the aforementioned metrics to compare the outcomes of various election rules on the data from *Pabulib*. In our plots we show only cities for which we have data for at least 3 years (Warsaw, Kraków, and Wrocław). The analysis of the remaining cities is publicly available through our web application, *Pabustats* (`http://pabulib.org/pabustats`) and it leads to similar conclusions.

5.2 The Comparison of Different Completions of Equal Shares

We present results for each city averaged over all years, with figures showing error bars corresponding to standard deviations over the years. While we consider averages, the conclusions of our analysis also hold for each year separately. The results for separate editions can be checked though *Pabustats*.

In our first set of simulations, we compare the three aforementioned variants of making Method of Equal Shares exhaustive. We observe that the completion strategy plays a critical role. For example, in citywide and districtwise elections, Equal Shares without completion uses, on average, only 32% and 50% of the available funds, respectively. The ES-Inc rule after the first step (increasing the budget) uses on average 98% and 88%, and after the second one (completion by UG)—99.9% and 95% of the funds, respectively.

In Figures 13 and 14, we compare our metrics for the three completion strategies, within all the four variants. These figures are quite involved, so let us provide some guidance on how to read it. First, we have a separate plot for each of the metrics, that is, for (1) average utility, (2) the improvement margin over ES-Inc, (3) power inequality and (4) exclusion ratio. Within each of the plots, different shades of each color correspond to different completion types (darkest for ES-Inc, middle for ES-U, and lightest for ES-Eps). The green and blue shades correspond to the districtwise elections, and red and yellow shades to citywide elections. Costwise-mapping variants correspond to green and red, whereas direct-mapping variants correspond to blue and yellow. We conclude the following:

1. The results for exclusion ratio are the best for elections in Kraków and the worst for elections in Wrocław. It follows from the fact that Kraków uses ranking ballots truncated to top 3 projects, by which it enforces voters to assign a nonzero score to at least 3 projects, while Wrocław, on the other hand, enforces all the voters to approve only a single project.
2. ES-Eps gives the worst results in terms of the average utility, power inequality, exclusion ratio, and improvement over ES-Inc.

3. ES-U (middle shades) gives a bit higher average utility than ES-Inc (darkest shades), but also a worse power inequality and exclusion ratio.
4. Among the direct-mapping variants (blue and yellow), the ES-U (middle shade) and ES-Inc (darkest shade) are comparably good. Among costwise-mapping variants (green and red), ES-Inc (darkest shade) has a large advantage, especially for improvement margins (the other rules have negative values in almost all settings).

Based on these observations, we advise using primarily the ES-Inc completion of Equal Shares.

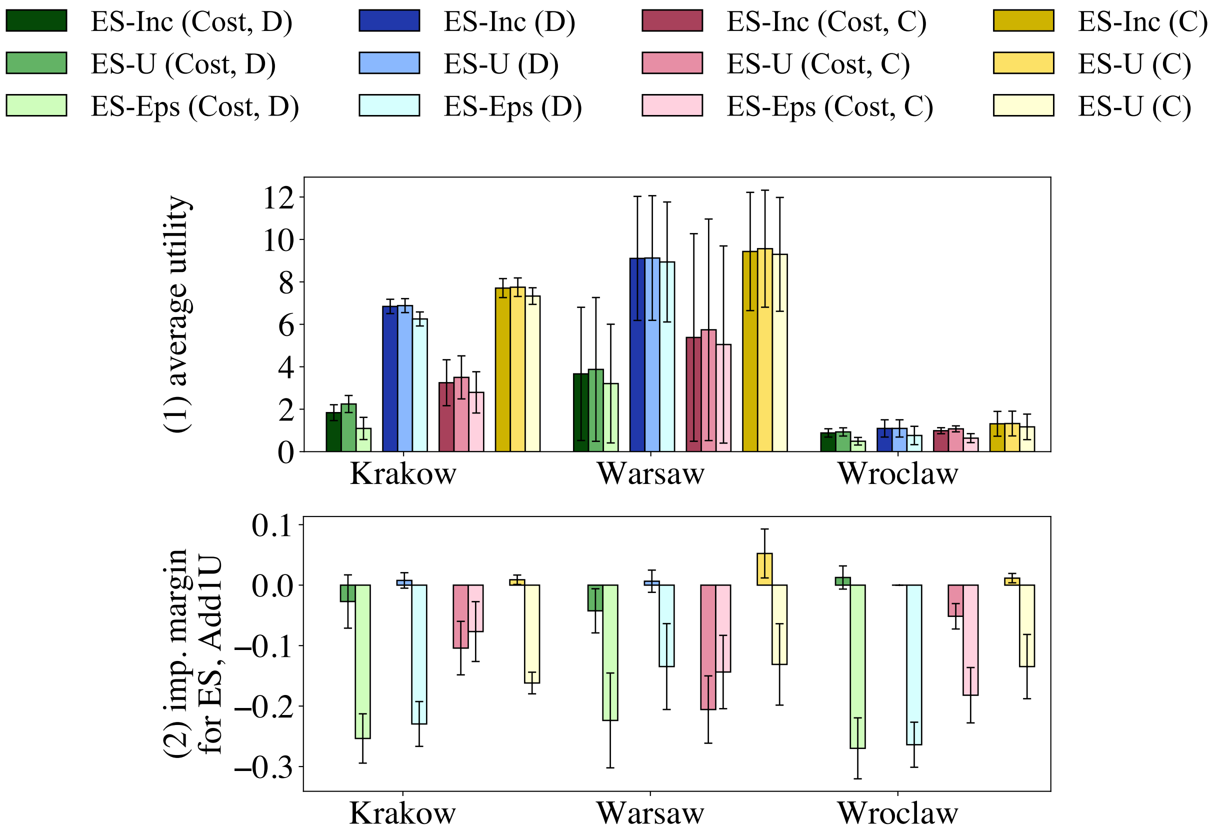


Figure 13: Comparison of different completion variants for Equal Shares with respect to the (1) average utility (for costwise-mapping variants it is presented in millions) and (2) the improvement margin for all the rules compared to ES-Inc. The label "Cost" means that we are referring to the costwise-mapping variant of the method; otherwise we are referring to its direct-mapping variant. The symbols "D" and "C" stand for the districtwise and citywide variants, respectively.

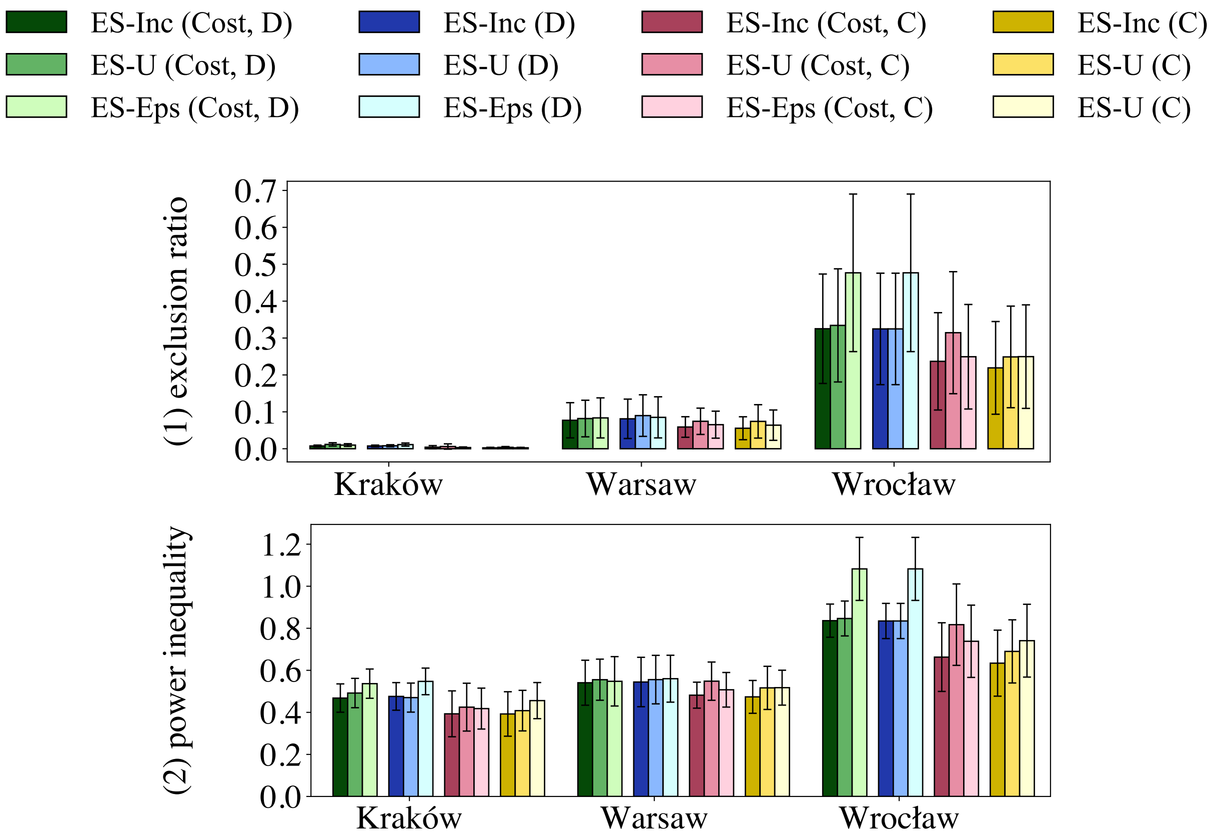


Figure 14: Comparison of different completion variants for Equal Shares with respect to (1) exclusion ratio, (2) power inequality. The label "Cost" means that we are referring to the costwise-mapping variant of the method; otherwise we are referring to its direct-mapping variant. The symbols "D" and "C" stand for the districtwise and citywide variants, respectively.

5.3 The Comparison of Equal Shares and Utilitarian Greedy

From now (until the end of the chapter) let us fix ES-Inc as the default variant of Method of Equal Shares. In this section we compare it to Utilitarian Greedy (especially to its costwise-mapping districtwise variant).

5.3.1 Basic Metrics

Let us start by comparing the ES-Inc with Utilitarian Greedy according to the metrics presented in Section 5.1. The results are depicted in Figure 15. In these figures, each scenario corresponds to a color. The darker shades represent Equal Shares and the lighter ones represent Utilitarian Greedy. Our findings can be summarized as follows:

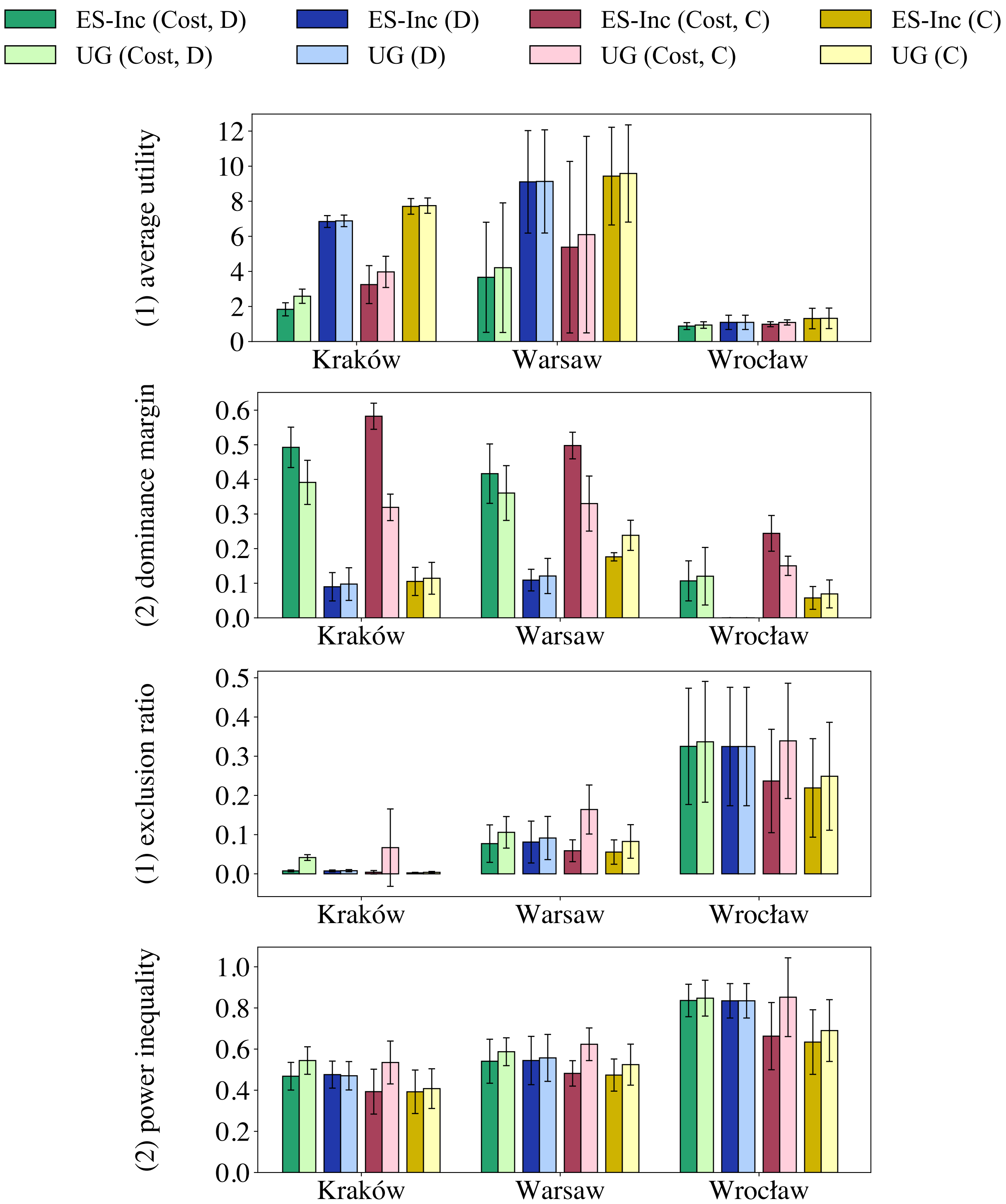


Figure 15: Comparison of ES-Inc and Utilitarian Greedy (UG). The presentation is analogous to Figure 13 and Figure 14.

1. Among direct-mapping variants of the rules (blue and yellow), the results of Method of Equal Shares and of Utilitarian Greedy are comparable. Equal Shares selects outcomes with slightly lower exclusion ratio as well as lower power inequality, at the cost of lower average utility. However, these differences are relatively small.

2. For costwise-mapping variants (green and red), we see a significant difference between the two rules. Unsurprisingly, the outcomes of Utilitarian Greedy have better average utility, but the difference is relatively small (for example, it is the largest in Warsaw for citywide elections, 13%). For all other metrics, Equal Shares outperforms Utilitarian Greedy by a large margin. For example, for citywide elections in Warsaw, the average score utility of the outcomes of Equal Shares is 43% higher, and using Equal Shares would result in a drop of the exclusion ratio from 16% to 6% . The improvement margin over UG is on average, respectively, 59% and 17% .

3. We observe a significant difference between the citywide (red and yellow) and districtwise (blue and green) variants of the rules. Citywide elections result in a much higher average utility and much lower exclusion ratio, for Equal Shares. The large difference between districtwise and citywide elections arises because in some districts no popular projects are submitted, and their residents would prefer to fund citywide projects instead (recall the example of the two projects on Modlińska Street in Warsaw 2021, from Chapter 1).

Finally, since our findings suggest to replace the current districtwise elections with citywide ones, it is interesting to see how fairly the budgets were distributed among districts when using the citywide variant. Let $\mathcal{D} = \{D_1, D_2, \ldots, D_t\}$ be the set of districts, which is formally a partition of N. Ideally, the voters from a district $D \in \mathcal{D}$ should get a share of the budget that is proportional to the size of D.[2] Now we introduce the fairness metric, capturing this intuition:

Dispersion of the budget allocation This metric captures the average relative difference between how much money the district got and how much we would expect it to get (defined as the ℓ_1 distance, similarly as in the power inequality metric). Formally:

$$\frac{1}{|\mathcal{D}|} \cdot \sum_{D \in \mathcal{D}} \frac{\left| \sum_{i \in D} \mathrm{share}_i(W) - {}^{|D|}\!/_{n} \cdot b \right|}{{}^{|D|}\!/_{n} \cdot b}.$$

Table 5.1 shows average dispersion values, which are lower for Equal Shares than for Utilitarian Greedy.

[2]This assumes that the budget should be divided proportionally to the number of voters and not to the number of residents of a district. If turnout varies between districts, the difference matters. Being proportional to the number of voters promotes participation, incentives districts to encourage their residents to vote, and follows the "one person, one vote" principle. If the city prefers being proportional to the number of residents, the citywide variant of Equal Shares can be adapted by giving voters from districts with lower turnout a larger initial endowment.

City	ES-Inc, C	Util. G, D	Util. G, C
Częstochowa	0.23	0.28	0.39
Gdańsk	0.27	0.33	0.46
Katowice	0.19	0.26	0.51
Kraków	0.08	0.24	0.23
Warsaw	0.20	0.41	0.41
Wrocław	0.15	0.26	0.22
Zabrze	0.38	1.24	0.41

Table 5.1: Average dispersion of the budget allocation. We compare costwise-mapping variants of ES-Inc and Utilitarian Greedy (both the districtwise and the citywide variants).

5.3.2 Budget Distribution among Categories

Cities often organize projects by topics (such as public space, environment, education) to make browsing the list of projects easier. Warsaw, for example, categorizes projects using 10 different tags (where projects can get multiple tags). This allows us to ask whether voter preferences across categories are well-reflected by the spending of the rules.

We focus on Warsaw district elections (2020–23), which use approval voting. Denote by A_i the set of projects approved by $i \in N$. For each project p, denote by $\mathrm{tags}(p)$ the subset of tags assigned to p. For each tag t, we can then compute its *vote share*:

$$\frac{1}{n}\sum_{i\in N}\sum_{p\in A_i : t\in \mathrm{tags}(p)} \frac{1}{|A_i|\cdot|\mathrm{tags}(p)|}$$

This intuitively counts the fraction of the votes that went to projects with tag t, in a way that each voter contributes the same amount to the vote share, and for projects with multiple tags, splitting their contribution equally between them. Note that the vote shares of all the tags sum to 1. For an outcome W, we can similarly define the spending share of the tag:

$$\frac{1}{\mathrm{cost}(W)}\sum_{p\in W : t\in \mathrm{tags}(p)} \frac{\mathrm{cost}(p)}{|\mathrm{tags}(p)|}.$$

We can now compute the ℓ_2 distance between the vector of vote shares and the spending shares of all the tags, to see how well they align. While it is not necessarily desirable for the two vectors to perfectly coincide, a large distance indicates that an outcome neglects certain categories.

We find that for 93% of districts, Utilitarian Greedy gives outcomes with a larger distance to the vote shares than the (costwise) ES-Inc outcome, see Figure 16. In some cases, the distance is much larger, like in the district Bielany, where in each year, Utilitarian Greedy spends much less on education projects than suggested by the vote shares. For example, in 2020, when education had a vote share of 33% , Utilitarian Greedy spent only 5% of the budget on these projects (Equal Shares spent 26%), see Figure 17.

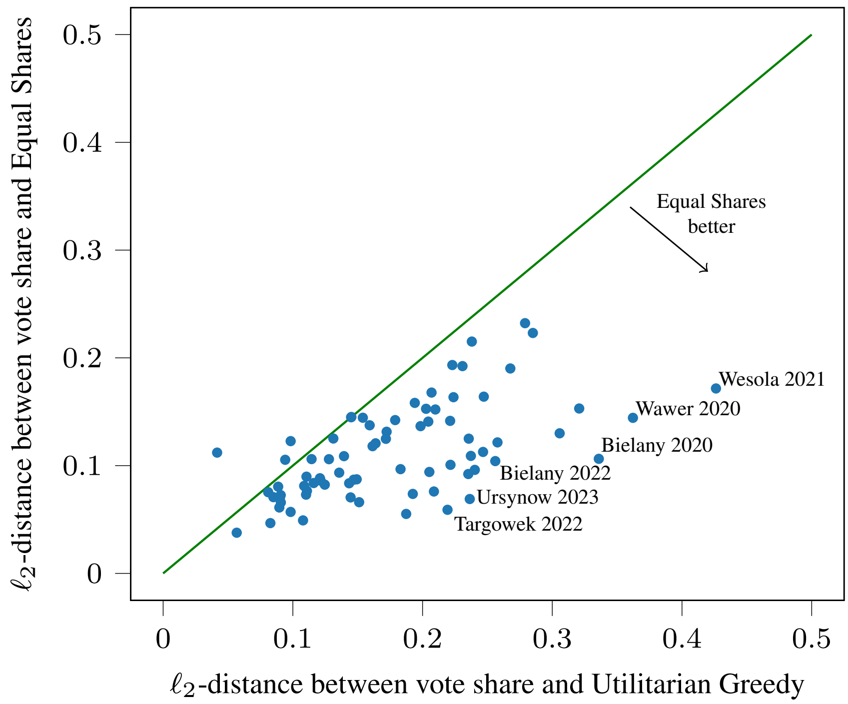


Figure 16: Comparison of costwise-mapping variants of ES-Inc and Utilitarian Greedy with respect to how well the voters' preferences across categories are reflected by the spending of the rules. Each Warsaw district in each year 2020–23 is represented by a blue point, placed according to the ℓ_2-distance between the vote share vector and the spending shares for the two rules. For points below the green line, Utilitarian Greedy has a higher ℓ_2-distance. Points with a particularly large imbalance are labelled.

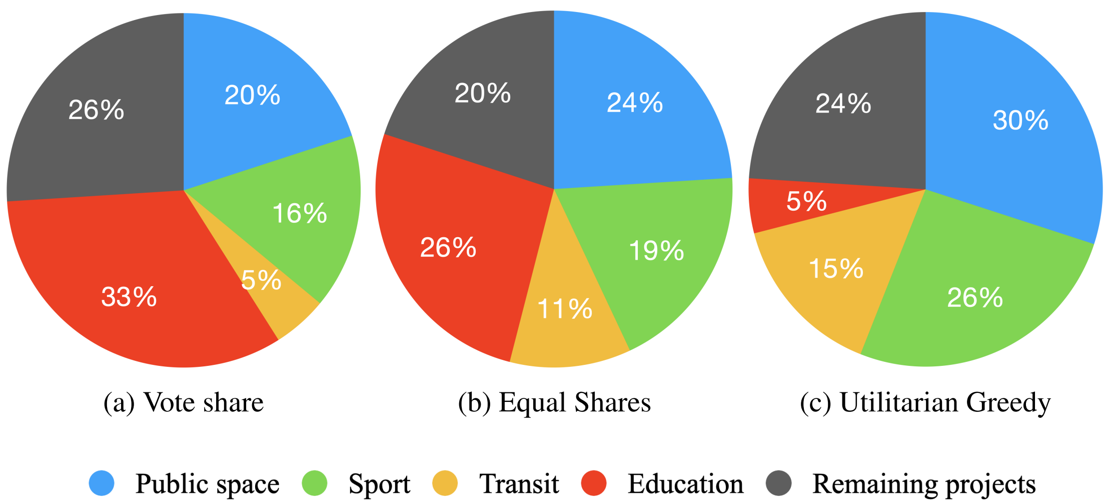


Figure 17: The vote share and the spending share of different tags in district Bielany, Warsaw 2020. The picture would be similar for 2021–23.

5.3.3 Maps of Participatory Budgeting Elections

We provide easy-to-interpret visualisations of the outcomes of different voting rules. For elections that were carried out in Warsaw between 2020 and 2023, most of the projects (but not all) were associated with their GPS locations. Thus, we can depict those submitted projects that have GPS data in such a way that their relative locations correspond to their physical locations in the respective districts. We present such visualisations for Warsaw 2023.[3] in Figure 18 The legend of the maps is the following: each project is represented by two glued-together half-discs. The size of the left half is proportional to the project's cost, whereas the size of the right half is proportional to the total number of votes the project received. The figures compare the outcomes of the costwise-mapping variant of ES-Inc, with the outcomes of the Utilitarian Greedy rule. Specifically, gray projects were not selected by either of the rules, green projects were selected by both, blue projects were selected only by Equal Shares, and red projects were selected only by Utilitarian Greedy.

A different approach is to create a map that illustrates voters' preferences rather than geographic locations of the projects. Here, for a given approval PB election we first compute the Jaccard distances between all pairs of projects. Recall that for two projects, p_1 and p_2, their Jaccard distance is

$$\frac{|N(p_1) \triangle N(p_2)|}{|N(p_1) \cup N(p_2)|},$$

where $N(p)$ denotes the set of voters who support project p (in other words, we assume that two projects are similar if similar groups of voters voted for them). Next, based on these distances, we create a two-dimensional embedding, using the Multidimensional Scaling Algorithm MDS [Kruskal, 1964, de Leeuw, 2005] (all the distances lie between 0 and 1, but most of them are relatively high, with very few being below 0.5—thus, we normalize the distances by subtracting 0.5, that is, $d' = max(0, d - 0.5)$). . Intuitively, we obtain a plot where the closer two projects are, the larger is the fraction of their common supporters (however, note that MDS is only a heuristic and, more importantly, a perfect embedding may not exist). This type of maps for Warsaw 2023 (see Footnote 3) is depicted in Figure 19.

We observe that Equal Shares selects more diverse and more representative sets of projects both in terms of their geographic locations and in terms of their supporters. We further observe that Utilitarian Greedy mainly selects large and expensive projects, whereas Equal Shares selects a mixture of projects, including some large and some small ones.

[3] The results for other Warsaw elections are consistent and are publicly available in [Faliszewski et al., 2023].

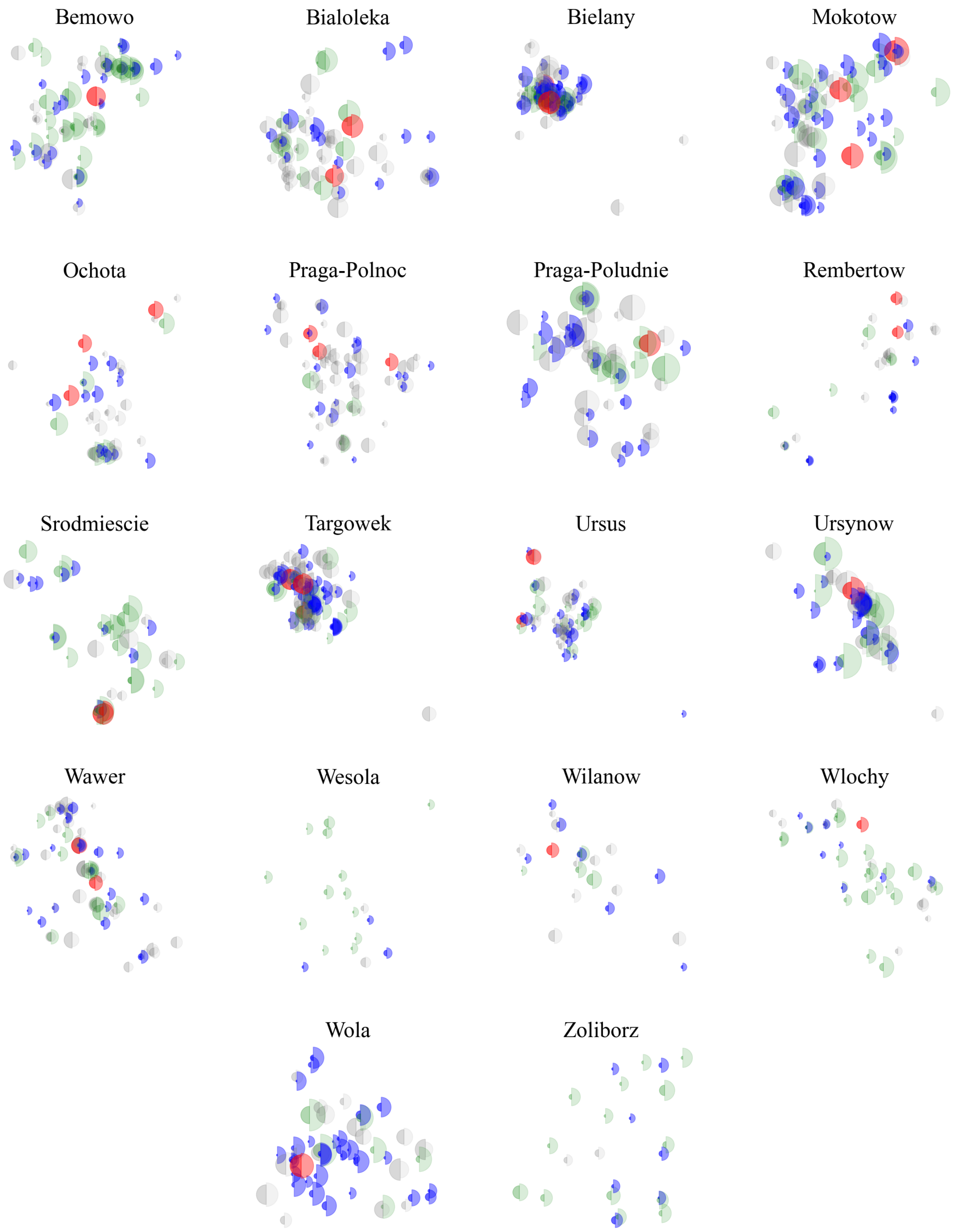


Figure 18: Visualization of projects in PB elections from Warsaw 2023 using the GPS data.

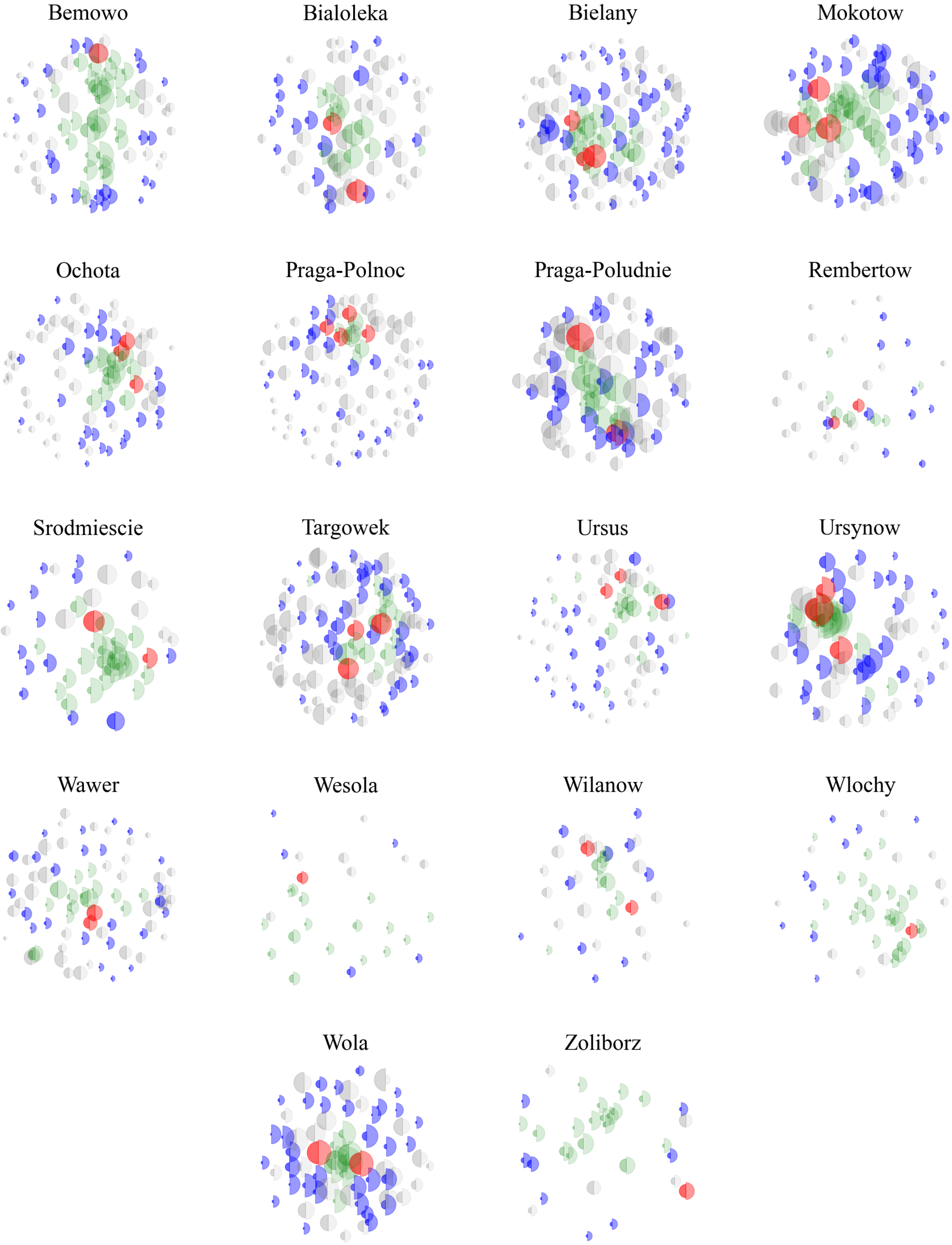


Figure 19: Visualization of projects in PB elections from Warsaw 2023 using the Jaccard distance.

5.4 Discussion over the Types of Ballots

The results from the previous section support our recommendation for the cities that Equal Shares (specifically, ES-Inc) in its costwise-mapping citywide variant should be used instead of the currently used costwise-mapping districtwise Utilitarian Greedy. Let us now complete our analysis with a short discussion over the types of ballots. One could wonder which one should be used in practice.

Answering this question requires performing lab experiments. Recently, such research has been done by Benade et al. [2023]. The participants of their experiments found approval ballots superior (simpler and more expressive) to all the other types. At the same time, one of their findings was that Method of Equal Shares is robust to changing the type of ballots (checked on synthetic data). In this section we reinforce this conclusion by analyzing data from real PB elections.

For each election where voters used cardinal ballots, we construct a corresponding approval election by letting a voter approve all the projects to which she assigned a positive score. Then, we compare the outcomes of different rules for these two elections. Let W and W_{appr} be the outcomes of a given voting rule for the original and the approval elections, respectively. We define the *robustness ratio* as $\text{cost}(W_{\text{appr}} \cap W)/\text{cost}(W)$. Table 5.2 summarizes the results of our analysis. We can see that the outcomes of Equal Shares change much less after switching to approval compared to Utilitarian Greedy. For users of Equal Shares, this provides an argument in favor of approval ballots.

City	ES-Inc, C	Util. G, D	Util. G, C
Częstochowa	0.80	0.35	0.39
Gdańsk	0.87	0.26	0.39
Katowice	0.83	0.56	0.42
Kraków	0.78	0.52	0.41

Table 5.2: Robustness ratio for different voting rules.

Recall now from Section 2.2 that the approval ballots are typically modified so that the maximal possible number of approved projects is limited. It is then interesting to check if we can see from the data analysis how it affects the voting process. We believe that approval ballots modified in a way it is done in Wrocław or Zabrze (that is, in a way that a voter can approve only one project) lead to discourage submitting smaller and cheaper projects, in contrast to less restricted approval ballots (for example, to Warsaw ballots, where the voters can approve up to 15 districtwise projects and up to 10 citywide ones). In Figures 20 and 21, we illustrate the distribution of project costs and the total numbers of votes per project for the cities collected in Pabulib.

To sum up, our findings show that if Equal Shares is used, the problem of the choice of the ballot type becomes less critical. Based on the experiments of Benade et al. [2023] we recommend to use rather approval ballots, yet we advise not to restrict too much the maximal number of projects one can approve.

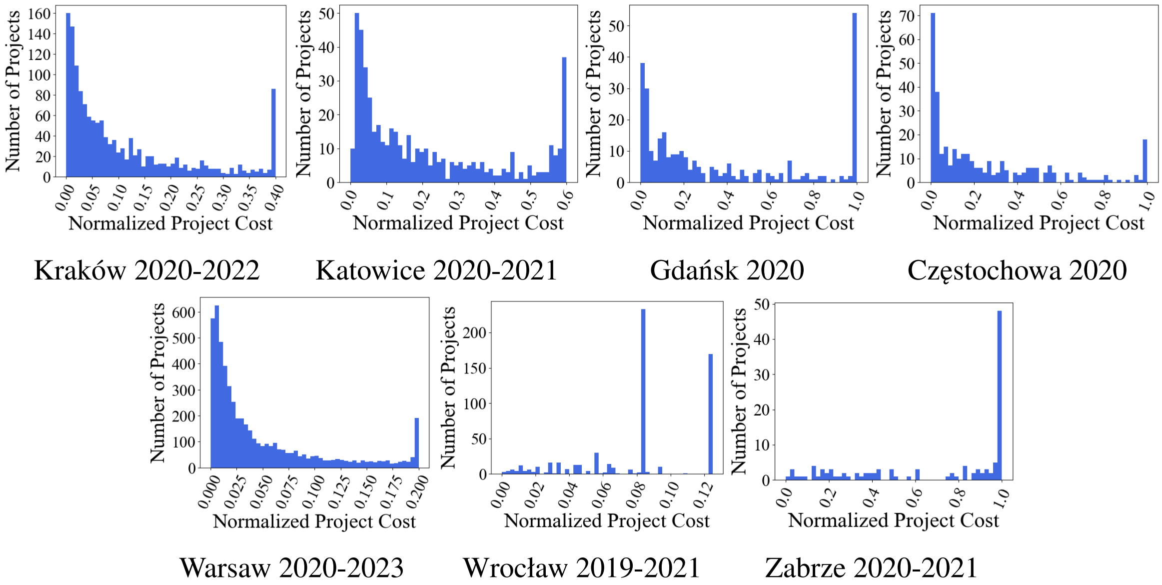


Figure 20: The distribution of costs per project. The histogram for each city contain aggregated data from different years. For each PB election we normalize the costs by dividing them by the total budget.

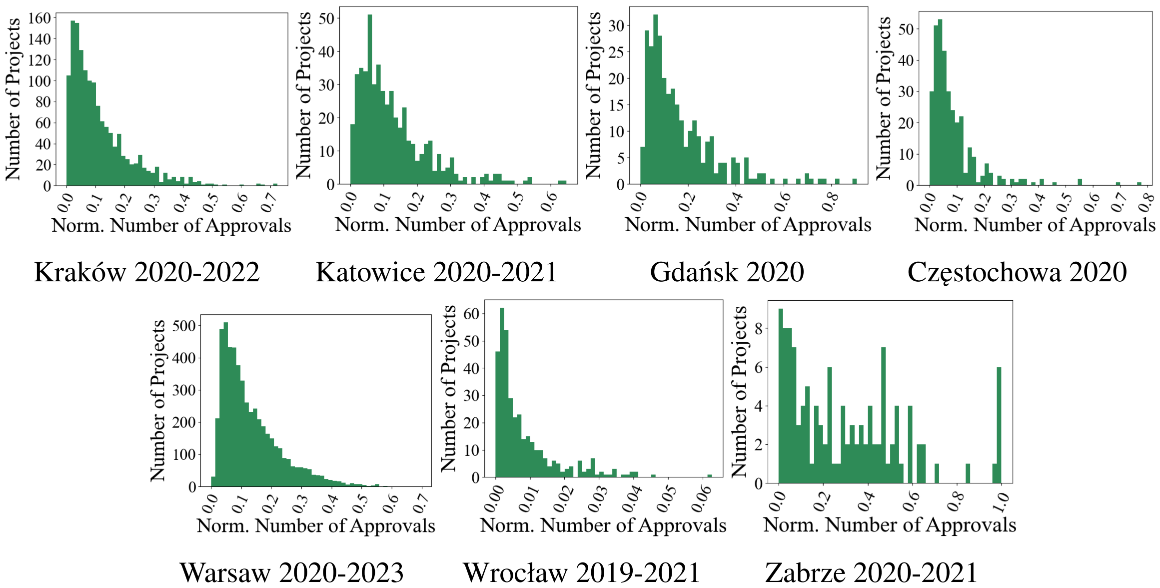


Figure 21: The distribution of the total number of approvals per project. The histograms for each city contain aggregated data from different years. For each PB election we normalize the number of approvals by dividing it by the number of voters in a given election.

Chapter 6

Method of Equal Shares for Ordinal Preferences

In this chapter we discuss how Method of Equal Shares can be adapted for the setting where voters have *ordinal preferences*, that is, voters express their preferences by ranking the projects. Recall from Chapter 2 that in this case we assume that the voters' preferences over outcomes are lexicographical with respect to their rankings—-which intuitively means that the voters care *infinitely more* about projects ranked higher than the ones ranked lower. We will show that in this model Equal Shares satisfies an axiom called Inclusion Proportionality for Solid Coalitions (IPSC) introduced by Aziz and Lee [2021], which is a recent extension of a well-established Proportionality for Solid Coalitions (PSC) axiom for the committee election setting, introduced by Dummett [1984]. This result is interesting, because it shows that Equal Shares is a very flexible method—it can be reasonably defined for all the most important types of preferences studied in the literature and satisfies strong notions of proportionality in a number of various settings.

To adapt Equal Shares to ordinal preferences, we need to somehow cast them to cardinal ones. We follow here two principles:

Assumption 1 If two voters $i, j \in N$ have at the kth position in their rankings ($k \leq m$) projects c_1 and c_2 respectively, the utility of i from c_1 is the same as the utility of j from c_2.[1]

Assumption 2 If a voter i prefers project a to project b, she gains *infinitely more* utility from a than from b. This is the general intuition behind the lexicographical comparison of outcomes.

By combining these two assumptions, we obtain the following observation:

Observation 6.1. If two voters $i, j \in N$ have at the kth and k'th position in their rankings ($k < k' \leq m$) projects c_1 and c_2 respectively, the utility of i from c_1 is *infinitely greater* than the utility of j from c_2.

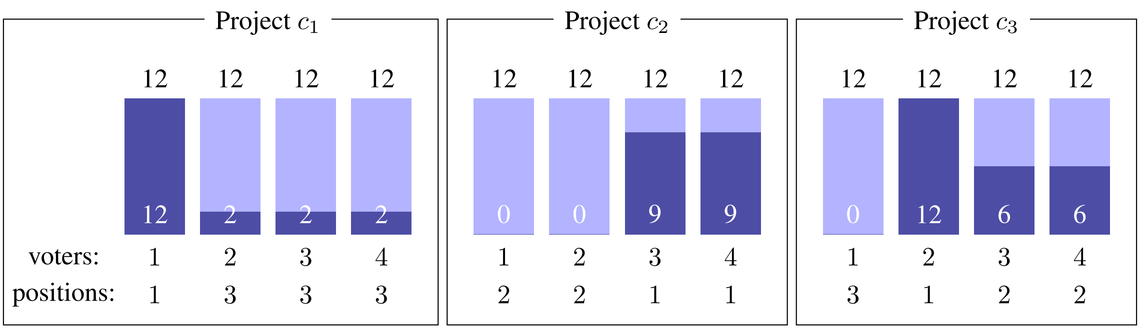
[1] An alternative reasonable assumption would be to assume that the relative comparison of utilities of i and j is proportional to the costs of c_1 and c_2. We consider this alternative in Section 6.2.

The phrase "infinitely greater" may seem vague, yet it is enough to understand it on an intuitive level.

These two assumptions are sufficient to determine which project should be elected by Equal Shares in a specific round and how the payments should be distributed—even though it does not determine the exact value of its "payment-per-utility". For example, let us consider the following voters' preferences:

$$\begin{aligned} 1:&\quad c_1 \succ c_2 \succ c_3 \\ 2:&\quad c_3 \succ c_2 \succ c_1 \\ 3:&\quad c_2 \succ c_3 \succ c_1 \\ 4:&\quad c_2 \succ c_3 \succ c_1 \end{aligned}$$

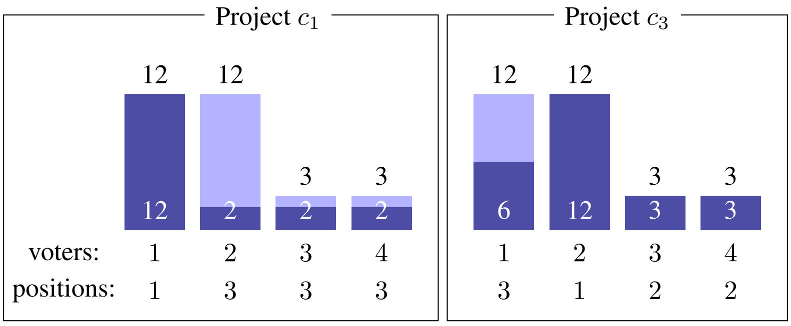
Here, we have 4 voters are we are interested in determining which project out of c_1, c_2, c_3 should be elected. Assume that projects c_1 and c_2 cost 18 dollars each and project c_3 costs 24 dollars. Let $b = 48$—consequently, every voter is initially endowed with 12 dollars. Now when distributing payments among the voters, Assumption 1 implies that voters $2, 3, 4$ should pay equal amounts of money for project c_1 and voters $3, 4$ should pay equal amounts of money for projects c_1, c_2 and c_3. On the other hand, Observation 6.1 implies that voter 1 should pay infinitely more for c_1 than voters $2, 3, 4$. It is naturally not possible—in this case, the standard Equal Shares mechanism allows voter 1 to pay less, that is, her whole remaining money. Therefore, the fair distribution of payments for c_1 should be the following: voter 1 pays 12 dollars, voters $2, 3, 4$ pay 2 dollars each. For project c_2, we can see that voters $3, 4$ actually have enough money to afford c_2 on their own (paying 9 dollars each), hence the remaining voters—having infinitely smaller utility from c_2—should not pay anything. For project c_3, we can see that voters $2, 3, 4$ have enough money to afford c_3 on their own, and the payments, analogously to c_1, should be divided so that voter 2 pays 12 dollars and voters $3, 4$ pay 6 dollars each. These payment distributions are depicted below (in a way similar to the examples in Section 3.2).



Here we cannot determine the exact values of "payment-per-utility" ρ for each project, yet we can determine for which one it is the lowest. The explanation is the following: for each project c, we should find a voter i such that (1) i pays a nonzero amount of money for c, (2) among all the voters satisfying the previous condition, i has c at the lowest position in her ranking (let it be k), (3) among all the voters satisfying previous two conditions, i pays the most for c. Note that only voters satisfying all these conditions are allowed to pay less for c than their whole remaining endowment. In the standard Equal Shares mechanism, the payment of i for c would be therefore equal to the maximal "payment-per-utility" multiplied by i's utility from c.

Now consider the another project c' (for c', we denote the analogues of i and k by i' and k' respectively) such that $k' < k$. From Observation 6.1, voter i' has infinitely greater utility from c' than i from c, hence the maximal "payment-per-utility" is greater for c than for c'. In such a case, c' should be elected instead of c.

In our example, this means we should elect project c_2 in the first round (the second-best option was c_3, the least one—c_1). Voters 3 and 4 pay 9 dollars for c_2 each. Now let us split the payments again in the second round:



Now we can see that voters $2, 3, 4$ no longer have enough money to afford c_3 on their own and voter 1 needs to participate in paying for c_3 (paying 6 dollars). Since for c_1 voters $2, 3, 4$ pay only 2 dollars each and their utilities from c_1 are equal to the utility of voter 1 from c_3 (from Assumption 1), we can see that the "payment-per-utility" ratio is now lower for c_1. Hence, c_1 is elected and the outcome $\{c_1, c_2\}$ is returned.

The detailed pseudocode of the algorithm is presented in Algorithm 2—by $\mathrm{pos}_i(c)$ we denote here the position of project c in the ranking of voter i. Note that the algorithm is exhaustive—indeed, we assume here that every voter gains here a nonzero utility from every project and is able to eventually pay for it (recall Proposition 3.6).

Algorithm 2: Implementation of Method of Equal Shares for Ordinal Utilities

$W \leftarrow \emptyset$.
for $i \in N$ **do**
    $b_i \leftarrow b/n$
while *true* **do**
    for $k \in [m]$ **do**
        for $c \in C \setminus W$ **do**
            if $\sum_{i \in N : \mathrm{pos}_i(c) \le k} b_i < \mathrm{cost}(c)$ **then**
                $p(c) \leftarrow \infty$ (project c is not affordable by voters ordering it on at most kth position)
            else
                $\mathrm{price} \leftarrow \mathrm{cost}(c) - \sum_{i \in N :\, \mathrm{pos}_i(c) < k} b_i$
                Let $1, \dots, t$ be a list of all voters $i \in N$ with $\mathrm{pos}_i(c) = k$, ordered so that $b_1 \le \dots \le b_t$.
                for $s = 1, \dots, t$ **do**
                    $p(c) \leftarrow (\mathrm{price} - (b_1 + \dots + b_{s-1}))/(t - s + 1)$
                    if $p(c) \le b_s$ **then**
                        break (the maximal payment among the voters ranking c at the kth position is $p(c)$)
        if $\min_{c \in C \setminus W} p(c) = \infty$ **then**
            return *W*
        $c \leftarrow \mathrm{argmin}_{c \in C \setminus W} p(c)$ (break ties arbitrarily)
        $W \leftarrow W \cup \{c\}$
        for $i \in N$ *such that* $\mathrm{pos}_i(c) < k$ **do**
            $b_i \leftarrow 0$
        for $i \in N$ *such that* $\mathrm{pos}_i(c) = k$ **do**
            $b_i \leftarrow b_i - \min\{b_i, p(c)\}$

6.1 Equal Shares and Proportionality for Solid Coalitions

Now we will show that the presented algorithm provides good proportional properties in the considered setting. For simplicity, in this section we will write $X \succ_i Y$ for a voter $i \in N$ and sets $X, Y \subseteq C$, to denote that for each $x \in X$, $y \in Y$, voter i prefers x to y. The formal definition of the Proportionality for Solid Coalitions axiom for the committee election model is the following:

Definition 6.2 (Proportionality for Solid Coalitions (PSC))**.** An outcome W satisfies *Proportionality for Solid Coalitions (PSC)* if for each $\ell \in [b]$, each subset of voters $S \subseteq N$ with $|S| \ge \ell \cdot n/b$, and each subset of candidates T such that $T \succ_i C \setminus T$ for all $i \in S$, it holds that $|W \cap T| \ge \min(\ell, |T|)$.

Intuitively, each group of voters S should be able to "control" the fraction of the committee that is proportional to its size. To better understand this axiom, consider the following simple example:

Example 6.3. Consider three voters with the following preference orders over the set of candidates $C = \{c_1, c_2, \dots, c_{100}\}$:

$$\begin{aligned}
1&\colon c_1 \succ c_2 \succ c_3 \succ \dots\\
2&\colon c_2 \succ c_3 \succ c_1 \succ \dots\\
3&\colon c_4 \succ c_5 \succ \dots\\
4&\colon c_5 \succ c_4 \succ \dots
\end{aligned}$$

The omitted parts of the rankings $(\dots)$ are arbitrary. Assume $b = 2$. In this example PSC would require that one candidate from $\{c_1, c_2, c_3\}$ and one candidate from $\{c_4, c_5\}$ are elected. ⌟

We can easily see that the definition of PSC assumes lexicographic preferences over committees: indeed, we do not even need to know the missing parts of the rankings to know that they would not affect the PSC guarantees.

The problem of extending PSC to the PB setting has been considered by [Aziz and Lee, 2021].[2] The particularly interesting extension (to the best of our knowledge, the only proposed one that is always satisfiable and possible to be found in polynomial time) is the following one:

Definition 6.4 (Inclusion Proportionality for Solid Coalitions (IPSC))**.** An outcome W satisfies *Inclusion Proportionality for Solid Coalitions (IPSC)* if for each subset of voters $S \subseteq N$, each subset of projects $T \subseteq C$ such that $T \succ_i C \setminus T$ for all $i \in S$, and each project $c \in T \setminus W$, it holds that:

$$\mathrm{cost}(W \cap T) + \mathrm{cost}(c) > \frac{|S|}{n} \cdot b.$$

Note that this approach is similar in spirit to our Extended Justified Representation up to one project—in both axioms the fraction of the budget "controlled" by the group of voters can be smaller than the fraction proportional to its size only with respect to one project.

For the committee election model, Method of Equal Shares as defined in Algorithm 2 is an example of the class of *Expanding Approvals* rules [Aziz and Lee, 2020, 2021]. All such rules satisfy PSC [Aziz and Lee, 2021], and thus Equal Shares satisfies PSC. For the general PB model and IPSC it is also the case.

Theorem 6.5. *Method of Equal Shares for ordinal preferences satisfies IPSC.*

Proof. Consider an outcome W returned by Equal Shares for an election (N, C, b). Suppose that IPSC is violated and let $S \subseteq N$, $T \subseteq C$ and $c \in T \setminus W$ witness this violation. The voters in S initially have the endowment of $|S| \cdot b/n$ dollars.

Since the voters initially had $|S| \cdot b/n$ dollars and $\mathrm{cost}(W \cap T) + \mathrm{cost}(c) \leq \mathrm{cost}(W \cap T) + \mathrm{cost}(c) \leq |S|/n \cdot b$, we can see that the voters had initially enough money to afford c. Now it

[2] The definition presented in the dissertation is a much simpler reformulation of the definition presented in [Aziz and Lee, 2021], defined originally for non-strict voters' rankings.

is enough to note that no project ranked below the $|T|$th position by voters from S was paid for them in any round. Indeed, in every round of the algorithm, voters from S were able to purchase project c on their own so that no voter would pay for a project ranked below the $|T|$th position. Since Equal Shares while choosing a project, minimizes its worst-case position in the rankings of its payers, the rule prefers c to any project from $C \setminus T$ that would require a voter from S to pay for it. Since the rule stops when no project is affordable, c was purchased, a contradiction. □

One may wonder, whether satisfying IPSC is just a consequence of Equal Shares satisfying EJR (as defined in Definition 4.10, adapted to ordinal utilities with respect to Assumptions 1 and 2). Example 6.6 below shows that this is not the case even for the committee election model and that the two axioms are logically incomparable in this context.

Example 6.6 (PSC is logically incomparable to EJR)**.** Consider a committee election and two voters with the following preferences:

$$1\colon c_1 \succ c_2 \succ c_3 \succ c_4$$
$$2\colon c_4 \succ c_1 \succ c_3 \succ c_2.$$

Assume $b = 1$. Here, group $\{1, 2\}$ agrees that candidate c_2 is worth at least the utility derived from the second position in a ranking. Hence, EJR requires that at least one voter would gain this utility, which would be the case if c_1, c_2, or c_4 was selected. On the other hand, $\{c_3\}$ is a committee that satisfies PSC. Thus, PSC does not imply EJR.

Consider now three voters with the following preferences:

$$1\colon c_1 \succ c_2 \succ c_3 \succ c_4$$
$$2\colon c_2 \succ c_3 \succ c_1 \succ c_4$$
$$3\colon c_3 \succ c_1 \succ c_2 \succ c_4.$$

Assume $b = 2$. In this example the group $\{1, 2\}$ again agrees that candidate c_2 is worth at least the utility derived from the second position in a ranking. It is also the case for groups $\{1, 3\}$ and $\{2, 3\}$ and candidates c_1, c_3, respectively. All these groups can be satisfied by electing c_1, hence the committee $\{c_1, c_4\}$ satisfies EJR. However, PSC requires here that two candidates from $\{c_1, c_2, c_3\}$ are elected. Thus, EJR does not imply PSC. ⌟

6.2 The Costwise Ordinal Variant of Equal Shares

In this section let us consider the alternative version of Assumption 1, mentioned at the beginning of the chapter.

Assumption 1* If two voters $i, j \in N$ have at the kth position in their rankings ($k \leq m$) projects c_1 and c_2 respectively, the relation between the utility of i from c_1 and the utility of j from c_2 is the same as the relation between $\mathrm{cost}(c_1)$ and $\mathrm{cost}(c_2)$.

In fact, the major part of the algorithm does not change after this modification. First, note that Observation 6.1 still holds as a conclusion from Assumptions 1^* and 2. Moreover, for every project $c \in C$, the voters ranking it at the same position still gain the same utility from it. Hence, the way of dividing payments for c among the voters remains unchanged. Besides, the rule in each round still chooses the project minimizing its worst-case position in the rankings of its payers.

Now, the only place in the algorithm where the difference appears is the choice of a project when for several projects this worst-case position k is the same. In the original algorithm we choose the project c minimizing the maximal payment of a voter ranking c at the kth position. Under Assumption 1^*, we would rather choose the project c minimizing the maximal payment of a voter ranking c at the kth position divided by $\mathrm{cost}(c)$ (in line 18 of Algorithm 2, there would be $c \leftarrow \mathrm{argmin}_{c \in C \setminus W} p(c)/\mathrm{cost}(c)$). Note that it is an analogous difference to the one between Equal Shares for approval-based and approval-based cost utilities (presented in Section 3.2). Therefore, we call this variant of Equal Shares for ordinal preferences the *costwise* variant.

The considered change does not affect the proof of Theorem 6.5. The further exploration of differences in axiomatic properties between the two ordinal variants of Equal Shares is left for future research.

As we can see, Equal Shares is a flexible voting rule and it can be adapted to various settings, satisfying the strongest proportionality notions in each of them.

Part II

Beyond Equal Shares: Stronger Notions of Proportionality

Chapter 7

Full Justified Representation and the Greedy Cohesive Rule

In Chapter 4, we discussed the EJR axiom for the PB model, and saw that in its "up-to-one" variant it is satisfied by Method of Equal Shares.

We will now propose a strengthening of EJR, called Full Justified Representation (FJR), which is substantially stronger even in the approval-based committee election model. We also prove that this axiom is satisfiable, by presenting an voting rule satisfying it. This new rule, called Greedy Cohesive Rule (GCR) provides the strongest known proportionality guarantees. On the other hand, we will argue that, compared to Equal Shares, it is computationally more expensive, less flexible and less natural.

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 7.1 (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

$$\frac{|S|}{n} \geq \frac{\mathrm{cost}(T)}{b} \quad \text{and} \quad u_i(T) \geq \beta$$

for every voter $i \in S$.

A rule $\mathcal{R}$ satisfies *Full Justified Representation (FJR)* if for each election E and each weakly (β, T)-cohesive group of voters S there exists a voter $i \in S$ such that $u_i(\mathcal{R}(E)) \geq \beta$.

For a better understanding, let us first consider the above definition in the approval-based committee election model. Then 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 $|S| \geq \ell/b \cdot n$. Then at least one voter from S must have at least β representatives in the committee. The difference between EJR and FJR can be seen in Example 7.2.

Example 7.2. Consider an approval-based committee election with $b = 10$ and a group S of 30% of voters. All the members of S approve some candidate c_1. Additionally, half of group S approve a candidate c_2 and the other half approve candidate c_3. Now according to EJR, at least one voter from S should have at least one representative in the elected committee (because even though they are large enough to claim 3 candidates, they jointly approve only one). On the other hand, according to FJR, at least one voter from S should have at least two representatives—because all the voters from S approve some 2 candidates in the three-element set $\{c_1, c_2, c_3\}$, the group is weakly (2, $\{c_1, c_2, c_3\}$)-cohesive.

It is clear that in the special case of $\beta = \ell$, we obtain Definition 4.3, hence FJR implies EJR in the approval-based committee election model. The same implication holds in the general PB model.

Proposition 7.3. *FJR implies EJR in the general PB model.*

Proof. Suppose that rule $\mathcal{R}$ satisfies FJR and take an (α, T)-cohesive group of voters S for some $\alpha \in \mathbb{N}$, $T \subseteq C$. We set $\beta = \alpha$; clearly, we have also $u_i(T) \geq \beta$, thus S is weakly cohesive. As $\mathcal{R}$ satisfies FJR, we have that $u_i(\mathcal{R}(E)) \geq \beta = \alpha$, which completes the proof. □

In turn, it is easy to see that FJR is implied by core-stability (cf. Definition 4.14). Figure 22 shows the logical relationships between these axioms. On the other hand, FJR is stronger than some other relaxations of core-stability discussed by Peters and Skowron [2020].

Core	$\longrightarrow$	FJR Full Justified Representation	$\longrightarrow$	EJR Extended Justified Representation	$\longrightarrow$	EJR-1 EJR up to one project

Figure 22: Implications between the axioms considered in the dissertation.

To the best of our knowledge, no rule known before our work satisfies FJR for approval-based committee elections, let alone for the general PB model. In particular, it is the case of Method of Equal Shares.

Example 7.4 (Equal Shares fails FJR). Consider the following approval-based committee election for $n = 22$ voters, $m = 13$ candidates, and where the goal is to select a committee of size $b = 11$:

voters 1-3:	$\{c_1, c_2, c_3, c_4, c_8\}$	voters 13-15:	$\{c_1, c_2, c_3, c_4, c_{12}\}$
voters 4-6:	$\{c_1, c_2, c_3, c_4, c_9\}$	voters 16-18:	$\{c_5, c_6, c_7, c_8, c_9, c_{10}, c_{11}, c_{12}\}$
voters 7-9:	$\{c_1, c_2, c_3, c_4, c_{10}\}$	voters 19-21:	$\{c_5, c_6, c_7\}$
voters 10-12:	$\{c_1, c_2, c_3, c_4, c_{11}\}$	voter 22:	$\{c_{13}\}$.

In the first 4 steps, Equal Shares chooses candidates c_1, c_2, c_3, c_4 (this happens for $\rho = 1/15$). After that, each of the first 15 voters has a remaining budget of $11/22 - 4/15$. In next 3 steps, for $\rho = 1/6$, candidates c_5, c_6, c_7 are chosen: the 6 voters who support them spend all their money ($11/22 - 3 \cdot 1/6 = 0$). After that, the algorithm stops. Each of the first 15 voters has 4 candidates she approves; voters 16-18 approve 3 selected candidates. Thus, no member of the weakly $(5, \{c_1, c_2, c_3, c_4, c_8, c_9, c_{10}, c_{11}, c_{12}\})$-cohesive group of the first 18 voters has 5 representatives.

⌟

The provided example works for the base version of Equal Shares, yet we conjecture that the answer is negative for any completion variant of this rule. This is one of the questions that we leave for the future research.

Still, it turns out that FJR can always be satisfied: we present a (somewhat artificial) rule satisfying this strong notion of proportionality.[1]

Definition 7.5 (Greedy Cohesive Rule (GCR))**.** The *Greedy Cohesive Rule* (GCR) is defined sequentially as follows: we start with an empty outcome $W = \emptyset$ and label all voters as *active*. At each step, we search for a value $\beta > 0$, a set of voters $S \subseteq N$ who are all active, and a set of candidates $T \subseteq C \setminus W$ 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 $\mathrm{cost}(T)$.[2] We add all the candidates from T to W, and then label all voters in S as inactive.

Let us first check that the Greedy Cohesive Rule always selects an outcome 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 $|S|/n \geq \mathrm{cost}(W)/b$, and hence the algorithm labels at least $\mathrm{cost}(W)/b \cdot n$ voters as inactive after this step. Thus, if GCR selects an outcome with total cost $\mathrm{cost}(W)$, then it must have inactivated at least $\mathrm{cost}(W)/b \cdot n$ voters during its execution. Because there are n voters, we have $\mathrm{cost}(W)/b \cdot n \leq n$, and hence $\mathrm{cost}(W) \leq b$.

In Example 7.4, GCR explicitly finds in the first step the weakly $(5, T)$-cohesive group of the first 18 voters where $T = \{c_1, c_2, c_3, c_4, c_8, c_9, c_{10}, c_{11}, c_{12}\}$ and elects outcome T. After that, the rule marks these voters as inactive and, since there are no more weakly cohesive groups among active voters, terminates. As we can see, Greedy Cohesive Rule is designed specifically to satisfy FJR.

Theorem 7.6. *The Greedy Cohesive Rule satisfies FJR.*

Proof. Assume for a contradiction that there exists a weakly (β, T)-cohesive group S which witnesses that FJR is not satisfied by the outcome selected by GCR. Consider the voter $i \in S$ who was first labeled inactive by GCR, and the outcome W right after that step (since S is weakly cohesive, such an i always exists). Since $i \in S$ and S witnesses the FJR failure, we

[1]A similar rule is used by Aziz and Lee [2021] to prove that there always exists an outcome satisfying their axiom BPJR-L for approval-based elections. That axiom is weaker than FJR even when applied to a utility profile where $u_i(c) = \mathrm{cost}(c)$ whenever i approves c, and $u_i(c) = 0$ otherwise.

[2]This way of breaking ties matters in Section 7.1 (to show that GCR can be extended to a priceable outcome), but it does not matter how we break ties for the proof of Theorem 7.6 (to show that GCR satisfies FJR).

have $u_i(W) < \beta$. We know that i was inactivated as a member of some weakly (β', T')-cohesive group S'. Just before S' was labeled inactive, all of the members of S were active. Thus, we have $\beta' \geq \beta$, as GCR maximizes this value. However, since $T' \subseteq W$, we have

$$\beta' \leq u_i(T') \leq u_i(W) < \beta,$$

a contradiction to $\beta' \geq \beta$. Hence, such a group S does not exist. □

What is particularly interesting, GCR provides proportionality guarantees also for much stronger model than the one considered in the dissertation.

Observation 7.7 (GCR and FJR do not require additivity)**.** In this dissertation, we have assumed additive utility functions throughout. But the definitions of FJR and of GCR make sense for any utility functions $u_i : 2^C \to \mathbb{R}_{\geq 0}$ over subsets of C that are *monotone* in the sense that $u_i(T_1) \geq u_i(T_2)$ whenever $T_1 \supseteq T_2$. Monotone utility functions allow us to encode, for example, complementarities and substitutes. The proof of Theorem 7.6 only used monotonicity, and hence GCR satisfies FJR for all monotone utility functions. In contrast, the definition of EJR relies on additivity. ⌟

7.1 Priceability of the Greedy Cohesive Rule

GCR satisfies neither priceability (Definition 4.18) nor exhaustiveness (Definition 3.2). However, we will prove that an outcome elected by this rule can always be completed to a priceable one; this suggests that GCR never elects outcomes that are "too unbalanced". In the proof of Theorem 7.10 we describe precisely how such a completion can be implemented. Using a somewhat different completion scheme, we can complete GCR to an exhaustive outcome. This way we obtain an outcome that is both exhaustive and also close to being priceable.

We start by proving two useful lemmas, which establish a kind of "Hall condition" for priceability which may of independent interest.

Lemma 7.8. *Let S be a (β, T)-cohesive group which is selected in some step of GCR. For every subset $A \subseteq T$, the size of the set of voters $S' := \{i \in S \colon u_i(A) > 0\}$ is at least* $\mathrm{cost}(A) \cdot {}^{n}\!/_{b}$.

Proof. The statement is trivial for $\mathrm{cost}(A) = 0$, so assume that $\mathrm{cost}(A) > 0$. Assume for a contradiction that the set $S' \subseteq S$ defined above has size $|S'| < \mathrm{cost}(A) \cdot {}^{n}\!/_{b}$. Then the group $S \setminus S'$ together with the set $T \setminus A$ is $(\beta, T \setminus A)$-cohesive because

$$|S \setminus S'| > |S| - \mathrm{cost}(A) \cdot \frac{n}{b} \geq \mathrm{cost}(T) \cdot \frac{n}{b} - \mathrm{cost}(A) \cdot \frac{n}{b} = \mathrm{cost}(T \setminus A) \cdot \frac{n}{b}.$$

Further, as $\mathrm{cost}(A) > 0$, we have $\mathrm{cost}(T \setminus A) < \mathrm{cost}(T)$. Thus, GCR would select $S \setminus S'$ instead of S, a contradiction. □

Lemma 7.9. *For every outcome W elected by GCR, there always exists a price system $p = ({}^{b}\!/_{n}, W, \{p_i\}_{i \in N})$, possibly not satisfying (P).*

Proof. Consider a single step of GCR and let S be a (β, T)-cohesive group considered in that step. We will prove that there exists a price system in which voters from S pay $\mathrm{cost}(c)$ dollars for each candidate $c \in T$. By combining these price systems for all the (pairwise disjoint) groups S selected by GCR, we obtain a price system for the outcome of GCR.

Without loss of generality, we assume that b/n is an integer and that $\mathrm{cost}(c)$ is an integer for all $c \in C$ (we can just appropriately scale up all costs and the budget, because we assumed that the budget and the costs are integers).

We now imagine that each candidate $c \in T$ is split into $\mathrm{cost}(c)$ many *pieces*, each with cost 1. Let A_T be the set of all pieces. We also imagine that each voter $i \in S$ has b/n many *coins*, each worth 1, and let A_S be the set of all coins. Note that one coin can pay for one piece.

Consider the bipartite graph $G = (A_S + A_T, E)$, where there is an edge between a coin belonging to voter $i \in S$ and a piece of a candidate $c \in T$ if and only if $u_i(c) > 0$.

Now, consider any subset of pieces $A \subseteq A_T$, and let us assess the size of the neighborhood $N_G(A) \subseteq A_S$. Let $C(A)$ denote the set of candidates who have at least one piece in A. Then $N_G(A)$ consists of all the coins of those voters who assign a positive utility to some candidate from $C(A)$. By Lemma 7.8 there are at least $\mathrm{cost}(C(A)) \cdot n/b$ such voters, each of whom have b/n many coins. Thus

$$|N_G(A)| \geq \mathrm{cost}(C(A)) \cdot \frac{n}{b} \cdot \frac{b}{n} = \mathrm{cost}(C(A)) \geq |A|.$$

Hence, by Hall's theorem, there is a one-to-one mapping between coins and pieces. This allows us to construct payment functions as follows: For every voter $i \in S$ and candidate $c \in T$, if exactly q coins of i are mapped to some parts of c, then $p_i(c) = q$. It is straightforward to check that such a payment function satisfies conditions for a price system for initial endowment equal to b/n and supported outcome W, which completes the proof. □

Finally, we can state the main result of this section.

Theorem 7.10. *Every outcome W elected by GCR can be completed to a priceable outcome.*

Proof. From Lemma 7.9, we know that there exists a price system $p = (b/n, W, \{p\}_{i \in N})$, possibly not satisfying (P). Now, to complete W to a priceable outcome $W' \supseteq W$, it is enough to run Equal Shares for this election with the initial elected outcome set to W and the initial endowment of each voter $i \in N$ set to $b/n - \sum_{c \in W} p_i(c)$. □

7.2 Drawbacks of the Greedy Cohesive Rule

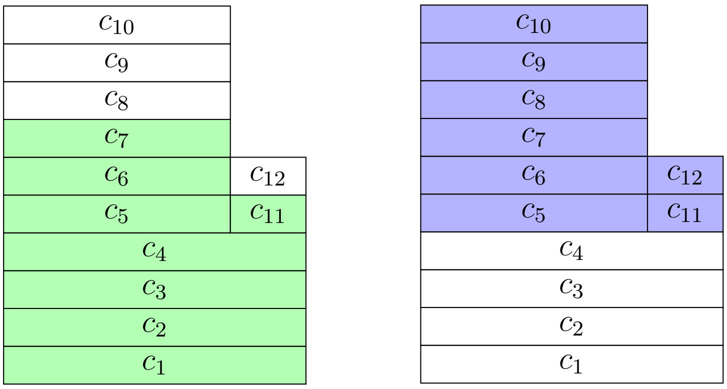
Since Greedy Cohesive Rule with a proper completion is priceable and, contrary to Equal Shares, satisfies FJR, we may conclude that GCR is a better voting rule. However, the fact that GCR is custom-engineered to satisfy FJR, makes it deficient in other dimensions. In this section we consider a few drawbacks of this rule compared to Equal Shares.

7.2.1 Inefficiency on (Nearly) Laminar Profiles

We begin by discussing a property that Peters and Skowron [2020] call *laminar proportionality*. The axiom is defined for the case of approval-based committee election model.

This property identifies a family of well-behaved preference profiles and specifies the outcome on those profiles. Equal Shares satisfies it; Example 7.11 shows that GCR does not.

Example 7.11 (GCR fails laminar proportionality)**.** Let $N = \{1, 2, 3, 4\}$ and $b = 8$, and introduce the candidate sets $X = \{c_1, \ldots, c_4\}$, $Y = \{c_5, \ldots, c_{10}\}$, and $Z = \{c_{11}, c_{12}\}$. The first three voters approve $X \cup Y$, and the fourth one approves $X \cup Z$. Two copies of the profile are depicted below. The candidates are represented by boxes; each candidate is approved by the voters who are below the corresponding box.



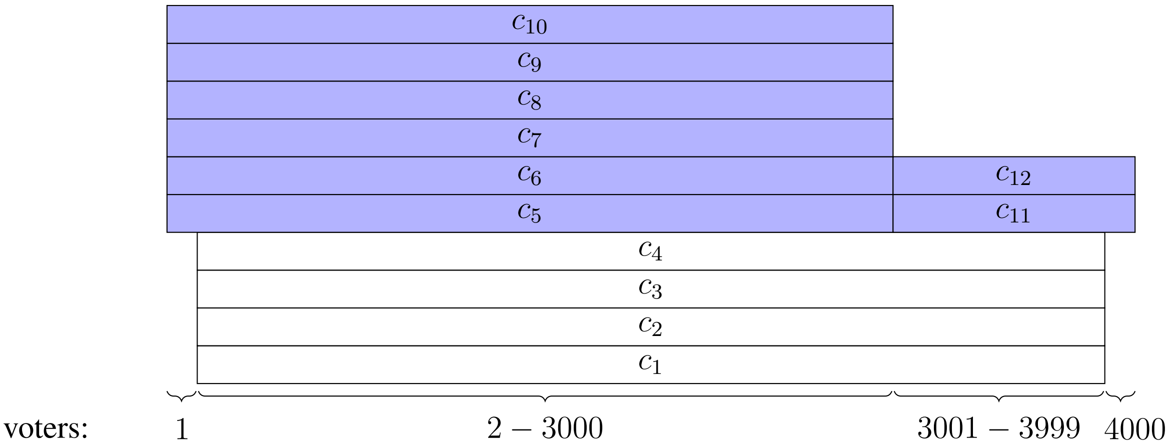
voters:	1	2	3	4	1	2	3	4
utilities:	7	7	7	5	6	6	6	2

In this election, laminar proportionality would require that the voting rule selects all the candidates from X since they are approved by everyone. After electing the candidates in X, four seats are left to fill. Since the group $\{1, 2, 3\}$ the three times as large as the group $\{4\}$, laminar proportionality requires that we elect three candidates from Y and one candidate from Z. Thus, the committee indicated by the green boxes on the left-hand figure is laminar proportional.

On the other hand, in the first step GCR can choose the weakly $(6, Y)$-cohesive group $\{1, 2, 3\}$ and in the second step it can select the weakly $(2, Z)$-cohesive group $\{4\}$. This results in the blue committee depicted in the right-hand figure; this committee fails laminar proportionality. ⌟

As we can see, GCR on laminar profiles can be severly inefficient—instead of electing an outcome providing utility 7 for the first three voters and utility 5 for the last one, it elects an outcome providing them utility 6 and 2 respectively. However, this example is not fully satisfactory, as it depends on tie-breaking. For example, in the first step we could choose the weakly $(6, \{c_2, c_3, c_4, c_5, c_6\})$-cohesive group containing the first three voters, and in the second step the weakly $(2, \{c_1, c_{11}\})$-cohesive group containing the last voter. An open question is whether GCR can always elect a committee satisfying laminar proportionality (among others). However, the following example shows that for some "nearly laminar" elections, GCR does not match the general intuition standing behind this axiom.

Example 7.12. Modify the election described in Example 7.11 in the following way: we have now 4000 voters. Voter 1 approves only candidates from Y, voters 2 to 3000 approve $X \cup Y$, voters 3001 to 3999 approve $X \cup Z$ and voter 4000 approves Z.



This election is not laminar (because of the two voters not approving X), but it is close to being laminar and it is reasonable to expect that the elected committee should be the same as the green one from Example 7.11. Equal Shares uniquely elects that committee. On the other hand, GCR selects first the weakly $(6, Y)$-cohesive group containing the first 3000 voters and in the second step the weakly $(2, Z)$-cohesive group containing the last 1000 voters. After that the algorithm stops, electing committee $Y \cup Z$, as depicted above. Note that in this case, the choice of weakly cohesive groups is unique. ⌟

Examples 7.11 and 7.12 do not rule out the existence of an FJR rule that is also laminar proportional; the existence of a natural example of such a rule is an interesting open problem.

7.2.2 Unproportionality for Ordinal Preferences

Let us now adapt GCR to ordinal preferences. Similarly to Method of Equal Shares (recall Chapter 6), the idea is to convert rankings to utilities. In case of GCR, to force the voters to compare outcomes lexixographically, it is sufficient to assume that the utilities are exponentially decreasing with the positions: for each $i \in N$ and $c \in C$ we set $u_i(c) = m^{-\mathrm{pos}_i(c)}$. Then, for each c we have that $u_i(c) > \sum_{c' \prec_i c} u_i(c')$, and so the utility a voter assigns to a project in position p is higher than the utility that it would assign to any outcome all of whose members are ranked below p.

This is sufficient, because in GCR utilities are only used to compare sets of candidates in order of preference. For Equal Shares, we needed to use infinite utilities because these values were also used to decide how much each voter pays.

Now we will show that GCR in the model of ordinal preferences fails Proportionality for Solid Coalitions (Definition 6.2) for the model of committee elections (which implies it fails IPSC for the general PB model).

Example 7.13. Consider the following preference profile:

$$
\begin{aligned}
1\colon & c_1 \succ c_7 \succ c_8 \succ c_6 \succ c_4 \succ c_5 \succ c_2 \succ c_3 \succ c_9 \succ c_{10} \succ c_{11} \succ c_{12} \\
2\colon & c_1 \succ c_7 \succ c_8 \succ c_6 \succ c_4 \succ c_5 \succ c_2 \succ c_3 \succ c_9 \succ c_{10} \succ c_{11} \succ c_{12} \\
3\colon & c_1 \succ c_2 \succ c_3 \succ c_6 \succ c_4 \succ c_5 \succ c_7 \succ c_8 \succ c_9 \succ c_{10} \succ c_{11} \succ c_{12} \\
4\colon & c_1 \succ c_2 \succ c_3 \succ c_6 \succ c_4 \succ c_5 \succ c_7 \succ c_8 \succ c_9 \succ c_{10} \succ c_{11} \succ c_{12} \\
5\colon & c_1 \succ c_2 \succ c_3 \succ c_6 \succ c_4 \succ c_5 \succ c_7 \succ c_8 \succ c_9 \succ c_{10} \succ c_{11} \succ c_{12} \\
6\colon & c_1 \succ c_2 \succ c_3 \succ c_6 \succ c_4 \succ c_5 \succ c_7 \succ c_8 \succ c_9 \succ c_{10} \succ c_{11} \succ c_{12} \\
7\colon & c_2 \succ c_3 \succ c_1 \succ c_7 \succ c_8 \succ c_4 \succ c_5 \succ c_6 \succ c_9 \succ c_{10} \succ c_{11} \succ c_{12} \\
8\colon & c_3 \succ c_2 \succ c_1 \succ c_7 \succ c_8 \succ c_4 \succ c_5 \succ c_6 \succ c_9 \succ c_{10} \succ c_{11} \succ c_{12} \\
9\colon & c_4 \succ c_5 \succ c_9 \succ c_7 \succ c_8 \succ c_1 \succ c_2 \succ c_3 \succ c_6 \succ c_{10} \succ c_{11} \succ c_{12} \\
10\colon & c_5 \succ c_4 \succ c_9 \succ c_7 \succ c_8 \succ c_1 \succ c_2 \succ c_3 \succ c_6 \succ c_{10} \succ c_{11} \succ c_{12} \\
11\colon & c_{10} \succ c_{11} \succ c_{12} \succ c_7 \succ c_8 \succ c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_9 \\
12\colon & c_{11} \succ c_{10} \succ c_{12} \succ c_7 \succ c_8 \succ c_1 \succ c_2 \succ c_3 \succ c_4 \succ c_5 \succ c_6 \succ c_9.
\end{aligned}
$$

Assume $b = 4$. Here, GCR will first pick $S = \{1, \dots, 6\}$ as a weakly cohesive group, with the corresponding set of candidates $T = \{c_1, c_6\}$. Indeed, if T consisted of 3 candidates, then S would need to have at least 9 voters. However, any 9 voters rank at least 4 different candidates at the top position, thus at least one of them would have a lower satisfaction than the voters from S have from T. By the same argument, T cannot consist of 4 candidates. If T consisted of 2 candidates but S included one voter from $7, \dots, 12$, then the satisfaction of voter 1 or 2 would also be lower. Indeed, these two voters rank c_2, c_3, c_4, c_5, c_{10}, and c_{11} (that is candidates that appear in the top positions) below c_6.

Hence, GCR picks c_1 and c_6, and labels the first 6 voters as inactive. In the second step, the rule picks c_7 and c_8. This is because each other candidate appears at most twice before c_7 and c_8 in the remaining voters' rankings. Thus, the rule picks c_1, c_6, c_7 and c_8.

On the other hand, by looking at voters $3, \dots, 8$ we observe that PSC requires that two candidates from c_1, c_2, c_3 should be selected. ⌟

Chapter 8

The Core under Restricted Domains

In this chapter we focus on the proportionality in the committee election model, specifically on notion of the *core*, first defined in Definition 4.14. Let us recall this definition in a slightly adapted way:

Definition 8.1 (The Core)**.** For a committee election $E = (N, C, b)$, a committee W is *core-stable* if for every $S \subseteq N$ and $T \subseteq C$ with

$$\frac{|S|}{n} \geq \frac{\mathrm{cost}(T)}{b}$$

there exists a voter $i \in S$ who weakly prefers W to T. The set of all core-stable committees is called *the core* of E.

Formulated in this way, this definition applies to all the types of voters' preferences considered in the dissertation, in particular to *ordinal* and *approval-based* ones. From now, we focus on these two types only. The main question is the following: is core-stability always satisfiable in these settings? For general cardinal utilities, the answer was negative, as we could see in Example 4.15. The same answer holds for ordinal preferences:

Example 8.2 (The core may be empty for ordinal preferences)**.**[1] Consider a committee election with 6 voters and 6 candidates. Voters' preferences are the following:

$$\begin{array}{ll} 1\colon c_1 \succ c_2 \succ c_3 \succ \dots & 4\colon c_4 \succ c_5 \succ c_6 \succ \dots \\ 2\colon c_3 \succ c_1 \succ c_2 \succ \dots & 5\colon c_6 \succ c_4 \succ c_5 \succ \dots \\ 3\colon c_2 \succ c_3 \succ c_1 \succ \dots & 6\colon c_5 \succ c_6 \succ c_4 \succ \dots \end{array}$$

[1]This example is a simplified version of the one from [Aziz et al., 2017b]. In this paper, the authors implicitly prove that the core may be empty for ordinal preferences, since the notion of *local stability* they consider is very closely related to core-stability (see the discussion in Section 8.5.1).

Fix $b = 3$. Now let us prove that at least two candidates out of $\{c_1, c_2, c_3\}$ need to be elected. Indeed, suppose that only one candidate is elected. Without loss of generality (the election is symmetrical) let it be c_1. Then voters 2 and 3, forming together a group of n/b voters, and candidate c_3 witness the violation of core-stability. Analogously, we can prove that at least two candidates out of $\{c_4, c_5, c_6\}$ need to be elected. This is a contradiction, since we have only place for 3 candidates in the committee. $\lrcorner$

For approval-based elections, the question whether core-stability is satisfiable, is one of the major open problems in computational social choice. Given the property is so demanding, so far the literature focused on its relaxed versions—either the weaker properties which we have analyzed in previous chapters, or the approximate [Jiang et al., 2020, Peters and Skowron, 2020] and the randomized [Cheng et al., 2019] variants of the core.

In this chapter we explore a different, yet related approach. Our point is that before we look at how a voting rule works in the general case, at the very minimum we shall ensure that it behaves well on well-structured preferences. Thus, the main question that we state is whether core-stability can be satisfied for certain natural restricted domains of voters' preferences, and what is the computational complexity of finding committees that are core-stable, given elections where the voters' preferences come from restricted domains. The idea to restrict the scope only to elections in which the preferences are somehow well-structured is not new [Elkind et al., 2017], yet to the best of our knowledge it has never been considered in the context of the core.

Our contribution is the following: first, we prove a number of structural theorems that describe existing domain restrictions. In particular, our results give a more intuitive explanation of the class of top-monotonic preferences. The original definition of this class is somewhat cumbersome. We show two independent conditions that provide alternative characterizations of top-monotonic preferences provided the voters' preference rankings have no ties.

We also introduce two new domain restrictions which are natural, and which provide sufficient conditions for the existence of core-stable rules. One of our new classes generalizes voter-interval and candidate-interval domains [Elkind and Lackner, 2015], and the other class is a weakening of the domain of top-monotonic preferences; yet our class still includes single-peaked [Black, 1948] and single-crossing preferences [Mirrlees, 1971, Roberts, 1977].

Second, we prove the existence of core-stable rules under the assumption that the voters' preferences come from certain restricted domains, in particular from domains of voter-interval, candidate-interval, single-peaked, and single-crossing preferences. Interestingly, we show a single algorithm that is core-stable for all four aforementioned domains. At the same time, we show that if we restrict our attention to top-monotonic elections no core-stable rule exists.

8.1 Restricted domains

A voting rule specifies an outcome of an election independently of how the voters' preferences look like. Similarly, core-stability puts certain structural requirements on the selected committees that should be satisfied in every possible election. However, the space of all elections is reach and it might be too demanding to expect a voting rule to satisfy a strong property in each possible case. For example, this is the case for the core: there are elections with strict rankings where no committee belongs to the core [Fain et al., 2018, Cheng et al., 2019]; the question whether core-stability is satisfiable assuming approval preferences is still open. Instead, what is often desired is that a voting rule should satisfy strong notions of proportionality when the voters' preferences are in some sense logically consistent. This motivates focusing primarily on elections where the voters' preferences are well-structured, or—in other words—come from certain restricted domains.

In this section we describe a few known and introduce one new preference domain. We show that the commonly known voting rules are not core-stable even for the most restricted preference domains. We also provide alternative conditions characterizing some of the considered domains. These results will help us in our further analysis of voting methods, but are also interesting on their own.

8.1.1 Ordinal preferences

For ordinal preferences we start by recalling the definitions of the following two known preference classes.

Definition 8.3 (Single-crossing preferences)**.** Given an election $E = (N, C, b)$, we say that E has single-crossing preferences if there exists a linear order $\sqsupset$ over voters such that for each voters $x \sqsupset y \sqsupset z$ and candidates $a, b \in C$ such that $a \succ_y b$ we have that $b \succ_x a \implies a \succ_z b$.

Intuitively, we say that preferences are single-crossing if the voters can be ordered in such a way that for each pair of candidates, $a, b \in C$, the relative order between a and b changes at most once while we move along the voters. For the definition of single-peaked preferences, let us denote by top_i the top-preferred option of voter $i \in N$.

Definition 8.4 (Single-peaked preferences)**.** Given an election $E = (N, C, b)$, we say that E has single-peaked preferences if there exists a linear order $\sqsupset$ over candidates such that for each voter $i \in N$ and candidates $a \sqsupset b \sqsupset c$ we have that $\mathrm{top}_i = a \implies b \succ_i c$ and $\mathrm{top}_i = c \implies b \succ_i a$.

Definition 8.5 (1D-Euclidean preferences)**.** Given an election $E = (N, C, b)$, we say that E has 1D-Euclidean preferences if there exists a 1D-Euclidean metric space in which both voters and candidates are located, such that each voter $i \in N$ prefers a candidate a to a candidate b if and only if a is closer to i than b.

Every 1D-Euclidean election is both single-peaked and single-crossing.

We will now present the definition of the *top monotonic* domain [Barberà and Moreno, 2011], generalizing both the single-peaked and single-crossing domains. This domain is defined assuming the voters submit their preferences as weak orders $\{\succsim_i\}_{i \in N}$—in exception, note that a voter $i \in N$ may here have a few top-preferred candidates. We call all candidates that are ranked top by at least one voter *top candidates*.

Definition 8.6 (Top monotonicity (TM))**.** [2] Given an election $E = (N, C, b)$, we say that E has *top monotonic* preferences if there exists a linear order $\sqsupset$ over candidates such that the two following conditions hold:

- for each candidates a, b, c and voters i, j such that a is top-preferred by i and b is top-preferred by j, it holds that:

$$(a \sqsupset b \sqsupset c \text{ or } c \sqsupset b \sqsupset a) \implies \begin{array}{l} b \succsim_i c \text{ if } c \text{ is top-preferred by both } i \text{ and } j \\ b \succ_i c \text{ otherwise} \end{array}$$

- the same implication holds also for each top candidates a, b, c and voters i, j such that $a \succsim_i b, c$ and $b \succsim_j a, c$.

The definition of top-monotonic preferences is complex and somewhat counterintuitive. We will first show that for strict ordinal preferences this definition can be equivalently characterized by two much simpler and more intuitive conditions.

Definition 8.7 (Single-top-peaked (STP) preferences)**.** Given an election $E = (N, C, b)$, we say that E has *single-top-peaked* preferences if there exists a linear order $\sqsupset$ over candidates such that for each candidates $a \sqsupset b \sqsupset c$ such that b is a top candidate, and a voter i it holds that $\mathrm{top}_i = a \implies b \succ_i c$ and $\mathrm{top}_i = c \implies b \succ_i a$.

Proposition 8.8. *In the strict model, single-top-peakedness is equivalent to top-monotonicity.*

Proof. Observe that the first condition in the definition of TM implies STP. Now, we will show the reverse implication. Consider an STP election. We will show that it satisfies the two conditions specified in Definition 8.6.

Note that in the strict model if the premise of the first condition is satisfied, then $\mathrm{top}_i = a$ and $\mathrm{top}_j = b$ and hence c is not top-ranked by neither i nor j. Hence, the first condition follows from the condition for STP.

Consider now the second condition. If $a \succsim_i b, c$ and $b \succsim_j a, c$, then in the strict model it holds that $a \succ_i b, c$ and $b \succ_j a, c$. Let us consider two cases: first assume that $a \sqsupset b \sqsupset c$. We know that $\mathrm{top}_i = \{d\}$ for some $d \in C \setminus \{b, c\}$. If $d \sqsupset b$, then $b \succ_i c$ follows from the definition of STC (for voter i and candidates d, b, c). Suppose now that $b \sqsupset d$. But then from the definition of STP (for voter i and candidates a, b, d) we obtain that $b \succ_i a$, a contradiction. The reasoning for the case when $c \sqsupset b \sqsupset a$ is analogous. □

It is clear that the definition of STP is closely connected to the definition of single-peaked preferences (only the condition is partially weakened to the candidates that are ranked top by some voter). One could also consider the analogous weakening for single-crossing preferences.

Definition 8.9 (Single-top-crossing (STC) preferences)**.** Given an election $E = (N, C, b)$, we say that E has *single-top-crossing* preferences if there exists a linear order $\sqsupset$ over voters such that for each voters $x \sqsupset y \sqsupset z$ and a candidate $a \in C$, we have that $a \succ_x \mathrm{top}_y \implies \mathrm{top}_y \succ_z a$.

[2] This definition presented here is a slightly adapted, yet equivalent version of the one in [Barberà and Moreno, 2011].

Although the definitions of STC and STP look different, they are in fact equivalent.

Proposition 8.10. *In the strict model, single-top-peakedness is equivalent to single-top-crossingness.*

Proof. Consider an STC election E and a linear order $\sqsupset$ over voters given by the definition of STC. We say that i preceds j if $j \sqsupset i$. We construct the linear order over candidates as follows:

1. Consider some $a, b \in C$ such that a is the top preference for some voter $i \in N$. From the definition of STC, we know that voters preferring b to a can all either succeed or precede i. If they succeed i, then we add constraint $b \sqsupset a$, otherwise we add constraint $a \sqsupset b$. If there are no voters prefering b to a, we add no constraint. We repeat this step for each pairs $a, b \in C$.

2. Finally, if after the previous step some pairs are still uncomparable, we complete the order in any transitive way.

We will show that the constraints placed during the first step of the procedure are transitive. Indeed, consider (for the sake of contradiction) three candidates a, b, c such that the procedure placed constraints $a \sqsupset b$, $b \sqsupset c$ and $c \sqsupset a$. Hence, we know that at least two out of these three candidates are top candidates. Assume without the loss of generality that a and b are top candidates. Let i_a, i_b be voters ranking top respectively a and b (naturally, $i_a \sqsupset i_b$). We know that all the voters preceding i_a prefer a to c and all the voters preceding i_b prefer c to b. There exists at least one voter i preferring c to b (as otherwise constraint $b \sqsupset c$ would not be added) and $i_b \sqsupset i$. By transitivity of the preference relation, we know that i prefers a over b. Consequently, i_a, i_b and i together with candidate a witness STC violation. The obtained contradiction shows that the order $\sqsupset$ is indeed transitive.

We will now prove that such linear order $\sqsupset$ over candidates satisfies the conditions of STP. Indeed, consider any three candidates $a \sqsupset b \sqsupset c$ such that b is a top candidate and a voter $i \in N$. Let $\mathrm{top}_i = \{a\}$. As b is a top candidate, there exists a voter j such that $\mathrm{top}_j = \{b\}$. As $a \sqsupset b$, it holds that $i \sqsupset j$. Then if we had that $c \succ_i b$, our procedure would place constraint $c \sqsupset b$, a contradiction. Hence $b \succ_i c$. The proof for the case $\mathrm{top}_i = \{c\}$ is analogous.

Now we will prove the reverse implication. Let E be an STP election with a linear order $\sqsupset$ over the candidates. Consider the following linear order $\sqsupset$ over the voters: for each $i, j \in N$ we have that if $\mathrm{top}_i \sqsupset \mathrm{top}_j$ then $i \sqsupset j$. Now consider three voters x, y, z and a candidate a such that $a \succ_x \mathrm{top}_y$. Suppose that $a \sqsupset \mathrm{top}_y$. Then from the properties of top monotonocity and the fact that $\mathrm{top}_y \sqsupset \mathrm{top}_z$, we have that z has preference ranking $\mathrm{top}_z \succ_z \mathrm{top}_y \succ_z a$. Suppose now that $\mathrm{top}_y \sqsupset a$. But since $\mathrm{top}_x \sqsupset \mathrm{top}_y \sqsupset a$, the fact that $a \succ_x \mathrm{top}_y$ leads to the contradiction with the definition of top monotonicity, which completes the proof. $\square$

Recall that single-crossingness implies single-peakedness for narcissist domains, that is, under the assumption that each candidate is ranked top at least once [Elkind et al., 2014]. Since for narcissist domains a single peaked profile is also single-top-peaked, we get a related result: that single-peakedness is equivalent to single-top-crossingness assuming narcissist preferences.

The class of top monotonic preferences (TM) puts a focus on the top positions in the voters' preference rankings. For example, an election in which the voters unanimously rank a single candidate as their most preferred choice is top-monotonic, independently of how the other candidates are ranked. This suggests that TM offers a combinatorial structure that might be useful in the analysis of single-winner elections, but which might not help to reason about committees. Indeed, below we define a new class which is a natural strengthening of TM. In Section 8.4 we show that the core-stable committees always exist for elections belonging to our newly defined class, and we show that this is not the case for the original class of TM.

Definition 8.11 (Recursive single-top-crossing (r-STC) preferences)**.** Given an election $E = (N, C, b)$, we say that E has *recursive single-top-crossing* preferences if every subelection of E obtained by removing some candidates from E is STC.

Although r-STC is stricter than STC, it still contains both single-peaked and single-crossing preferences. This follows from the fact that both single peaked and single-crossing preferences are top monotonic [Barberà and Moreno, 2011], and that single-peakedness and single-crossingness is preserved under the operation of removing candidates from the election.

8.1.2 Approval-based preferences

In the approval model we first recall the definitions of two classic domain restrictions, the voter-interval and the candidate-interval models [Elkind and Lackner, 2015].

Definition 8.12 (Voter-interval (VI) preferences)**.** Given an election $E = (N, C, b)$, we say that E has *voter-interval* preferences if there exists a linear order $\sqsupset$ over N such that for all voters $i, j, k \in N$ and for each candidate $c \in A_i \cap A_k$, we have that $i \sqsupset j \sqsupset k \implies c \in A_j$. Intuitively, each candidate is approved by a consistent interval of voters.

Definition 8.13 (Candidate-interval (CI) preferences)**.** Given an election $E = (N, C, b)$, we say that E has *candidate-interval* preferences if there exists a linear order $\sqsupset$ over C such that for each voter $i \in N$ and all candidates $a, c \in A_i, b \in C$ we have that $a \sqsupset b \sqsupset c \implies b \in A_i$. Intuitively, each voter approves a consistent interval of candidates.

Below we introduce a new class that generalizes both CI and VI domains. In Section 8.4.2 we will prove that core-stable committees always exist if preferences come from our new restricted domain.

Definition 8.14 (Linearly consistent (LC) preferences)**.** Given an election $E = (N, C, b)$, we say that E has linearly consistent preferences, if there exists a linear order $\sqsupset$ over $N \cup C$ such that for each voters $i, j \in N$ ($i \sqsupset j$) and candidates $a, b \in C$ ($a \sqsupset b$), if $b \in A_i$ and $a \in A_j$, then $a \in A_i$ (as depicted in Figure 23).

Proposition 8.15. *Each VI election is LC. Each CI election is LC.*

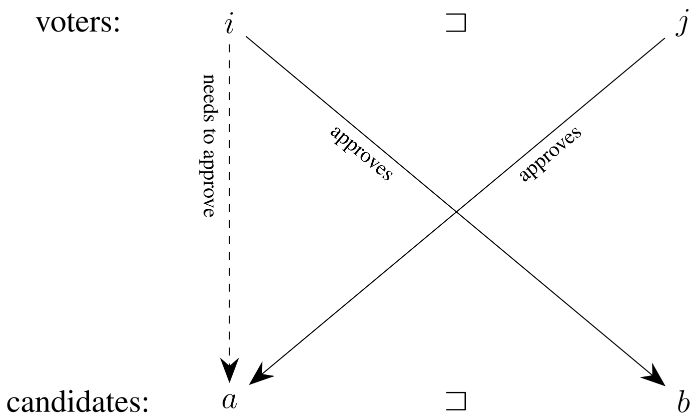


Figure 23: An illustration of the definition of linearly consistent preferences.

Proof. The case of voter-interval preferences. Let $\sqsupset$ be a linear order over N that witnesses that preferences are voter-interval. Let us sort N by this order. For each candidate c, by first_c we denote $\min\{i \in N : c \in A_i\}$. Let us now associate each candidate c to first_c (breaking the tie between c and first_c arbitrarily). If two candidates a, b are associated to the same point, we also break the tie between them arbitrarily. In such a way we obtained an order $\sqsupset$ over $N \cup C$. For simplicity, for each $x, y \in N \cup C$ by $x \sqsupseteq y$ we denote "$x \sqsupset y$ or $x = y$".

Consider two voters, i and j, with $i \sqsupset j$, and two candidates, a and b, with $a \sqsupset b$. Assume i approves b and j approves a. We will prove that i approves a. Since $a \sqsupset b$, by our definition $\mathrm{first}_a \sqsupseteq \mathrm{first}_b$. Since i approves b, $\mathrm{first}_b \sqsupseteq i$, and so $\mathrm{first}_a \sqsupseteq i$. If $i = \mathrm{first}_a$, then i approves a. Otherwise, $\mathrm{first}_a \sqsupset i$. Consequently, first_a, i, and j are three voters, such that $\mathrm{first}_a \sqsupset i \sqsupset j$. Since the preferences are voter-interval we infer that i approves a.

The case of candidate-interval preferences. Let $\sqsupset$ be a linear order on C witnessing the candidate-interval property. Let us sort C by this order. We associate each voter $i \in N$ with $(\min A_i)$, again breaking all the ties arbitrarily. Consider two voters, i and j with $i \sqsupset j$, and two candidates a and b, with $a \sqsupset b$. Further, assume that i approves b and j approves a. Since $i \sqsupset j$, we get that $(\min A_i) \sqsupseteq (\min A_j)$, and since j approves a, we have $(\min A_j) \sqsupseteq a$. Consequently, $(\min A_i) \sqsupseteq a$. If $(\min A_i) = a$, then i approves a. Otherwise, $(\min A_i)$, a and b are three candidates, such that $(\min A_i) \sqsupset a \sqsupset b$. Given that preferences are candidate-interval, and that i approves b, we get that i approves a. $\square$

8.2 The Analysis of Known Voting Rules

To the best of our knowledge, none of the known voting rules is core-stable, even under the more restricted domains than the ones considered in our work. For ordinal preferences, we show that it is the case for 1D-Euclidean elections, that is, elections where both voters and candidates can be placed on the line and voters prefer closer candidates to the further ones. This class is known to be more restrictive class than the intersection of single-peaked and single-crossing elections [Elkind et al., 2020]. First, we show this for the ordinal variant of Equal Shares presented in Chapter 6.

Example 8.16 (Equal Shares for ordinal preferences fails core-stability)**.** Let $b = 2$. Voters' preferences are the following:

$$\begin{aligned}
1&\colon c \succ b \succ a \succ d \succ e \succ f \succ g \succ h\\
2&\colon e \succ d \succ c \succ b \succ a \succ f \succ g \succ h\\
3&\colon f \succ e \succ g \succ h \succ d \succ c \succ b \succ a\\
4&\colon h \succ g \succ f \succ e \succ d \succ c \succ b \succ a
\end{aligned}$$

This election is 1D-Euclidean as presented in Figure 24.

Here in the first round nothing is selected (no project can be afforded if only top-ranked choices are taken into account). In the second round, the rule selects project e, afforded by voters 2 and 3. After that, these two voters lose all their money. The next round in which anything is purchased, is the 5th round, in which candidate d is bought by 1 and 4.

As a result, the committee $\{e, d\}$ is elected by Equal Shares, but group $S = \{3, 4\}$ and $T = \{f\}$ witness the violation of core-stability. ⌟

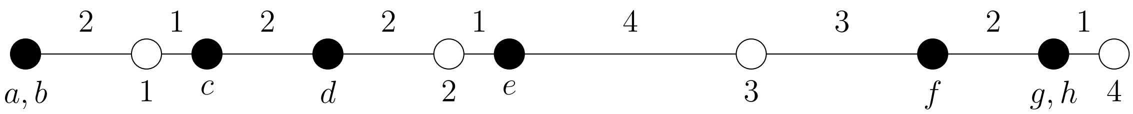


Figure 24: An illustration of Example 8.16. White and black points mean the positions of respectively the voters and the candidates.

The same example also works for the ordinal version of Greedy Cohesive Rule, presented in Section 7.2.2—note that there are no weakly cohesive groups of voters, hence GCR elects nothing and, for example, it can be further completed with Equal Shares (as proposed in Theorem 7.10) so that the core is violated. Next, let us consider two other archetypal proportional rules, the Monroe rule and STV.

Definition 8.17 (The Monroe Rule)**.** Consider an election E with ordinal preferences and assume that n/b is integral. For a committee $T \subseteq C$, a *balanced matching* is a collection of subsets of voters $\{N_c\}_{c \in T}$ such that for every $c \in T$, $|N_c| = n/b$. The *value* of a matching is the sum $\sum_{c \in T} \sum_{i \in N_c} \mathrm{pos}_i(\{c\})$. The matching $\{N_c\}_{c \in T}$ is minimal, if it has the minimal value among all matchings for T. The Monroe Rule returns the committee W minimizing the value of the minimal balanced matching.

Definition 8.18 (Single Transferable Vote (STV))**.** Consider an election E with ordinal preferences. STV proceeds sequentially: at each round we elect a candidate that is ranked top by at least $n/b{+}1 + 1$ voters and remove any $n/b{+}1 + 1$ of these voters from the election. If there are no such candidates, we remove from the election a candidate ranked top by the least number of voters.

Both the Monroe Rule and STV are not core-stable even for 1D-Euclidean elections, as shown in Example 8.19 and Example 8.20, respectively.

Example 8.19 (The Monroe Rule fails core-stability)**.** Let $b = 2$. Voters' preferences are the following:

$$\begin{aligned} 1&\colon b \succ a \succ c \succ d \succ e \\ 2&\colon c \succ b \succ d \succ a \succ e \\ 3&\colon c \succ d \succ b \succ e \succ a \\ 4&\colon d \succ e \succ c \succ b \succ a \end{aligned}$$

This election is 1D-Euclidean as presented in Figure 25.

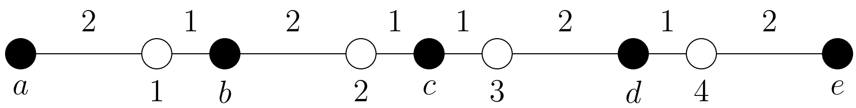


Figure 25: An illustration of Example 8.19. White and black points mean the positions of respectively the voters and the candidates.

Here the committee $\{b, d\}$ is elected by Monroe, but group S consisting of middle voters $\{2, 3\}$ and $T = \{c\}$ witness the violation of core-stability. ⌟

Example 8.20 (STV fails core-stability)**.** Let $n = 60$, $b = 2$. The value of the STV quota is $n/b+1 + 1 = 21$. Voters' preferences are divided into 5 groups:

$$\begin{aligned} G_1 \text{ (18 voters)}&\colon a \succ b \succ c \succ d \succ e \\ G_2 \text{ (7 voters)}&\colon b \succ c \succ d \succ e \succ a \\ G_3 \text{ (5 voters)}&\colon c \succ d \succ e \succ b \succ a \\ G_4 \text{ (16 voters)}&\colon d \succ e \succ c \succ b \succ a \\ G_5 \text{ (14 voters)}&\colon e \succ d \succ c \succ b \succ a \end{aligned}$$

This election is 1D-Euclidean as presented in Figure 26.

Here candidate c is eliminated at the first round and all votes for her are transferred to d. Second, candidate d is elected (in the second round she gains exactly 21 votes) and the votes from groups 3 and 4 are removed. Third, candidate b is eliminated and all votes for her are transferred to e. Fourth, candidate e is elected (gaining in the final round exactly 21 votes) and the committee $\{d, e\}$ is returned. However, 30 voters from the three first groups and candidate c witness the violation of core-stability. ⌟

In the case of approval-based preferences, the fact that GCR is not core-stable follows from the fact it violates laminar proportionality—in Example 7.11 the preferences of the voters clearly belong to the intersection of VI and CI.

Now let us show the same fact for Method of Equal Shares.

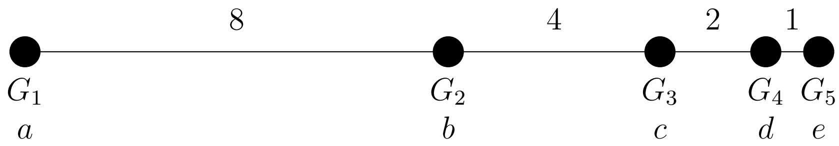


Figure 26: An illustration of Example 8.20. Black points mean the positions of both candidates and groups of voters (ties can be broken arbitrarily).

Example 8.21 (Equal Shares for approval-based preferences fails core-stability)**.** Let $n = 42$, $b = 14$. Voters' preferences are divided into the following groups:

$$\begin{aligned}
&G_1 \text{ (1 voter)}: && \{c_1, c_2, c_3, x_1, x_2\} \\
&G_2 \text{ (8 voters)}: && \{c_1, c_2, c_3, x_1, x_2, a_1, a_2, a_3, a_4\} \\
&G_3 \text{ (12 voters)}: && \{c_1, c_2, c_3, a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4, e_1, e_2\} \\
&G_4 \text{ (12 voters)}: && \{d_1, d_2, d_3, b_1, b_2, b_3, b_4, a_1, a_2, a_3, a_4, e_1, e_2\} \\
&G_5 \text{ (8 voters)}: && \{d_1, d_2, d_3, y_1, y_2, b_1, b_2, b_3, b_4\} \\
&G_6 \text{ (1 voter)}: && \{d_1, d_2, d_3, y_1, y_2\}
\end{aligned}$$

Let us denote the groups of candidates by $X = \{x_1, x_2\}$, $Y = \{y_1, y_2\}$, $A = \{a_1, a_2, a_3, a_4\}$, $B = \{b_1, b_2, b_3, b_4\}$, $C = \{c_1, c_2, c_3\}$, $D = \{d_1, d_2, d_3\}$ and $E = \{e_1, e_2\}$.

Assuming

$$G_1 \sqsupset G_2 \sqsupset \ldots \sqsupset G_6$$

and

$$X \sqsupset C \sqsupset A \sqsupset E \sqsupset B \sqsupset D \sqsupset Y$$

(voters and candidates within each group can be ordered arbitrarily), it is clear that the election is both VI and CI.

At the beginning each voter has 1 dollar and the price for candidates is $p = n/b = 3$. First Equal Shares elects candidates from $A \cup B$; for each of them 32 out of 40 middle voters pay (each of them pays $3/32$). Second, Equal Shares elects candidates from E, and the middle 24 voters run out of money; indeed, each of them pays $8 \cdot 3/32 + 2 \cdot 3/24 = 1$ for the so far elected candidates. Next Equal Shares elects candidates from $X \cup Y$. We have elected the committee $A \cup B \cup E \cup X \cup Y$ which is not even Pareto-optimal as the committee $A \cup B \cup C \cup D$ is better for every voter. Thus, in particular, the elected committee does not belong to the core. ⌟

Besides, we show that it is also the case for Proportional Approval Voting (PAV) (recall Definition 4.12).

Example 8.22 (PAV fails core-stability)**.** Let $n = 3$, $b = 8$. Voters' preferences are the following:

$$\begin{aligned}
1&\colon \{b_1, b_2, b_3, b_4, a\}\\
2&\colon \{b_1, b_2, b_3, b_4, c\}\\
3&\colon \{d_1, d_2, d_3, d_4\}
\end{aligned}$$

Assuming

$$1 \sqsupset 2 \sqsupset 3$$

and

$$a \sqsupset b_1 \sqsupset \ldots \sqsupset b_4 \sqsupset c \sqsupset d_1 \sqsupset \ldots \sqsupset d_4,$$

it is clear that the election is both VI and CI.

Here PAV elects candidates $\{b_1, \ldots, b_4, d_1, \ldots, d_4\}$. However, this committee does not belong to the core, which is witnessed by the groups $S = \{1, 2\}$ and $T = \{a, b_1, \ldots, b_4, c\}$. ⌟

As we can see, no well-established voting rule is core-stable even for very restricted preference domains. In the next section, we will present the algorithm that has this property.

8.3 The Description of the Main Algorithm

In this section we describe our main algorithm, called *Quantile Rule*. Its name refers to the fact that it chooses b candidates that are top-preferred by voters 1st, $(n/b + 1)$st, . . ., $(b - 1) \cdot n/b + 1$st if these candidates are unique for each voter and all different. However, in order to handle the case where these candidates repeat, its definition is slightly more complex.

The rule works in polynomial time, assuming that we are given the linear order $\sqsupset$ over $N \cup C$, existence of which is ensured by the definitions of these preference classes. For approval-based preferences, such an order can be found in polynomial time for candidate-interval and voter-interval domains [Elkind et al., 2017]. It is not known whether it is the case for general LC preferences—hence, for this class we show only that core-stability is satisfiable. However, LC is mainly a technical domain, allowing us to present a coherent algorithm for both voter-interval and candidate-interval preferences. In case of r-STC preferences, the linear order witnessing this class is the same as the one witnessing top monotonicity which can be found in polynomial time [Magiera and Faliszewski, 2019].

Since Quantile Rule works both for ordinal and approval-based preferences, we need to start by introducing a common framework capturing both types of preferences (namely weak ordinal preferences, briefly mentioned before in the context of Definition 8.6). Next, we will present the main idea of *fractional committees* that the algorithm is based on. Finally, we will present the formal definition of Quantile Rule and illustrative examples.

8.3.1 Weak Ordinal Preferences

From now, we assume that each voter $i \in N$ submits a weak ranking $\succsim_i$ over the candidates—for each $i \in N$ and $a, b \in C$, we say that voter i weakly prefers candidate a over candidate b if $a \succsim_i b$. We set $a \sim_i b$ if $a \succsim_i b$ and $b \succsim_i a$, and we write $a \succ_i b$ if $a \succsim_i b$ and $a \not\sim_i b$. For a voter $i \in N$ and $j \in [m]$, by $\mathrm{pos}_i(j)$ we denote the equivalence class of candidates ranked at the jth position by voter i. Formally, a candidate c belongs to $\mathrm{pos}_i(j)$ if there are $(j-1)$ candidates $a_1, \dots, a_{j-1}$ such that $a_1 \succ_i a_2 \succ_i \dots \succ_i a_{j-1} \succ_i c$ and if there exist no j candidates $a_1, \dots, a_j$ for which $a_1 \succ_i a_2 \succ_i \dots \succ_i a_j \succ_i c$. By d_i we denote the number of the nonempty positions in the i-th voter's preference list. For each $j \in [d_i]$, by $\mathrm{pos}_i([j])$ we denote $\bigcup_{q \le j} \mathrm{pos}_i(q)$. For each $i \in N$, by top_i and bot_i we denote the sets of candidates ranked respectively at the highest and the lowest position (note that $\mathrm{top}_i = \mathrm{pos}_i(1)$ and $\mathrm{bot}_i = \mathrm{pos}_i(d_i)$).

Now, both types of voters' preferences that we focus on, can be defined as follows:

Approval-based preferences. The preferences are *approval-based*, if for each candidate $c \in C$ and each voter $i \in N$ either $c \in \mathrm{top}_i$ or $c \in \mathrm{bot}_i$. We say that i *approves* c if $c \in \mathrm{top}_i$.

Ordinal preferences. The preferences are (strictly) *ordinal*, if for all $a, b \in C$ ($a \neq b$) and each $i \in N$ it holds that $a \not\sim_i b$.

The extension of the preferences over candidates to the preferences over committees is done is the following way: voters compare committees lexicographically with respect to the numbers of candidates in specific equivalence classes. It generalizes both the lexicographical extension for ordinal preferences and additive extension for approval-based ones. Formally:

$$W \rhd_i T \iff \exists \sigma \in [d_i].\ |\mathrm{pos}_i(\sigma) \cap W| > |\mathrm{pos}_i(\sigma) \cap T| \\ \text{and } \forall \varrho < \sigma.\ |\mathrm{pos}_i(\varrho) \cap W| = |\mathrm{pos}_i(\varrho) \cap T|. \tag{8.1}$$

An alternative preference extension is considered in Section 8.5.

8.3.2 Fractional Committees

Further, we extend the definition of a committee to the continuous model as follows: a *fractional committee* is a function $p\colon C \to [0,1]$ that assigns to each candidate from $c \in C$ a value $p(c)$ such that $0 \le p(c) \le 1$; intuitively $p(c)$ can be thought of as the probability that candidate c is a member of the selected committee. We extend this notation to sets, defining $p(T) = \sum_{c \in T} p(c)$ for each $T \subseteq C$. The value of $p(C)$ is the *size* of the fractional committee. If for a candidate c it holds that $p(c) = 1$, then we say that c is *elected*, otherwise she is *unelected*. If for an unelected candidate c it holds that $p(c) > 0$, then c is *partially elected*. If there are no partially elected candidates in p, then we say that p is a *discrete committee* (or simply a committee) and associate it with the set $\{c \in C\colon p(c) = 1\}$.

The notion of a fractional committee is similar to several probabilistic concepts considered in the literature. For instance, in probabilistic social choice (see the book chapter [Brandt, 2017]) we also assign fractional values to candidates. The main difference is that in probabilistic social choice, the whole value that we want to divide among the candidates can be assigned to fewer than b candidates; in particular it is feasible to set $p(c) = b$ for one candidate and $p(c') = 0$ for all c', $c' \neq c$. Thus, intuitively, in probabilistic social choice each candidate is divisible and appears in an unlimited quantity. Viewed from this perspective, probabilistic social choice extends the discrete model of approval-based apportionment [Brill et al., 2022]. Several works have considered axioms of proportionality for probabilistic social choice [Aziz et al., 2019, Fain et al., 2016], yet unfortunately their results do not apply to fractional committees.

Another concept related to fractional committees is where we assign probabilities to committees instead of individual candidates. The notions of proportionality in this setting have been considered, for example, by Cheng et al. [2019]. It is worth noting that fractional committees can induce probability distributions over committees, for example, by applying sampling techniques, such as dependent rounding [Srinivasan, 2001], that ensure we always select b candidates. Yet, there is no one-to-one equivalence between the two settings, thus the results of Cheng et al. [2019] do not apply to fractional committees.

Now, note that Definition 8.1 naturally extends to fractional committees:

Definition 8.23 (Core-stability (for fractional committees))**.** Given an election $E = (N, C, b)$, we say that a fractional committee p is core-stable, if for each $S \subseteq N$ and each fractional committee p' with $p'(C) \leq b \cdot {}^{|S|}/_{n}$, there exists a voter $i \in S$ such that i weakly prefers p over p'.

So does the lexicographical extension of weak rankings (equation (8.1)).

$$p \rhd_i p' \iff \exists \sigma \in [d_i].\ p(\mathrm{pos}_i(\sigma)) > p'(\mathrm{pos}_i(\sigma)) \text{ and } \forall \varrho < \sigma.\ p(\mathrm{pos}_i(\varrho)) = p'(\mathrm{pos}_i(\varrho)).$$

8.3.3 Quantile Rule

We can now formally describe the algorithm. Hereinafter we assume that the fraction ${}^{n}/_{b}$ is integral—if it is not the case, we multiply each voter b times (note that if a committee is in the core of the modified election, it is also in the core of the original one).

Intuitively, it consists of two phases: first we construct a fractional committee and then we discretize it. The first part of the algorithm (Phase 1) is the following: imagine that each voter has an equal probability portion ${}^{b}/_{n}$ to distribute, and that we want to choose one candidate (her *representative*) who gets this portion. Initially, the fractional committee p is empty. We iterate over the set of voters, sorted according to the relation $\sqsupset$. Let us denote by P_i the set of unelected candidates at the moment of considering voter $i \in N$. The representative of i is defined as a candidate $r_i \in P_i$ such that for each $c \in P_i$ it holds that either $r_i \succ_i c$ or that $r_i \sim_i c$ and $r_i \sqsupset c$. Next, $p(r_i)$ is increased by ${}^{b}/_{n}$. Note that, as ${}^{n}/_{b}$ is integral, the election probability of each candidate does not exceed 1.

In Section 8.4.1 we prove that after this phase the obtained fractional committee p is in the core for all strict elections and all LC approval elections. Denote by W_1 the set of candidates c

such that $p(c) = 1$. Before the second phase of the algorithm, remove candidates from W_1 from the election together with the voters who are represented by them, obtaining a smaller election E_2. By b' we denote $b - |W_1|$ (remaining seats in the committee) and by n' we denote $n - |W_1| \cdot {}^{n}/_{b}$ (remaining voters). Renumerate the voters so that they are numbers from $[n']$ (and in case of r-STC elections, resort them so that E_2 is still r-STC). Note that if $b' \neq 0$ then by definition:

$$\frac{n'}{b'} = \frac{n - |W_1| \cdot {}^{n}/_{b}}{b - |W_1|} = \frac{n}{b} \cdot \frac{b - |W_1|}{b - |W_1|} = \frac{n}{b}.$$

Phase 2 is simple: for each $q \in [b']$ denote by m_q the voter $(q-1) \cdot {}^{n}/_{b} + 1$. Further we will refer to these voters as *quantile voters*. Then elect committee $W_2 = \{r_{m_q} : q \in [b']\}$.

Finally, we return the committee $W = W_1 \cup W_2$. This algorithm is presented in Algorithm 3.

Algorithm 3: Implementation of Quantile Rule

`If` ${}^{n}/_{b}$ `is not integral, multiply each voter` b `times.`
`Each time we iterate over sets` N `and` C`, we assume they are sorted according to the order` $\sqsupset$ `witnessing the membership of the election to a specific class.`
`Phase 1:`
for $c \in C$ **do**
    $p[c] \leftarrow 0$
for $i \in N$ **do**
    $P_i \leftarrow \{c \in C : p[c] < 1\}$
    $r_i \leftarrow$ the first $c \in P_i$ such that $p[r_i] < 1$
    for $c \in C$ **do**
        if $c \in P_i$ *and* $c \succ_i r_i$ **then**
            $r_i \leftarrow c$
    $p[r_i] \leftarrow p[r_i] + {}^{b}/_{n}$
$W_1 \leftarrow \{c : p[c] = 1\}$
$C \leftarrow C \setminus W_1$
$N \leftarrow N \setminus \{i : r_i \in W_1\}$
`Phase 2:`
`Resort the voters if needed, so that` $\sqsupset$ `still witness the membership of the election to a specific class.`
$W_2 \leftarrow \emptyset$
$t \leftarrow 0$
for $i \in N$ **do**
    if $t \cdot b \bmod n = 0$ **then**
        $W_2 \leftarrow W_2 \cup \{r_i\}$
    $t \leftarrow t + 1$
return $W_1 \cup W_2$

For better understanding of the rule, let us now analyze the behavior of Quantile Rule on the five examples presented in Section 8.2 (we assume that the linear orders over voters and candidates are the same as in the descriptions/illustrations of these examples).

Example 8.16 Here, $b = 2$ and every candidate is ranked top by only one out of 4 voters, hence after Phase 1 we have that $p(c) = 1/4$ for each $c \in C$ and $W_1 = \emptyset$. After Phase 2, candidates c and f are elected.

Example 8.19 Here, after Phase 1 candidate c is elected. For Phase 2, we remove c and two middle voters from the election and, since no resorting is needed, candidate b is additionally elected.

Example 8.20 Here, after Phase 1 we have that $p(a) = 3/5$, $p(b) = 7/30$, $p(c) = 1/6$, $p(d) = 8/15$ and $p(e) = 7/15$. Hence, $W_1 = \emptyset$. After Phase 2, the representatives of the first voter (one from the group G_1) and 31st voter (one from the group G_4) is elected, namely a and d.

Example 8.21 Let us assume here that the candidates each group are sorted by indexes. Consider Phase 1. The 9 voters from $G_1 \cup G_2$ make candidates $\{x_1, x_2, c_1\}$ elected. The next 12 voters from G_3 make candidates $\{c_2, c_3, a_1, a_2\}$ elected. The next 12 voters from G_4 make candidates $\{a_3, a_4, e_1, e_2\}$ elected. The next 6 voters from G_5 make candidates $\{b_1, b_2\}$ elected. Finally, b_3 becomes the representative of the last two voters from G_5, and d_1 becomes the representative of the last voters from G_6—yet these two candidates are only partially elected after Phase 1. After Phase 2, b_3 is elected and the committee $X \cup C \cup A \cup E \cup \{b_1, b_2, b_3\}$ is returned by Quantile Rule.

Example 8.22 Here, n/b is not integral, therefore each voter is replaced by 8 voters with the same preferences. After Phase 1, candidates $\{a, b_1, b_2, b_3, b_4, d_1, d_2\}$ are elected and candidates c (the representative of one voter from the middle group) and d_3 (the representative of two voters from the last group) are partially elected. After Phase 2, c is elected and the committee $\{a, b_1, b_2, b_3, b_4, d_1, d_2, c\}$ is returned.

It is straightforward to check that for all these examples, the committees returned by Quantile Rule are core-stable. Now it is time to prove that it is true in the general case.

8.4 The Analysis of Quantile Rule

In this section we prove that Quantile Rule is core-stable for all the restricted domains mentioned in Section 8.1. We divide its content into three parts: first, we prove that after Phase 1, the fractional committee p is in the core for LC approval elections and for all elections with ordinal preferences. The proof is the same for those two models; we will refer only to the following property:

Definition 8.24. Given an election $E = (N, C, b)$, we say that E is *well-ordered*, if there exists a linear order $\sqsupset$ over $N \cup C$ such that for each voters $i, j \in N$ $(i \sqsupset j)$ and candidates $a, b \in C$ $(a \sqsupset b)$, if $a \sim_j b$ and $a, b \notin \mathrm{bot}_j$, then $a \succsim_i b$.

It is clear that every strict election is well-ordered for every order $\sqsupset$ (the premise is never satisfied). For approval elections this definition is a weakening of Definition 8.14 (because for approval elections $a, b \notin \mathrm{bot}_j \implies a \in A_j$ and the condition $a \succsim_i b$ boils down to "if i approves b, then she approves also a"), hence every LC election is well-ordered.

Second, we prove that after Phase 2, the committee W elected by Quantile Rule is in the core for approval LC preferences. Third, we prove the same fact for ordinal r-STC preferences. Contrary to the first part, here the proofs that the rule satisfies core-stability for both preference types differ significantly.

8.4.1 Core-Stability for Fractional Committees

For convenience, for $i \in N$ by p_i we denote the fractional committee p after considering voter i. Let $\sigma_i \in [m]$ be the number such that $r_i \in \mathrm{pos}_i(\sigma_i)$. From how Phase 1 works, we have that for every voter $i \in N$ and a candidate $c \in \mathrm{pos}_i([\sigma_i - 1])$ it holds that $p_{i-1}(c) = 1$ (and also $p_j(c) = 1$ for every $j \geq i$).

Before proving that Phase 1 returns committees belonging to the core, let us start from the following observation.

Observation 8.25. For each $i \in N$ and $c \in C$, there exists $q \in [{}^n/_b]$ such that $p_i(c) = q \cdot {}^b/_n$. In particular, q is the number of voters for whom c is a representative.

Theorem 8.26. *Each fractional committee elected by Phase 1 belongs to the core for well-ordered elections.*

Proof. We will prove the following invariant: for each $i \in N$, p_i satisfies the condition of the fractional core (see Definition 8.23) with the additional restriction that $S \subseteq [i]$. We will prove the invariant by induction.

For the first voter the invariant is clearly true. Assume, there exists $i \in N$ satisfying the invariant. We will prove that the invariant holds also for voter $(i + 1)$.

For the sake of contradiction suppose that there exists a group $S \subseteq [i + 1]$ and a fractional committee p'_{i+1} such that for each $v \in S$ we have that v prefers lexicographically p'_{i+1} to p_{i+1}.

First, note that if $(i+1) \notin S$, then the invariant does not hold also for i, a contradiction. This is the case becasue the election probability of no candidate is decreased during a loop iteration. Hence, $(i + 1) \in S$.

By the definition of Phase 1, we have that for each $\varrho < \sigma_{i+1}$ and $c \in \mathrm{pos}_{i+1}(\varrho)$ it holds that $p_{i+1}(c) = 1$. From that, in particular we have the following equation:

$$\forall \varrho < \sigma_{i+1}.\ p'_{i+1}(\mathrm{pos}_{i+1}(\varrho)) \leq |\mathrm{pos}_{i+1}(\varrho)| = p_{i+1}(\mathrm{pos}_{i+1}(\varrho))$$

Hence, as $(i + 1)$ prefers lexicographically p' to p:

$$\forall \varrho < \sigma_{i+1}.\ p'_{i+1}(\mathrm{pos}_{i+1}(\varrho)) = p_{i+1}(\mathrm{pos}_{i+1}(\varrho)) \tag{8.2}$$

It also needs to hold that:

$$p'_{i+1}(\mathrm{pos}_{i+1}([\sigma_{i+1}])) > p_{i+1}(\mathrm{pos}_{i+1}([\sigma_{i+1}])) \tag{8.3}$$

We can conclude that $\sigma_{i+1} < d_{i+1}$, as otherwise voter $(i+1)$ could not prefer p'_{i+1} over p_{i+1}. Consequently:

$$r_{i+1} \notin \mathrm{bot}_{i+1} \tag{8.4}$$

Suppose that $p'_{i+1}(r_{i+1}) = 0$. From (8.3) and the fact that for all $c \in \mathrm{pos}_{i+1}(\sigma_{i+1})$ with $c \sqsupset r_{i+1}$ we have $p(c) = 1$, we infer that there exists $a \in \mathrm{pos}_{i+1}(\sigma_{i+1})$ such that $r_{i+1} \sqsupset a$ and $p'_{i+1}(a) > 0$. From Observation 8.25 we have that $p'_{i+1}(a) \geq {}^{b}/_{n}$. Now we modify p'_{i+1} by moving the fraction of ${}^{b}/_{n}$ from a to r_{i+1}. By Definition 8.24 and (8.4) we have that for every $v \in S$ (naturally, $v \sqsupset (i+1)$) it holds that $r_{i+1} \succsim_v a$. Thus, after the change p'_{i+1} still witnesses core violation for S.

Now consider a fractional committee p'_i obtained from p'_{i+1} by decreasing the probability portion of r_{i+1} by ${}^{b}/_{n}$. We will show that p'_i together with $S \setminus \{(i+1)\}$ witness the core violation for p_i. Indeed, the election probability of no candidate except r_{i+1} changed, and the election probability of r_{i+1} changed in the same way: in p_{i+1} and p'_{i+1} it is higher by ${}^{b}/_{n}$ than in p_i and p'_i, respectively. Hence, if for a voter $v \in S$ it holds that $p'_{i+1} \rhd_v p_{i+1}$, then also $p'_i \rhd_v p_i$. Besides, we have that $p'_i(C) \leq b \cdot {}^{|S-1|}/_{n}$, so we obtain a contradiction with our inductive assumption. $\square$

8.4.2 Discrete Core-Stability for Approval LC Elections

Let us continue with the following observation.

Observation 8.27. Consider an approval LC election E and two voters i, j who were not removed from the election after the first phase, such that $i \sqsupset j$. Then either $r_i = r_j$ or $r_i \sqsupset r_j$.

Proof. Towards a contradiction assume that $r_j \sqsupset r_i$. From LC we have that i approves r_j and r_j should be i's representative. $\square$

Next, we prove that Quantile Rule elects exactly b candidates.

Lemma 8.28. *Quantile Rule for an approval LC election E elects exactly b candidates.*

Proof. We will show that Phase 2 elects exactly b' candidates. Suppose for the sake of contradiction that there are two quantile voters i, j in E_2 such that $r_i = r_j$. Without loss of generality assume $i \sqsupset j$. Consider now any voter v between these quantile voters. If $r_v \sqsupset r_i$ then from the definition of LC, i approves r_v, and so r_v should be selected as i's representative, a contradiction. If $r_j = r_i \sqsupset r_v$, then from the definition of LC, v approves r_j, and so r_j should be selected as v's representative, a contradiction. Hence, $r_v = r_i$. But then we have that after running Phase 1, r_i was a representative for at least ${}^{n}/_{b}$ voters and was not elected, a contradiction. $\square$

Note that every LC election remains LC for the same order $\sqsupset$ after removing any number of voters and candidates.

Finally, we are ready to prove the main technical lemma together with the main result.

Lemma 8.29. *For each voter $i \in N$ it holds that $|W \cap \mathrm{top}_i| + 1 > p(\mathrm{top}_i)$.*

Proof. Consider a voter $i \in N$. Define part_i as $p(\mathrm{top}_i) - |W_1 \cap \mathrm{top}_i|$. As W_1 contains all candidates c such that $p(c) = 1$, then part_i is intuitively the joint sum of election probabilities of partially elected candidates in top_i. From Observation 8.25 we have that:

$$\mathrm{part}_i = q \cdot {}^{b}/_{n} \tag{8.5}$$

where q is the number of voters for whom a candidate from $\mathrm{top}_i \setminus W_1$ is a representative. Naturally, such voters could not be removed from the election after the execution of Phase 1.

We will prove that $\mathrm{part}_i < |W_2 \cap \mathrm{top}_i| + 1$. From the fact that $W = W_1 \cup W_2$ and $W_1 \cap W_2 = \emptyset$, it will imply the desired statement. We will now focus on upper-bounding q from (8.5).

Consider three voters x, y, z such that $x \sqsupset y \sqsupset z$ and $r_x, r_z \in \mathrm{top}_i$. We will prove that then also $r_y \in \mathrm{top}_i$. Indeed, from Observation 8.27 we have that either $r_y \in \{r_x, r_z\}$ (and the statement is true) or $r_x \sqsupset r_y \sqsupset r_z$. First, consider the case, when $y \sqsupset i$. Since i approves r_x by LC applied to voters y, i and candidates r_x and r_y, we get that also y approves r_x, a contradiction with Observation 8.27. Second, we look at the case when $i \sqsupset y$. From LC applied to y, i and candidates r_y and r_z and by the fact that i approves r_z we get that i also approves r_y, which is what we wanted to prove.

Hence, these q voters from (8.5) need to form a consistent interval among all non-removed voters. Besides, we know that there is no more than $|W_2 \cap \mathrm{top}_i|$ quantile voters inside this interval and that between each two quantile voters there is ${}^{n}/_{b} - 1$ non-removed voters. Hence:

$$q \leq (|W_2 \cap \mathrm{top}_i| + 1) \cdot ({}^{n}/_{b} - 1) + |W_2 \cap \mathrm{top}_i| = (|W_2 \cap \mathrm{top}_i| + 1) \cdot {}^{n}/_{b} - 1$$

and:

$$\mathrm{part}_i = q \cdot {}^{b}/_{n} < (|W_2 \cap \mathrm{top}_i| + 1) \cdot {}^{n}/_{b} \cdot {}^{b}/_{n} = |W_2 \cap \mathrm{top}_i| + 1$$

which completes the proof. □

Theorem 8.30. *For approval LC elections, Quantile Rule elects committees from the core.*

Proof. We know that fractional committee p elected by Phase 1 belongs to the core. Suppose now that W is not in the core. Hence, there exists a nonempty set $S \subseteq N$ and a committee T of size $|S| \cdot {}^{b}/_{n}$ such that $|W \cap \mathrm{top}_i| < |T \cap \mathrm{top}_i|$ for each $i \in S$—alternatively, $|W \cap \mathrm{top}_i| + 1 \leq |T \cap \mathrm{top}_i|$.

From Lemma 8.29 we know that for each voter $i \in S$ we have $p(\mathrm{top}_i) < |W \cap \mathrm{top}_i| + 1 \leq |T \cap \mathrm{top}_i|$. Let us define a fractional committee p' such that $p'(c) = 1$ for $c \in T$ and $p'(c) = 0$ otherwise. Hence, S and p' witness also the violation of the core for p, which is contradictory with Theorem 8.26. □

8.4.3 Discrete Core-Stability for Ordinal r-STC Elections

We will now assume that E is a strict r-STC election. Similarly as in case of approval preferences, we start by proving that the Quantile Rule elects exactly b candidates.

Lemma 8.31. *Quantile Rule for strict r-STC election E elects exactly b candidates.*

Proof. We need to show that Phase 2 elects exactly b' candidates. Suppose for the sake of contradiction that there are two quantile voters i, j in E_2 such that $r_i = r_j$. From STC it follows that $r_v = r_i$. But this means that after running Phase 1, r_i was a representative for at least n/b voters and was not elected, a contradiction. $\square$

Now we prove a general statement about the application of Phase 2 to STC elections.

Lemma 8.32. *Consider an STC election $E = (N, C, b)$ and apply Phase 2 to E to obtain the committee W. If $|W| = b$, then W is in the core.*

Proof. Towards a contradiction suppose that the statement of the lemma is not true. Without loss of generality, assume that E is an election with the smallest b among those for which the statement of the lemma does not hold. Let S and T be subsets of voters and candidates, respectively, that witness that the committee returned by Phase 2 does not belong to the core.

Observe that there are at least two candidates from W that do not belong to T. Indeed, if there were only one such candidate, we would have that $|T| = |W|$ (as $T \setminus W$ is nonempty) and $|S| = n$. In particular, in such a case all quantile voters would belong to S. Consequently, the most preferred candidates of the quantile voters would belong to T, hence $W \subseteq T$, a contradiction.

Let us fix a candidate $a \in W \setminus T$ that is elected by the greatest quantile voter $(i \cdot n/b + 1)$. In particular, $i \neq 0$. For a candidate $b \in T$ by $S_b \subseteq S$ we denote the subset of voters in S preferring b to a. Since E is single-top-crossing, it holds that either $S_b \subseteq [i \cdot n/b]$ or $S_b \subseteq N \setminus [i \cdot n/b]$.

Now we split E into two smaller elections $E_{\text{low}} = ([i \cdot n/b], C, i)$ and $E_{\text{grt}} = (N \setminus [i \cdot n/b], C, b-i)$. By W_{low} and W_{grt} we denote the committees elected by Phase 2 for E_{low} and E_{grt}, respectively. Observe that $W_{\text{low}} \sqcup W_{\text{grt}} = W$.

Let us also split S and T into two parts, as follows:

$$S_{\text{low}} = S \cap [i \cdot n/b], \qquad S_{\text{grt}} = S \cap (N \setminus [i \cdot n/b]),$$
$$T_{\text{low}} = \{c \in T \colon S_c \subseteq [i \cdot n/b]\}, \qquad T_{\text{grt}} = \{c \in T \colon S_c \subseteq N \setminus [i \cdot n/b]\}.$$

Note that $S_{\text{low}} \cup S_{\text{grt}} = S$ and $T_{\text{low}} \cup T_{\text{grt}} = T$. Hence, if we had that both $|T_{\text{low}}| > |S_{\text{low}}| \cdot n/b$ and $|T_{\text{grt}}| > |S_{\text{grt}}| \cdot n/b = (|S| - |S_{\text{low}}|) \cdot n/b$, then we would have also $|T| > |S| \cdot n/b$, a contradiction. Hence, for at least one of the pairs $(S_{\text{low}}, T_{\text{low}}), (S_{\text{grt}}, T_{\text{grt}})$ the opposite inequality holds. Without the loss of generality, assume that $|T_{\text{low}}| \leq |S_{\text{low}}| \cdot n/b$.

We claim that the pair $(S_{\text{low}}, T_{\text{low}})$ witnesses the core violation for E_{low} and committee W_{low}.

Consider a voter $j \in S_{\text{low}}$. We know that there exists a candidate $c \in T \setminus W$ such that $c \succ_j W \setminus T$. First observe that W_{low} and T_{grt} are disjoint—indeed, for every candidate $b \in T_{\text{grt}}$ we have that $S_b \subseteq N \setminus [i \cdot n/b]$. As a result, there is no quantile voter in $[i \cdot n/b]$ who prefers b to a, hence $b \notin W_{\text{low}}$. From this fact we conclude that $W_{\text{low}} \setminus T_{\text{low}} = W_{\text{low}} \setminus T \subseteq W \setminus T$. Consequently, $c \succ_j W_{\text{low}} \setminus T_{\text{low}}$.

Further, observe that $c \in T_{\text{low}}$. Indeed, voter j prefers c to $W \setminus T$, thus in particular j prefers c to a. Consequently, $j \in S_c$, and thus $S_c \subseteq S_{\text{low}}$, from which we get that $c \in T_{\text{low}}$. Since $c \in T_{\text{low}}$ and $c \succ_j W_{\text{low}} \setminus T_{\text{low}}$, we get that j prefers lexicographically T_{low} to W_{low}.

Finally, we obtain that if the core was violated for E, it also needs to be violated for E_{low}, which is contradictory to our assumption that E minimizes the value of b. $\square$

Corollary 8.33. *In Phase 2, the committee W_2 is in the core for election E_2.*

Now we are ready to prove the main theorem in this subsection:

Theorem 8.34. *Committees elected by Quantile Rule are core-stable.*

Proof. For the sake of contradiction suppose that the statement of the theorem is not true. Then there exist a set $S \subseteq N$ and a set $T \subseteq C$ witnessing the violation of the condition of the core. For every candidate $c \in C$, by $R(c)$ we denote set $\{i \in N \colon r_i = c\}$. Note that for a candidate $c \in W_1$ and a voter $i \in S$ such that $i \in R(c)$, we have $c \in T$. Hence,

$$S \cap \bigcup_{c \in T \cap W_1} R(c) = S \cap \bigcup_{c \in W_1} R(c).$$

Consider now sets $S \cap N'$ and $T \cap C'$ (recall that $E_2 = (N', C', b')$ is the election obtained after the first step of our algorithm). It holds that:

$$\begin{aligned}
|T \cap C'| &= |T| - |T \cap W_1| \\
&\leq |S| \cdot {}^{b}\!/\!_{n} - \left| \bigcup_{c \in T \cap W_1} R(c) \right| \cdot {}^{b}\!/\!_{n} \leq |S| \cdot {}^{b}\!/\!_{n} - \left| S \cap \bigcup_{c \in T \cap W_1} R(c) \right| \cdot {}^{b}\!/\!_{n} \\
&\leq |S \setminus \bigcup_{c \in W_1} R(c)| \cdot {}^{b}\!/\!_{n} = |S \cap N'| \cdot {}^{b}\!/\!_{n} = |S \cap N'| \cdot {}^{b'}\!/\!_{n'}.
\end{aligned}$$

Further, for each voter $i \in S \cap N'$ we have that:

$$T \rhd_i W \implies (T \setminus W_1) \rhd_i (W \setminus W_1) \implies (T \cap C') \rhd_i W_2.$$

Consequently, $S \cap N'$ and $T \cap C'$ witness the violation of core-stability for committee W_2, which is contradictory to Corollary 8.33. $\square$

Corollary 8.35. *The core is always non-empty and committees satisfying this condition can be found in polynomial time for the following classes of voters' preferences: (1) voter-interval, (2) candidate-interval, (3) single-peaked, and (4) single-crossing preferences.*

In Theorem 8.36 below we show that the condition of recursiveness in the definition of the class of r-STC preferences is necessary for the existence of the core. Thus, in a way Theorem 8.34 gives a rather precise condition on the satisfiability of core-stability for strict voters' preferences. For approval preferences one cannot easily argue that the conditions are precise, since it is still a major open question whether a core-stable committee exists in each approval election.

Theorem 8.36. *There is a top-monotonic election with ordinal preferences, where no committee is core-stable.*

Proof. Let A be a Condorcet cycle consisting of $r = 100$ candidates:

$$\begin{aligned}
&a_1 \succ a_2 \succ \ldots \succ a_r \\
&a_2 \succ a_3 \succ \ldots \succ a_r \succ a_1 \\
&\ldots \\
&a_r \succ a_1 \succ a_2 \succ \ldots \succ a_{r-1}
\end{aligned}$$

Now let B, C, D, E and F be five clones of A. Thus in $A \cup B \cup \ldots \cup F$ we have $6r = 600$ candidates. We add two more candidates, namely g and h.

Consider the following profile with 600 voters:

$$\begin{aligned}
&g \succ A \succ B \succ C \succ D \succ E \succ F \succ h \\
&g \succ B \succ C \succ A \succ E \succ F \succ D \succ h \\
&g \succ C \succ A \succ B \succ F \succ D \succ E \succ h \\
&h \succ D \succ E \succ F \succ A \succ B \succ C \succ g \\
&h \succ E \succ F \succ D \succ B \succ C \succ A \succ g \\
&h \succ F \succ D \succ E \succ C \succ A \succ B \succ g
\end{aligned}$$

For example, the first two votes in this profile are:

$$\begin{aligned}
&g \succ a_1 \succ a_2 \succ \ldots \succ a_r \succ b_1 \succ b_2 \succ \ldots \succ b_r \succ \ldots\ldots \succ f_1 \succ f_2 \succ \ldots \succ f_r \succ h \\
&g \succ a_2 \succ a_3 \succ \ldots \succ a_1 \succ b_2 \succ b_3 \succ \ldots \succ b_1 \succ \ldots\ldots \succ f_2 \succ f_3 \succ \ldots \succ f_1 \succ h
\end{aligned}$$

The above profile is single-top-crossing since there are only two top-candidates, g and h, and each of them crosses with each other candidate only once.

Let $b = 7$, and consider a committee W. We will show that W does not belong to the core. Without loss of generality, we can assume that $g, h \in W$, as there exists more than $600/7$ voters who rank each of these candidates as their favorite one. Further, since the profile is symmetric, without loss of generality we can also assume that it contains at most two candidates from $A \cup B \cup C$. If the two candidates belong to the same clone, say A, then we take a candidate $c \in C$, and observe that 200 voters (the second and the third group) prefer $\{c, g\}$ over W. Otherwise, if the two candidates are from two different clones, say A and B (the situation is symmetric), then we take the clone which is preferred by the majority (in this context A) and select the candidate $a \in A$ that is preferred by $r-1$ voters to the member of $W \cap A$. There are $2r-2 = 198$ voters who prefer $\{g, a\}$ to W. Thus, W does not belong to the core. $\square$

8.5 Extensions, Discussion and Open Questions

Let us conclude the results presented in this chapter with two interesting observations, relating our definitions to the ones already existing in the literature, and discuss the most important open questions.

8.5.1 Core-Stability versus (Full) Local Stability

Aziz et al. [2017b] proposed the concept of *full local stability*, which is equivalent to the definition of core-stability for ordinal preferences. Interestingly, while the concept of the core has been studied before in the context of ordinal committee elections, the equivalence of the two concepts has never been claimed so far. Yet, most of the results in the work of Aziz et al. [2017b] are formulated for the concept of *local stability*. Interestingly, this concept is also equivalent to the core-stability for ordinal preferences, but for a different preference extension: we say that voter i weakly prefers W over T according to $\rhd^{\max}$ preference extension if and only if she ranks her most preferred candidate in W as high as her most preferred candidate in T. In words, according to the $\rhd^{\max}$ extension we focus only on the single top preferred candidate in the committee and do not break ties lexicographically.

Definition 8.37 (Local stability)**.** Consider an election E and a value $q \in \mathbb{Q}$. A committee W violates local stability for quota q if there exists a group $S \subseteq N$ with $|S| \geq q$ and a candidate $c \in C \setminus W$ such that each voter from S prefers c to each member of W.

Proposition 8.38. *Local stability for quota $\lceil n/b \rceil$ is equivalent to core-stability for the $\rhd^{\max}$ preference extension.*

Proof. The fact the core-stability with $\rhd^{\max}$ implies local stability is straightforward—local stability is a special case of the core condition for $|T| = 1$. Now consider any election E and a committee W that is not core-stability with $\rhd^{\max}$. Let $S \subseteq N$ and $T \subseteq C$ be the witness that W is not in the core. For a candidate $c \in T$ let $R_c \in S$ denote a set of voters i such that $c \succ_i W$. Since for every $i \in S$ there exists $c \in T$ with $i \in R_c$:

$$|S| \leq \sum_{c \in T} |R_c|$$

Hence, there exists a candidate $c \in T$ such that $|R_c| \geq |S|/|T| \geq n/b$. Yet, R_c together with the candidate c witness the violation of local stability, which completes the proof. $\square$

8.5.2 Linearly Consistent versus Seemingly Single-Crossing Preferences

Let us now compare the domain of linearly consistent preferences with the one of *seemingly single-crossing (SSC)* preferences [Elkind et al., 2017]—another known class that generalizes VI and CI domains. We say that preferences are seemingly single-crossing if there is a linear order over voters such that for each $a, b \in C$, the voters approving a and not b either all succeed or all precede the voters approving b and not a.[3]

Observe that LC implies SSC. Indeed, consider an LC election, two candidates $a, b \in C$ such that $a \sqsupset b$, and two voters $i, j \in N$. Let i approve a but not b, and j approve b, but not a.

[3]There is also another class, generalizing both VI and CI—namely, the class of *possibly single-crossing* (PSC) preferences. This is the class of approval preferences that can be obtained from some strict single-crossing ones (assuming that every voter approves a consistent prefix of her ranking). Interestingly, PSC is equivalent to the class of seemingly single-crossing preferences [Elkind et al., 2017].

Then, if $j \sqsupset i$, we would obtain a contradiction with the condition for linear consistence (LC would require that j approves a). Hence, $i \sqsupset j$, and the same linear order over voters witnessing linear consistence witnesses also seemingly single-crossingness.

The reverse implication does not hold, as we show in Example 8.39 below.

Example 8.39. Consider the election with 3 voters and the following preferences:

$$1\colon \{a, c\}$$
$$2\colon \{a, b\}$$
$$3\colon \{b, c\}$$

It is straightforward to check that these preferences are SSC for all pairs of candidates and any linear order over voters.

Suppose that this election is LC and let $\sqsupset$ be the required linear order over $N \cup C$. Without the loss of generality, let $a \sqsupset b$. Then we have that $1, 2 \sqsupset 3$ (otherwise, LC would be violated for voter 3, a voter $j \in \{1, 2\}$ such that $3 \sqsupset j$, and candidates a, b).

Further, suppose that $b \sqsupset c$. Then, voters 1 and 3 together with candidates b, c witness the violation of LC, a contradiction. Hence, $c \sqsupset b$. But then, voters 2 and 3 together with candidates b, c witness the violation of LC. The obtained contradiction completes the proof. ⌟

8.5.3 Open Questions

In Section 8.2 we have shown that the classic committee election rules that are commonly considered proportional are not core-stable even if the voters' preferences come from certain restricted domains. Since these domains are natural and can be intuitively explained, one would expect a good rule to behave well for such well-structured elections. On the other hand, we often require a rule which is well-defined for all preference profiles. This leads us to the following important open question.

Question 8.40. *Is there a natural voting rule that satisfies the strongest axioms of proportionality, and which at the same time satisfies core-stability for restricted domains.*

The requirement that a rule should be "natural" says in particular that its definition cannot conditionally depend on whether the election at hand comes from a restricted domain or not. Question 8.40 is valid for both approval and ordinal voters' preferences.

Additionally, it would be interesting to check how often the classic rules violate the core, especially in the case of restricted domains. One can make such a quantitative comparison via experiments. This however raises the algorithmic questions of how hard it is to verify if a given committee (in our case the committee returned by the particular rule) belongs to the core. This question is easy for the $\rhd^{\max}$ preference extension.

Proposition 8.41. *There exists a polynomial-time algorithm for deciding whether a given committee is core-stable for the $\rhd^{\max}$ preference extension.*

Proof. Given a committee W it is sufficient to iterate over all candidates $c \in C \setminus W$ and check if the number of voters who prefer c over W is no-greater than $\lceil n/b \rceil$. $\square$

However, for ordinal preferences the question is much less obvious.

Question 8.42. *What is the computational complexity of deciding whether a given committee belongs to the core?*

This question has been already answered for general elections with approval-based preferences by Brill et al. [2022]—the problem is CoNP-complete. It still remains open for general elections with ordinal preferences as well as for both types of preferences under each restricted domain studied in this work.

Chapter 9

Market-Based Axioms

Let us now go back to the general PB model, with approval-based preferences. Recall the idea of priceability introduced in Definition 4.18, in particular, the condition (P) for a price system $p = (\mathrm{end}, C_p, \{p_i\}_{i \in N})$:

(P) For each project $c \notin C_p$, the unspent budget of her supporters is at most $\mathrm{cost}(c)$:

$$\sum_{i \in N : u_i(c) > 0} \left(\mathrm{end} - \sum_{c' \in C_p} p_i(c') \right) \leq \mathrm{cost}(c) \quad \text{for each } c \notin C_p.$$

Intuitively, priceability seeks outcomes which can be explained via the process of paying for the projects by voters with the virtual money corresponding to their voting power. This way, priceability not only is more restrictive than core-stability in some cases like Example 4.19, but also provides an explicit and intuitive explanation why the considered outcome is fair: voters' payments serve as a mapping of the elected projects to the voters they are meant to represent. We believe this is an appealing and prospective idea.

However, the explanation provided by priceability is weak: it only requires that voters have limited leftover money, but not that their money is spent *wisely*. Consider the following simple example:

Example 9.1. Consider a committee election with b voters $1, 2, \ldots, b$. Each voter $i \in N$ approves a different candidate a_i. Additionally, all the voters approve candidates $c_1, \ldots, c_b$. Consider now a committee $\{a_1, \ldots, a_b\}$, depicted in Figure 27b. Such a committee is priceable, as witnessed by the price system with $b' = b$ and each voter $i \in N$ spending all her endowment on a candidate a_i. On the other hand, a committee $\{c_1, \ldots, c_b\}$ depicted in Figure 27a, would be a much better choice here (every voter would gain utility b instead of 1). ⌟

This is why priceability, on its own, does not imply strong fairness guarantees. In fact, this is the case even if we exclude the most unreasonable priceable outcomes. For example, consider the *Utilitarian Priceable Rule* (UPR) which picks, among priceable outcomes, those that maximize the utilitarian social welfare (total number of approvals from voters). In Example 9.1, UPR

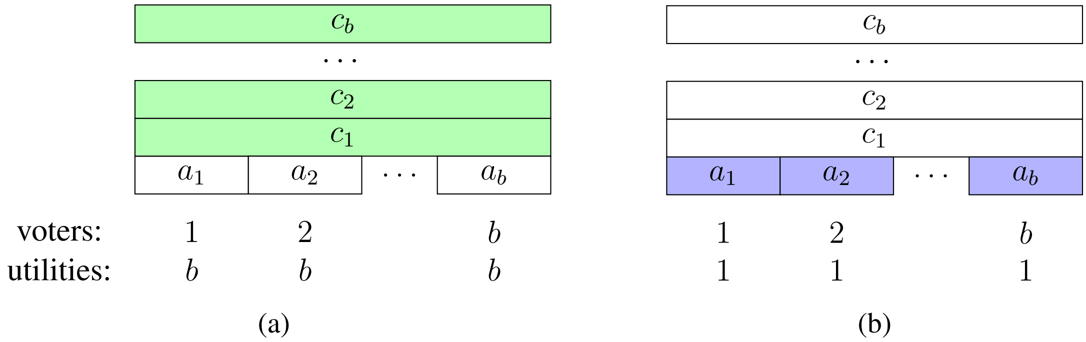


Figure 27: An illustration of Example 4.19.

chooses the committee depicted in Figure 27a. However, UPR fails EJR, even under the unit cost assumption. In fact, as we show in Proposition 9.2, UPR does not even approximate EJR up to any constant number of projects (in a way similar to PAV in Example 4.13). Intuitively, this means priceability provides a very weak proportionality guarantee for cohesive groups of voters. Consequently, UPR also violates all stronger notions than EJR, like FJR or the core.

Proposition 9.2. *In the model of committee elections, the Utilitarian Priceable Rule does not satisfy EJR up to r projects for any $r \geq 0$.*

Proof. Consider the following committee election. Let $x \geq 2$ be a natural number. We introduce $n = 4x^3$ voters, and $4x^2 + x$ candidates; the cost of each candidate equals to one, and $b = 2x^2$. The voters are divided into two groups.

1. The first group consists of $2x^3$ voters—we divide these voters into $2x^2$ equal-size subgroups. Each subgroup (of size x) approves a single different candidate. Let A denote the set of candidates approved by these voters; clearly $|A| = 2x^2$.

2. The second group (containing also $2x^3$ voters) is constructed as follows. We divide these voters again into x^2 subgroups, each of size $2x$. Each such a subgroup approves 2 common candidates—let B be the set of candidates approved by these voters; $|B| = 2x^2$. Additionally, from each subgroup we take one voter—let V denote the set of these voters; clearly $|V| = x^2$. The voters from V approve some common x candidates; let C denote the set of these candidates.

Let W be a committee elected by UPR here and let $p = (\text{end}, C_p, \{p_i\}_{i \in N})$ be a priceable price system. Note that the value of end cannot be higher than $1/x$. If it were, all the candidates from A would need to be members of the winning committee, leaving no room for the candidates from B; though the voters have money to pay for these candidates. First, consider the case when the value of end is lower than $1/x$. Then, at most half of the candidates from B could be members of the winning committee. The maximum possible utility would be then:

$$\underbrace{x^2 \cdot 2x \cdot 1}_{\text{from } B} + \underbrace{x \cdot x^2}_{\text{from } C} + \underbrace{(x^2 - x) \cdot x}_{\text{from } A} = 2x^3 + x^3 + x^3 - x^2 < 4x^3.$$

Now assume that $\mathrm{end} = {}^1\!/x$. If the committee contains at least one candidate from C, then at most half of the candidates from B can get to the committee. By the same reasoning as above we get that the total utility obtained in such a case is lower than $4x^3$.

On the other hand, B could be purchased by a priceable system with equal payments, inducing the total utility of $x^2 \cdot 2x \cdot 2 = 4x^3$. Consequently, $W = B$. Each voters from V has 2 representatives in this committee. Yet, group V is $({}^x\!/2)$-cohesive. Thus, since x can be arbitrarily large, this induces that UPR fails EJR up to r projects for any fixed $r \geq 0$. $\square$

In this chapter, we try to improve priceability by adding a *stability* condition: informally, voters should not want to change how their money is spent. We borrow the idea of making payments stable from the classic economic concept of Lindahl equilibrium for public economies [Foley, 1970], which in a fractional (divisible) model ensures fair outcomes [Fain et al., 2016]. In fact, we show that one of the two notions we propose is closely related to (a discrete version of) Lindahl equilibrium.

We introduce two concepts: stable priceability (SP) and balanced stable priceability (BSP).

SP strengthens the concept of priceability by Peters and Skowron [2020]. We show that it is a strong fairness notion. It logically implies both the core and priceability, and for committee elections it guarantees a higher average satisfaction of members of cohesive groups (so called *proportionality degree*; see [Skowron, 2021]) than both. In contrast to the core, it can be checked in polynomial time whether an outcome satisfies SP. We also present a compact integer linear program for finding SP outcomes. We adapt the notion of Lindahl equilibrium to the context of PB, and show that SP is closely related to this notion.

One potential source of unfairness under stable priceability is that two voters may be paying different amounts of virtual money for the same candidate that they both approve. Our notion of balanced stable priceability (BSP) addresses this by requiring that any two voters paying for a candidate must pay the same amount. We uniquely characterize BSP outcomes as those returned by a rule similar to Method of Equal Shares.

Unfortunately, even in the committee election model, exhaustive SP and BSP committees do not always exist. However, through a series of experiments, we argue that SP/BSP committees whose size is very close to b often do.

9.1 Payment Systems

In order to extend the axiom of priceability, we need first to introduce one more term, the *payment system*, which intuitively corresponds to the price system without the upfront endowment restrictions. To distinguish payment systems from price systems, we use the symbol q for them instead of p.

Definition 9.3 (Payment system)**.** Fix an election $E = (C, N, b)$. A *payment system* is a pair $q = (C_q, \{q_i\}_{i \in N})$, where $C_q \subseteq C$ is the *supported outcome* and $\{q_i\}_{i \in N}$ is the collection of voters' *payment functions*. For each $i \in N$, $q_i \colon C_q \to \mathbb{R}_{\geq 0}$ specifies the amount of money that voter i spends on particular projects. We require that a payment system satisfies the following conditions:

1. Each voter can pay only for a project that she approves of, that is $q_i(c) = 0$ if $c \notin A_i$.
2. The cost of every project from C_q is fully paid:

$$\forall c \in C_q. \quad \sum_{i \in N} q_i(c) = \mathrm{cost}(c).$$

This term will further appear useful for defining our axioms—intuitively, there we will consider situations where voters are able to resign from paying for some projects and use their saved money to buy another projects. Note that we allow that two voters may pay different amounts of money for the same candidate—in Section 9.3 we consider price and payment systems with additional restrictions that prohibit such cases. For the payment functions from both price and payment systems, let us use the following convention: $q_i(T) = \sum_{c \in T} q_i(c)$, for each $i \in N$ and $T \subseteq C$.

9.2 Stable Priceability

Let $\succ$ be a linear order over $\mathbb{N} \times \mathbb{R}_{\geq 0}$ defined as follows:

$$(x_1, y_1) \succ (x_2, y_2) \iff x_1 > x_2 \text{ or } (x_1 = x_2 \text{ and } y_1 < y_2). \tag{9.1}$$

We will use $(x_1, y_1) \succ (x_2, y_2)$ to model a voter who "prefers" to pay y_1 dollars for an outcome where she approves x_1 candidates than pay y_2 dollars for outcome with x_2 approved members. Thus, under this linear order, the voter "prefers" to maximize her utility for the outcome, and only in case of a tie, prefers to pay less. We note that these are not the true preferences of the voters, but rather an artificial relation that helps us formulate our definition of stable priceability.

A price system $p = (\mathrm{end}, C_p, \{p_i\}_{i \in N})$ is said to be *stable priceable (SP)* if the following condition is satisfied:

(SP) There exists no collection $\{R_i\}_{i \in N}$ (with $R_i \subseteq C_p$ for $i \in N$) and no payment system $q = (C_q, \{q_i\}_{i \in N})$ that would satisfy the following two conditions:

(a) each voter $i \in N$ can additionally spend $q_i(C_q)$ dollars using her remaining money, after resigning from paying for candidates from R_i:

$$p_i(C_p \setminus R_i) + q_i(C_q) \leq \mathrm{end},$$

Besides, for at least one voter $i \in N$ the above inequality is strict.[1]

(b) for each $i \in N$, i is at least as "happy", according to $\succ$, with the outcome C_p after exchanging R_i to C_q as she was before:

$$\Big(u_i\big((C_p \setminus R_i) \cup C_q\big), p_i(C_p \setminus R_i) + q_i(C_q)\Big) \succeq \Big(u_i\big(C_p\big), p_i(C_p)\Big).$$

Besides, for at least one voter $i \in N$ the above inequality is strict.

[1] The role of this requirement is to allow tie-breaking in simple symmetrical elections, analogously as in the case of condition (P). In Section 9.2.2 we will consider SP without this requirement.

We say that an outcome W is stable priceable, if there exists a stable priceable price system $p = (\text{end}, C_p, \{p_i\}_{i \in N})$ such that $C_p = W$.

In order to better understand the condition, assume (SP) is not satisfied. Then, each voter $i \in N$ can find a set R_i of currently approved candidates such that she would "prefer" to stop paying for R_i and to pay for C_q instead (that is, no voter would lose, and at least one voter would strictly benefit according to $\succ$).

It is clear that stable priceability implies priceability: indeed, after fixing $|C_q| = 1$ and $R_i = \emptyset$ for every $i \in N$, the condition (SP) boils down to (P). However, (SP) provides much better fairness guarantees.

For example, a price system p defined in Example 9.1 violates (SP) for $b > 1$. Indeed, fix $R_i = \{a_i\}$ for each voter $i \in N$ and consider a payment system $q = \{q_i\}_{i \in N}$ where $C_q = \{c_1\}$, $q_i(c_1) = {}^1\!/_n$ for each $i \in N$. Then:

(a) $\sum_{i \in N} q_i(c_1) = 1 = \text{cost}(c_1)$ (the price for C_q is paid),

(b) for each $i \in N$, $q_i(C_q) = {}^1\!/_n < {}^b\!/_n = p_i(R_i)$ ((SP-a) is satisfied),

(c) each voter gains utility 1 from both $(C_p \setminus R_i) \cup C_q$ and C_p and pays less for the outcome $(C_p \setminus R_i) \cup C_q$ ((SP-b) is satisfied).

Consequently, $\{R_i\}_{i \in N}$ and q witness the (SP) violation (we additionally depict this fact in Figure 28). By the analogous reasoning, it is straightforward to check that the only stable priceable outcome here is $\{c_1, \ldots, c_k\}$.

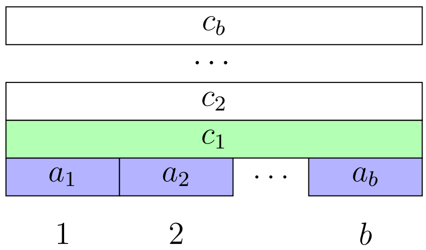


Figure 28: An illustration that the price system from Example 9.1 is not stable. Now all the voters can resign from paying for blue candidates and pay equally for the green one instead. Their utility would be the same, yet all of them would pay less.

In case of priceability, only a single project was sufficient to witness the violation of (P) and it is clear that the possible extending this condition to sets would not increase its strength. Indeed, if the voters could additionally buy a subset of candidates, then in particular they could additionally buy every single member of the subset.

What is more unexpected, the same holds also for stable priceability. As a result, the condition (SP) boils down to a simpler and more concise form, more similar to the one from condition (P):

(SP1) For each $c \notin C_p$:

$$\sum_{i \in N(c)} \max\left(\max_{a \in C_p}(p_i(a)), r_{p,i}\right) \leq \mathrm{cost}(c) \qquad \text{where} \quad r_{p,i} = \mathrm{end} - p_i(C_p).$$

Intuitively, condition (SP1) is a simplified form of (SP) for the case where $C_q = \{c\}$ and $|R_i| \leq 1$ for each $i \in N$. Then, every voter i approving the project from c can either pay for it using only her remaining money (the case of $R_i = \emptyset$) or can resign from paying for at most one project $a \in C_p$. In the latter case, since the utilities are approval-based (each project for which i pays provides her the utility of 1), the most profitable choice for i is to resign from the project for which she pays most. Then, she can use only $p_i(a)$ dollars to pay for c—otherwise, if i spent more money, she would be less happy according to $\succ$. Finally, since we require the weak inequality in (SP1), we ensure that after the change at least one voter would spend less than end dollars and at least one voter would be strictly more satisfied according to $\succ$.

From now, we will often use symbol $r_{p,i}$ to denote the remaining money of voter i under price system p.

Theorem 9.4. *Conditions (SP1) and (SP) are equivalent.*

Proof. Suppose first that $p = (\mathrm{end}, C_p, \{p_i\}_{i \in N})$ is a price system satisfying (SP) and not satisfying (SP1). Then there exists $c \notin C_p$ such that:

$$\sum_{i \in N(c)} \max\left(\max_{a \in C_p}(p_i(a)), r_{p,i}\right) > \mathrm{cost}(c).$$

However, then a collection $\{R_i\}_{i \in N}$ and a payment system $q = (\{c\}, \{q_i\}_{i \in N})$ such that:

(a) $R_i = \emptyset$ and $q_i(c) = r_{p,i}$ for each $i \in N(c)$ such that $r_{p,i} \geq \max_{a \in W}(p_i(a))$,

(b) $R_i = \{\mathrm{argmax}_{a \in W}(p_i(a))\}$ and $q_i(c) = \max_{a \in C_p}(p_i(a))$ otherwise,

witness violating condition (SP).

For the other direction the proof proceeds as follows. Let $p = (\mathrm{end}, C_p, \{p_i\}_{i \in N})$ be a price system satisfying (SP1). Assume towards a contradiction that there exist a payment system $q = (C_q, \{q_i\}_{i \in N})$ and a collection of sets $\{R_i\}_{i \in S}$ with $R_i \subseteq C_p$ for all $i \in N$, witnessing the violation of (SP).

Since each voter pays only for the candidates from C_p she approves (from the definition of a price system), without loss of generality, we can assume that $R_i \subseteq A_i \cap C_p$ for each voter $i \in N$.

Fix a voter $i \in N$. Note that:

$$\begin{aligned}
u_i((C_p \setminus R_i) \cup C_q) - u_i(C_p) &= u_i(C_p \setminus R_i) + u_i(C_q \setminus (C_p \setminus R_i)) - u_i(C_p) \\
&= u_i(C_q \setminus (C_p \setminus R_i)) - u_i(R_i) \\
&= u_i(C_q \setminus C_p) + u_i(C_q \cap R_i) - u_i(R_i) \\
&= u_i(C_q \setminus C_p) - u_i(R_i \setminus C_q).
\end{aligned}$$

and:

$$\begin{aligned}
p_i(C_p) - (p_i(C_p \setminus R_i) + q_i(C_q)) &= p_i(R_i) - q_i(C_q) \\
&= p_i(R_i \cap C_q) + p_i(R_i \setminus C_q) - q_i(C_q \setminus C_p) - q_i(C_q \cap C_p) \\
&\leq p_i(C_p \cap C_q) + p_i(R_i \setminus C_q) - q_i(C_q \setminus C_p) - q_i(C_q \cap C_p) \\
&\overset{(*)}{\leq} p_i(R_i \setminus C_q) - q_i(C_q \setminus C_p).
\end{aligned}$$

The inequality $(*)$ follows from the fact that $\sum_{i \in N} p_i(c') = \sum_{i \in N} q_i(c') = \mathrm{cost}(c')$ for all $c' \in C_p \cap C_q$.

Hence:

(a) inequality $u_i((C_p \setminus R_i) \cup C_q) > u_i(C_p)$ boils down to the inequality:

$$u_i(C_q \setminus C_p) > u_i(R_i \setminus C_q),$$

(b) inequality $p_i(C_p \setminus R_i) + q_i(C_q) < p_i(C_p)$ boils down to the inequality:

$$q_i(C_q \setminus C_p) < p_i(R_i \setminus C_q).$$

.

If $u_i(C_q \setminus C_p) = u_i(R_i \setminus C_q)$, then $|(C_q \setminus C_p) \cap A_i| = |R_i \setminus C_q|$ and:

$$q_i(C_q \setminus C_p) \leq p_i(R_i \setminus C_q) \leq |R_i \setminus C_q| \cdot \max_{a \in C_p}(p_i(a)) = |(C_q \setminus C_p) \cap A_i| \cdot \max_{a \in C_p}(p_i(a)).$$

On the other hand, if $u_i(C_q \setminus C_p) > u_i(R_i \setminus C_q)$, then $|(C_q \setminus C_p) \cap A_i| > |R_i \setminus C_q|$ and:

$$\begin{aligned}
q_i(C_q \setminus C_p) &\leq r_{p,i} + p_i(R_i) - q_i(C_q \cap C_p) \\
&\leq r_{p,i} + p_i(R_i \setminus C_q) + p_i(C_p \cap C_q) - q_i(C_q \cap C_p) \\
&\leq (|R_i \setminus C_q| + 1) \cdot \max\left(\max_{a \in C_p}(p_i(a)), r_{p,i}\right) \\
&\leq |(C_q \setminus C_p) \cap A_i| \cdot \max\left(\max_{a \in C_p}(p_i(a)), r_{p,i}\right).
\end{aligned}$$

Hence, for each voter $i \in N$ we have that

$$q_i(C_q \setminus C_p) \leq |(C_q \setminus C_p) \cap A_i| \cdot \max\left(\max_{a \in C_p}(p_i(a)), r_{p,i}\right).$$

Besides, we know that for at least one voter this inequality is strict.

We have that:

$$\begin{aligned}
\mathrm{cost}(C_q \setminus C_p) = q_i(C_q \setminus C_p) &< \sum_{i \in N}\left(|(C_q \setminus C_p) \cap A_i| \cdot \max\left(\max_{a \in C_p}(p_i(a)), r_{p,i}\right)\right) \\
&= \sum_{c \in C_q \setminus C_p} \sum_{i \in N(c)} \max\left(\max_{a \in C_p}(p_i(a)), r_{p,i}\right).
\end{aligned}$$

Finally:

$$0 < \sum_{c \in C_q \setminus C_p} \left(\sum_{i \in N(c)} \max\left(\max_{a \in C_p}(p_i(a)), r_{p,i} \right) - \mathrm{cost}(c) \right).$$

Thus, there must exist $c \in C_q \setminus C_p$ such that:

$$\sum_{i \in N(c)} \max\left(\max_{a \in C_p}(p_i(a)), r_{p,i} \right) > \mathrm{cost}(c).$$

Which gives a contradiction and completes the proof. □

An important consequence of Theorem 9.4 is that one can formulate a compact integer linear program for finding SP outcomes. Let us briefly describe it below.

For each project $c \in C$ we have a binary variable x_c which indicates whether c is a part of the SP outcome W. Inequalities (9.3) and (9.4) encode the feasibility and exhaustiveness constraints for the outcome W, respectively. For each $c \in C$ and $i \in N$ we have a variable $p_{i,c}$ which denotes the amount of money that voter i pays for c. Besides, we have a variable end for the voters' endowment. Then inequality (9.5) ensures that a voter will not spend more than its initial budget. Finally, (9.6) ensures that the total payment equals $\mathrm{cost}(c)$ dollars for every selected project and 0 dollars for every non-selected project. For each voter $i \in N$ we also have a variable m_i, which intuitively equals to $\max\left(\max_{a \in W}\left(p_{i,a}\right), r_{p,i}\right)$—this interpretation is encoded in (9.7) and (9.8). The last inequality (9.9) encodes the constraint of stability (SP1). Here we use the fact that in a feasible solution it holds that $\mathrm{end} \le b$ (otherwise, every candidate is bought).

constraints:

$$x_c \in \{0, 1\} \qquad \text{for } c \in C \tag{9.2}$$

$$\sum_{c \in C} x_c \cdot \mathrm{cost}(c) \le b \tag{9.3}$$

$$\sum_{c \in C} x_c \cdot \mathrm{cost}(c) > (1 - x_{c'})(b - \mathrm{cost}(c')) \qquad \text{for } c' \in C \tag{9.4}$$

$$\sum_{c \in C} p_{i,c} \le \mathrm{end} \qquad \text{for } i \in N \tag{9.5}$$

$$\sum_{i \in N} p_{i,c} = x_c \qquad \text{for } c \in C \tag{9.6}$$

$$0 \le p_{i,c} \le m_i \qquad \text{for } i \in N, c \in C \tag{9.7}$$

$$\mathrm{end} - \sum_{c \in C} p_{i,c} \le m_i \qquad \text{for } i \in N \tag{9.8}$$

$$\sum_{i \in N(c)} m_i \le \mathrm{cost}(c) + x_c \cdot b \qquad \text{for } c \in C \tag{9.9}$$

Further, for a given election and an outcome W, the problem of finding a price system supporting W and satisfying (SP1) can be formulated as a linear program, similar to the above one (yet with constant values of x_c for $c \in C$). Then we can efficiently check whether W is SP [Khachiyan, 1979].

Corollary 9.5. *Given an election and an outcome W, it can be checked in polynomial time whether W is stable priceable.*

This is in contrast to many other group fairness properties, which are CoNP-hard to check (see the works of Aziz et al. [2018] and Brill et al. [2022]).

9.2.1 Axiomatic Properties of Stable Priceability

Let us now prove several fairness guarantees provided by stable priceability. Theorem 9.6 shows that for exhaustive outcomes, stable priceability implies core-stability, and therefore, in turn, EJR.

Theorem 9.6. *Every exhaustive stable priceable outcome is core-stable.*

Proof. Suppose for the sake of contradiction that it does not hold for some election and an exhaustive outcome W. Let $p = (\mathrm{end}, C_p, \{p_i\}_{i \in N})$ be a stable priceable price system such that $W = C_p$. Assume towards a contradiction that there exists a group of voters S and a set of projects T such that: (i) $\mathrm{cost}(T) \leq b \cdot {}^{|S|}/_{n}$, and (ii) $|A_i \cap T| \geq |A_i \cap W| + 1$ for each $i \in S$.

Since the condition (SP1) is satisfied, for each project $c \in T \setminus W$ we have that:

$$\sum_{i \in N(c)} \max\left(\max_{c' \in W}\left(p_i(c')\right), r_{p,i}\right) \leq \mathrm{cost}(c) \tag{9.10}$$

Also, from the feasibility of p, for each $c \in W$ (in particular, for each $c \in T \cap W$) we have that:

$$\sum_{i \in N(c)} p_i(c) = \mathrm{cost}(c). \tag{9.11}$$

Now, let us sum equalities (9.10) and (9.11) over all $c \in T$, using inequality (9.10) whenever $c \in T \setminus W$, and using equality (9.11), for $c \in T \cap W$:

$$\sum_{c \in T \cap W} \sum_{i \in N(c)} p_i(c) + \sum_{c \in T \setminus W} \sum_{i \in N(c)} \max\left(\max_{c' \in W}\left(p_i(c')\right), r_{p,i}\right) \leq \mathrm{cost}(T).$$

Let us regroup the terms in the left-hand side of the above inequality:

$$\sum_{i \in N}\left(p_i(A_i \cap T \cap W) + |A_i \cap (T \setminus W)| \cdot \max\left(\max_{c' \in W}\left(p_i(c')\right), r_{p,i}\right)\right) \leq \mathrm{cost}(T).$$

Next, note that since $|A_i \cap T| \geq |A_i \cap W| + 1$ for each $i \in S$, we also have that $|A_i \cap (T \setminus W)| \geq |A_i \cap (W \setminus T)| + 1$. As a result, we get that:

$$
\begin{aligned}
\mathrm{cost}(T) &\overset{(*)}{\geq} \sum_{i \in S} \left(p_i(A_i \cap T \cap W) + (|A_i \cap (W \setminus T)| + 1) \cdot \max\left(\max_{c' \in W} (p_i(c')), r_{p,i} \right) \right) \\
&\geq \sum_{i \in S} \left(p_i(A_i \cap T \cap W) + p_i(A_i \cap (W \setminus T)) + \max\left(\max_{c' \in W} (p_i(c')), r_{p,i} \right) \right) \\
&= \sum_{i \in S} \left(p_i(A_i \cap W) + \max\left(\max_{c' \in W} (p_i(c')), r_{p,i} \right) \right) \\
&= \sum_{i \in S} \left(p_i(W) + \max\left(\max_{c' \in W} (p_i(c')), r_{p,i} \right) \right).
\end{aligned}
$$

From the definition of price system, we know that $\mathrm{end} \geq {}^{b}/_{n}$. Hence:

$$
\begin{aligned}
\mathrm{cost}(T) &\geq \sum_{i \in S} \left(p_i(W) + \max\left(\max_{c' \in W} (p_i(c')), r_{p,i} \right) \right) \\
&\overset{(**)}{\geq} \sum_{i \in S} (p_i(W) + r_{p,i}) \overset{(***)}{\geq} |S| \cdot {}^{b}/_{n}.
\end{aligned}
$$

If any of the above weak inequalities is strict, then we have a contradiction. Else we have that inequalities (9.10), $(*)$, $(**)$ and $(***)$ are equalities. From that we conclude the following facts:

1. If $(*)$ is an equality, then that is,$|A_i \cap (T \setminus W)| \cdot \max\left(\max_{c' \in W} (p_i(c')), r_{p,i}\right) = 0$ for each $i \in N \setminus S$, hence $A_i \cap (T \setminus W)| = 0$ for each $i \in N \setminus S$. As a result, $N(c) \subseteq S$ for each $c \in T \setminus W$.

2. If $(**)$ is an equality, then $\max\left(\max_{c' \in W} (p_i(c')), r_{p,i}\right) = r_{p,i}$ for each $i \in S$.

3. Inequality (9.10) needs to be an equality that is,for all candidates from $T \setminus W$. Fix any $c \in T \setminus W$. From the above remarks, we can rewrite (9.10) for c as follows:

$$\sum_{i \in N(c)} r_{p,i} = \mathrm{cost}(c)$$

Further:

$$\mathrm{cost}(c) \leq \sum_{i \in N} r_{p,i} = \sum_{i \in N} (\mathrm{end} - p_i(C_p)) = n \cdot \mathrm{end}(p) - \mathrm{cost}(C_p).$$

Finally, as $(***)$ is an equality ($\mathrm{end} = {}^{b}/_{n}$):

$$\mathrm{cost}(c) \leq b - \mathrm{cost}(C_p).$$

As C_p needs to be exhaustive, the last statement is a contradiction, which completes the proof. $\square$

Corollary 9.7. *Every exhaustive stable priceable outcome satisfies EJR.*

The core on its own is already a formidable axiom, and not known to be achievable in all elections. Are there any advantages of considering an axiom that further strengthens the core, and also strengthens priceability? We argue that there are several ones.

As already mentioned, a first advantage that SP has over the core is that whether an outcome is SP can be checked in polynomial time (Corollary 9.5), whereas the same question is known to be difficult for the core [Brill et al., 2022]. Additionally, in elections like Example 4.19 presented in the introduction, the core allows apparently unfair solutions, while SP rules them out and allows only fairer ones.

An advantage that SP has over priceability is that SP implies the core, and in turn, EJR (Theorem 9.6), whereas priceability does not even imply EJR (Proposition 9.2).

Finally, one advantage that SP has over both the core and priceability is that SP implies a high *proportionality degree* [Skowron, 2021]. This notion, defined for the committee election model, measures how much members of cohesive groups (defined in the same way as in case of EJR) are satisfied on average.

Definition 9.8 (Proportionality Degree [Skowron, 2021])**.** A group of voters S is *ℓ-cohesive* for $\ell \in \mathbb{N}$ if

$$|S| \geq \frac{\ell}{b} \cdot n \quad \text{and} \quad |\bigcap_{i \in S} A_i| \geq \ell.$$

Let $f \colon \mathbb{N} \to \mathbb{R}_{\geq 0}$. We say that a rule $\mathcal{R}$ has the proportionality degree of f, if for every committee election E, each winning committee $W \in \mathcal{R}(E)$, and each ℓ-cohesive group of voters S, the average number of committee members a voter from S approves is at least $f(\ell)$, that is, $(1/|S|) \cdot \sum_{i \in S} u_i(W) \geq f(\ell)$.

Theorem 9.9. *In the committee election model, every exhaustive stable priceable outcome provides a proportionality degree of $\ell - 1$.*

Proof. Fix an election, and consider a size-b SP committee W. Let S be an ℓ-cohesive group of voters and let T be a set of ℓ candidates who are approved by all members of S. We will show that an average number of representatives that the voters from S have in W equals at least $\ell - 1$.

Without loss of generality, let us assume that there exists a not-elected candidate $c \notin W$ that is approved by all members of S (as, otherwise, the average number of representatives for voters from S would be at least ℓ).

Note that, by the pigeonhole principle, for each voter $i \in N$ (i has the right to spend end dollars and can either save them or spend on at most $|A_i \cap W|$ candidates) we have that:

$$\max\left(\max_{c' \in W}(p_i(c')), r_{p,i}\right) \geq \frac{\mathrm{end}(p)}{|A_i \cap W| + 1} \geq \frac{b}{n \cdot (|A_i \cap W| + 1)}.$$

By condition (SP1) applied to c, we get that:

$$1 = \mathrm{cost}(c) \geq \sum_{i \in S} \max\left(\max_{c' \in W}\left(p_i(c')\right), r_{p,i}\right) \geq \sum_{i \in S} \frac{b}{n \cdot (|A_i \cap W| + 1)}.$$

By the inequality between the harmonic and arithmetic mean, we get that:

$$\frac{\sum_{i \in S}(|A_i \cap W|)}{|S|} = \frac{\sum_{i \in S}(|A_i \cap W| + 1)}{|S|} - 1 \geq \frac{|S|}{\sum_{i \in S} \frac{1}{|A_i \cap W| + 1}} - 1$$
$$\geq |S| \cdot \frac{b}{n} - 1 \geq \frac{n\ell}{b} \cdot \frac{b}{n} - 1 = \ell - 1.$$

This completes the proof. □

In contrast, priceability does not imply a proportionality degree better than 2 (which follows from the proof of Proposition 9.2). As we have mentioned in Section 4.1, it is known that EJR only implies a proportionality degree of $\ell-1/2$. We can easily check that the construction showing that this bound is tight, provided by Skowron [2021], works also for the core:

Example 9.10. Consider a committee election; let $n = b^2$ and consider the following preferences of the voters:

$$\begin{array}{ll} b \text{ voters}: & \{a_1, \ldots, a_k, d_1\} \\ b \text{ voters}: & \{a_1, \ldots, a_k, d_1, d_2\} \\ \ldots & \ldots \\ b \text{ voters}: & \{a_1, \ldots, a_k, d_1, \ldots, d_k\} \end{array}$$

We will prove that the committee $W = \{d_1, \ldots, d_k\}$ is in the core; yet the average satisfaction of the b-cohesive group of all voters will indeed be equal to

$$\frac{(1 + 2 + \ldots + b) \cdot b}{b^2} = \frac{b-1}{2}.$$

Consider any group $S \subseteq N$ of size $\ell \cdot n/b$ and any committee T such that every voter from S has more approved candidates in T than in W. Let $i \in S$ be the voter from S who is the most satisfied from W—note that group S needs to contain voters from at least ℓ different groups of voters (each one approving a different positive number of candidates in W), hence $|A_i \cap W| \geq \ell$. It means that committee T contains at least $\ell + 1$ candidates and is too large to witness the core violation. ⌟

Hence, core-stability also implies proportionality degree of only $\ell-1/2$.

One could ask whether similar proportionality guarantees (for example, EJR) are provided also by non-exhaustive SP outcomes. We leave this question for future research. However, if a large part of the budget is spent, non-exhaustive stable priceable outcomes provide good approximations of the axioms presented above—for example, from Theorem 9.6 we can obtain the following corollary (recall the idea of α-core-stability, Definition 4.16):

Corollary 9.11. *Every stable priceable outcome W is $(b/\mathrm{cost}(W))$-core-stable.*

Proof. Directly from Theorem 9.6 and the fact that after reducing the budget from b to $\mathrm{cost}(W)$, W becomes exhaustive. $\square$

In an analogous way we obtain that also other properties of SP outcomes W regarding EJR or proportionality degree hold for cohesive groups that are $b/\mathrm{cost}(W)$ times larger than in the original definitions of these axioms. Hence, every stable priceable outcome W such that $b/\mathrm{cost}(W)$ is close to 1, provides good axiomatic guarantees.

9.2.2 Stable Priceability versus Lindahl equilibrium

The concept of stable priceability suggests there might exist a relation between our voting model and classic market models for economies with public goods. In this section we explain this relation in more detail, focusing on the most influential equilibrium concept from the literature on public goods—the Lindahl equilibrium, which was formalized by Foley [1970]. The relation that we explain in this section: (1) gives additional insights into the concept of SP, and (2) explains the key differences that prohibit one to use the concepts from the public economics directly for designing voting systems.

The public economics (PE) model for participatory budgeting (PB), adapted from the work of Foley [1970], is set up as follows. We imagine that there is a *producer* who will set up the outcome in exchange for money. The *production function* $\pi\colon 2^C \to \mathbb{R}_{\geq 0}$ assigns to each outcome $W \subseteq C$ the cost to the producer of producing W.[2] We assume the cost of producing a candidate is $\pi(W) = \mathrm{cost}(W)$ for all $W \subseteq C$.

A *PE price system* is a pair $\gamma = (\mathrm{end}, \{\gamma_i\}_{i \in N})$, where each $\gamma_i\colon C \to \mathbb{R}_{\geq 0}$ is a *PE payment function*. In contrast to price systems defined in Section 4.3 (which we will further call PB price systems), PE price systems on their own do not immediately justify purchasing any set of candidates—that is, they are defined for all $c \in C$ and we do not require that $\sum_{i \in N} \gamma_i(c) = \mathrm{cost}(c)$ for any $c \in C$. However, we still keep the assumption that voters pay only for the candidates they approve (for each $i \in N$ and $c \notin A_i$, $\gamma_i(c) = 0$). To avoid confusion we use different symbols to denote PE and PB price systems (γ and p, respectively).

Intuitively, in a PE price system, the purchased outcome is not given upfront—it is the producer who decides which candidates are produced. To produce an outcome W, she needs to spend $\mathrm{cost}(W)$ dollars and after that she if given $\sum_{i \in N} \gamma_i(W)$ dollars from the voters. Note that, as a result, PE payments do not refer to the amount of money which is *actually spent* by

[2]In Foley's model (Foley [1970]), the production function specifies how private goods can be transformed into public goods. In our case, we assume there is only one private good, money (which represents voting power); the candidates are the public goods. Thus, as in Foley's model, the production function describes how private goods can be transformed into public goods.

A crucial difference to Foley's model is that we use an indivisible model, where each candidate can be either bought (elected) or not, and there are no intermediate states. Due to indivisibilities, Foley's existence proof does not apply. Further, in our model each candidate is available in a single copy, which can affect decisions of the producers, and thus the prices.

the voters, as they are charged only for produced candidates—payments for the other ones do not affect their endowment end.

An exhaustive outcome W is in *Lindahl equilibrium* if there exists a PE price system $\gamma = (\mathrm{end}, \{\gamma_i\}_{i \in N})$ such that the following conditions hold:

(L1) Profit maximization: producer chooses the outcome that maximizes her profit (the difference between the total payment for the outcome and its cost).

For each outcome $W' \subseteq C$ we have that:

$$\sum_{i \in N} \gamma_i(W) - \pi(W) \geq \sum_{i \in N} \gamma_i(W') - \pi(W').$$

Note that since $\pi(\emptyset) = 0$ the above condition implies that $\sum_{i \in N} \gamma_i(W) \geq \pi(W)$ (the total payments payed to W are sufficient to produce W).

(L2) Utility maximization: voters spend the money they are entitled to so that they maximize their utility.

For each voter i, there is no outcome W' with $\gamma_i(W') \leq \mathrm{end}$ and:

$$(u_i(W'), \gamma_i(W')) \succ (u_i(W), \gamma_i(W)).$$

In the divisible PE model (where we can elect candidates fractionally) the conditions (L1) and (L2) are always satisfiable, and the resulting outcome is guaranteed to be in the core [Foley, 1970]. For us, neither is true. We start by providing an example of a profile where a Lindahl equilibrium is not Pareto optimal, and thus is not in the core.

Example 9.12. Consider a committee election with $b = 3$, 3 voters and the following preferences:

$$\begin{aligned} 1&\colon \{a, e, d_1, d_2, d_3\} \\ 2&\colon \{e, c, d_1, d_2, d_3\} \\ 3&\colon \{c, a, d_1, d_2, d_3\} \end{aligned}$$

Let $\epsilon = 1/1000$. Consider a PE price system in which the payments are the following:

$$\begin{aligned} \gamma_i(x) &= 1 - 4 \cdot \epsilon && \text{for } i \in N \text{ and } x \in \{a, c, e\} \\ \gamma_i(d_i) &= 1 - 3 \cdot \epsilon && \text{for } i \in N \\ \gamma_i(d_j) &= \epsilon && \text{for } i \in N \text{ and } i \neq j \end{aligned}$$

This price system witnesses that $W = \{a, c, e\}$ is a Lindahl equilibrium. Indeed:

(a) for every $x \in W$ we have that $\sum_{i \in N} \gamma_i(x) > \pi(W)$ and for every $x \notin W$ we have that $\sum_{i \in N} \gamma_i(x) < \pi(W)$ (hence, (L1) is satisfied)

(b) each voter gains utility 2 from W, and to gain a higher utility she would need to pay more than $\gamma_i(W)$ ((L2) is satisfied).

Yet, W is Pareto-dominated by $\{d_1, d_2, d_3\}$. ⌟

The problem underlying Example 9.12 is that the producer gets paid less than the cost of d_1, d_2 and d_3 if the producer chooses to produce these candidates. In contrast, the producer receives a payment of almost double the cost of a, b, c for producing these candidates. Thus, in this equilibrium, the producer is better off at the cost of consumers. In the divisible model this issue never appears: in every Lindahl equilibrium the total payment to the producer for producing a unit of candidate c is always equal to the cost of producing that unit. (Otherwise, the producer would want to produce a higher quantity of c.) Since this equality is implied in the divisible model, it is natural to add it as an additional property to our definition of Lindahl equilibrium in the indivisible model.

We say that an outcome W is a *cost-efficient Lindahl equilibrium (CELE)* if there exists a PE price system $\gamma = (\mathrm{end}, \{\gamma_i\}_{i\in N})$ that satisfies (L1), (L2), and:

(L3) Cost-Efficiency:

$$\sum_{i\in N} \gamma_i(W) \le \pi(W).$$

By (L1), the condition in (L3) could also be written as an equality. Further, by (L1) and (L3) we can infer a seemingly stronger condition, that for each $c \in W$:

$$\sum_{i\in N} \gamma_i(c) = \pi(c).$$

Theorem 9.13, below, shows a close relationship between stable priceability and Lindahl equilibrium. Let us slightly adapt condition (SP) by removing the tie-breaking requirement from condition (SP-a): we call the resulting solution concept *strict stable priceability (SSP)*. Note that from the same argument as in Theorem 9.4 we have that this definition can be written as follows: for each $c \notin W$:

$$\forall \epsilon > 0. \sum_{i\in N(c)} \max\left(\max_{a\in C_p}(p_i(a)) - \epsilon, r_{p,i}\right) < \mathrm{cost}(c) \qquad \text{where} \quad r_{p,i} = \mathrm{end} - p_i(C_p).$$

As one of our main results, we can prove that SSP coincides with cost-efficient Lindahl equilibrium.

Theorem 9.13. *Cost-efficient Lindahl equilibrium results in the same fairness notion as SSP.*

Proof. We first prove that the outcomes that are in a cost-efficient Lindahl equilibrium are SSP. Consider an outcome $W \subseteq C$ that is in the cost-efficient Lindahl equilibrium, and let γ be the corresponding PE price system. From γ we construct the price system $p = (\mathrm{end}, W, \{p_i\}_{i\in N})$ witnessing SSP as follows. For each $i \in N$ and $c \in W$ we set $p_i(c) = \gamma_i(c)$; for $c \notin W$ we set $p_i(c) = 0$. Note that the fact that every candidate $c \in W$ is paid exactly $\mathrm{cost}(c)$ dollars follows from (L1) and (L3).

We now verify that p satisfies SSP: suppose it is not the case. Then, there exists a candidate $c \notin W$ and $\epsilon > 0$ such that:

$$\sum_{i \in N(c)} \max\left(\max_{a \in W}(p_i(a)) - \epsilon, r_{p,i}\right) \geq \mathrm{cost}(c).$$

Consider any voter $i \in N(c)$. Consider two cases:

1. $\max_{a \in W}(p_i(a)) - \epsilon \leq r_{p,i}$. Then let $W' = W \cup \{c\}$. Clearly, we have that $u_i(W') > u_i(W)$, hence, from (L2), we have that:

$$\gamma_i(W') > \mathrm{end} = \mathrm{end} = p_i(W) + r_{p,i} = \gamma_i(W) + r_{p,i}, \tag{9.12}$$

   hence $\gamma_i(c) > r_{p,i} = \max\left(\max_{a \in W}(p_i(a)) - \epsilon, r_{p,i}\right)$.

2. $\max_{a \in W}(p_i(a)) - \epsilon > r_{p,i}$. Let $a = \mathrm{argmax}_{a \in W}(p_i(a))$. Then let $W' = (W \setminus a) \cup \{c\}$. From (L2), we have that either (9.12) holds and then:

$$\gamma_i(c) > \gamma_i(a) + r_{p,i} \geq \max\left(\max_{a \in W}(p_i(a)) - \epsilon, r_{p,i}\right)$$

   or it needs to hold $\gamma_i(W') \geq \gamma_i(W)$ and hence

$$\gamma_i(c) \geq \gamma_i(a) = p_i(a) > \max\left(\max_{a \in W}(p_i(a)) - \epsilon, r_{p,i}\right).$$

In any case we obtain that $\gamma_i(c) > \max\left(\max_{a \in W}(p_i(a)) - \epsilon, r_{p,i}\right)$. Summing this inequality over all $i \in N(c)$:

$$\sum_{i \in N(c)} \max\left(\max_{a \in W}(p_i(a)) - \epsilon, r_{p,i}\right) < \sum_{i \in N(c)} \gamma_i(c) \overset{(L1)}{\leq} \mathrm{cost}(c),$$

a contradiction.

Second, we show that an outcome W that is SSP is in a cost-efficient Lindahl equilibrium. Consider an outcome $W \subseteq C$ that is SSP and let p be the corresponding price system. We know that there exists $\epsilon > 0$ such that for all $c \notin W$ we have that:

$$\sum_{i \in N(c)} \max\left(\max_{c' \in W}(p_i(c')) - \epsilon, r_{p,i}\right) < \mathrm{cost}(c).$$

We construct a price system γ that will witness that W is in the cost-efficient Lindahl equilibrium. For each $i \in N$ and each $c \in W$ we set $\gamma_i(c) = p_i(c)$. For $c \notin W$ and $i \in N(c)$ we set:

$$\gamma_i(c) = \max\left(\max_{c' \in W}(p_i(c')) - \epsilon, r_{p,i}\right).$$

First let us see that $\sum_{i \in N} \gamma_i(c) < \mathrm{cost}(c)$ for all $c \notin W$; hence (L1) is satisfied. Second, observe that for each voter i and each $c \notin W$ buying c costs at least the same as buying any candidate from W; thus, (L2) is satisfied. From the definition of price system we get cost-efficiency and profit maximization for $c \in W$. $\square$

Based on this equivalence, we can immediately deduce several other properties of cost-efficient Lindahl equilibria.

Corollary 9.14. *Cost-efficient Lindahl Equilibria are stable priceable.*

Corollary 9.15. *Every outcome that is in a cost-efficient Lindahl equilibrium is in the core.*

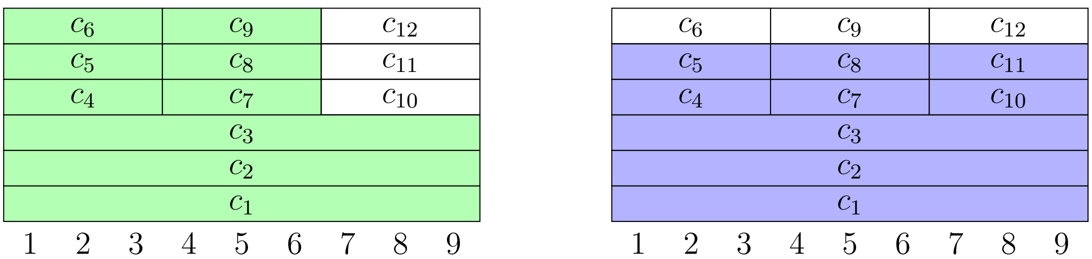
The latter result mirrors Foley's theorem in the classical model [Foley, 1970].

Summarizing, the idea of SP is very close to the idea of Lindahl equilibrium. The key conceptual difference is that in the public economics model, the prices of candidates and voters' payments are fixed elements of the model. In our case, they are adjustable parts of the outcome justification—the voters do not truly have money, they only have preferences, and spending money is a virtual concept that we use to ensure that public decisions are fair.

9.3 Balanced Stable Priceability

So far we have considered priceability notions where two voters could face significantly different prices for the same project. This can seem unnatural—why does one voter need to pay much more for the same thing as another?—and might thereby limit the usefulness of using these price systems as explanations. Consider the following example:

Example 9.16. Consider a committee election with 12 candidates and 9 voters. The voters have the following approval sets. All 9 voters approve candidates c_1, c_2, and c_3. Further, voters $1, 2, 3$ approve c_4, c_5, and c_6; voters $4, 5, 6$ approve c_7, c_8, and c_9; and voters $7, 8, 9$ approve c_{10}, c_{11}, and c_{12}. The committee size is $b = 9$. The election is depicted below.



Here, the committee marked green in the left-hand side of the figure is SP. The corresponding price system can be the following: each from the last three voters (7, 8 and 9) pays $1/3$ for each commonly approved candidate (c_1, c_2 and c_3). The voters $1, 2, 3$ pay $1/3$ for candidates c_4, c_5, and c_6; the voters $4, 5, 6$ pay $1/3$ for candidates c_7, c_8, and c_9. However, the committee is arguably not fair. A much better choice would be to pick the committee marked blue in the right-hand side part of the figure. $\lrcorner$

The reason why the SP solution from Example 9.16 is not fair is that the candidates who are approved by all the voters (candidates c_1, c_2, and c_3), are paid for by only a small subset of them. Example 9.16 shows that the properties of stable priceable outcomes very much depend on the structure of payment functions. Specifically, in Example 9.16 the payment functions were very unbalanced. Even though all voters approved c_1, only $7, 8$, and 9 payed for it. In a way, the mechanism 'stole money' from $7, 8$, and 9, depriving them the possibility of paying for other candidates.

This example suggests that in an ideally-fair price system, all voters who enjoy the same satisfaction from the same project should pay the same amount of money for it. We call such price systems *balanced*.

9.3.1 Formal Definition

The notion of balanced stable priceability differs from the notion that we considered in Section 9.2 in two main aspects. First, we require that the payments of any two voters, i and j, who decide to pay for a given project c must be the same, that is, $p_i(c) = p_j(c)$. Second, we allow a voter not to pay for elected projects—but then, when considering possible deviations, the voter takes no utility from an approved project, even if the project is elected.

Formally, we say that a price system p is *balanced* if the following condition is satisfied for each $i, j \in N$:

$$p_i(c) > 0 \implies p_i(c) = p_j(c).$$

For each voter $i \in N$ and a balanced price system $p = (\mathrm{end}, C_p, \{p_i\}_{i \in N})$, by $C_{p,i} \subseteq C_p$ we denote the set of projects that can be used by i (projects $c \in C_p$ such that $p_i(c) > 0$). Analogously we define the set $C_{q,i}$ for a balanced payment system $q = (C_q, \{q_i\}_{i \in N})$.

We say that a feasible balanced price system $p = (\mathrm{end}, C_p, \{p_i\}_{i \in N})$ is *balanced stable priceable* (BSP) if the following condition is satisfied:

(BSP) There exists no collection $\{R_i\}_{i \in N}$ (with $R_i \subseteq C_{p,i}$ for $i \in N$) and no balanced payment system q such that:

(a) For each voter $i \in N$:

$$p_i(C_p \setminus R_i) + q_i(C_q) \leq \mathrm{end}.$$

Besides, for at least one voter $i \in N$ the above inequality is strict.

(b) For each voter $i \in N$:

$$\Big(u_i((C_{p,i} \setminus R_i) \cup C_{q,i}), p_i(C_p \setminus R_i) + q_i(C_q)\Big) \succeq \Big(u_i(C_{p,i}), p_i(C_p)\Big),$$

Besides, for at least one voter $i \in N$ the above inequality is strict.

We say that an outcome W is *balanced stable priceable* if there exists a balanced stable price system $p = (\mathrm{end}, C_p, \{p_i\}_{i \in N})$ such that $W = C_p$.

Intuitively, (BSP) is similar to (SP) with the additional requirement that both the original payments and the ones witnessing the deviation need to be balanced. Besides, every voter i gains utility only from the projects she pays for.

The green committee in Example 9.16 is not BSP. The price system given in the example violates condition (BSP): all the voters would prefer to share the cost of project c_1 ($q_i(c_1) = 1/9$ for each $i \in N$). The first three voters would prefer to pay for c_1 instead of c_4 ($C_{q,i} = \{c_1\}$, $R_i = \{c_4\}$), since then the number of their representatives would not change—recall that according to our definition of stability, a voter cannot be represented by a project for whom she does not pay—but they would need to pay for them a smaller amount of money (they would need to pay $1/9$ dollars for c_1 versus $1/3$ dollars for c_4). Similarly, voters $4, 5$, and 6 would prefer to pay for c_1 instead of c_7 ($C_{q,i} = \{c_1\}$, $R_i = \{c_7\}$). Finally, the last 3 voters would be happy with the change ($C_{q,i} = \{c_1\}$, $R_i = \{c_1\}$) since the individual price they would need to pay for c_1 would be lower ($1/9$ instead of $1/3$).

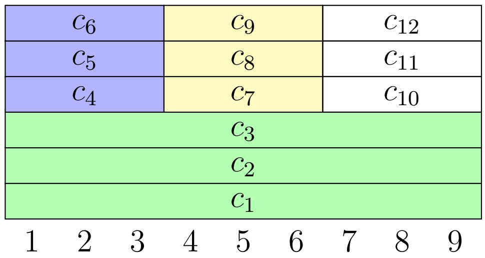


Figure 29: An illustration that a price system justifying the green committee in Example 9.16 is not balanced stable priceable. Now, when checking whether the price system is stable, we assume that voters $1-3$ gain satisfaction only from blue candidates, voters $4-6$ only from the yellow ones and voters $7-9$ only from the green ones. Therefore, voters from the first and second group now have an incentive to participate in paying for green candidates.

9.3.2 Axiomatic Properties of Balanced Stable Priceability

Unlike stable priceability, BSP does not imply priceability (it is possible now that the voters have enough unused money in total to buy a project, but cannot deviate if payments need to be equal).

Like in the case of SP, imposing $|C_q| \leq 1$ in the definition of BSP does not reduce the strength of the notion. Indeed, below we present the analogue of Theorem 9.4 for (BSP) and a suitably modified inequality (SP1):

(BSP1) For all projects $c \in C$ and all groups of voters $S \subseteq N(c)$:

$$|S| \min_{i \in S} s_i \leq \mathrm{cost}(c)$$

where for each $i \in N(c)$:

$$s_i = \begin{cases} p_i(c) & \text{if } p_i(c) > 0, \\ \max\left(\max_{a \in W}\left(p_i(a)\right), r_{p,i}\right) & \text{otherwise.} \end{cases}$$

Theorem 9.17. *Conditions (BSP1) and (BSP) are equivalent.*

Proof. We first prove that condition (BSP) implies (BSP1). Indeed, assume first that there exists a BSP price system p and a project $c \in C$ such that for some $S \subseteq N(c)$ condition (BSP1) does not hold. As a result, it holds that:

$$\frac{\mathrm{cost}(c)}{|S|} < \min_{i \in S} s_i.$$

Consequently, for some $\epsilon > 0$:

$$\frac{\mathrm{cost}(c)}{|S|} < \min_{i \in S} s_i - \epsilon.$$

Set $C_q = \{c\}$. For each $i \in S$, $q_i(c) = s_i - \epsilon$ and R_i is set as follows:

(a) $R_i = \emptyset$ if $s_i = r_{p,i}$,

(b) $R_i = \{c\}$, if $s_i = p_i(c)$ (note that in such case $c \in W$),

(c) $R_i = \mathrm{argmax}_{a \in W}\left(p_i(a)\right)$ if $s_i = \max_{a \in W}\left(p_i(a)\right)$.

Then, we can clearly see that $\{R_i\}$ and q satisfy all the conditions of (BSP), hence they witness the lack of stability.

To prove that (BSP1) implies (BSP), assume that for some balanced price system p, there exist a balanced price system q and $\{R_i\}_{i \in S}$, witnessing the violation of (BSP). Among all possible witnesses, consider the one minimizing $|C_q|$. We will prove that in this case condition (BSP1) is not satisfied.

Set c to the project from C_q purchased by the largest group of voters (denote that group by S). Assume that $c \notin W$. Fix any $i \in S$—note that, as payments are balanced, every voter from S pays $q_i(c)$ dollars for c, hence:

$$|S| \cdot q_i(c) > \mathrm{cost}(c) \tag{9.13}$$

We know that $u_i(R_i) = |R_i| \leq |C_{q,i}|$. If we have that $|R_i| = |C_{q,i}|$, then

$$q_i(C_{q,i}) < p_i(R_i).$$

and:

$$q_i(c) = \frac{q_i(c) \cdot |C_{q,i}|}{|C_{q,i}|} \leq \frac{\sum_{a \in C_{q,i}} q_i(a)}{|C_{q,i}|} = \frac{q_i(C_{q,i})}{|C_{q,i}|} \leq \frac{p_i(R_i)}{|C_{q,i}|} \leq \frac{|R_i| \cdot (\max_{a \in W}(p_i(a)))}{|C_{q,i}|}$$
$$= \max_{a \in W}(p_i(a)) \leq \max\left(r_{p,i}, \max_{a \in W}(p_i(a))\right) = s_i$$

Otherwise, we have that $|R_i| = u_i(R_i) < u_i(C_{q,i}) = |C_{q,i}|$ and $q_i(C_{q,i}) \leq r_{p,i} + p_i(R_i)$. In this case:

$$\begin{aligned}
q_i(c) &= \frac{q_i(c) \cdot |C_{q,i}|}{|C_{q,i}|} = \frac{\sum_{a \in C_{q,i}} q_i(a)}{|C_{q,i}|} \leq \frac{\sum_{a \in C_{q,i}} q_i(a)}{|C_{q,i}|} \\
&= \frac{q_i(C_{q,i})}{|C_{q,i}|} \leq \frac{r_{p,i} + p_i(R_i)}{|C_{q,i}|} \leq \frac{r_{p,i} + |R_i| \cdot \max_{a \in W}(p_i(a))}{|C_{q,i}|} \\
&\leq \frac{(|R_i| + 1) \cdot \max\left(r_{p,i}, \max_{a \in W}(p_i(a))\right)}{|C_{q,i}|} \leq \max\left(r_{p,i}, \max_{a \in W}(p_i(a))\right) = s_i.
\end{aligned}$$

Hence, in both cases we have that:

$$q_i(c) \leq \min_{i \in S} s_i. \tag{9.14}$$

Combining (9.13) and (9.14) we obtain:

$$|S| \cdot \min_{i \in S} s_i > \mathrm{cost}(c),$$

which completes the proof of the case when $c \notin W$.

Now assume that $c \in W$. Let $S' = \{i \in N : c \in C_{p,i}\}$. We have three cases now:

Case 1: It holds that $q_i(c) \leq p_i(c) = s_i$ for every $i \in S$. Then the condition (BSP1) is clearly violated for S and c.

Case 2: $S \subseteq S'$ and it holds that $q_i(c) > p_i(c)$ for every $i \in S$. Then we can remove c from every set R_i and fix $q_i(c) = 0$ for each $i \in N$—and the collection $\{R_i\}_{i \in N}$ together with q would still violate condition (BSP), a contradiction (as the witness was chosen to minimize $|C_q|$).

Case 3: $S \nsubseteq S'$ and it holds that $q_i(c) > p_i(c)$ for every $i \in S$. Consider a voter $j \in S \setminus S'$ and set $X = \{i \in N : c \in C_{p,i}\}$. Set payments $q'_i(c) = 1/|S'|+1$ for each $i \in X \cup \{j\}$. It is clear that $q'_i(c) = 1/|S'|+1 < 1/|S'| = p_i(c) = s_i$ for $i \in S'$. For voter j, we can repeat the reasoning from the case $c \notin W$ to obtain that $q_j(c) \leq s_j$, hence also $q'_j(c) \leq s_j$. Finally, set $S' \cup \{j\}$ and project c witness the violation of (BSP1).

□

This result allows us to characterize BSP outcomes as the ones that are returned by the polynomial-time algorithm presented in Algorithm 4. This algorithm is a variant of Method of Equal Shares, with three modifications. First, we assume that we are given the initial endowment end of each voter as a part of the input. Second, the payments of the voters need to *exactly* equal; if a voter cannot afford paying an equal part of a price for a candidate c she support, she pays nothing (and hence, gains no satisfaction from c even if elected). The third modification distinguishes *demanding* and *strongly demanding* groups of voters and only a latter is guaranteed to be satisfied.

Algorithm 4: Algorithm for BSP price systems

Input: An election $E = (N, C, b)$, initial endowment end.
Output: A BSP price system p

for $i \in N$ **do**
    for $c \in C$ **do**
        $p_i(c) \leftarrow 0$;
        $r_{p,i} \leftarrow \mathrm{end}$;
while *there exists a demanding group* **do**
    $Q_a \leftarrow$ the demanding group minimizing $cost(a)/|S|$;
    $Q_b \leftarrow$ the strongly demanding group minimizing $cost(b)/|S|$;
    $Q_c \leftarrow$ either Q_a or Q_b, depending on some tie-breaking over demanding groups;
    for $i \in Q_c$ **do**
        $p_i(c) \leftarrow \mathrm{cost}(c)/|Q_c|$;
        $r_{p,i} \leftarrow r_{p,i} - p_i(c)$;
return $\{p_i\}_{i \in N}$;

We say that a group of voters Q_c is *demanding* if every member of Q_c approves some project $c \notin W$ and has at least $^{\mathrm{cost}(c)}/_{|Q_c|}$ dollars left. We say that Q_c is *strongly demanding* if it demanding and every voter from Q_c has strictly more than $^{\mathrm{cost}(c)}/_{|Q_c|}$ dollars left. Intuitively, every demanding group can afford to additionally buy a new project to the outcome. Strongly demanding groups can afford to buy a project even exceeding the price. Algorithm 4 can be viewed as an algorithm greedily satisfying these demands, starting from the ones minimizing $^{\mathrm{cost}(c)}/_{|Q_c|}$. If a group is not strongly demanding, it can be either considered or skipped. The algorithm stops when there are no nonempty, non-skipped demanding groups.

Theorem 9.18. *Every price system elected by Algorithm 4 is BSP. Every BSP price system is possible to be elected by Algorithm 4.*

Proof. First we show that every price system elected by Algorithm 4 is BSP. It is clear that this price system is feasible and balanced. Now, we show that it also satisfies (BSP1). Indeed, for the sake of contradiction suppose that it does not hold and let $c \in C$, and $S \subseteq N(c)$ be witnessing the violation of stable priceability. Hence, every voter $v \in S$ has more than $^{\mathrm{cost}(c)}/_{|S|}$ money left or pays for some project from W (possibly c) more than $^{\mathrm{cost}(c)}/_{|S|}$. If for all the voters the first option is true, then at the end of execution of the algorithm S is a strongly demanding group, thus the algorithm would not stop. Otherwise, let c' be the project added to W at the earliest step such that each voter pays for c' more than $^{\mathrm{cost}(c)}/_{|S|}$. As the individual price for c' is greater than for c, it needs to be the case that c' was added to W for some demanding group of size smaller than $|S|$. However, before that step, group S was a strongly demanding group. Hence, the demanding group supporting c' was not the largest one, a contradiction.

Now we show that every BSP price system is possible to be elected by Algorithm 4. For the sake of contradiction suppose $p = (\mathrm{end}, W, \{p_i\}_{i \in N})$ is a BSP price system that cannot be elected by Algorithm 4. Let $|W| = \ell$. Enumerate the projects $c_1, \ldots, c_\ell$ in W by number of voters who pay for them in the descending order and denote these groups by $Q_{c_1}, \ldots, Q_{c_\ell}$. Consider a rule, which is similar to Algorithm 4, but instead of taking the demanding group Q_c minimizing $^{\mathrm{cost}(c)}/_{|Q_c|}$, considers only groups $Q_{c_1}, \ldots, Q_{c_\ell}$. The only difference between such a rule and Algorithm 4 may appear if at some ith iteration, there exists some strongly demanding group Q_c which has strictly lower value of $^{\mathrm{cost}(c)}/_{|Q_c|}$ than Q_{c_i}. Such a group, if ignored, either remains strongly demanding at the end of the execution of the algorithm (and then is a witness for violating stability of W) or stops being strongly demanding at some step (after some voters from Q_c have their initial budgets decreased). However, individual payments for the projects added to W in further steps need to be strictly higher than $^{\mathrm{cost}(c)}/_{|Q_c|}$. Hence, Q_c is still the witness for violating stability of W. □

This characterization allows us to use BSP as an explanation of the output of Method of Equal Shares. The characterization also makes it easy to establish logical relations to other properties; for instance, BSP implies EJR; the proof is the same as the proof that Equal Shares satisfies EJR for approval-based utilities (which follows from Theorem 4.11). Note that here, in contrast to SP, we do not need to require exhaustiveness.

Proposition 9.19. *BSP implies EJR.*

Based on the characterization from Theorem 9.18, we can design a polynomial-time heuristic algorithm for finding BSP outcomes of a specified cost b. We use binary search to find the endowment for which Algorithm 4 (as defined in Section 9.3.2) finds the outcomes of the closest cost to b as possible. Besides, we do not skip demanding groups which are not strongly demanding, until the outcome is exhaustive.

This algorithm is heuristic, for two reasons: (1) our adopted tie-breaking over demanding groups is not the only possible one, and (2) the size of outcomes elected by Algorithm 4 is not monotonous with respect to the initial endowment, as it has been shown in Example 3.5.

However, our experiments show that in the committee election model, this algorithm very often successfully manages to find BSP committees.

9.4 Satisfiability of Market-Based Axioms

The most pressing question is whether exhaustive SP and BSP outcomes exist for all elections. The answer is negative for both notions, even for the committee election model.

Theorem 9.20. *There exists a committee election for which there exists no exhaustive stable priceable committee.*

Proof. Consider the following committee election with 6 voters and candidates $\{a, c, d, e\}$:

$$1\colon \{a, c\} \qquad 2\colon \{a, d\} \qquad 3\colon \{a, e\} \qquad 4\colon \{c, d\} \qquad 5\colon \{c, e\} \qquad 6\colon \{d, e\}$$

Let $b = 2$. As the election is symmetrical, without the loss of generality we can fix $W = \{a, c\}$.

Suppose that there exists a price system $p = (\mathrm{end}, C_p, \{p_i\}_{i \in N})$ with $C_p = W$ satisfying (SP1). As voters pay only for approved candidates, we have that $p_2(c) = p_3(c) = p_4(a) = p_5(a) = p_6(a) = p_6(c) = 0$. From that we obtain:

$$p_1(a) + p_1(c) + p_2(a) + p_3(a) + p_4(c) + p_5(c) = 2.$$

From (SP1) we have that:

$$r_{p,6} + p_4(c) + p_2(a) \leq 1, \text{ because of candidate } d,$$
$$r_{p,6} + p_5(c) + p_3(a) \leq 1, \text{ because of candidate } e.$$

Hence:

$$2 \cdot r_{p,6} + p_2(a) + p_3(a) + p_4(c) + p_5(c) \leq 2 = p_1(a) + p_1(c) + p_2(a) + p_3(a) + p_4(c) + p_5(c)$$
$$2 \cdot \mathrm{end} = 2 \cdot r_{p,6} \leq p_1(a) + p_1(c) \leq \mathrm{end}$$

a contradiction. □

Theorem 9.21. *There exists a committee election for which there exists no exhaustive balanced stable priceable committee.*

Let us first prove the following lemma:

Lemma 9.22. *Consider a committee election, two voters i, j and a BSP price system p. If $A_i = A_j$, then for each $c \in C$ we have that $p_i(c) = p_j(c)$.*

Proof. Suppose for the sake of contradiction that there exists a candidate a for which i pays $1/x_a$ dollars and j pays nothing. Then either j has at least $1/x_a$ dollars left or there exists a candidate b for which j pays and i does not. In the first case, j prefers to join the other x_a voters paying for a (then the individual price for a decreases to $1/x_a+1$, which is strictly less than the amount of the savings of j). In the second case, compare the individual payment for b (equal to $1/x_b$) to $1/x_a$. Without loss of generality, assume that $1/x_a \leq 1/x_b$. Then j has an incentive to stop paying for b and join other q_a voters paying for a—then again the individual price for a decreases to $1/x_a+1$, which is strictly less than $1/x_b$. We obtained a contradiction with the assumption that the committee is BSP, which proves that for each $c \in C$ it holds that $p_i(c) = p_j(c)$. □

Proof of Theorem 9.21. Let us consider the following election for $n = 49, m = 46, b = 44$:

$$
\begin{aligned}
&\text{Group 1 (16 voters)}: && \{x_1, \ldots, x_{18}\}\\
&\text{Group 2 (8 voters)}: && \{y_1, \ldots, y_{12}\}\\
&\text{Group 3 (3 voters)}: && \{z_1, \ldots, z_{12}\}\\
&\text{Group 4 (4 voters)}: && \{a, x_1, \ldots, x_{18}\}\\
&\text{Group 5 (6 voters)}: && \{a, b, y_1, \ldots, y_{12}\}\\
&\text{Group 6 (4 voters)}: && \{c, z_1, \ldots, z_{12}\}\\
&\text{Group 7 (4 voters)}: && \{d, z_1, \ldots, z_{12}\}\\
&\text{Group 8 (2 voters)}: && \{b, c, z_1, \ldots, z_{12}\}\\
&\text{Group 9 (2 voters)}: && \{b, d, z_1, \ldots, z_{12}\}
\end{aligned}
$$

In total, 20 voters approve candidates $\{x_1, \ldots, x_{18}\}$, 14 voters approve candidates $\{y_1, \ldots, y_{12}\}$, 15 voters approve candidates $\{z_1, \ldots, z_{12}\}$, 10 voters approve a, 10 voters approve b, 6 voters approve c and 6 voters approve d. Further, the sets of voters who approve x-, y-, and z-candidates are disjoint. Hence, Algorithm 4 will elect candidates $x_1, \ldots, x_{18}, z_1, \ldots, z_{12}, y_1, \ldots, y_{12}$ first. After that, we have at most 42 candidates elected. As our goal is to elect a committee of size 44, the value of the voter's endowment in a BSP price system end should allow all those 42 candidates to be elected. After that:

1. Voters from groups 1-3 run out of approved candidates.
2. Each voter from group 4 has $\mathrm{end} - {}^{9}\!/_{10}$ dollars left.
3. Each voter from group 5 has $\mathrm{end} - {}^{6}\!/_{7}$ dollars left.
4. Each voter from groups 6-9 has $\mathrm{end} - {}^{4}\!/_{5}$ dollars left.

We need to elect exactly 2 more candidates. Assume for the sake of contradiction that it is possible and let us consider all possible pairs of candidates from $\{a, b, c, d\}$ as the ones that can be included in the final committee.

Case 1: $\{a, b\}$. As we have at least 4 voters (group 6) who approve c and have $\mathrm{end} - {}^{4}\!/_{5}$ dollars left, the following inequality holds:

$$\frac{1}{4} \geq \mathrm{end} - \frac{4}{5}$$

$$\mathrm{end} \leq \frac{21}{20} = 1.05 \tag{9.15}$$

Suppose that there exist a voter paying for both a and b (from group 5). As each voter from this group has $\mathrm{end} - {}^{6}\!/_{7}$ dollars left, at most 10 voters pay for a (groups 4 and 5) and at most 10 voters pay for b (groups 5, 8, 9) we would have the following inequality:

$$\frac{1}{10} + \frac{1}{10} \leq \mathrm{end} - \frac{6}{7}$$

$$14p \leq 70 - 60p$$

$$\mathrm{end} \geq \frac{74}{70} \approx 1.057 \tag{9.16}$$

which contradicts (9.15). Hence, we need to assume that no voter pays for both a and b. By Lemma 9.22 we have only two cases: either no voters from group 5 pays for b or no voters from group 5 pays for a.

If no voters from group 5 pays for b, then only at most 4 voters do so. Suppose that at least one voter from group 8 pays for b (the case of group 9 is analogous) the individual price which is at least $\frac{1}{4}$ dollars. Then, $\frac{1}{4} \leq \mathrm{end} - \frac{4}{5}$. In this case, all 6 voters approving c would prefer to pay $\frac{1}{6}$ dollars for c instead of paying for b or (in case of voters not paying for b) from their savings—which is sufficient, as $\frac{1}{6} < \frac{1}{4} \leq \mathrm{end} - \frac{4}{5}$. We obtain a contradiction.

Now assume that no voters from group 5 pay for a. Then only at most 4 voters from group 4 do so. Then we have a following inequality:

$$\frac{1}{4} \le \text{end} - \frac{9}{10}$$
$$\text{end} \ge \frac{23}{20} = 1.15 \tag{9.17}$$

which also contradicts (9.15).

Case 2: {a,c}. In this case, 6 voters from groups 7 and 9 shall not be able to pay for d. Hence we have the following inequality:

$$\frac{1}{6} \ge \text{end} - \frac{4}{5}$$
$$\text{end} \le \frac{29}{30} < 1 \tag{9.18}$$

Suppose now that only voters from group 5 pay for a. Then the following inequality needs to hold:

$$\frac{1}{6} \le \text{end} - \frac{6}{7}$$
$$\text{end} \ge \frac{43}{42} > 1 \tag{9.19}$$

which contradicts (9.18).

Now suppose that at least one voter from group 4 pays for a. As in total there are 10 voters approving a, the following inequality holds:

$$\frac{1}{10} \le \text{end} - \frac{9}{10}$$
$$\text{end} \ge 1 \tag{9.20}$$

which also contradicts (9.18).

Case 3: {a,d}. The reasoning here is analogous as in Case 2 (inequality (9.18) still holds because of candidate c and voters from groups 6 and 8).

Case 4: {b,c}. The reasoning in this case is similar to the one in Case 1. First, note that 10 voters approving a shall not be able to pay for this candidate. Hence:

$$\frac{1}{10} \ge \text{end} - \frac{9}{10}$$
$$\text{end} \le 1 \tag{9.21}$$

(an opposite inequality to (9.20)).

Suppose that there exist a voter paying for both b and c (from group 8). As each voter from this group has ${}^{5-4p}/_{5}$ dollars left, at most 10 voters pay for b (groups 5, 8, 9) and at most 6 voters pay for c (groups 6, 8) we have:

$$\frac{1}{6} + \frac{1}{10} \le \text{end} - \frac{4}{5}$$

$$\text{end} \geq \frac{128}{120} > 1 \tag{9.22}$$

which contradicts (9.21). Hence, we assume that no voter pays for both c and b. Note that by Lemma 9.22 we have only two cases: either voters from group 8 pay for b or for c. The second option is not possible—if the voters pay for c, then they pay at least $1/6$, while voters paying for b pay $1/8$. Hence, they have an incentive to stop paying for c and start paying for b.

Now assume that voters from group 8 pay for b. Then only at most 4 voters from group 6 pay for c. Then we have a following inequality:

$$\frac{1}{4} \leq \text{end} - \frac{4}{5}$$

$$\text{end} \geq \frac{21}{20} > 1 \tag{9.23}$$

(the opposite inequality to (9.15)), which also contradicts (9.21).

Case 5: $\{b, d\}$. This case is analogous to Case 4 (swapping candidates c and d).

Case 6: $\{c, d\}$. In this case the voters from groups 6 and 8 pay for c and the voters from groups 7 and 9 pay for d.

Consider the group of all 10 voters approving b who may want to start paying for b (and in case of groups 8 and 9, stop paying for c and d—as $1/10 < 1/6$, it is always profitable for them). To prevent them, the price needs to be high enough so that voters from group 5 do not have enough money to spend. Hence we have that:

$$\frac{1}{10} > \text{end} - \frac{6}{7}$$

$$\text{end} \leq \frac{67}{70} \approx 0.957 \tag{9.24}$$

On the other hand, as at most 6 voters pay for c, we have that:

$$\frac{1}{6} \leq \text{end} - \frac{4}{5}$$

$$\text{end} \geq \frac{29}{30} \approx 0.967 \tag{9.25}$$

which contradicts (9.24). □

9.4.1 Experiments on Synthetic Data

Let us now describe the experiments that we conducted for synthetic distributions of voters' preferences. We have checked only the committee election model—we leave the experiments with various costs of the projects for the future research. We found that it is almost always possible to find an SP/BSP outcome spending a large part of the available budget (which in case of SP, as we noted before, guarantees good proportionality properties). Then we can spend the remaining part in a different way or complement the outcome with a different algorithm, obtaining an outcome large part of which is SP/BSP.

Datasets

We generated committee elections randomly from the following models:

1D-Euclidean model. In this model the voters and candidates are represented as points in the one-dimensional Euclidean space. The points were selected uniformly at random form the interval $[0, 1]$. The approval ballots were created in one of the following two ways:

1. For each candidate we chose uniformly at random the length of the radius of the approval ball—every candidate was approved only by the voters inside her ball.
2. The radii were chosen for each voter, and every voters approved the candidates inside her ball.

In our results, we refer to these two models as "1D Euclidean 1" and "1D Euclidean 2".

2D-Euclidean model. Here, we represent the voters and the candidates as points in the Euclidean plane $[0, 1] \times [0, 1]$. The points were generated as follows. We first generated between 1 and 5 points of concentration of the voters and the candidates. Next, we randomly divided the voters and the candidates so that each of them was assigned to one point of concentration. Finally, for each voter and candidate we selected their position from the normal distribution with the center at the corresponding point of concentration and with the standard deviation set to 0.2.

We generated the approval-ballots from the positions of the voters and the candidates similarly as in the first case in the 1D-Euclidean model.

Impartial Culture model. Here, each candidate was approved by each voter with probability $1/2$.

Mallow's model. Here we first generate a ordinal preference profile according to the mixture of three Mallow's models [Mallows, 1957]. The parameters ϕ for the three models were generated uniformly at random from $[0, 1]$; the reference rankings were also selected uniformly at random. Next, for each voter we selected uniformly at random a position $\varrho \in [0; 0.25 \cdot m]$, and made this voter approve the first ϱ candidates in her ranking.

Pólya-Eggenberger urn model. Here our model is parameterized by the size of the approval sets α and the replacement value β. At first, we consider an urn containing all the candidates; for each voter we draw α candidates from the urn uniformly at random, and each time we return to the urn β copies of the selected candidate—increasing the probability that next time the same candidate would be chosen again. In our tests parameter α was chosen uniformly at random from interval $[1; 0.1 \cdot m]$, and parameter β was chosen in two ways:

1. uniformly at random from interval $[0; 0.1 \cdot m]$,
2. uniformly at random from interval $[0; 0.25 \cdot m]$.

In our results, we refer to these two models as "Urn 1" and "Urn 2".

Results for BSP

Let us first describe the results of the experiments for BSP committees for $n = 100$ voters and $m = 30$ candidates. The value of the b was sequentially increased, covering all the values from $[m]$. For each b and each model with specific values of the parameters was run at least 1000 experiments. Thus, in total we checked $1000 \cdot 30 \cdot 7 = 210000$ elections.

In Table 9.3 we present the results for existence for every model with fixed parameters and every value of b. Summarizing, in most cases BSP committees exist— 2D Euclidean model appeared to be the worst one, especially when $b \geq 20$. Therefore, in Table 9.4 we present the additional results for that model, showing that in most cases, even if BSP committees do not exist, they exist for slightly decreased committee size b.

Results for SP

In case of SP, we were able to perform tests for $n = 100$ voters and $m = 30$ candidates (analogous as in case of BSP) for all models except for Impartial Culture. In case of this model, because of complexity issues we decided to reduce the number of voters to 50 and the number of candidates to 15 (the number of experiments for each $b \in [m]$ remained 1000). Furthermore, the results (when b is small) were the worst for this model, even on such reduced elections (which is somehow surprising, as they were the best in case of BSP). The detailed results—in particular, checking how much we need to deviate b to obtain SP—are presented in Table 9.1.

For other models—also for Euclidean 2D, which appeared to be the worst for the existence of BSP—it was hard to even find any election not satisfying SP, as we can see in Table 9.2.

b	Existence	Max non-exhaustiveness	Average non-exhaustiveness
2	587	1	1.00
3	649	2	1.76
4	842	3	2.58
5	948	4	3.46
6	968	5	4.41
7	999	6	6.00

Table 9.1: The detailed results of the experiments for the existence of SP committees under the Impartial Culture model. We present here only the values of b for which exhaustive outcomes do not exist in some of the sampled elections. We present (1) the number of elections (out of 1000) where SP committees exist, (2) the maximal difference between the size of SP committee and b among the elections where SP committees do not exist, (3) the average difference between the size of SP committee and b among the elections where SP committees do not exist.

b	1D Euclidean 1	1D Euclidean 2	2D Euclidean	IC	Urn 1	Urn 2	Mallows
1	1000	1000	1000	1000	1000	1000	1000
2	1000	1000	996	587	1000	990	996
3	1000	1000	1000	649	1000	1000	997
4	1000	1000	1000	842	1000	1000	1000
5	1000	1000	1000	948	1000	1000	1000
6	1000	1000	1000	968	1000	1000	1000
7	1000	1000	1000	999	1000	1000	1000
8	1000	1000	1000	1000	1000	1000	1000
9	1000	1000	1000	1000	1000	1000	1000
10	1000	1000	1000	1000	1000	1000	1000
11	1000	1000	1000	1000	1000	1000	1000
12	1000	1000	1000	1000	1000	1000	1000
13	1000	1000	1000	1000	1000	1000	1000
14	1000	1000	1000	1000	1000	1000	1000
15	1000	1000	1000	1000	1000	1000	1000
16	1000	1000	1000	-	1000	1000	1000
17	1000	1000	1000	-	1000	1000	1000
18	1000	1000	1000	-	1000	1000	1000
19	1000	1000	1000	-	1000	1000	1000
20	1000	1000	1000	-	1000	1000	1000
21	1000	1000	1000	-	1000	1000	1000
22	1000	1000	1000	-	1000	1000	1000
23	1000	1000	1000	-	1000	1000	1000
24	1000	1000	1000	-	1000	1000	1000
25	1000	1000	1000	-	1000	1000	1000
26	1000	1000	1000	-	1000	1000	1000
27	1000	1000	1000	-	1000	1000	1000
28	1000	1000	1000	-	1000	1000	1000
29	1000	1000	1000	-	1000	1000	1000
30	1000	1000	1000	-	1000	1000	1000

Table 9.2: Existence of SP committees for various values of b under various sampling models. For the Impartial Culture model, the number of candidates was reduced to 15 for complexity reasons. The case where for all 1000 sampled elections we have found exhaustive SP committees was marked with green, while the case where we have found exhaustive SP committees for less than 95% of sampled elections was marked with red. In particular, we can easily see that the worst results were obtained for Impartial Culture.

b	**1D Euclidean 1**	**1D Euclidean 2**	**2D Euclidean**	**IC**	**Urn 1**	**Urn 2**	**Mallows**
1	1000	1000	1000	1000	1000	1000	1000
2	1000	1000	1000	1000	1000	1000	1000
3	1000	999	999	1000	1000	1000	1000
4	996	998	999	1000	995	1000	1000
5	998	999	993	1000	999	1000	1000
6	995	999	996	1000	999	1000	999
7	991	998	989	1000	992	999	999
8	996	998	991	1000	991	998	998
9	991	999	993	1000	986	999	998
10	995	997	985	1000	981	998	998
11	997	999	992	1000	976	998	995
12	995	999	994	1000	972	999	999
13	994	999	988	1000	978	997	995
14	993	999	989	1000	981	999	998
15	994	998	983	1000	971	999	997
16	995	1000	980	1000	980	999	996
17	994	1000	970	1000	976	999	994
18	992	998	962	1000	979	998	996
19	994	997	960	1000	994	997	998
20	994	998	930	1000	983	1000	998
21	995	998	905	1000	998	1000	997
22	995	999	887	1000	997	999	996
23	989	999	872	1000	1000	1000	995
24	994	998	855	1000	1000	1000	996
25	997	997	865	1000	1000	1000	997
26	992	998	856	1000	1000	1000	1000
27	994	998	891	1000	1000	1000	999
28	998	1000	891	1000	1000	1000	1000
29	1000	994	934	1000	1000	1000	1000
30	1000	1000	1000	1000	1000	1000	1000

Table 9.3: Existence of BSP committees for various values of b under various sampling models. The case where for all 1000 sampled elections we have found exhaustive BSP committees was marked with green, while the case where we have found exhaustive BSP committees for less than 95% of sampled elections was marked with red. In particular, we can easily see that the best results were obtained for Impartial Culture and the worst ones for 2D Euclidean.

b	**Existence**	**Max non-exhaustiveness**	**Average non-exhaustiveness**
3	999	1	1.00
4	999	1	1.00
5	993	1	1.00
6	996	1	1.00
7	989	1	1.00
8	991	1	1.00
9	993	1	1.00
10	985	2	1.20
11	992	1	1.00
12	994	1	1.00
13	988	1	1.00
14	989	1	1.00
15	983	1	1.00
16	980	4	1.30
17	970	2	1.30
18	962	3	1.13
19	960	3	1.30
20	930	3	1.17
21	905	3	1.22
22	887	4	1.28
23	872	3	1.23
24	855	5	1.22
25	865	5	1.30
26	856	4	1.27
27	891	3	1.21
28	891	4	1.17
29	934	3	1.15

Table 9.4: The detailed results of the experiments for the existence of BSP committees under 2D-Euclidean model. We present here only the values of b for which exhaustive outcomes do not exist in some of the sampled elections. We present (1) the number of elections (out of 1000) where BSP committees exist, (2) the maximal difference between the size of BSP committee and b among the elections where BSP committees do not exist, (3) the average difference between the size of BSP committee and b among the elections where BSP committees do not exist.

9.4.2 Experiments on the Real-Life Data

Let us now present the results of the experiments using the real-life data from *Pabulib*. Here for complexity reasons we present only the results for balanced stable priceability. The development of faster ways of computing SP outcomes so that they can be applied to the data from *Pabulib* is left for future research.

We have performed experiments for the citywide elections in all the cities that are currently available on *Pabulib*. Here we present only the results for the greatest ones (Gdańsk, Kraków, Warsaw, Wrocław), yet the remaining results are consistent. As we can see, only in two presented elections (Gdansk 2020, Kraków 2019) we have found exhaustive BSP outcomes. However, in all of them we have found BSP outcomes spending over 92% of the available budget, and in multiple cases we have spent nearly 100% of the budget. It shows that even if we cannot find an exhaustive BSP outcome, it is easy to find the one spending a large fraction of the available budget.

Election	**Exhaustive**	**Budget used (PLN)**	**Available budget (PLN)**	**Ratio**
Gdańsk 2020	✓	17,408,200	18,543,608	94%
Kraków 2018		11,467,289	12,455,000	92%
Kraków 2019	✓	29,978,087	30,000,000	100%
Kraków 2020		31,938,940	32,000,000	100%
Kraków 2021		34,928,590	34,999,991	100%
Kraków 2022		37,829,973	37,999,992	100%
Warsaw 2017		57,953,933	58,046,682	100%
Warsaw 2018		61,237,180	61,294,114	100%
Warsaw 2019		64,372,671	64,647,778	100%
Warsaw 2020		82,598,915	83,111,363	99%
Warsaw 2021		82,819,186	83,111,363	100%
Warsaw 2022		93,458,653	93,575,094	100%
Warsaw 2023		100,135,016	101,130,815	99%
Wroclaw 2015		59,897,250	60,000,000	100%
Wroclaw 2016		17,771,000	18,000,000	99%
Wroclaw 2017		17,370,000	18,000,000	96%
Wroclaw 2018		18,176,000	18,250,000	100%
Wroclaw 2019		24,682,000	25,000,000	99%
Wroclaw 2020		24,300,000	25,000,000	97%
Wroclaw 2021		24,285,000	25,000,000	97%

Table 9.5: The existence of BSP outcomes on the citywide data from *Pabulib*. We present (1) the information whether the found BSP committee W is exhaustive, (2) the cost of W, (3) the available budget b, (4) the value of the fraction ${}^{\mathrm{cost}(W)}/_{b}$, rounded to the nearest whole percent (in particular, the value "100%" means here over 99.5%).

9.5 Market-Based Axioms for Cardinal Utilities

Our notions extend naturally to the more general setting of cardinal utilities. The definitions of stable priceability (SP, condition (SP)), strict stable priceability (SSP, condition (SP) without the tie-breaking condition in (SP-a)), and Lindahl equilibrium remain unchanged.

Note that the original definition of a price system required that no voter pays for projects she gains no utility from. This restriction was well-justified in the approval-based setting, but in the general PB model it would have a significantly limited scope of impact (it would not put any restrictions on the payments when the utilities $u_i(c)$ are very small, yet still positive). Besides, it is implied by (SP):

Observation 9.23. Consider a price system p, a project $c \in C$ and a voter $i \in N$ such that $u_i(c) = 0$. If p is stable priceable, then $p_i(c) = 0$.

Proof. Indeed, otherwise it would be enough to consider a payment system q with an empty C_q and set $R_i = \{c\}$ ($R_j = \emptyset$ for all $j \neq i$). It is clear that such a construction witness the violation of (SP). □

For balanced stable priceability (BSP), we only change the definition of a balanced price system. Now instead of requiring that the payments for each purchased project c are equal for all the voters, we require that they are proportional to their utilities from c. Formally, for each project c, there exists a value $\rho(c)$ such that for each voter $i \in N$:

$$p_i(c) > 0 \implies p_i(c) = \rho(c) \cdot u_i(c).$$

Except for that change, the definition of BSP remains unchanged.

Unfortunately, the simpler formulation of both SP and BSP ((SP1) and (BSP1), respectively) does not hold in the cardinal model. It implies that our ILP formulation of SP is no longer valid. Besides, now BSP cannot be characterized by a Equal-Shares-like algorithm (electing in each step project c minimizing $\rho(c)$), as we can see in the following example:

Example 9.24. Consider an election with one voter i and two projects c_1, c_2. We have that $\mathrm{cost}(c_1) = 1$, $u_i(c_1) = 1$, $\mathrm{cost}(c_2) = 3$, $u_i(c_2) = 2$. Let $b = 3$. Now Equal Shares would elect $\{c_1\}$ (with $\rho(c_1) = 1$), while $\{c_2\}$ witness the violation of BSP. ⌟

On the other hand, an SP price system p now does not need to imply the core (unless we require $\mathrm{end} > b/n$), as it can be seen by the following example:

Example 9.25. Consider a committee election with four voters, $b = 2$ and the following preferences:

$$\begin{array}{lll} u_1(a) = 1 & u_1(b) = 0 & u_1(c) = 0 \\ u_2(a) = 1 & u_2(b) = 2 & u_2(c) = 0 \\ u_3(a) = 0 & u_3(b) = 2 & u_3(c) = 1 \\ u_4(a) = 0 & u_4(b) = 0 & u_4(c) = 1 \end{array}$$

Now the committee $W = \{a, c\}$ is SP as witnessed by the following price system with $\mathrm{end} = {}^1\!/_2$: $p_1(a) = p_2(a) = p_3(c) = p_4(c) = {}^1\!/_2$. However, the group $S = \{2, 3\}$ and a committee $T = \{b\}$ witness the core violation. ⌟

The main idea of the above example is that the middle voters have only the enough amount of money to buy $\{b, c\}$, but not to overpay for it (as required by the condition (SP)). Actually, it turns out that this is the only reason why SP committees may fail core-stability; consequently, it is still implied by SSP.

Proposition 9.26. *Strict stable priceability implies core-stability.*

Proof. Consider an outcome W that is SSP, and for the sake of contradiction assume W is not in the core. Then, there exists $S \subseteq N$ and $T \subseteq C$ with $\mathrm{cost}(T) \le b \cdot {}^{|S|}\!/_n$ such that $u_i(W) < u_i(T)$ for all $i \in S$. Let us choose the witness so that $|T|$ is minimized. We set $R_i = W$ for all $i \in S$; for $i \notin S$ we set $R_i = \emptyset$.

Now the only issue is whether we can construct a payment system $q = (W, \{q_i\}_{i\in N})$ supporting W. Note that for every project $c \in T$, the number of voters in group $S' \subseteq S$ gaining positive utility from c is at least ${}^n\!/_b \cdot \mathrm{cost}(c)$. Indeed, if $T = \{c\}$ it is clearly the case; otherwise the set $T \setminus \{c\}$ together with group $S \setminus S'$ would be the smaller witness.

Note that this is an analogous result to Lemma 7.8 proved in Section 7.1 for weakly cohesive groups chosen by Greedy Cohesive Rule; consequently, we obtain the required payment system via the analogous construction to Lemma 7.9. □

Theorem 9.27. *Every outcome that is in a cost-efficient Lindahl equilibrium is strictly stable priceable. The other implication does not hold.*

Proof. Consider an outcome $W \subseteq C$ that is in the cost-efficient Lindahl equilibrium, and let γ be the corresponding price system. From γ we construct the price system p witnessing stable priceability as follows. The initial endowment in p is the same as in γ. For each $i \in N$ and $c \in W$ we set $p_i(c) = \gamma_i(c)$; for $c \notin W$ we set $p_i(c) = 0$. Note that from (L1) and (L3) the price system is feasible.

Let us now prove that p is SSP. Suppose it is not and let us fix a payment system q and a collection $\{R_i\}_{i\in N}$ witnessing the SSP violation. Fix a voter $i \in N$ and let $W' = (W \setminus R_i) \cup C_q$. Observe that if $u_i(W') > u_i(W)$, then by (L2) $\gamma_i(W') > \max_{j\in N} \gamma_j(W)$ and so:

$$
\begin{aligned}
q_i(C_q) &\le r_{p,i} + p_i(R_i) = \mathrm{end} - p_i(W \setminus R_i) = \max_{j\in N} \gamma_j(W) - \gamma_i(W \setminus R_i) \\
&\le \max_{j\in N} \gamma_j(W) - \gamma_i(W') + \gamma_i(C_q) < \gamma_i(C_q).
\end{aligned}
$$

On the other hand, if $u_i(W') = u_i(W)$ then either $\gamma_i(W') > \max_{j\in N} \gamma_j(W)$ (and we obtain the estimation as above), or $\gamma_i(W) < \gamma_i(W')$. In the latter case we get that:

$$
\begin{aligned}
q_i(C_q) &< p_i(R_i) = \gamma_i(R_i) = \gamma_i(W) - \gamma_i(W \setminus R_i) \\
&< \gamma_i(W') - \gamma_i(W \setminus R_i) \le \gamma_i(C_q).
\end{aligned}
$$

In any case, we get that $q_i(C_q) \leq \gamma_i(C_q)$. By (L1) we get that for each $c \in C_q$ we have $\sum_{i \in N} \gamma_i(c) \leq \mathrm{cost}(c)$. Thus, we can continue as:

$$\mathrm{cost}(C_q) \leq \sum_{i \in N} q_i(C_q) < \sum_{i \in N} \gamma_i(C_q) \leq \mathrm{cost}(C_q).$$

Which proves that p is indeed SSP.

Second, we show a committee election and an outcome that is strictly stable priceable, but which is not in a Lindahl equilibrium. We have four candidates, a_1, a_2, d_1, d_2, and two voters. The budget is $b = 2$. The voters' preferences and the costs of the candidates are summarized in the table below.

candidate	cost	$u_1(\cdot)$	$u_2(\cdot)$	$p_1(\cdot)$	$p_2(\cdot)$
a_1	ϵ	6	0	ϵ	0
a_2	$2 - \epsilon$	1	3	$1 - \epsilon$	1
d_1	1	2	2	0	0
d_2	1	2	2	0	0

For this election, outcome $A = \{a_1, a_2\}$ is stable priceable with the initial endowments $\mathrm{end} = 1$. The corresponding price system p is also given in the above table. To see that p is SSP, consider all possible values of C_q—namely $\{d_1\}$, $\{d_2\}$ and $\{d_1, d_2\}$.

If $C_q = \{d_1\}$ or $C_q = \{d_2\}$, then it needs to hold that $S = \{1\}$ and $R_1 \subseteq \{a_2\}$ (voter 2 does not pay for a_1 and will decrease her utility if she resigned from a_2; voter 1 will decrease her utility if she resigned from a_1). Then $q_1(C_q) \leq p_1(a_2) \leq 1 - \epsilon < \mathrm{cost}(C_q)$, a contradiction.

If $C_q = \{d_1, d_2\}$ and $S = \{1, 2\}$, then $R_1 \subseteq \{a_2\}$ (voter 1 will decrease her utility if she resigned from a_1) and $R_2 \subseteq \{a_2\}$ (voter 2 does not pay for a_1). Then $\sum_{i \in S} q_i(C_q) \leq \sum_{i \in S} p_i(a_2) \leq 2 - \epsilon < \mathrm{cost}(C_q)$, a contradiction. Naturally, taking smaller set S will even decrease value $\sum_{i \in S} q_i(C_q)$.

Yet, given the initial endowment $\mathrm{end} = 1$, the outcome $\{a_1, a_2\}$ is not in the cost-efficient Lindahl equilibrium. For the sake of contradiction, assume that $\{a_1, a_2\}$ is in the cost-efficient Lindahl equilibrium. Then, $\gamma_1(a_1) = \epsilon$ and $\gamma_2(a_1) = 0$. Further, it must also be the case that $\gamma_1(a_2) = 1 - \epsilon$ and $\gamma_2(a_2) = 1$. Further, since voter 1 prefers d_1 to a_2 and d_2 to a_2, it must hold that $\gamma_1(d_1) > 1 - \epsilon$ and that $\gamma_1(d_2) > 1 - \epsilon$. Also, $\gamma_2(d_1) + \gamma_2(d_2) > 1$. This means that the sum of the prices for d_1 and d_2 is at least $3 - 2\epsilon$, thus it exceeds the cost of producing d_1 and d_2, and so it violates (L1). This gives a contradiction, and completes the proof. □

9.6 Conclusion

To conclude, in this chapter we have introduced two market-based solution concepts for the PB setting with approval-based utilities that allow to reason about, explain, and justify fairness of the outcome of an election to voters. We have shown relations between our notions of (balanced) stable priceability and known concepts of fairness and stability from the literature, such as EJR, core, proportionality degree, and Lindahl equilibrium. We have characterized the outcomes satisfying market-based axioms using simpler formulas, which allowed us to obtain more efficient algorithms for finding SP and BSP outcomes. Notably, we have characterized a close variant of Method of Equal Shares as the only rule that returns BSP outcomes. Although exhaustive SP/BSP outcomes do not always exist, through experiments we have shown that our algorithms can effectively find (balanced) stable priceable outcomes that are almost exhaustive. Both market-based axioms can be extended to be model with arbitrary cardinal utilities, yet in this model they lose some of their good properties. Therefore, the problem of designing proportionality notions for this setting based on the idea of a fair market remains open.

Summary

Within the dissertation, we have studied the problem of participatory budgeting. We have focused on the algorithms for counting votes that aim to proportionally represent the preferences of the voters. We have studied both the special case of the committee election model, as well as the model with arbitrary project costs. We have considered several types of voters' preferences, with different levels of complexity, showing how the difficulty of certain problems depends on the type of voters' preferences and how certain concepts designed for one model can be applied to another one. In particular, we have considered the model in which voters' preferences are given in a form of cardinal utilities, as well as the model in which we have access only to ordinal preferences. Within the model of cardinal utilities, we have also considered the restricted domains of approval-based and approval-based cost utilities.

Our most important contribution is Method of Equal Shares, a simple and intuitive polynomial-time computable algorithm. We have presented a versatile argumentation in favor of this method. In particular, in Chapter 4 we have shown that this rule provides strong fairness guarantees in a very general model of PB with arbitrary costs and arbitrary cardinal utilities of the voters. The rule provides a high level of satisfaction for cohesive groups of voters and balances equally the influence of the voters on the elected outcome. Both properties were formalized by us through two axioms: Extended Justified Representation up to one project (EJR-1) and priceability. Additionally, we have shown that Equal Shares provides a good approximation of the core, one of the strongest known notions of proportionality considered in the literature.

Moreover, those good theoretical properties of Method of Equal Shares are also preserved when the rule is adapted to ordinal preferences (see Chapter 6). The variant of the rule for ordinal preferences also satisfies the strongest known axiom that has been proposed specifically for this setting, namely Inclusion Proportionality for Solid Coalitions.

Apart from good behavior in theory, in Chapter 5 we have argued that Equal Shares can be successfully used in real-life PB elections. We have provided a large dataset of elections from the Polish cities of various size, *Pabulib*, and conducted experiments on this dataset. The results appeared to be very consistent, showing undoubtedly that using Equal Shares instead of current solutions results in electing outcomes that provide a comparable level of total voters' satisfaction (efficiency) and at the same time spread this satisfaction much more equally among the voters. This was not a foreseeable result, since in theory proportionality can enforce us to significantly decrease the efficiency of an outcome.

Based on our experiments, we have prepared a detailed recommendation for the cities which exact variant of Equal Shares should be used. We have also shown that if a city uses approval ballots the definition of the rule can be presented in a more straightforward way. This is an interesting observation, since our experiments suggest that outcomes elected by Equal Shares under approval ballots should not differ much from the ones elected under more general score ballots.

For the time being, our method has been used by two Polish communities (Wieliczka and Świecie) and one Swiss city (Aarau). While the results for Świecie will be only available in 2024, we can already say that using Equal Shares in Wieliczka and Aarau led to improving the quality of elected outcomes compared to the alternative scenario in which the traditional method was used (we have discussed this in Chapter 1).

Method of Equal Shares is already well-established in the research community, which manifests itself in the number of various research groups that study and further extend our concept (see, for example, [Benade et al., 2023, Lu et al., 2023, Boehmer et al., 2023, Brill and Peters, 2023, Brill et al., 2023, Maly et al., 2023, Aziz et al., 2023]).

The second part of the dissertation has been devoted to the study the notions of proportionality stronger than the ones satisfied by Method of Equal Shares. We have started by proposing an axiom called Full Justified Representation (FJR) that provides stronger proportionality guarantees for cohesive groups than EJR-1 (see Chapter 7). To the best of our knowledge, this is the strongest currently known axiom of this type that is satisfiable. However, the only algorithm that is known to satisfy FJR (Greedy Cohesive Rule, GCR) is exponential-time computable. Besides, it is custom-engineered to satisfy this particular axiom which makes it less flexible and natural than Equal Shares.

In Chapter 8 we have studied the notion of the core in the approval-based and ordinal committee election model under specific restricted domains of voters' preferences. Intuitively, under such domains voters and candidates can be ordered in a way preserving some specific structure (for example, corresponding to their placement in one-dimensional Euclidean space). We have shown that then core-stability is satisfiable in polynomial time, contrary to the case of unrestricted cardinal utilities (where core-stability is not satisfiable) or approval-based ones (where the problem of satisfiability of the core is still open). However, the rule satisfying this notion is custom-engineered to work for elections with restricted domains and cannot be easily extended in a reasonable way to the general model. On the other hand, no other well-established proportional rule (including Equal Shares) is core-stable even under restricted domains. Therefore, the problem of designing such a rule (well-defined and proportional in general case and satisfying the core for well-structured preferences) remains open.

Finally, in Chapter 9 we have extended the intuitively appealing axiom of priceability so that it provides much stronger fairness guarantees, in particular, stronger than the core or EJR. This approach resulted in two mutually incomparable axioms, stable priceability (SP) and balanced stable priceability (BSP). The outcomes satisfying the first axiom are always core-stable, provide high utility on average to cohesive groups of voters. Moreover, stable priceability is very closely related to the well-established idea of Lindahl equilibrium. On the other hand, the guarantees provided by BSP to cohesive groups are also strong (this axiom implies EJR) and the justification of outcomes provided by this axiom is more intuitively appealing, ruling out some unbalanced stable priceable outcomes. Both for SP and BSP axioms, exhaustive outcomes satisfying them may not exist, yet our preliminary experiments suggest the SP/BSP outcomes close to being exhaustive often do.

Concluding the second part of the dissertation, we do not know a voting rule that is as strong and universal as Method of Equal Shares. Most of the known notions of proportionality that are violated by this rule are unsatisfiable in general. Even if for some axioms it is not the case, the only known algorithms satisfying them lack the versatility of Equal Shares—they are specifically designed to satisfy only one precise proportionality axiom.

In the future, our research may lead to designing proportionality definitions and algorithms for even more complex settings. Let us complete the dissertation by discussing one of the possible extensions of our model.

In the dissertation we have assumed that every subset of projects whose total cost fits in the available budget can be selected. Sometimes it is not the case, for example if multiple projects are planned to be conducted in the same location. Moreover, the election designer might want to add some constraints, for example, requiring that at least 30% of the budget should be spent on building green areas. In general, instead of just one feasibility constraint (the budget), we can have several ones. Designing and analyzing proportional algorithms for this setting appears now to be an interesting challenge for the future research.

Bibliography

Elliot Anshelevich and Wennan Zhu. Ordinal approximation for social choice, matching, and facility location problems given candidate positions. *ACM Transactions on Economics and Computation*, 9, 2021.

Elliot Anshelevich, Onkar Bhardwaj, Edith Elkind, John Postl, and Piotr Skowron. Approximating optimal social choice under metric preferences. *Artificial Intelligence*, 264:27–51, 2018.

Elliot Anshelevich, Aris Filos-Ratsikas, Nisarg Shah, and Alexandros A. Voudouris. Distortion in social choice problems: The first 15 years and beyond. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI-2021)*, pages 4294–4301, 2021.

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

Haris Aziz and Barton E. Lee. Proportionally representative participatory budgeting with ordinal preferences. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI-2021)*, volume 35, pages 5110–5118, 2021.

Haris Aziz, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh. Justified representation in approval-based committee voting. *Social Choice and Welfare*, 48 (2):461–485, 2017a.

Haris Aziz, Edith Elkind, Piotr Faliszewski, Martin Lackner, and Piotr Skowron. The condorcet principle for multiwinner elections: From shortlisting to proportionality. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI-2017)*, pages 84–90, 08 2017b.

Haris Aziz, Edith Elkind, Shenwei Huang, Martin Lackner, Luis Sánchez-Fernández, and Piotr 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, 2018.

Haris Aziz, Anna Bogomolnaia, and Hervé Moulin. Fair mixing: the case of dichotomous preferences. In *Proceedings of the 20th ACM Conference on Economics and Computation (EC-2019)*, pages 753–781, 2019.

Haris Aziz, Xinhang Lu, Mashbat Suzuki, Jeremy Vollen, and Toby Walsh. Best-of-both-worlds fairness in committee voting. *arXiv preprint arXiv:2303.03642*, 2023.

Michel Balinski and Rida Laraki. *Majority Judgment: Measuring, Ranking, and Electing*. The MIT Press, 2010.

Salvador Barberà and Bernardo Moreno. Top monotonicity: A common root for single peakedness, single crossing and the median voter result. *Games and Economic Behavior*, 73(2): 345–359, 2011.

Gerdus Benade, Roy Fairstein, and Kobi Gal. Participatory budgeting design for the real world. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI-2023)*, 2023.

Jean-Pierre Benoit and Lewis Kornhauser. *Voting simply in the election of assemblies*. C.V. Starr Center for Applied Economics, New York University, 1991.

Duncan Black. On the rationale of group decision-making. *Journal of Political Economy*, 56 (1):23–34, 1948.

Duncan Black. Partial justification of the borda count. *Public Choice*, pages 1–15, 1976.

Niclas Boehmer, Piotr Faliszewski, Łukasz Janeczko, and Andrzej Kaczmarczyk. Robustness of participatory budgeting outcomes: Complexity and experiments. *arXiv preprint arXiv:2305.08125*, 2023.

Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D. Procaccia, and Or Sheffet. Optimal social choice functions: A utilitarian view. *Artificial Intelligence*, 227:190–213, 2015.

Shaun Bowler and Bernard Grofman. *Elections in Australia, Ireland, and Malta under the Single Transferable Vote: Reflections on an embedded institution*. University of Michigan Press, 2000.

Felix Brandt. Rolling the dice: Recent results in probabilistic social choice. In Ulle Endriss, editor, *Trends in Computational Social Choice*, chapter 1, pages 3–26. AI Access, 2017.

Markus Brill and Jannik Peters. Robust and verifiable proportionality axioms for multiwinner voting. In *Proceedings of the 24th ACM Conference on Economics and Computation (EC-2023)*, page 301. Association for Computing Machinery, 2023.

Markus Brill, Paul Gölz, Dominik Peters, Ulrike Schmidt-Kraepelin, and Kai Wilker. Approval-based apportionment. *Mathematical Programming*, pages 1–29, 2022.

Markus Brill, Stefan Forster, Martin Lackner, Jan Maly, and Jannik Peters. Proportionality in approval-based participatory budgeting. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI-2023)*, 2023.

Jeff Burdges, Alfonso Cevallos, Peter Czaban, Rob Habermeier, Syed Hosseini, Fabio Lama, Handan Kilinc Alper, Ximin Luo, Fatemeh Shirazi, Alistair Stewart, et al. Overview of polkadot and its design considerations. *arXiv preprint arXiv:2005.13456*, 2020.

Alfonso Cevallos and Alistair Stewart. A verifiably secure and proportional committee election rule. In *Proceedings of the 3rd ACM Conference on Advances in Financial Technologies*, pages 29–42, 2021.

Yu Cheng, Zhihao Jiang, Kamesh Munagala, and Kangning Wang. Group fairness in committee selection. In *Proceedings of the 20th ACM Conference on Economics and Computation (EC-2019)*, pages 263–279, 2019.

Vincent Conitzer and Toby Walsh. Barriers to manipulation in voting. In *Handbook of Computational Social Choice*. 2016.

George B. Dantzig. Discrete-variable extremum problems. *Operations Research*, 5(2):266–288, 1957.

Jan de Leeuw. Modern multidimensional scaling: Theory and applications. *Journal of Statistical Software*, 14:1–2, 2005.

Michiel S. De Vries, Juraj Nemec, and David Špaček. International trends in participatory budgeting. *Cham: Palgrave Macmillan*, 2022.

Michael Dummett. *Voting Procedures*. Oxford University Press, 1984.

Cynthia Dwork, Ravi Kumar, Moni Naor, and Dandapani Sivakumar. Rank aggregation methods for the web. In *Proceedings of the 10th International World Wide Web Conference*, pages 613–622, 2001.

Edith Elkind and Martin Lackner. Structure in dichotomous preferences. In *Proceedings of the 24rd International Joint Conference on Artificial Intelligence (IJCAI-2015)*, pages 2019–2025, 2015.

Edith Elkind, Piotr Faliszewski, and Piotr Skowron. A characterization of the single-peaked single-crossing domain. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI-2014)*, 2014.

Edith Elkind, Martin Lackner, and Dominik Peters. Structured preferences. *Trends in computational social choice*, pages 187–207, 2017.

Edith Elkind, Piotr Faliszewski, and Piotr Skowron. A characterization of the single-peaked single-crossing domain. *Social Choice and Welfare*, 54:167–181, 2020.

Brandon Fain, Ashish Goel, and Kamesh Munagala. The core of the participatory budgeting problem. In *Proceedings of the 8th Workshop on Internet & Network Economics (WINE-2016)*, pages 384–399, 2016.

Brandon Fain, Kamesh Munagala, and Nisarg Shah. Fair allocation of indivisible public goods. In *Proceedings of the 19th ACM Conference on Economics and Computation (EC-2018)*, pages 575–592, 2018.

Piotr Faliszewski, Jakub Sawicki, Robert Schaefer, and Maciej 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.

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

Piotr Faliszewski, Jarosław Flis, Dominik Peters, Grzegorz Pierczyński, Piotr Skowron, Dariusz Stolicki, Stanisław Szufa, and Nimrod Talmon. Participatory budgeting: Data, tools and analysis. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI-2023)*, pages 2667–2674, 2023.

Michal Feldman, Amos Fiat, and Iddan Golomb. On voting and facility location. In *Proceedings of the 17th ACM Conference on Economics and Computation (EC-2016)*, page 269–286, 2016.

Duncan K. Foley. Lindahl's solution and the core of an economy with public goods. *Econometrica*, 38(1):66–72, 1970.

Allan Gibbard. Manipulation of voting schemes. *Econometrica*, 41(4):587–601, 1973.

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

Leonid G. Khachiyan. A polynomial algorithm in linear programming (english translation). In *Soviet Mathematics Doklady*, volume 20, pages 191–194, 1979.

Joseph Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964.

Martin Lackner and Piotr Skowron. Approval-based committee voting. In *Multi-Winner Voting with Approval Preferences*, pages 1–7. Springer, 2022.

Martin Lackner, Jan Maly, and Simon Rey. Fairness in long-term participatory budgeting. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI-2021)*, pages 299–305, 2021.

Dean Lacy and Emerson M. S. Niou. A problem with referendums. 12(1):5–31, 2000.

Jérôme Lang and Lirong Xia. Voting in combinatorial domains. In *Handbook of Computational Social Choice*. 2016.

Jean-François Laslier. The strange "majority judgment". *Revue économique*, 70(4):569–588, 2019.

Xinhang Lu, Jannik Peters, Haris Aziz, Xiaohui Bei, and Warut Suksompong. Approval-based voting with mixed goods. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI-2023)*, pages 5781–5788, 2023.

Krzysztof Magiera and Piotr Faliszewski. Recognizing top-monotonic preference profiles in polynomial time. *Journal of Artificial Intelligence Research*, 66:57–84, 2019.

Colin L. Mallows. Non-null ranking models. I. *Biometrika*, 44(1-2):114–130, June 1957.

Jan Maly, Simon Rey, Ulle Endriss, and Martin Lackner. Fairness in participatory budgeting via equality of resources. In *Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2023)*, pages 2031–2039, 2023.

Borys Martela, Liliana Janik, and Kamil Mróz. Barometr budżetu obywatelskiego. 2023.

James A. Mirrlees. An exploration in the theory of optimal income taxation. *Review of Economic Studies*, 38:175–208, 1971.

Kamesh Munagala, Yiheng Shen, Kangning Wang, and Zhiyi Wang. Approximate core for committee selection via multilinear extension and market clearing. In *Proceedings of the 33rd ACM-SIAM Symposium on Discrete Algorithms (SODA-2022)*, pages 2229–2252, 2022.

Dominik Peters and Piotr Skowron. Proportionality and the limits of welfarism. In *Proceedings of the 21st ACM Conference on Economics and Computation (EC-2020)*, pages 793–794, 2020. Extended version arXiv:1911.11747.

Dominik Peters, Grzegorz Pierczyński, Nisarg Shah, and Piotr Skowron. Market-based explanations of collective decisions. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI-2021)*, 2021a.

Dominik Peters, Grzegorz Pierczyński, and Piotr Skowron. Proportional participatory budgeting with additive utilities. *Advances in Neural Information Processing Systems (NeurIPS-2021)*, 34:12726–12737, 2021b.

Lars E. Phragmén. Sur une méthode nouvelle pour réaliser, dans les élections, la représentation proportionnelle des partis. *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar*, 51 (3):133–137, 1894.

Grzegorz Pierczyński and Piotr Skowron. Core-stable committees under restricted domains. In *Proceedings of the 18th Conference on Web and Internet Economics (WINE-2022)*, pages 311–329, 2022.

Ariel D. Procaccia and Jeffrey S. Rosenschein. The distortion of cardinal preferences in voting. In *International Workshop on Cooperative Information Agents*, pages 317–331. Springer, 2006.

Kevin W. S. Roberts. Voting over income tax schedules. *Journal of Public Economics*, 8(3): 329–340, 1977.

Piotr Skowron. Proportionality degree of multiwinner rules. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 820–840, 2021.

Piotr Skowron, Martin Lackner, Markus Brill, Dominik Peters, and Edith Elkind. Proportional rankings. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI-2017)*, pages 409–415, 2017.

Aravind Srinivasan. Distributions on level-sets with applications to approximation algorithms. In *Proceedings of the 42th Annual IEEE Symposium on Foundations of Computer Science (FOCS-2001)*, pages 588–597, 2001.

Nicolaus Tideman. The single transferable vote. *Journal of Economic Perspectives*, 9(1):27–38, 1995.

Brian Wampler, Stephanie McNulty, and Michael Touchton. *Participatory budgeting in global perspective*. Oxford University Press, 2021.

Douglas R. Woodall. Monotonicity of single-seat preferential election rules. *Discrete Applied Mathematics*, 77(1):81–98, 1997.

William S. Zwicker. Introduction to the theory of voting. In *Handbook of Computational Social Choice*. 2016.

Appendix A

Pabulib: Data Format

In this section we define the *.pb* format, which we recommend for storing PB elections.

The data concerning a single PB election is stored in a single UTF-8 text file with the extension `.pb`. The file should consists of three sections:

META section containing general information about the election, such like the country, the budget, and the number of votes.

PROJECTS section specifying the costs of the projects and optionally providing additional information about the projects, such as their categories.

VOTES section listing all votes cast in the election, optionally with additional information about the respective voters (for example, their age, sex, etc.). We support four types of ballots: approval, score, ranking (ordinal) and cumulative.

Below we present a simple example of a PB file.

```
META
key; value
description; Municipal PB in Wieliczka
country; Poland
unit; Wieliczka
instance; 2020
num_projects; 5
num_votes; 10
budget; 2500
rule; greedy
vote_type; approval
min_length; 1
max_length; 3
PROJECTS
project_id; cost; category
1; 600; culture, education
2; 800; sport
4; 1400; culture
5; 1000; health, sport
7; 1200; education
VOTES
voter_id; age; sex; vote
1; 34; f; 1,2,4
2; 51; m; 1,2
3; 23; m; 2,4,5
4; 19; f; 5,7
5; 62; f; 1,4,7
6; 54; m; 1,7
7; 49; m; 5
8; 27; f; 4
9; 39; f; 2,4,5
10; 44; m; 4,5
```

Now let us describe the format in more detail.

Detailed Description

The fields marked with the **bold** font are obligatory.

META

- **`key`**
  - **`description`**
  - **`country`**
  - **`unit`**: the name of the municipality, region, organization, etc., holding the PB process.
  - `subunit`: the name of the sub-jurisdiction or category of the particular election.
    - *Example*: in Paris, a single edition of participatory budgeting consists of 21 independent elections—there is one election concerning citywide projects and 20 local elections, one per each district. For the election with citywide projects, the field `unit` is set to Paris, and `subunit` is undefined; for the district elections, the field `unit` is also Paris, and the `subunit` is the name of the respective district (for example, IIIe arrondissement).
    - *Example*: before 2019, in Warsaw there were two types of local elections: district elections and neighborhood elections. For all of them, the field `unit` is set to Warsaw; the field `subunit` is the name of the district (for district elections) or the name of the neighborhood (for neighborhood elections). In order to connect neighborhoods with their districts, an optional field `district` can be used.
    - *Example*: suppose that in a given city, there is a separate election for each of $n > 1$ categories (for example, environmental projects, transportation projects, cultural projects, etc.). For each such an election the field `unit` is set to the city name; the field `subunit` is set to the name of the respective category.
  - **`instance`**: a unique identifier of the participatory budgeting edition (for example, year, edition number, etc.). Note that the year specified in the field `instance` does not necessarily correspond to the year in which the elections were held—some organizers identify the edition by the fiscal year in which the projects are carried out.
  - **`num_projects`**
  - **`num_votes`**
  - **`budget`**: the total amount of funds
  - **`vote_type`**: the type of ballots used in the election. The library currently supports four types of ballots:
    - `approval`: each vote is a vector of Boolean values, $\mathbf{v} \in \{0,1\}^{|P|}$, where P is the set of all projects,
    - `ordinal`: each vote is a permutation of a subset $Q \subseteq P$ such that $|Q| \in [\texttt{min\_length}, \texttt{max\_length}]$, corresponding to a strict preference order over Q,

– `cumulative`: each vote is a vector $\mathbf{v} \in \mathbb{R}_+^{|P|}$ such that $\sum_{p \in P} v[p] \leq$ **`max_sum_points`** $\in \mathbb{R}_+$,
– `scoring`: each vote is a vector $\mathbf{v} \in I^{|P|}$, where $I \subseteq \mathbb{R}$.

∗ **`rule`**: the name of the rule that was used in the election. Currently we support the following rules:
  – `greedy`: this corresponds to the costwise Utilitarian Greedy rule,
  – other rules will be defined in the future versions.

∗ `date_begin`: the date when the process of collecting ballots started.

∗ `date_end`: the date when the process of collecting ballots ended.

∗ `language`: the language of the descriptions of the projects (that is, full names of the projects)

∗ `edition`

∗ `district`

∗ `comment`

∗ if `vote_type` = `approval`:
  – `min_length` [default: 1]
  – `max_length` [default: num_projects]
  – `min_sum_cost` [default: 0]
  – `max_sum_cost` [default: ∞]

∗ if `vote_type` = `ordinal`:
  – `min_length` [default: 1]
  – `max_length` [default: num_projects]
  – `scoring_fn` [default: Borda]

∗ if `vote_type` = `cumulative`:
  – `min_length` [default: 1]
  – `max_length` [default: num_projects]
  – `min_points` [default: 0]
  – `max_points` [default: max_sum_points]
  – `min_sum_points` [default: 0]
  – **`max_sum_points`**

∗ if `vote_type` = `scoring`:
  – `min_length` [default: 1]
  – `max_length` [default: num_projects]
  – `min_points` [default: $-\infty$]
  – `max_points` [default: ∞]

    – `default_score` [default: 0]

  ∗ `non-standard fields`

- **`value`**: the value of the corresponding field.

Section 2: PROJECTS

- **`project_id`**
- **`cost`**
- `name`: the full name of the project.
- `category`: the list of tags depscribing the project, separated with commas; for example: education, sport, health, culture, environmental protection, public space, public transit, roads.
- `target`: type voters that might be especially interested in the project. For example: adults, seniors, children, youth, people with disabilities, families with children, animals.
- `non-standard fields`

Section 3: VOTES

- **`voter_id`**
- `age`
- `sex`
- `voting_method` (for example, paper, Internet, mail)
- if `vote_type` = `approval`:
  ∗ **`vote`**: identifiers of the approved projects, separated with commas.
- if `vote_type` = `ordinal`:
  ∗ **`vote`**: identifiers of the selected projects, from the most preferred to the least preferred one, separated with commas.
- if `vote_type` = `cumulative`:
  ∗ **`vote`**: identifiers of the projects, separated with commas; projects not listed are assumed to get 0 points. Projects are listed in the decreasing order of the number of points they got from the voter,

* **`points`**: points given to the projects, listed in the same order as the project identifiers in the field **`vote`**.

- if `vote_type` = `scoring`:

  * **`vote`**: project identifiers, separated with commas; projects not listed are assumed to get `default_score` points. Projects are listed in the decreasing order of the number of points they got from the voter.
  * **`points`**: points given to the projects, listed in the same order as the project identifiers in the field **`vote`**.

- `non-standard fields`