
Cost Considerations in Participatory Budgeting: An Empirical Comparison of Knapsack and Approval Ballots

Oliwia Szczupska¹

Abstract

This article examines the role of cost considerations in shaping the selection of participatory budgeting projects through an empirical comparison between knapsack and k-approval ballot types. Using data from the 2018–2021 elections in the city of Warsaw (Poland), we analyzed how these two methods – each with a fundamentally different approach to cost – affect individual decision-making and collective outcomes. We find that knapsack voting, with its predefined budget constraint, encourages economically mindful behavior and leads to the funding of a greater number of lower-cost initiatives. In contrast, k-approval voting, which limits the number of projects rather than their costs, results in the selection of relatively fewer but more expensive initiatives. To validate these findings, we conducted an online voting experiment simulating both methods. The results show a consistent bias toward more affordable projects when cost consideration mechanisms are embedded in either the ballot type (knapsack) or the aggregation rule (cost-proportional algorithm). Our findings highlight the importance of voting method design in the participatory budgeting process, provide an empirical contribution to the understanding of how different rules function in practice, and offer crucial insights for guiding future implementations and development.

Keywords: collective decision-making, participatory budgeting, voting behavior, cost-benefit analysis, knapsack voting, approval voting.

¹ Oliwia Szczupska – SWPS University, Institute of Social Sciences in Warsaw, Poland, e-mail: oszczupska@swps.edu.pl; ORCID: 0009-0007-7702-0543.

© Oliwia Szczupska. Published in *Collective and Individual Decisions*. This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

Introduction

Participatory budgeting (PB) allows for citizen engagement in public funds' allocation to better represent the needs and preferences of the communities (Cabannes, 2004). Although a PB process can be executed in many ways, it often includes a voting stage in which the residents cast ballots to select a winning set of projects subject to a budget constraint. The question of how to determine a collective outcome with a variety of people's possibly diverse needs and interests is central to social choice theory (Sen, 1970). In fact, Arrow (1951) proved that no voting procedure can simultaneously satisfy certain plausible fairness criteria, and it is well-established that different ballot types and aggregation rules can lead to different outcomes (Nurmi, 2010). Thus, designing a voting method is a high-complexity task and often incorporates making difficult tradeoffs.

The related literature has proposed a wide variety of voting methods appropriate for PB settings (Benade *et al.*, 2017; Elkind *et al.*, 2017; Aziz *et al.*, 2018). First, the choice of a ballot type determines how voters can express their preferences. In real-world PB elections, voters are often asked to approve (or rank) a limited number of initiatives. In contrast to ballots that restrict the number of selected projects, other types – such as knapsack which requires voters to select projects within a total budget limit – are specifically designed to increase voters' awareness of project costs. Once the ballots are cast, the aggregation rule indicates how they are translated into the election results. Here, the project selection process can consider solely the number of votes each initiative receives or both the number of votes and the costs of the projects, i.e., require that a high-cost project needs to garner more support than a lower-cost one.

This article aims to address how mechanisms for cost consideration embedded within participatory budgeting voting rules influence individual behaviors and shape the resulting social outcomes. Our study focuses on the empirical evaluation of two ballot types: knapsack and k-approval. We study the change of a voting system in an actual large-scale participatory budgeting election using data from real Warsaw PB throughout its four consecutive editions. In 2018–2019, under knapsack voting, the residents were asked to select as many projects as they wished as long as their total cost would not exceed the dedicated budget. In 2020–2021, under k-approval voting, the residents were asked to select up to a certain fixed number of projects, without any limitations imposed on their costs. We validate these observations by conducting an online voting experiment, where we ask participants to vote under both tested methods. The integration of insights coming from real-world and experimental settings offers a unique contribution to understanding how alternative rules for participatory budgeting operate in practice. Thus, our empirical findings offer valuable insights for designing and implementing PB solutions with desirable outcomes.

Participatory Budgeting Problem

Participatory budgeting is a process that allows for the allocation of public funds by collecting and aggregating the preferences of residents over potential projects. Initiated in southern Brazil in the 1980s, it has gradually spread across the world as a significant area of innovation

in democracy and local development (Cabannes, 2004). Participatory budgeting processes can vary widely in their implementation, but they generally follow several key stages (Wampler, 2007). First, authorities divide regions into districts and establish the budget for each. Residents then generate and develop project ideas, often through public discussions or online tools. The feasible proposals undergo expert evaluation, which leads to cost estimation and project categorization (e.g. thematically). Next, final selection of verified projects is prepared for the voting stage, where eligible residents cast their votes to determine how the available budget should be spent. Two key design considerations at this stage are the ballot type, which determines how residents' preferences are collected, and the aggregation rule, which specifies how the votes are translated into the final budget allocation.

Let us properly formulate the participatory budgeting problem (Aziz & Shah, 2021). We have a set V of n residents (voters), sometimes divided across different districts, and a set P of m projects (alternatives), which may be categorized into various themes. There is a fixed total budget B . Each project $p \in P$ has an associated cost $c(p)$ representing the amount of money needed to fully implement that project. A budgeting scenario can be defined as a tuple (V, P, c, B) . The participatory budgeting problem then involves selecting an optimal subset of projects $S \subseteq P$ such that the total cost does not exceed the budget, formally expressed as $\sum_{p \in S} c(p) \leq B$. Determining which feasible budget allocations are optimal depends on how we model the residents' preferences and social welfare.

Let each voter $i \in V$ have a utility function $u_i: P \rightarrow \mathbb{R}_+ \cup \{0\}$, where $u_i(p)$ represents the utility that voter i derives from project p . Since different project proposals are typically independent of one another, we assume that these utilities are additive across projects. Thus, the utility of voter i for a subset $S \subseteq P$ is defined as $u_i(S) = \sum_{p \in S} u_i(p)$. To ensure fairness among voters, we further assume that $u_i(P) = u_j(P)$ for all voters $i, j \in N$. We refer to the vector $\vec{u} = \{u_1, \dots, u_n\}$ of voters' utility functions as a utility profile.

We employ a social welfare function to aggregate the individual utility functions of residents into a collective social utility function. Given a utility profile u , we denote the social welfare of a project p as $SW(p, \vec{u})$, and that of a subset of projects $S \subseteq P$ as $SW(S, \vec{u})$. Formally, a solution to participatory budgeting is a subset of projects S^* that maximizes social welfare subject to a budget constraint, that is selects $S^* \in \arg \max_{S \subseteq P} \{SW(S, \vec{u}) : \sum_{p \in S} c(p) \leq B\}$. We adopt a utilitarian approach to welfare maximization (for alternatives see Aziz & Shah, 2021), meaning that the larger the sum of individual utilities, the better the outcome. Formally, the utilitarian welfare of a budget allocation is defined as the total utility it provides to residents: $SW^{SUM}(S, \vec{u}) = \sum_{i \in V} u_i(S)$ for $S \subseteq P$.

While having access to detailed utility functions would be ideal for modeling purposes, it is practically impossible to ask residents to report complex preferences. As a result, real-world PB systems typically use simpler input formats, collecting only partial information about voter preferences. We denote the vote cast by voter i as ρ_i and refer to the vector $\vec{\rho} = \{\rho_1, \dots, \rho_n\}$ as a vote profile. The utility function of a voter can only be inferred from their vote ρ_i , which is induced by u_i . In this article, we define two specific input formats being the focus of our study, while acknowledging other commonly used alternatives (Benade *et al.*, 2017; 2018).

A k -approval vote of voter i is represented as a binary vector $\alpha_i \in \{0, 1\}^m$, where $\alpha_i(p) = 1$ if project p is approved, $\alpha_i(p) = 0$ otherwise, and subject to the constraint $\sum_{p \in P} \alpha_i(p) \leq k$ that limits the voter to approving at most k projects. This voting method serves as a cognitive-

ly less demanding, truncated version of the approval vote, where voters would otherwise approve all projects they favor. A *knapsack* vote of size B of voter i is expressed as a binary vector $\kappa_i \in \{0,1\}^m$, subject to the constraint $\sum_{p \in P: \kappa_i(p)=1} c(p) \leq B$, representing an optimal set of projects while remaining within the total cost limit of B . Goel *et al.* (2015) introduced this approach as particularly well-suited for the participatory budgeting problem as it requires voters to consider both the costs of projects and the overall budget constraint when determining their preferred allocations.

Next, for a bundle S and a voter i with a selected set of projects A_i we assume that a derived utility is the number of projects approved and funded. Thus, $u_i(S)$ is represented by a satisfaction function: $f(A_i, S) = \sum_{p \in S} \mathbb{1}_{p \in A_i}$, where $\mathbb{1}_{p \in A_i} = 1$ if $p \in A_i$ and $\mathbb{1}_{p \in A_i} = 0$ otherwise. This specific translation of a voter ballot and a project bundle into the expected utilities was proposed for approval and knapsack type of ballots by Talmon and Faliszewski (2019), together with alternative approaches (e.g., dichotomous utilities where utility is zero if none of the voter's approved items are funded, and one if at least one is funded). For ranking-based ballots, see Laruelle (2020).

The final step is to identify the subset of projects that maximizes social welfare within the budget constraint. This selection process is typically NP-hard, making it impractical to find an exact solution. Therefore, heuristic greedy algorithms are often applied in place of the maximum rules. While they do not guarantee an optimal solution, they are well-defined through an iterative process and can be executed in polynomial time, providing a significant advantage, especially when the number of projects is large.

Talmon and Faliszewski (2019) proposed two greedy selection rules that are particularly effective for approximating utilitarian welfare under budget constraints. The *greedy algorithm* proceeds iteratively, maintaining a partial bundle of selected projects Q . It begins with an empty set, assuming all projects are initially admissible, and computes their total utilities. At each step, the algorithm selects project q with the highest welfare score to maximize the overall value: $\sum_{i \in V} u_i(Q \cup q)$. After each selection, the set of admissible projects is updated by removing the selected project and any others whose inclusion would exceed the remaining budget. The procedure continues until no further admissible projects remain. The second rule, the *proportional greedy algorithm*, follows a similar process but incorporates cost efficiency into the decision. At each iteration, it selects the project q that offers the highest utility gain per unit cost, maximizing the ratio: $[\sum_{i \in V} u_i(Q \cup q) - \sum_{i \in V} u_i(Q)]/c(q)$.

In simple terms, these two algorithms offer practical ways to select projects that bring the most benefit to residents within a limited budget. The basic greedy approach picks the most popular projects one by one, while the proportional greedy version also considers how much each project costs, prioritizing those that deliver the most value for the money. Although both methods are efficient and easy to implement, they may lead to different project selections and outcomes.

Related Research

From a social choice perspective, the existing research on participatory budgeting focuses mostly on the computational complexity of the aggregation methods and their desired axi-

omatic properties. Below, we present key findings on the comparison of various input formats and aggregation rules in participatory budgeting, emphasizing experimental and empirical evidence related to the impact of voting rules and embedded cost consideration mechanisms on individual behavior and election outcomes. For a comprehensive review of participatory budgeting literature, refer to Aziz and Shah (2021).

Goel *et al.* (2019) made a substantial contribution to the discussion on participatory budgeting by combining theoretical insights with practical applications. The authors highlighted k-approval's shortcomings, notably its failure to incorporate project costs into voter considerations. They proposed a knapsack vote and value-for-money ranking that were specifically designed to nudge voters to naturally consider cost-benefit tradeoffs – either optimize over the total budget or compare projects based on their value per dollar spent. The implementation of these voting mechanisms on a digital platform in partnership with local governments enabled them to collect empirical data from various PB elections. In most of the studied elections, k-approval served as the official ballot, and either knapsack vote, value-for-money comparisons, or both were experimentally added. The results demonstrated that the knapsack voting method led to more economically considerate selections than k-approval, with an average cost reduction of approximately 30% for winning projects. Additional analysis based on the information on voters' preferences between pairs of projects (value-for-money comparisons) showed that the knapsack vote better aligned with voter preferences compared to k-approval.

Gelauff and Goel (2024) validated the findings that voters are more cost-conscious when faced with explicit budget constraints. By utilizing ballot pairs, they compared project costs selected by the same voters under different voting methods, isolating the effects of the voting interface and method. In most of the 17 elections, where voters submitted both k-approval and knapsack votes, they found that the former generally results in a higher average cost for selected projects (normalized by the available budget).

Laruelle (2020) also examined voting methods in participatory budgeting from both theoretical and practical perspective, using data from the 2018 PB elections in Portugalete (Spain). Her study compared the official rank-based Dowdall method with three alternative approaches. In the Dowdall method, one selects projects based on the highest total scores, where each score is the reciprocal of the project's rank, continuing until the budget is depleted. The tested alternatives included: a greedy algorithm applied to data with dichotomous preferences; an approach that prioritizes residents who have none of their preferred projects selected, by maximizing the number of voters with at least one approved project; and a cost-benefit method (the greedy cost-proportional algorithm), which selects projects based on the lowest average cost per vote – that is, the project's cost divided by the number of votes it receives. The resulting winning sets varied significantly among the methods, with the cost-benefit algorithm selecting more projects at lower costs, consistent with the previous findings that methods incorporating cost considerations tend to favor more economical projects.

To analyze how different voting rules reacted to projects with varying costs, Talmon and Faliszewski (2019) conducted numerical experiments testing approval-based rules across nine configurations based on different satisfaction functions and aggregation algorithms. They manipulated either the cost of expensive items to determine when each rule would cease selecting costly projects or the proportion of voters supporting expensive projects to see when rules would favor high-cost items over cheaper alternatives. The results showed that maximum

algorithms were highly responsive to costly items, even with lower voter support, whereas proportional rules tended to prioritize cheaper projects to maximize satisfaction per unit of cost. Greedy approaches adopted a moderate solution between the two.

Benade *et al.* (2018) examined participatory budgeting by assessing the efficiency and usability of various ballot formats. They aimed to balance welfare optimization with the cognitive burden on voters. An experimental scenario with over 1,200 Amazon Mechanical Turk participants revealed that k-approval and value-for-money ranking achieved near-optimal welfare, while value ranking incurred a significant efficiency loss. Simultaneously, k-approval was the fastest format, whereas value-for-money ranking was the hardest to learn and longest to execute. Respondents rated a simple value ranking as the most user-friendly.

Rosenfeld and Talmon (2021) explored how non-experts perceived the fairness of PB voting rules. In their experiment, they asked the participants (Israeli and Polish students) to select the winning project based on outcomes derived from different aggregation methods. The results showed that utilitarian-based methods aligned most closely with participants' choices and received the highest preference, highlighting their practicality and ease of understanding. In a separate experiment, participants preferred outcomes that relied primarily on the number of projects approved and funded, challenging the authors' initial assumption that any reasonable utility function in a PB setting must account for project costs, suggesting instead that perceived fairness may be driven more by the number of selected projects than by precise cost-based utility calculations.

The reported experimental results are relatively recent and still scarce in the social choice participatory budgeting literature. They offer valuable insights into how different voting methods and ballot designs influence outcomes, showing that methods incorporating cost considerations – such as knapsack voting or cost-benefit algorithms – tend to lead to more economically efficient project selections compared to simpler rules like k-approval. Simultaneously, these findings emphasize the importance of usability, perceived fairness, and voters' cognitive burden, with participants generally favoring simpler, more intuitive methods, even if they are less optimal. Although the data from real participatory budgeting elections remains limited – aside from the notable Pabulib open library (Stolicki *et al.*, 2020; Faliszewski *et al.*, 2023) – there is a clear need for further empirical studies to better understand the effects of different input formats and aggregation rules. Our article contributes to this growing body of research.

Field Study Results

Warsaw's Participatory Budget, launched in 2014, has been held annually since then. It follows a typical PB process, including project submission, deliberation, assessment, voting, implementation, and evaluation (Sudolska-Bytof & Janiak, 2019). While there is a unified process for all 18 districts, each district manages its own set of proposed projects (specific to the area), budget (based on population), group of voters (mostly local residents), and selects projects for implementation separately. Since 2020, all residents have also been able to vote for city-wide projects.

Formally, we have a set V of n residents (voters), a set P of m projects (alternatives) with costs $c(p)$ indicating the amount of money needed to fully implement project p , and a fixed budget

B. We also have a set L of l elections. In each election k , a subset of residents votes over a set of projects P_k with a given total available budget B_k . The projects and the total budget are divided across the elections in a way that $P = P_1 \cup \dots \cup P_l$ and $B = \sum_{k \in L} B_k$. Given the elections are conducted at different levels, a voter can cast a ballot in more than one election. Therefore, for each election, we distinguish a set W_k of n_k ballots, where $W = W_1 \cup \dots \cup W_l$ is the set of all ballots cast in a given edition and $|W| = \sum_{k \in L} n_k$. Within a given edition, there are multiple parallel participatory budgeting problems. For each election k , a winning subset of projects is selected subject to $\sum_{p \in P_k} c(p) \leq B_k$ with a greedy utilitarian approach.

In the 2020 edition of the Warsaw PB, significant structural adjustments accompanied the switch to a new voting method. Alongside replacing the knapsack vote with k-approval, a key change was the elimination of sub-district projects, merging them into district-wide projects and adding a city-wide level. Previously, residents voted for projects within smaller, localized sub-districts, leading to 92 distinct elections in 2018 and 90 in 2019. By contrast, the 6th and 7th editions had 19 elections: 18 district-wide and one city-wide. As a consequence of the new structure, the project lists grew in length, often exceeding 100 projects, and the scale and cost of projects increased. Average project costs rose from approximately 95,000 PLN in the earlier editions to 150,000 PLN for district-level and 750,000 PLN for city-level projects in more recent editions. Readers may find other minor changes that were not critical to the study’s experimental analysis in the legal documentation (Directive No. 1660/2017, Resolution No. XI/218/2019).

Given significant differences across editions, we establish some key assumptions to allow a fair comparison. First, since some residents voted in more than one election within the same edition, we treat each ballot as a separate observation, regardless of whether it was cast by the same or different individuals. Second, we exclude empty ballots – those without any project selected – as evaluation reports suggest they typically reflect disinterest in available options rather than intentional abstention (Ośrodek Ewaluacji, 2019). Third, we exclude city-level projects from 2020–2021 from the comparative analysis due to their higher costs and different selection rules under k-approval. Table 1 outlines key differences across editions, showing variation in project scope and budget scale.

TABLE 1. Comparison of key summary indicators across the four studied editions of Warsaw PB

	Voters	Non-empty ballots	Total budget (PLN)	Eligible projects	Average cost of eligible projects (PLN)
2018 (total)	116 438	163 052	61 419 912	1 808	96 292
district-wide	62 934	50 326	7 156 371	306	101 887
sub-district	113 124	112 726	54 263 541	1 502	95 153
2019 (total)	88 977	128 682	64 784 230	1 628	94 964
district-wide	53 270	43 584	10 653 648	361	76 722
sub-district	85 277	85 098	54 130 582	1 267	100 161
2020 (total)	105 534	191 662	83 111 363	1 425	185 166
district-wide	104 997	104 941	58 177 954	1 324	148 588
city-wide	105 534	86 721	24 933 409	101	664 666
2021 (total)	108 376	203 631	83 111 363	1 503	204 680
district-wide	107 758	107 732	58 177 954	1 397	158 610
city-wide	108 376	95 899	24 933 409	106	811 852

Source: own elaboration.

Individual Behavior

The analysis of individual voting behavior in the Warsaw PB elections revealed distinct strategies among voters under knapsack versus k-approval ballots, particularly regarding the number and cost of projects that the residents were voting for. Table 2 presents the summary results. The average number of projects selected by a voter increased significantly between the knapsack editions (with around 6 projects in both 2018 and 2019) and the k-approval ones (with 9 and 10 in 2020 and 2021 respectively). Furthermore, a closer examination of the data distribution revealed that the new k-approval format seemed to motivate residents to choose the maximum number of projects allowed. For example, in the 2021 edition, almost half of the voters opted for 15 district-wide projects.

TABLE 2. Average number and cost of selected projects across the four studied editions of Warsaw PB

	2018	2019	2020	2021
	knapsack	knapsack	district 15-approval	district 15-approval
Average number of selected projects	5.9	6.4	9.0	10.2
Average total cost of projects selected on a ballot (PLN)	577 154	627 917	1 719 222	2 278 686
Average mean cost of an individual project selected on a ballot (PLN)	142 069	143 791	210 453	233 176

Source: own elaboration.

The average total cost of selected projects in 2020–2021 was about three times higher than in the earlier editions. Given that respondents selected more projects in the latter editions, the average cost per selected project showed a smaller, yet still meaningful difference – with voters selecting on average around 50% more expensive projects under k-approval compared to knapsack. Moreover, we observed that nearly half of the knapsack ballots had more than 80% of the available budget allocated. According to PB evaluation reports, many voters reported a conscious selection of one key project and then a strategic addition of smaller, less expensive projects to fully utilize their budget (Makurat *et al.*, 2018). Such behavior may have had unintended consequences, as voters – by supporting multiple cheaper initiatives – weakened the position of their key projects (Huras *et al.*, 2017). Furthermore, under k-approval, the average proportion of the total budget allocated to selected projects increased from 42% in 2020 to 57% in 2021, suggesting a behavioral adjustment as voters became more familiar with the new method.

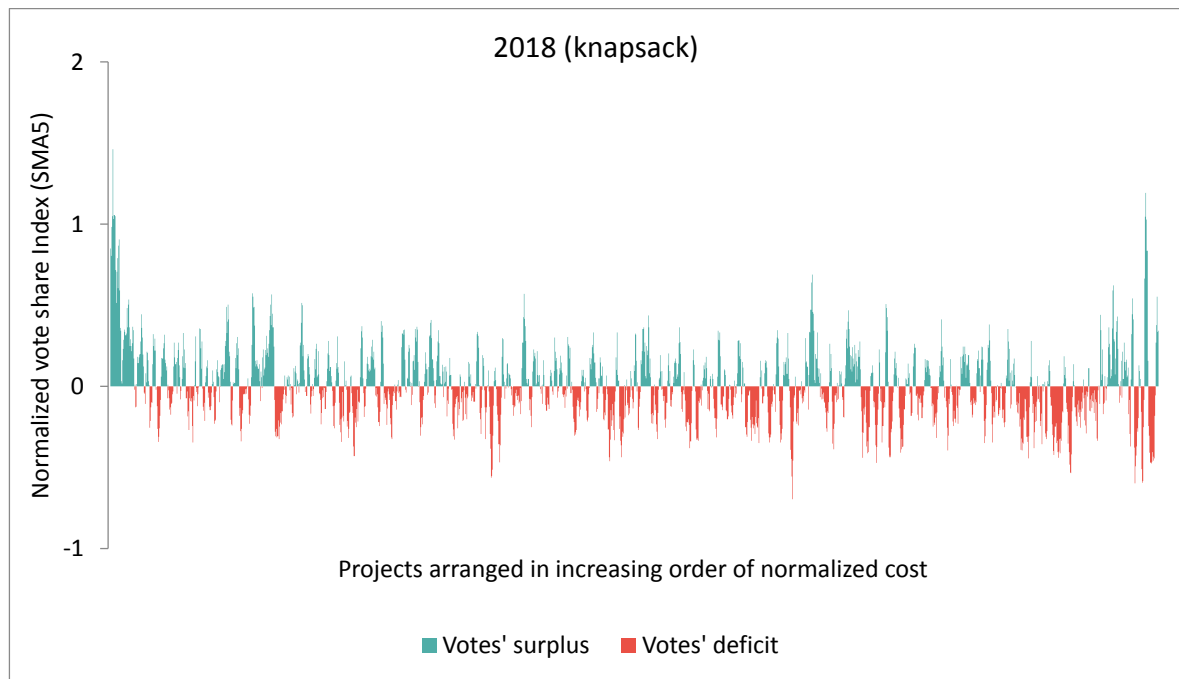
These results support the intuition that without a budget constraint, people are more likely to choose more expensive items. However, the average cost of projects eligible for voting also increased with the change in voting method, leading to potentially overlapping effects. To isolate the impact of the voting method itself, we apply the following transformation to the data.

We introduce a normalized project cost as a ratio of the individual cost of project p to the total budget available for an election k : $c^N(p) = \frac{c(p)}{B_k}$, where B_k is a total budget available for the k -th election, and $c^N(p) \in [0,1]$. We define a normalized share of votes as a ratio of the number of votes received by project p to the total number of votes cast for any project in the

election k : $w^N(p) = \frac{w(p)}{\sum_{q \in P_k} w(q)}$ and $w^N(p) \in [0,1]$. Let us denote a uniform share of votes for a project p as $w^U(p) = \frac{1}{m_k}$, where m_k is the number of projects on the k -th list.

We further assume that a project is considered more popular than expected if its normalized share of votes is higher than the share of votes coming from a uniform distribution, and less popular than expected otherwise. Thus, we define an index $I(p)$ as the ratio of the normalized share of votes to the uniform share of votes minus one: $I(p) = \frac{w^N(p)}{w^U(p)} - 1$. We have a “surplus” of votes if $I(p) > 0$, and a “deficit” of votes if $I(p) < 0$. We then proceed by plotting the indices against the projects arranged in increasing order of their normalized costs $c^N(p)$ to validate whether there is a linkage between the relative cost and popularity of the projects. To highlight the trends, we smooth out some of the naturally occurring popularity fluctuations from project to project by applying a simple 5-period moving average (SMA5). It means that we calculate the reported values as follows: $I_{SMA5}(p_i) = \frac{1}{5} \sum_{i=n-4}^n p_i$.

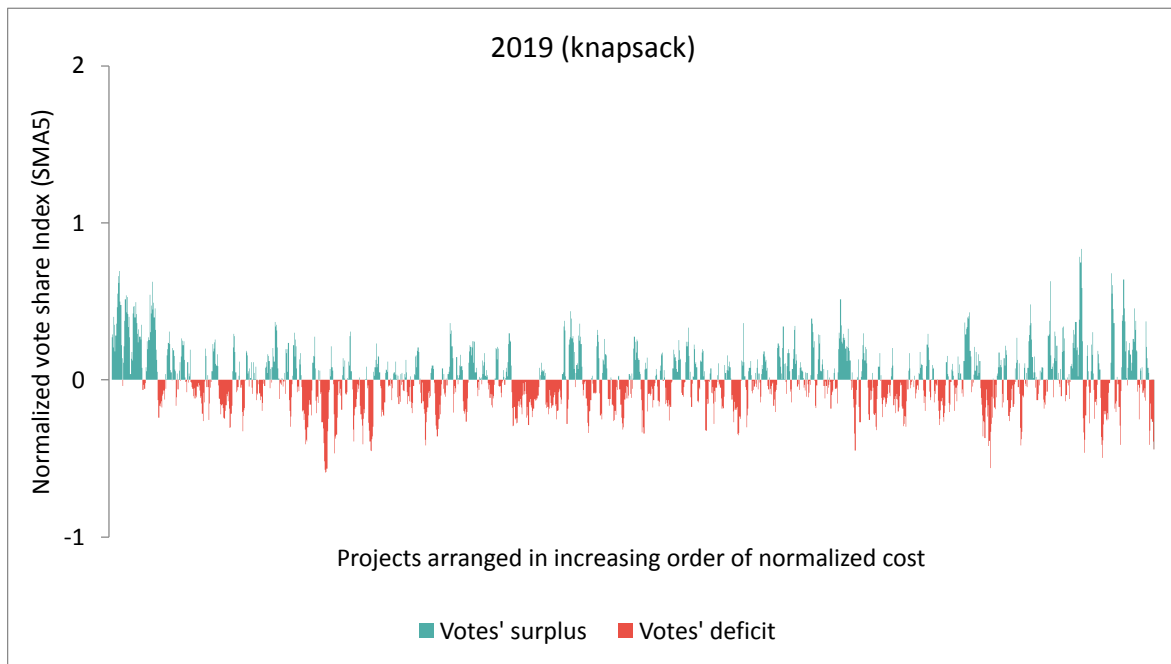
FIGURE 1. The index of normalized to uniform vote shares versus the normalized cost of projects in the 2018 edition



Note. Spearman's and Kendall's coefficients ($\rho = -0.13$, $\tau = 0.08$) show a weak but significant negative association between the normalized cost of the project and the normalized vote share over the uniform share ratio, $p < 0.05$.

Source: own elaboration.

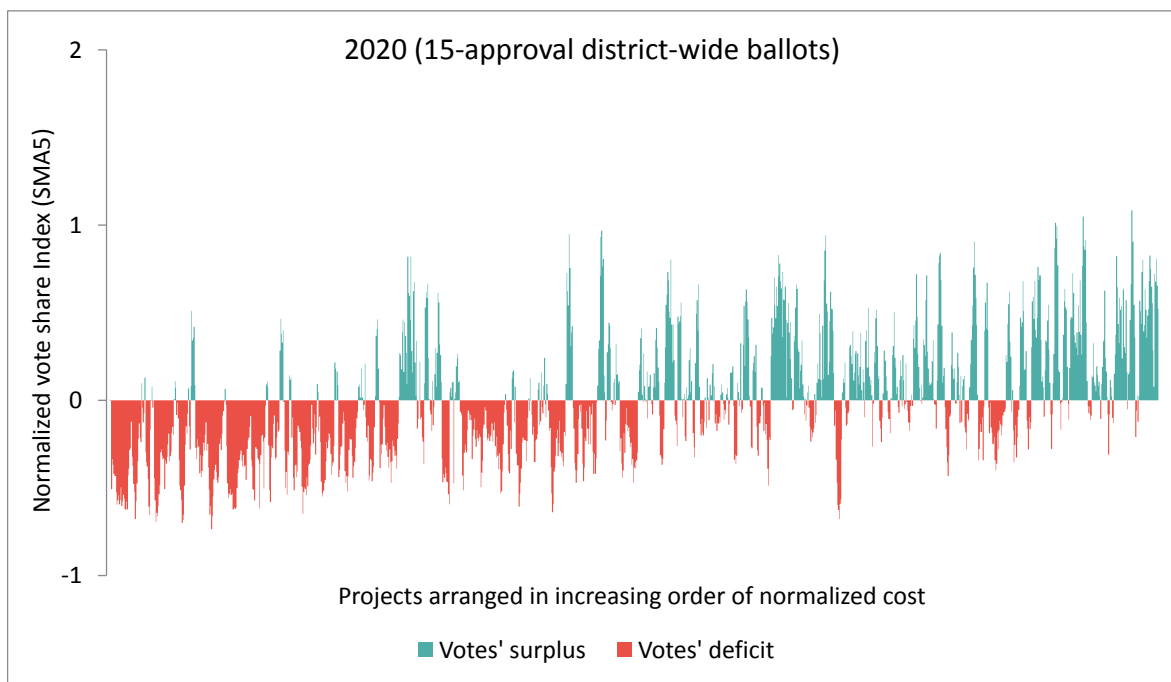
FIGURE 2. The index of normalized to uniform vote shares versus the normalized cost of projects in the 2019 edition



Note. Spearman's and Kendall's coefficients ($\rho = -0.02$, $\tau = 0.01$) show no significant association between the normalized cost of the project and the normalized vote share over the uniform share ratio, $p < 0.05$.

Source: own elaboration.

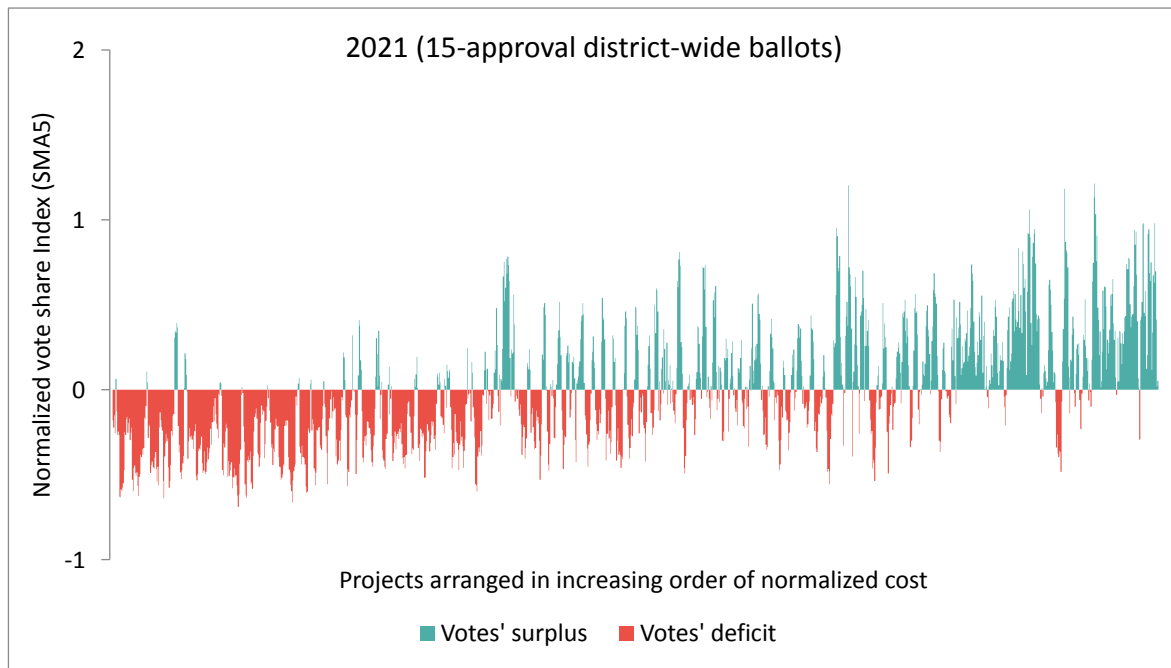
FIGURE 3. The index of normalized to uniform vote shares versus the normalized cost of projects in the 2020 edition



Note. Spearman's and Kendall's coefficients ($\rho = 0.33$, $\tau = 0.23$) show a significant positive association between the normalized cost of the project and the normalized vote share over the uniform share ratio, $p < 0.05$.

Source: own elaboration.

FIGURE 4. The index of normalized to uniform vote shares versus the normalized cost of projects in the 2021 edition



Note. Spearman's and Kendall's coefficients ($\rho = 0.36$, $\tau = 0.24$) show a significant positive association between the normalized cost of the project and the normalized vote share over the uniform share ratio, $p < 0.05$.

Source: own elaboration.

Under k-approval (Figures 3–4), the calculated Spearman's ρ and Kendall's τ coefficients revealed a significant positive association between the normalized cost of the project $c^N(p)$ and the normalized vote share over the uniform share ratio $\frac{w^N(p)}{w^U(p)}$, i.e., the higher the relative cost the higher the relative popularity of the project. Under knapsack (Figures 1–2), the popularity of projects was quite equally distributed across the full spectrum of relative costs. There was a certain vote surplus around the cheapest projects, which confirms the idea of residents selecting additional cheap projects to use up the allowed budget to the fullest.

Collective Outcomes

These different individual strategies influenced the collective outcomes in terms of the number and cost of the projects to be funded. Table 3 presents the summary results. The average cost of a winning project was around 70,000 PLN for the knapsack editions and more than double the amount (over 150,000 PLN) for the k-approval editions. As a result, with total budgets across the editions being relatively stable at the respective levels, less than half as many projects were funded in the 2020–2021 editions compared to the preceding ones.

TABLE 3. Number and average cost of winning projects across the four studied editions of Warsaw PB

	2018 knapsack	2019 knapsack	2020 district 15-approval	2021 district 15-approval
Number of winning projects	881	850	385	341
Average cost of a winning project (PLN)	68 135	73 352	150 896	170 477

Source: own elaboration.

Moreover, under knapsack voting, the smallest, least costly projects (under 25,000 PLN) were greatly overrepresented among funded projects compared to unfunded ones, as voters tended to maximize their budgets with multiple inexpensive choices. In the following k-approval editions, larger projects gained favor, and the cost distribution of funded projects closely matched the distribution of unfunded ones, indicating a shift toward favoring higher-cost options.

Experimental Study Results

The following section presents an online voting experiment designed to complement insights from the Warsaw PB study by addressing limitations of the natural experiment. Unlike the real-world setting, this controlled experiment isolates the effect of the voting method by holding all other factors constant. It aims to assess how voting behavior and outcomes differ when using knapsack versus k-approval ballots. In particular, we seek to validate the following research questions: Is the number of projects supported by voters under knapsack voting different from k-approval? Are voters more economical in their project selection under knapsack voting compared to k-approval? Does the knapsack ballot produce a different winning set of projects than the k-approval one? Furthermore, we explore how a modification of the aggregation rule changes the collective outcomes.

We targeted the experiment at Warsaw residents and designed it to closely mimic real elections in scope and experience. Conducted via the online Stanford Participatory Budgeting Platform,² the experiment was authorized and supported by the Centre for Public Communication of the Capital City of Warsaw Municipal Office. We recruited the participants by posting information on the website of the Municipal Office, a Facebook group for the Warsaw Participatory Budget, and several other groups dedicated to Warsaw districts' communities, as well as the Warsaw School of Economics online newspaper and newsletter. We administered the survey between November 27 and December 31, 2020.

We invited the participants, verified as Warsaw residents, to cast their votes in fictional participatory budgeting elections using both the knapsack and k-approval methods. The knapsack vote used an interactive budget bar to show participants their spending against a set budget of 10 million, while k-approval tracked selections against a maximum of 10 projects. After completing their votes using the first method, participants voted again using the second method with the same project list, with the order of voting methods assigned randomly.

We selected the list of 50 eligible projects from city-level proposals featured in the official 2018–2019 Warsaw PB editions. Each project retained its original cost and category, though the length of the list and the total budget were reduced for simplicity. While the use of real projects from the past editions aimed to ensure realism and relevance, it may have also introduced some familiarity bias. Participants might have previously encountered, voted for, or seen the implementation of some projects, potentially reinforcing prior preferences or shaping cost-effectiveness perceptions. While we do not expect familiarity to have significantly

² The Stanford Participatory Budgeting Platform is an open-source, free tool developed by the Stanford Crowdsourced Democracy Team at Stanford University for running online participatory budgeting elections. It is available at pbstanford.org.

impacted results, we acknowledge it as a potential limitation. To further anchor the design, the list deliberately included both top- and bottom-ranked projects from Warsaw’s 6th and 7th PB editions. Their corresponding strong and weak performances in the experiment suggest that participants’ choices aligned with actual voter behavior, lending credibility to the sample despite its non-representative nature.

TABLE 4. Comparison of key summary indicators across the two experimental scenarios

	Voters	Non-empty ballots	Non-empty ballots by voting order	Total budget (PLN)	Eligible projects	Average cost of eligible project (PLN)
Knapsack	180	160	1 st – 83 2 nd – 77	10 000 000	50	919 857
10-approval	180	179	1 st – 96 2 nd – 83	10 000 000	50	919 857

Source: own elaboration.

Wherever statistical analysis is performed on paired samples to compare voter behavior depending on the ballot type used, we consider only the voters who cast non-empty ballots for both scenarios. In total, there were 159 such voters (that is, almost 90% of the total sample). Out of them, 77 cast k-approval ballots first, and 82 cast knapsack ballots first. Throughout the following sections, we apply non-parametric tests since we could not adequately assume the normality of the collected data.

Individual Behavior

To assess how consistently participants selected projects across the two voting methods, we used the Jaccard index³ to compare individual ballots. Only 1.9% of participants made entirely different selections, while 5.7% chose exactly the same projects under both methods. Most participants showed moderate to high consistency, suggesting that their choices reflected stable preferences rather than random selection. Table 5 presents the average number and cost of selected projects across the two tested scenarios. On average, voters supported around 12 projects under the knapsack method compared to 9 under k-approval (significant difference using Wilcoxon signed-rank test, $p < 0.05$). The distribution patterns of the number of supported projects further reveal that in the k-approval scenario, over 75% of participants opted for the maximum of 10 projects, which is an even higher proportion than observed in the real elections.

Both the average total cost of all the projects selected on the ballot and the mean cost of a single supported project were around one-third higher under 10-approval compared to the knapsack scenario. These differences were statistically significant (Wilcoxon signed-rank test, $p < 0.05$). In the knapsack scenario, almost 90% of the participants allocated 80% or more of the total budget, aiming to use as much as possible of the available funds.

³ Jaccard index is a ratio between the size of the intersection and the size of the union of two sets S_1 and S_2 , that is $J(S_1, S_2) = \frac{|S_1 \cap S_2|}{|S_1 \cup S_2|}$. The metric formally describes the idea that two sets’ similarity depends on the number of common versus distinct elements.

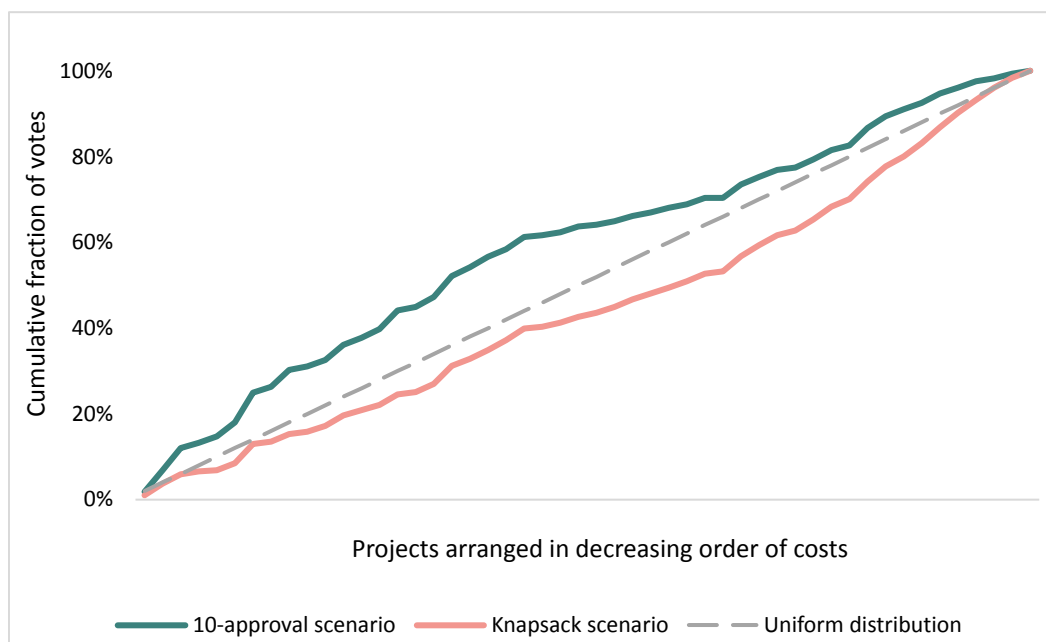
TABLE 5. Average number and cost of selected projects across the two experimental scenarios

	knapsack	10-approval
Average number of selected projects	11.7	9.0
Average total cost of projects selected on a ballot (PLN)	9 105 402	12 170 694
Average mean cost of an individual project selected on a ballot (PLN)	1 048 192	1 402 481

Source: own elaboration.

Following Goel *et al.* (2019), we compared cost considerations across knapsack and k-approval voters by plotting the cumulative fraction of votes against projects sorted in decreasing order of cost (Figure 5), thereby illustrating the greater cost-awareness in knapsack voting. The plot for k-approval dominates that of knapsack, indicating a relative overrepresentation of high-cost projects and suggesting lower cost-sensitivity among k-approval voters.

These results indicate that voters supported a different number of projects under knapsack and k-approval methods, with selections under knapsack voting reflecting more frugal and cost-conscious behavior.

FIGURE 5. The cumulative fraction of total votes plotted against project costs for the two experimental scenarios

Source: own elaboration.

Collective Outcomes

We determine the winning sets of projects for both the knapsack and 10-approval experimental ballots using a greedy utilitarian algorithm, consistent with the rules of the real Warsaw PB elections. The knapsack method resulted in 21 funded projects, while the 10-approval method selected only 10. Given the same available budget, the average cost of a project funded under knapsack was approximately half that of a project funded under 10-approval. Table 6 presents a comparison of the results.

While the knapsack and k-approval methods produced notably different winning sets, 9 k-approval winners also appeared in the knapsack set. Despite this overlap, the overall similarity was moderate (Jaccard index – 40.9%), given the larger size of the knapsack set. McNemar’s test confirmed a significant effect of the ballot type on whether a given project was funded ($p < 0.05$). However, despite the magnitude of the difference, with potentially meaningful real-life impact, the observed difference in project costs between the two methods was not statistically significant (Mann-Whitney U-test, $p > 0.05$).

These findings indicate that different ballot types yield different sets of winning projects. In particular, knapsack voting tends to favor less expensive projects compared to k-approval.

TABLE 6. Number and costs of winning projects across the two experimental scenarios

	knapsack	10-approval
Number of winning projects	21	10
Average cost of a winning project (PLN)	474 217	999 990

Note. In the 10-approval scenario, there was a tie and – broken randomly – it resulted in the reported winning set. While quantitatively different, the results based on the alternative outcome (had the tie been broken differently) were qualitatively consistent with those of the reported solution.

Source: own elaboration.

Alternative Aggregation Rules

This section investigates hypothetical “what-if” scenarios by modifying how individual inputs are aggregated into group decisions under the tested voting methods. Specifically, we replace the greedy utilitarian algorithm – currently used to determine winning projects – with its greedy cost-proportional variant. This alternative approach selects projects based on the highest vote-to-cost ratios (rather than simply the highest vote counts), thus emphasizing cost-efficiency in project selection.

The purpose of this exercise is not to evaluate the performance of aggregation rules per se, but rather to illustrate how – given the same voter input – applying different algorithms can lead to different collective outcomes. While our experimental setup was not designed for a definitive assessment of algorithmic effects on voter behavior, readers may approach this section as a demonstration of the often-overlooked yet critical role of aggregation rule design in shaping final results, even when individual preferences remain unchanged.

Applying this modification to the experimental data resulted in two new winning sets, both consisting of the same list of 37 projects, regardless of the original ballot type (with only minor shifts in the relative ranks). Notably, respondents consistently selected five projects across all four scenarios. These were generally mid- to low-cost initiatives that attracted sufficient support even when cost-efficiency was explicitly prioritized. Replacing the basic utilitarian algorithm with its cost-proportional counterpart had a significant impact on which projects were funded (McNemar’s test, $p < 0.05$). Table 7 compares the sizes and costs of the winning sets under the two original and the two what-if scenarios. The number of funded projects increased under the cost-proportional algorithm, which corresponded to the lowest average costs of funded projects. They were close to one-fourth of the highest ones under the original

10-approval scenario. However, the observed differences in winning projects' costs were not statistically significant (Mann-Whitney U-test, $p < 0.05$).

TABLE 7. Comparison of the winning sets of projects under different aggregation rules

	Winning projects	Average cost of a winning project (PLN)
Knapsack with utilitarian (original)	21	474 217
10-approval with utilitarian (original)	10	999 990
Knapsack with cost-proportional (what-if)	37	250 953
10-approval with cost-proportional (what-if)	37	250 953

Source: own elaboration.

Given that participants were unaware of the alternative aggregation method during the experiment, we make the necessary assumption that their voting behavior would have remained unchanged. However, similar to the ballot format, awareness of the aggregation rule could potentially influence voters' strategies. Future research could investigate how transparency regarding the aggregation method affects individual voting decisions and, consequently, collective outcomes.

By prioritizing vote-to-cost ratios rather than vote counts alone, the cost-proportional algorithm favors more cost-efficient projects, similar in spirit to the logic embedded in knapsack voting, albeit applied at a different stage. Using both simultaneously could risk over-favoring cheaper projects. However, in our case, the influence of the cost-proportional rule was so dominant that the initial ballot type had little impact on the final outcomes. The significant differences between knapsack and k-approval observed under the default utilitarian algorithm largely disappeared under the cost-proportional approach.

Voter Experience and Perceptions

We conducted a voluntary poll at the end of the experiment to gather participants' feedback on their experience with the two tested ballot formats, assessing the simplicity, effort, cost-benefit consideration, and preferences for real-world implementation. Out of 180 voters, 139 (over 75%) chose to provide feedback, indicating a high engagement level and enhancing the reliability of the collected data.

Participants answered a series of questions using 5-point scales designed to assess their experience with each voting method. First, they rated how well they understood the voting rules, with scale endpoints ranging from "simple" to "complicated". Respondents saw both methods as very intuitive, with over 90% of respondents selecting one of the two highest simplicity ratings, and no significant difference observed between them. When evaluating the effort involved in casting their vote (from "easy" to "hard"), participants found the 10-approval ballot to be less demanding, with 84.2% choosing one of the two lowest effort ratings, compared to 63.3% for knapsack. Furthermore, when asked about how much they considered the relationship between project costs and benefits (from "a lot" to "not at all"), respondents indicated that the knapsack method prompted significantly more cost-conscious behavior, with 58.3% selecting one of the two highest levels of consideration, compared to only 31.7% for 10-approval.⁴ These self-reported perceptions align with our experimental results showing knapsack voting encourages more

⁴ All reported differences were tested for statistical significance using the Wilcoxon signed-rank test, with $p < 0.05$

economical project selection, as well as with empirical observations coming from the studied editions of the real Warsaw PB.

Finally, when asked which method they would prefer for the real Warsaw PB, 46% favored knapsack, 28% preferred k-approval, about 24% liked both equally, and 2% selected neither. Thus, despite knapsack's higher perceived effort, often linked in open-ended feedback to more conscious consideration of project cost, many voters appreciated the more thoughtful decision-making it encouraged, indicating that they did not view the extra effort as a disadvantage.

Conclusions

This research highlights how cost consideration mechanisms embedded in participatory budgeting voting rules affect voters' behavior, showing that knapsack voting encourages more economically mindful project selection, while k-approval voting tends to favor higher-cost projects. By integrating Warsaw PB data across its four consecutive editions with a complementary online experiment, the study offers a comprehensive evaluation of how different voting rules shape both individual decisions and collective outcomes.

We reveal that voting strategies show a strong response to the specific incentives built into the voting rules. Under the budget-limited knapsack format, voters tend to maximize the use of available funds. Conversely, k-approval voting leads voters to select the maximum allowed number of projects with less regard to their costs. This results in a tendency for knapsack voters to be more cost-conscious, compared to k-approval. Consequently, we observe that the 2018–2019 Warsaw PB elections, using the k-approval, funded fewer but more expensive projects, while the 2020–2021 editions, held under the newly implemented knapsack voting, selected a greater number of smaller, lower-cost projects.

Our experimental survey results support findings from real-life Warsaw PB data. We confirm that voters are more economical in their project selection under knapsack voting compared to k-approval, leading to the selection of a greater number of projects at lower costs. Moreover, when we replace the standard greedy aggregation with a cost-proportional variant, the winning sets shift further toward lower-cost projects, suggesting that algorithmic design can influence collective outcomes, even when individual inputs remain constant. These insights align with and contribute to the findings from the existing literature (Goel *et al.*, 2019; Talmon & Faliszewski, 2019; Laruelle 2020), which indicates a bias toward cheaper projects when cost considerations are integrated into the voting method, either through the ballot type (knapsack vote) or the aggregation rule (cost-proportional algorithm). Finally, our user experience survey reveals that while respondents perceive knapsack voting as more demanding, it is often preferred due to the more deliberate, cost-aware decision-making it encourages.

Since different participatory budgeting voting methods can lead to different winning sets of projects – representing the interests of the same group of residents – the choice of ballot type (and aggregation rule) should receive serious consideration in the PB process. This work advances our understanding of the design complexities involved and offers evidence-based insights that can help policymakers make informed decisions about selecting voting methods to effectively meet community needs and priorities. While our findings can provide practical

indicating significant effects where noted.

guidelines for the design and implementation of PB solutions, we want to clarify that this article does not make normative recommendations. Instead, it outlines specific properties of different voting methods that one may see as advantageous in certain contexts and disadvantageous in others.

Acknowledgments

This article is a revised and extended version of selected chapters from my PhD dissertation. I am grateful to my advisors, Honorata Sosnowska and Paweł Pońsko, as well as to the dissertation reviewers, Mikołaj Jasiński and Tadeusz Trzaskalik, for their valuable guidance and feedback. I also wish to thank the Stanford Crowdsourced Democracy Team – Lodewijk Ge-lauff, Ashish Goel, and Sukolsak Sakshuwong – for their generous support and technical assistance with the Stanford Participatory Budgeting Platform. This research was supported by the Doctoral Scholarship of the Capital City of Warsaw (Grant No. CKS/B/V/8/8/U/247/2019). I gratefully acknowledge the City of Warsaw, especially Magdalena Stefańska, for providing data access and institutional support.

References

- Arrow, K. (1951). *Social Choice and Individual Values*. Wiley. <https://doi.org/10.12987/9780300186987>.
- Aziz, H., Lee, B.E., & Talmon, N. (2018). Proportionally Representative Participatory Budgeting: Axioms and Algorithms, *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '18)*, 23–31.
- Aziz, H. & Shah, N. (2021). Participatory Budgeting: Models and Approaches. In: T. Rudas & G. Péli (Eds.), *Pathways Between Social Science and Computational Social Science* (pp. 215–236), Springer. https://doi.org/10.1007/978-3-030-54936-7_10
- Benade, G., Itzhak, N., Shah, N., Procaccia, A.D., & Gal, Y. (2018). *Efficiency and usability of participatory budgeting methods* [Unpublished manuscript]. Available from: https://www.cs.toronto.edu/~nisarg/papers/pb_usability.pdf (accessed: 15.05.2025).
- Benade, G., Nath, S., Procaccia, A., & Shah, N. (2017). Preference Elicitation For Participatory Budgeting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 376–382. <https://doi.org/10.1609/aaai.v31i1.10563>
- Cabannes, Y. (2004). Participatory budgeting: a significant contribution to participatory democracy, *Environment and Urbanization*, 16(1), 27–46. <https://doi.org/10.1177/095624780401600104>.
- Directive No. 1660/2017 of the President of the Capital City of Warsaw of October 18, 2017 on social consultations with residents of the Capital City of Warsaw regarding the participatory budget. Available from: https://um.warszawa.pl/documents/57254/21498757/regulamin_przeprowadzania_budzetu_partycypacyjnego_w_warszawie_na_rok_2019.pdf/8de8aaed-3c79-6c20-27b6-952092f07207?t=1619801066481 (accessed: 15.05.2025).
- Elkind, E., Faliszewski, P., Skowron, P., & Slink, A. (2017). Properties of multiwinner voting rules. *Social Choice and Welfare*, 48(3), 599–632. <https://doi.org/10.1007/s00355-017-1026-z>.
- Faliszewski, P., Flis, J., Peters, D., Pierczyński, G., Skowron, P., Stolicki, D., Szufa, & Talmon, N. (2023). *Participatory budgeting: Data, tools, and analysis*. arXiv preprint. <https://doi.org/10.48550/arXiv.2305.11035>

- Gelauff, L., & Goel, A. (2024). Rank, pack, or approve: Voting methods in participatory budgeting. *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 18, pp. 448–461). <https://doi.org/10.1609/icwsm.v18i1.31326>
- Goel, A., Krishnaswamy, A., Sakshuwong, S., & Aitamurto, T. (2015). *Knapsack Voting: Voting mechanisms for Participatory Budgeting*. Presented at the Collective Intelligence Conference 2015, Santa Clara, CA. Available from: https://web.stanford.edu/~anilesh/files/knapsack_voting_full.pdf (accessed: 15.05.2025).
- Goel, A., Krishnaswamy, A., Sakshuwong, S., & Aitamurto, T. (2019). Knapsack Voting for Participatory Budgeting, *ACM Transactions on Economics and Computation*, 7, 1–27. <https://doi.org/10.1145/3340230>
- Huras, P., Przybył, C., Bienkowska, Z., Chojecki, J., Grajek, A., Kamińska, A., Matysiak, M., Nowińska, K., & Pałka, S. (2017). *Badanie ewaluacyjne IV edycji budżetu partycypacyjnego w kontekście mieszkańców zaangażowanych w proces*. Available from: https://um.warszawa.pl/documents/57254/21471167/raport_koncowy_-_ewaluacja_iv_edycji.pdf/fe16b8bd-93c7-ae88-7402-ab7a91b04db7?t=1634498675568 (accessed: 15.05.2025).
- Laruelle, A. (2020). Voting to select projects in participatory budgeting. *European Journal of Operational Research*, 288(2), 598–604. <https://doi.org/10.1016/j.ejor.2020.05.063>
- Makurat, M., Charchuła, Z., & Wałaszek, J. (2018). *Badanie ewaluacyjne V edycji budżetu partycypacyjnego w Warszawie. Raport końcowy*. Available from: https://um.warszawa.pl/documents/57254/21548037/raport_z_ewaluacji_v_edycji.pdf/5d4b9c67-083c-7d07-6864-28dd942b767e?t=1634498675316 (accessed: 15.05.2025).
- Nurmi, H. (2010). Voting Systems for Social Choice. In: D. Kilgour & C. Eden (eds.), *Handbook of Group Decision and Negotiation* (pp. 167–182). Springer. https://doi.org/10.1007/978-90-481-9097-3_11
- Ośrodek Ewaluacji. (2019). *Badanie ewaluacyjne VI edycji budżetu obywatelskiego w Warszawie. Raport końcowy*. Available from: https://um.warszawa.pl/documents/57254/21602385/raport_z_ewaluacji_vi_edycji_budzetu_obywatelskiego.pdf/28052dcc-223a-4d65-c296-4d0c21aa537a?t=1634498674593 (accessed: 15.05.2025).
- Resolution No. XI/218/2019 of the Council of the Capital City of Warsaw of April 11, 2019 on social consultations with residents of the Capital City of Warsaw regarding the civic budget. Available from: <https://um.warszawa.pl/documents/57254/21602391/Uchwa%C5%82a+Rady+m.st.+Warszawy+w+sprawie+bud%C5%BCetu+obywatelskiego.pdf/1e55f23d-7d64-4f30-1a3b-8def8be168fa?t=1620746261772> (accessed: 15.05.2025).
- Rosenfeld, A., & Talmon, N. (2021). *What Should We Optimize in Participatory Budgeting? An Experimental Study*. arXiv preprint. <https://doi.org/10.48550/arXiv.2111.07308>
- Sen, A. (1970). *Collective Choice and Social Welfare*. Elsevier. <https://doi.org/10.1016/C2009-0-12011-1>.
- Stolicki, D., Szufa, S., & Talmon, N. (2020). *Pabulib: A Participatory Budgeting Library*. ArXiv:2012.06539.
- Sudolska-Bytof, A. & Janiak, M. (2019). *Warsaw Participatory Budget: Case Study Report & Analysis*. Inicjatywy. Available from: https://bugetareparticipativa.viitorul.org/wp-content/uploads/2019/03/REPORT_WARSZAW-PB-1.pdf
- Talmon, N. & Faliszewski, P. (2019). A Framework for Approval-Based Budgeting Methods, *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 2181–2188. <https://doi.org/10.1609/aaai.v33i01.33012181>
- Wampler, B. (2007). A Guide to Participatory Budgeting. In: A. Shah (ed.), *Participatory Budgeting* (Chapter 1: pp. 21–54). The World Bank. <https://doi.org/10.1596/978-0-8213-6923-4>

