The Fortieth AAAI Conference on Artificial Intelligence (AAAI-26)

Understanding the Impact of Proportionality in Approval-Based Multiwinner Elections

Niclas Boehmer[1], **Lara Glessen**[1], **Jannik Peters**[2]
[1]Hasso Plattner Institute, University of Potsdam, Germany
[2]National University of Singapore
niclas.boehmer@hpi.de, lara.glessen@hpi.de, peters@nus.edu.sg

Abstract

Despite extensive theoretical research on proportionality in approval-based multiwinner voting, its impact on which committees and candidates can be selected in practice remains poorly understood. We address this gap by (i) analyzing the computational complexity of several natural problems related to the behavior of proportionality axioms, and (ii) conducting an extensive experimental study on both real-world and synthetic elections. Our findings reveal substantial variation in the restrictiveness of proportionality across instances, including previously unobserved high levels of restrictiveness in some real-world cases. We also introduce and evaluate new measures for quantifying a candidate's importance for achieving proportional outcomes, which differ clearly from assessing candidate strength by approval score.

Code — https://github.com/lara-glessen/impact-proportionality-mwav.git
Extended version — https://arxiv.org/abs/2511.09479

1 Introduction

Proportionality in multiwinner voting is one of the most actively studied topics in computational social choice. Given potentially conflicting preferences of voters over candidates, the goal is to select a fixed-size subset of candidates that proportionally represents the preferences of the voters. Applications of proportional multiwinner voting arise in contexts such as participatory budgeting (Peters, Pierczyński, and Skowron 2021), blockchain systems (Boehmer et al. 2024a), or civic participation platforms (Revel et al. 2025). One of the building blocks of proportional multiwinner voting are so-called proportionality axioms: properties that certify the proportionality of an outcome. Within the study of proportionality, *approval-based* multiwinner voting, in which each voter votes for a subset of the candidates they approve of, has emerged as a simple, yet expressive model. Accordingly, researchers have put significant effort into the development of axioms capturing various notions of proportionality, their theoretical analysis, and the design of voting rules that satisfy them (e.g., Aziz et al. 2017; Peters and Skowron 2020; Brill and Peters 2023; Halpern et al. 2025; Aziz et al. 2023; Kehne, Schmidt-Kraepelin, and Sornat 2025).

Despite the extensive theoretical coverage of proportionality axioms, research into their practical impact remains limited. We are only aware of two papers whose experiments explicitly focus on how proportionality axioms constrain the set of outcomes in approval-based multiwinner elections: Bredereck et al. (2019) and Brill and Peters (2023) (see Section 1.2 for further related work). Both of these papers measure the strength of proportionality axioms through their restrictiveness, i.e., they compute the fraction of committees satisfying an axiom in elections sampled from various synthetic models. The conclusions Bredereck et al. and Brill and Peters draw can be viewed as quite negative. Bredereck et al. remark that even for extended justified representation (EJR) (Aziz et al. 2017), the strongest axiom they study: "satisfying EJR is quite easy" (Bredereck et al. 2019, page 114), as in the majority of their elections, a majority of *all* committees satisfy EJR. Similarly, although the EJR+ axiom introduced by Brill and Peters is stronger than EJR, their findings suggest that committees satisfying EJR+ remain quite common in most of their synthetic sampling models.

This leaves us in an unsatisfying state. One could interpret the prior work as evidence that proportionality axioms impose only extremely mild constraints, raising doubts about their importance and practical relevance: If proportionality axioms were rarely restrictive, imposing them would not change the set of selectable committees. Thus, appealing to proportionality as a justification for a rule, as commonly done, would be of very limited practical significance. At the same time, both studies focus exclusively on synthetically generated elections, raising questions about the extent to which their conclusions translate to realistic instances and whether proportionality axioms might be even less powerful on real-world preference data. This motivates our *first research question*: to what extent do proportionality axioms constrain which committees can be selected in practice? To illustrate the significant, yet nuanced, practical impact proportionality may have, and to motivate further investigation, we present Figure 1. This plot is based on a participatory budgeting election conducted in the Praga-Północ district of Warsaw in 2022. Treating this election's votes and projects as an approval-based multiwinner election, we approximate, for each possible committee size k, the percentage of committees satisfying the weak proportionality axiom JR (red)

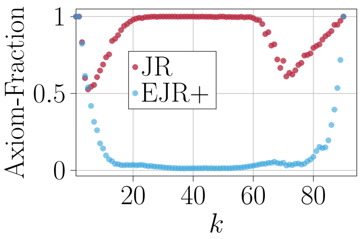


Figure 1: Percentage of committees satisfying JR and EJR+ depending on the committee size k, in the Warsaw Praga-Północ PB election 2022.

and the stronger axiom EJR+ (blue).[1] We observe that for certain values of k, only a small fraction of committees satisfy EJR+, and that the satisfaction rate for JR can fluctuate in seemingly unpredictable ways. This example also highlights a potential limitation of prior experimental work on proportionality, which typically evaluates restrictiveness only for a single, hand-picked value of k.

In addition to studying the impact of proportionality on the space of feasible committees, as our *second research question*, we investigate the impact of proportionality on the merit of different candidates. In elections based on the principle of individual excellence, in which candidates with the most approvals are selected, a candidate's merit is naturally quantified by their approval score. In contrast, when aiming to select a committee that satisfies a proportionality axiom, assessing a candidate's strength becomes more complex, as the proportionality of a committee containing the candidate always depends on which other candidates are present. Good measures for a candidate's strength under proportionality would contribute to improving the transparency and our understanding of proportional multiwinner elections.

1.1 Our Contributions

Our investigation consists of two parts: an algorithmic analysis (Section 3) and an experimental analysis (Section 4), in both of which we focus on the axioms JR and EJR+, which can be verified and satisfied in polynomial time (see Section 2 for definitions). Proofs and additional experimental results can be found in our full version (Boehmer, Glessen, and Peters 2025).

In the algorithmic part, we study the complexity of a range of computational problems aimed at understanding the impact of proportionality in a given instance. Our selection of questions serves two purposes. First, we study the problems that we would like to compute in our experimental analysis. Second, we provide new results on existing computational problems and introduce new problems related to the behavior of proportionality axioms, which might be of independent interest and help to understand the boundaries of tractability when dealing with (easy-to-verify) proportionality axioms. These problems include (i) deciding whether two proportionality axioms are equivalent in a given instance, (ii) finding a pair of distant proportional committees, (iii) finding a proportional committee containing a given subset of candidates, (iv) and counting proportional committees. We show that all of our problems are NP-hard (resp. #P-hard), highlighting that even for easy-to-verify and easy-to-construct axioms, slightly more nuanced computational questions immediately yield computational hardness. Complementing these results, we develop ILP- and sampling-based algorithms for our experiments and investigate our problem's parameterized complexity. While simple brute-force algorithms yield fixed-parameter tractability (FPT) with respect to the number of candidates m for all problems, we develop more involved algorithms showing FPT with respect to the number of voters n, a parameterization of interest in case only a few opinions need to be aggregated, e.g., in expert or criteria aggregation.

In our experimental analysis, we investigate the impact of proportionality on the space of feasible committees and on candidates' strength. Due to the general lack of real-world data on approval-based multiwinner elections (Boehmer et al. 2024b), we base our analysis on voting data from participatory budgeting instances available in pabulib (Faliszewski et al. 2023a), complemented by synthetic data generated using the resampling and Euclidean models. Our experimental findings include the following observations: (i) The restrictiveness of proportionality axioms can vary drastically and in unexpected ways with small changes in k, highlighting that the choice of k deserves more careful consideration than it typically receives in experimental studies. (ii) While JR and EJR+ often do not impose strong restrictions, there are some real-world instances where they can significantly constrain the outcome space. Even in such cases, the set of proportional committees tends to be highly diverse. (iii) We introduce two measures for candidates' importance for proportionality, one of which is inspired by the Banzhaf power index in cooperative game theory. Both measures turn out to be highly correlated. We demonstrate that using candidates' approval scores to assess their proportionality merit can be highly misleading.

1.2 Related Work

For an overview of experimental work in approval-based multiwinner voting (and beyond), we refer to the recent survey-style "guide" by Boehmer et al. (2024b).

The work closest to ours is the aforementioned paper by Bredereck et al. (2019), which was the first to experimentally investigate the impact of proportionality axioms. They sampled from different variants of the Impartial Culture and Euclidean model, evaluated how many committees satisfy JR, PJR, and EJR, analyzed the size of the smallest candidate set satisfying JR, and studied the satisfaction and coverage scores achievable by JR committees.[2] Bredereck et al. also introduced and studied related algorithmic questions, such as the complexity of counting JR, PJR, or EJR committees.

---

[1] Note that the slightly higher reported percentages for JR than EJR+ are due to small sampling errors.

[2] In a similar setup, satisfaction and coverage scores of JR committees were studied by Elkind et al. (2024), while the size of JR-satisfying candidate sets was examined by Elkind et al. (2023).

Our work differs from theirs in that we focus our experiments on real-world voting data and explore new aspects of the impact of proportionality.

In their work on verifiable proportionality axioms, Brill and Peters (2023) also conducted experimental studies on how often random committees satisfy different proportionality axioms, using synthetic elections from four models (not including the Euclidean model). While they observed that their new axiom EJR+ is significantly harder to satisfy than existing ones, such as EJR, for a large part of their parameter space, still over 50% of the committees satisfy it.

Faliszewski et al. (2023b) and Boehmer et al. (2024a) examined how frequently committees selected by various approval-based multiwinner voting rules satisfy proportionality axioms. Their main finding is that common proportionality axioms are frequently satisfied by a large array of rules. Faliszewski et al. (2023b) studied this question using elections sampled from various synthetic models (Szufa et al. 2022), along with some instances from pabulib, whereas Boehmer et al. (2024a) analyzed large real-world elections from the proof-of-stake Polkadot blockchain.

Finally, a recent study by Bardal et al. (2025) empirically examined proportionality and proportional rules in the setting of ordinal multiwinner voting, using a dataset from Scottish local government elections.

2 Preliminaries

Model and Notation. For $t \in \mathbb{N}$, we let $[t] := \{1, \dots, t\}$. Let $N = [n]$ be a set of *voters* and $C = \{c_1, \dots, c_m\}$ be a set of *candidates*. Each voter $i \in N$ possesses an *approval ballot* $A_i \subseteq C$ indicating the candidates approved by this voter. For a given candidate $c \in C$, we denote by $N_c := \{i \in N : c \in A_i\}$ the set of approvers of c and refer to $|N_c|$ as the *approval score* of c. Together, $A = (A_i)_{i \in N}$ forms an approval profile. We let $k \leq m$ denote the committee size and refer to any set $W \subseteq C$ of candidates of size $|W| = k$ as a *committee*. Together, the tuple $\mathcal{I} = (N, C, A, k)$ forms an *approval-based multiwinner election*.

Proportionality Axioms. The central objects we study in our work are *proportionality axioms*. We focus on two easy-to-verify axioms: *justified representation (JR)* (Aziz et al. 2017) and *extended justified representation plus (EJR+)* (Brill and Peters 2023).

Definition 2.1 (EJR+)**.** Given an election $\mathcal{I} = (N, C, A, k)$ and parameter $t \in [k]$, a subset $W \subseteq C$ satisfies *t-extended justified representation plus (t-EJR+)* if for any candidate $c \notin W$, $\ell \in [t]$, and set $N' \subseteq N_c$ of voters with $|N'| \geq \ell \frac{n}{k}$, there exists an $i \in N'$ with $|A_i \cap W| \geq \ell$.

A committee satisfies justified representation (JR) if it satisfies 1-EJR+, and EJR+ if it satisfies k-EJR+. By definition, any committee satisfying EJR+ also satisfies JR. EJR+ and JR committees are guaranteed to exist and can be computed in polynomial time (Brill and Peters 2023). For some election $\mathcal{I}$, we denote by $\mathrm{EJR+}(\mathcal{I})$, resp. $\mathrm{JR}(\mathcal{I})$, the set of committees satisfying EJR+, resp. JR, in $\mathcal{I}$. Moreover, we refer to $|\mathrm{EJR+}(\mathcal{I})|/\binom{m}{k}$, resp. $|\mathrm{JR}(\mathcal{I})|/\binom{m}{k}$, as the election's *EJR+-fraction*, resp. *JR-fraction*. Further, for some candidate $c \in C$, we let $\mathrm{EJR+}(\mathcal{I}, c) \subseteq \mathrm{EJR+}(\mathcal{I})$, resp. $\mathrm{JR}(\mathcal{I}, c) \subseteq \mathrm{JR}(\mathcal{I})$, be the set of committees satisfying EJR+, resp. JR, that contain c.

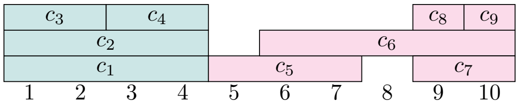


Figure 2: Example instance to illustrate JR and EJR+.

Example 2.1. *Consider the instance depicted in Figure 2: voters correspond to integers on the x-axis, while candidates correspond to boxes, with every voter approving the candidates above them, e.g., voter 1 approves c_1, c_2, and c_3. For $k = 5$, the committee $\{c_1, c_3, c_4, c_5, c_7\}$ satisfies JR, as the only "uncovered" voter 8 is on its own not enough to constitute a JR violation. It, however, does not satisfy EJR+ as witnessed by candidate c_6. This candidate is approved by five voters, but no voter approves two candidates in the committee. However, each one of them approves at most one candidate, leading to an EJR+ violation (as $n/k = 2$). The committee $\{c_1, c_2, c_3, c_4, c_6\}$ satisfies EJR+.*

Parameterized Algorithms. In our algorithmic analysis, we make use of the framework of parameterized algorithms (Cygan et al. 2015), focusing on the complexity classes FPT and W[1]. A parameterized problem X consists of an instance $\mathcal{I}$ together with a parameter $\ell \in \mathbb{N}$. A fixed-parameter tractable (FPT) algorithm for $(\mathcal{I}, \ell)$ is an algorithm running in time $\mathcal{O}(f(\ell) \cdot |\mathcal{I}|^{\mathcal{O}(1)})$ for some computable function f. Under standard complexity theoretical assumptions, W[1]-hard problems do not admit an FPT algorithm.

Measures of Candidate Importance for Proportionality. We introduce two ways in which we can quantify a candidate's importance for proportional representation. As our first measure, we define the *JR-prevalence*, resp., *EJR+-prevalence*, of some candidate $c \in C$ as the fraction of committees fulfilling the axiom that contain c, i.e., $|\mathrm{JR}(\mathcal{I},c)|/|\mathrm{JR}(\mathcal{I})|$, resp. $|\mathrm{EJR+}(\mathcal{I},c)|/|\mathrm{EJR+}(\mathcal{I})|$. For our second measure, we draw inspiration from the raw Banzhaf power index (Banzhaf 1965) from cooperative game theory, which counts how often an agent is "pivotal" for a coalition to be winning. Translating this idea, we count the number of proportional committees for which proportionality is violated after we remove the candidate: We let the *JR-power-index* (resp. *EJR+-power-index*) of candidate $c \in C$ be $|\{W \in \mathrm{JR}(\mathcal{I}, c) \colon W \setminus \{c\} \text{ does not satisfy JR}\}|$ (resp. $|\{W \in \mathrm{EJR+}(\mathcal{I}, c) \colon W \setminus \{c\} \text{ does not satisfy EJR+}\}|$).

3 Algorithmic Analysis

In this section, we present an algorithmic analysis of combinatorial problems we would like to solve to better understand the impact of proportionality axioms in practice.

3.1 Differentiating Proportionality Axioms

We start with a simple problem: can we determine whether, for a given instance, EJR+ is a strictly stronger property than

JR, i.e., is there a committee satisfying JR, but not EJR+?

JR-NOT-EJR+
Input: Election $\mathcal{I} = (N, C, A, k)$.
Question: Is there a committee $W \subseteq C$ satisfying JR but not EJR+?

Despite the significant conceptual differences between EJR+ and JR and the fact that both can be verified in polynomial time, it turns out that checking whether JR and EJR+ coincide on a given instance is NP-complete.

Theorem 3.1. JR-NOT-EJR+ *is NP-complete, and W[1]-hard when parameterized by the committee size k.*

In fact, our reduction from a variant of MULTICOLORED INDEPENDENT SET even shows that it is hard to check the existence of a committee satisfying JR, but not 2-EJR+.

We complement this negative result in two ways. Firstly, we present a simple polynomial-size ILP for JR-NOT-EJR+ in our full version (Boehmer, Glessen, and Peters 2025). Secondly, we show that JR-NOT-EJR+ is in FPT when parameterized by the number of voters n by introducing equivalence classes of candidates: We say that two candidates are equivalent if they are approved by the exact same set of voters. Since there are at most 2^n such classes, we can formulate an ILP where the number of variables is a function in n by replacing candidates by their equivalence class. Using Lenstra's algorithm, which solves ILPs in FPT time with respect to the number of variables (Lenstra 1983), we get an FPT algorithm for our problem.

Theorem 3.2. JR-NOT-EJR+ *is in FPT when parameterized by the number of voters n.*

3.2 Distances Between Proportional Committees

To measure the *richness* and *diversity* of the space of proportional committees, we want to compute how different they can be. For two committees W_1 and W_2, we define their distance as $d(W_1, W_2) := |W_1 \setminus W_2|$. Dong et al. (2026) prove that JR committees can be quite different and even isolated: there exists an instance with a JR-committee W_1 such that the closest other JR-committee W_2 has distance $k - 1$. This motivates us to study the following problem:

X-DIFF-COMMITTEES
Input: Election $\mathcal{I} = (N, C, A, k)$ and $k' \leq k$.
Question: Are there two committees W_1 and W_2, both satisfying axiom X, with $d(W_1, W_2) \geq k'$?

Bredereck et al. (2019) showed that for any election, every candidate that is approved by at least one voter belongs to at least one JR-committee. Thus, we conclude that any election with $m > k$ has two JR-committees W_1 and W_2 with $d(W_1, W_2) \geq 1$. We show that already for $k' \geq 2$ the problem becomes NP-complete.

Theorem 3.3. *For each $k' \geq 2$,* JR-DIFF-COMMITTEES *and* EJR+-DIFF-COMMITTEES *are NP-complete and W[1]-hard when parameterized by the committee size k.*

We again complement the negative result by giving a polynomial-size ILP formulation in our full version (Boehmer, Glessen, and Peters 2025). Using a similar approach as for Theorem 3.2, we additionally show that both JR- and EJR+-DIFF-COMMITTEES are in FPT when parameterized by n. For JR-DIFF-COMMITTEES we give an explicit combinatorial algorithm, while for EJR+-DIFF-COMMITTEES we again employ Lenstra's algorithm.

Theorem 3.4. JR-DIFF-COMMITTEES *and* EJR+-DIFF-COMMITTEES *are in FPT when parameterized by n.*

3.3 Candidate Containment

To measure the importance of (subsets of) candidates for proportionality, we aim to solve the following problem:

p-CANDIDATES-X
Input: Election $\mathcal{I} = (N, C, A, k)$ and set of candidates $C' \subseteq C$ with $|C'| = p$.
Question: Is there a committee $W \subseteq C$ with $C' \subseteq W$, satisfying the proportionality axiom X?

Indeed, as mentioned previously, for any single candidate c, there always exists a committee satisfying JR containing c, i.e., any instance of 1-CANDIDATES-JR is a 'Yes-Instance'. For EJR+ it follows from Dong et al. (2026, Theorem 4.14) that 1-CANDIDATES-EJR+ is also always a 'Yes-Instance'. Neither of these results transfers to pairs of candidates, for which the problem becomes computationally intractable.

Theorem 3.5. 2-CANDIDATES-JR *is NP-complete, and W[1]-hard when parameterized by the committee size k.*

Finally, we prove membership in FPT, when parameterized by n. For JR, this follows from a combinatorial algorithm, while for EJR+ we again employ Lenstra's algorithm.

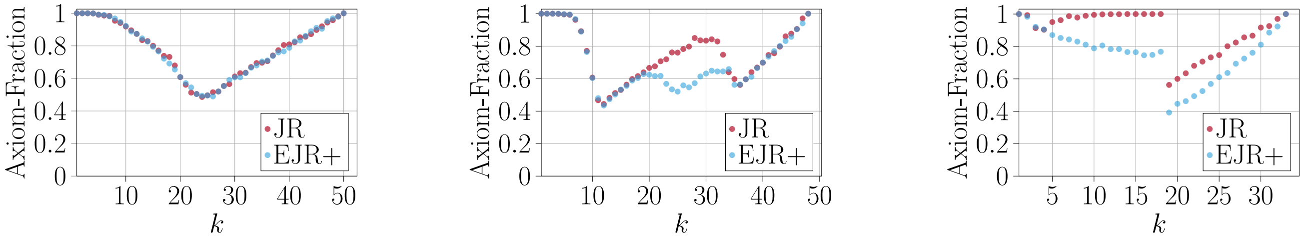
Theorem 3.6. p-CANDIDATES-JR *and* p-CANDIDATES-EJR+ *are in FPT when parameterized by n.*

3.4 Counting Committees

To measure the restrictiveness of proportionality axioms, we count the number of proportional committees. Bredereck et al. (2019) have shown that counting the number of JR committees is $\#P$-hard, and $\#W[1]$-hard when parameterized by k. We complement this result by showing that one can count the number of JR committees in FPT time in n. We achieve this via a dynamic programming approach, again by introducing equivalence classes of candidates.

Theorem 3.7. *For a given election $\mathcal{I}$, the number of JR committees $|\mathrm{JR}(\mathcal{I})|$ can be computed in FPT time in n.*

It remains an open problem whether these results generalize to EJR+. Although most of the measures relevant to our experiments are computationally intractable, we still aim to compute them. For counting problems, we turn to a sampling-based approach. We present an exemplary sampling-based approximation algorithm for the counting variant of the 1-CANDIDATES-JR problem; observe that the #P-hardness of 1-CANDIDATES-JR follows from the #P-hardness of counting JR committees. In particular, we give a sampling bound on the number of samples needed to accurately approximate the JR-prevalence of a given candidate. The sampling complexity of this problem is polynomial in the instance size, error, and in $\binom{m}{k}/|\mathrm{JR}(\mathcal{I})|$, i.e., the inverse of the fraction of JR committees in the given election.



(a) Warszawa – Obszar 3 (2019)

(b) Ursynów Wysoki Południowy (2018)

(c) Praga-Północ (2017)

Figure 3: JR-fraction (red) and EJR+-fraction (blue) for different committee sizes $k \in [m]$ across selected pabulib instances.

Proposition 3.8. *There is an algorithm that, given an election* $\mathcal{I}$*, candidate* c*, and two positive rational numbers* ε *and* δ*, outputs a value* $\tilde{\alpha}$ *such that the JR-prevalence of* c *lies, with probability at least* $1-\delta$*, in the interval* $[\tilde{\alpha}-\varepsilon, \tilde{\alpha}+\varepsilon]$*. This algorithm has an expected running time polynomial in* $|C|$*,* $|V|$*,* $1/\varepsilon$*,* $\ln(1/\delta)$*, and* $\binom{m}{k}/|\mathrm{JR}(\mathcal{I})|$*.*

To compute an instance's JR/EJR+-fraction, a simple Monte Carlo estimation with Hoeffding's inequality leads to a similar ε-approximation with probability at least $1-\delta$ in time polynomial in $|C|$, $|V|$, $1/\varepsilon$, $\ln(1/\delta)$, and $\binom{m}{k}/|\mathrm{JR}(\mathcal{I})|$.

4 Experiments

We present our experimental investigation. After introducing our dataset, we examine the restrictiveness of different proportionality axioms in Section 4.1 and the importance of individual candidates for proportionality in Section 4.2.

Dataset. We present our results on a real-world dataset here, and relegate results on two synthetic datasets based on the resampling and Euclidean models to our full version (Boehmer, Glessen, and Peters 2025). Given the scarcity of real-world data from approval-based multiwinner elections, we rely on pabulib, the primary source of participatory budgeting elections, as our dataset (Faliszewski et al. 2023a). Participatory budgeting generalizes multiwinner voting by assigning costs to candidates and selecting a subset that respects a total budget constraint. We include all pabulib instances with approval ballots and an average ballot size of at least four as of May 2nd, 2025. We do not apply any preprocessing, resulting in 369 instances with $n \in [69, 95899]$ and $m \in [7, 138]$. While these instances include voters' approval ballots over candidates, they do not specify a value for k (since the budget constraint takes the role of k). As we discuss in Section 4.1, we set $k = \lfloor \frac{m}{2} \rfloor$.[3]

4.1 How Restrictive Are Proportionality Axioms?

We analyze the restrictiveness of JR and EJR+. To shed light on the effect of the choice of k, and select a value for k, we analyze how the JR/EJR+-fraction changes with changing k. Subsequently, we take a closer look at the JR- and EJR+-fractions for $k = \lfloor \frac{m}{2} \rfloor$, and the average and maximum distances between JR, resp., EJR+-committees.

Dependency on k. We start by analyzing the JR- and EJR+-fraction for varying values of k to understand k's influence on the impact of proportionality and which values of k yield instances interesting from the perspective of proportionality.[4]

We begin by discussing some intuition. At first glance, increasing k has two opposing effects on the constraints imposed by JR (and analogously by EJR+). First, more candidates surpass the approval threshold n/k and may induce JR violations. Second, in a randomly selected committee, more voters approve at least one candidate and can therefore not be part of a JR violation. These two effects explain the very high JR-fraction at extreme k: at very low k, no candidate meets the approval threshold n/k; at very high k, nearly every voter approves someone in the committee. However, there is a third, more subtle effect: a candidate c with approval score above n/k can still impose additional constraints as k increases, since every n/k-subset of N_c must be "satisfied", and their number grows with k. Our results demonstrate that the interplay of these three effects can be quite intricate.

We present results for four instances in Figures 1 and 3. Figures 1 and 3a represent typical behaviors observed in our dataset, while Figures 3b and 3c illustrate more exceptional cases. The pattern in Figure 3a is relatively simple and aligns with basic intuition. Figure 1 already exhibits more complex behavior: JR and EJR+ begin to diverge. JR becomes easy to satisfy between $k = 20$ and $k = 60$, but is restrictive before and after, including sharp increases and drops. Figure 3b contains even more local extrema, and Figure 3c re-

---

[3] A similar approach is taken by Szufa et al. (2022) and Faliszewski et al. (2023b). The only large-scale, real-world approval-based datasets we are aware of are from two blockchains (Boehmer et al. 2024a), but their size (up to 10000 candidates and 40000 voters) is prohibitive for our experiments. Our combination of participatory budgeting data and synthetic approval-based multiwinner voting data represents the most informative and widely used approach in the literature (Boehmer et al. 2024b). While the cost of a project in a participatory budgeting instance certainly influences its approval score, the connection in Pabulib is much weaker and more heterogeneous than one might expect (see Faliszewski et al. (2023a, Fig. B.6)). Further, based on measured distances between instances, Szufa et al. (2022) argue that Pabulib instances without costs yield realistic approval elections.

[4] For each instance and each $k \in [m]$, we estimate the JR-fraction (resp. EJR+-fraction) by sampling until we found 1000 JR (resp. 1000 EJR+) committees. We apply a timeout of 15 minutes per k and instance, resulting in the exclusion of 9 instances that did not finish for all k and for which all computed EJR+-fractions exceeded 0.95.

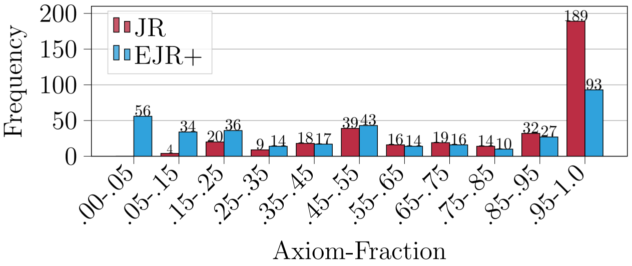


Figure 4: Number of pabulib instances (in total 360) with given JR-fractions (red) and EJR+-fractions (blue).

veals a striking non-continuity, where changing k by just one dramatically affects the behavior of the axiom-fractions.

One explanation for the sudden drops in the JR-fraction in Figures 1, 3b and 3c could be the first effect: new candidates surpass the n/k threshold as k increases. This explanation is at best partial. In Figure 3c, indeed, a new candidate passes the threshold at the jump at $k = 19$. In contrast, in Figure 3b, all candidates have an approval score of at least $n/23$, i.e., all are "relevant" for $k \geq 23$, and yet we still observe rapid changes and a non-monotonous behavior afterwards.

The key takeaway from this analysis is that the dependency of JR/EJR+-fractions on k is more complex and less predictable than one might expect. We observe sudden jumps, multiple local extrema, and differing behaviors between JR and EJR+. This calls for caution when conducting experiments: the choice of k, often treated as a generic hyperparameter, can substantially influence an instance's behavior from the perspective of proportionality.

In our remaining experiments, we aim to select a value of k that maximizes the number of instances that are non-trivial from the perspective of proportionality, i.e., have an EJR+-fraction below 95%. We pick $k = \lfloor \frac{m}{2} \rfloor$, as it maximizes the number of non-trivial instances among all $k = \lfloor \frac{m}{c} \rfloor$ for $c \in \{2, \dots, 10\}$; see Figure 9 in our full version (Boehmer, Glessen, and Peters 2025). While we have seen above that the chosen k can have a strong impact on the instance level, we verify the robustness of dataset-level trends. For this, we compare the distribution of JR- resp. EJR+-fractions for $k = \lfloor \frac{m}{2} \rfloor$ to the distributions for $k = \lfloor \frac{m}{c} \rfloor$ for $c \in \{2, \dots, 10\}$ and for k equal to the total budget of the instance divided by the average project cost (as an estimation for the average number of affordable projects). General trends turn out to be stable in both cases; for details, we refer again to our full version.

JR/EJR+-Fractions. In Figure 4, we show a frequency histogram for the JR- (red) and EJR+-fractions (blue) of the pabulib instances with $k = \lfloor \frac{m}{2} \rfloor$.[5] We observe a diverse picture: While many instances exhibit high fractions and thus confirm previous observations by Bredereck et al. (2019),

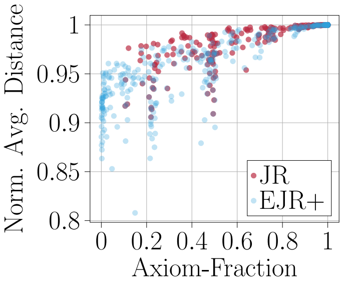


Figure 5: Each point represents one instance and one axiom (JR in red, EJR+ in blue). We plot the correlation between the axiom fraction and the normalized average distance among committees fulfilling the axiom.

and Brill and Peters (2023) that JR and EJR+ impose only extremely mild constraints on many instances, there are also numerous instances where only a small fraction of committees satisfy these axioms. This is especially pronounced for EJR+, with 56 instances exhibiting an EJR+-fraction of at most 5%. For JR, this effect is slightly less pronounced but still present, as 24 instances have a JR-fraction of at most 25%. These findings suggest that proportionality notions can matter in real-world voting instances, as they can significantly reduce the space of feasible committees.

Prior work suggests that different proportionality axioms tend to coincide in practice. For instance, Boehmer et al. (2024a) note that "proportionality axioms seem to lose their discriminative power in practice". However, Figure 4 shows that this is not generally the case, as the EJR+-fraction is often substantially lower than the JR-fraction for our chosen value of k. We shed further light on this question with the help of the JR-NOT-EJR+ problem:[6] For 275 pabulib instances, there exists a committee that satisfies JR but not EJR+, whereas in 70 instances, every JR committee also satisfies EJR+, rendering the two axioms effectively equivalent. These results draw a mixed picture: while there are cases where the distinction between proportionality axioms becomes irrelevant, there are also many instances where the "strength" of the considered axiom continues to matter.

Distances Between Committees. After establishing that EJR+ and even JR can impose quite strong restrictions on real-world instances, the next natural question is: Does each of these axioms lead to a solution space consisting of only similar committees? We answer this question in the negative. For this, we start by measuring the average distance between two JR, resp. EJR+, committees,[7] normalized by the expected distance between two randomly drawn committees of size k. In Figure 5, each red (resp. blue) dot represents an

---

[5] For every instance, we sample committees until we reach 5000 JR committees and return 5000 divided by the total number of samples as the JR-fraction. We parallelize instance-wise on 128 kernels with a total timeout of five days, resulting in 360/369 solved instances. We proceed analogously for EJR+.

[6] We use Gurobi with the ILP implementation for JR-NOT-EJR+ provided in our full version (Boehmer, Glessen, and Peters 2025) with a timeout of 30 minutes per instance, resulting in 24 unsolved instances.

[7] We approximate these values by calculating the average distance between pairs of committees from the set of 5000 sampled JR, resp. EJR+, committees from Footnote 5.

instance, with the x-coordinate indicating the instance's JR-fraction (resp. EJR+-fraction) and the y-coordinate the instance's normalized average distance between two JR (resp. EJR+) committees. We observe that the normalized average distance—even for EJR+—is typically above 0.9, suggesting that proportional committees are widely distributed over the space of all committees. Moreover, while there is some correlation between instances' normalized distance and their JR/EJR+-fraction, the general differences in terms of average distance across instances are small, further confirming the heterogeneity of proportional committees, even in cases where only a few committees fulfill proportionality.

Examining the maximum, instead of the average distance, between proportional committees paints a similar picture.[8] For JR, we find that 329 out of 369 instances admit disjoint JR committees, while for all other instances we can find two JR committees with an overlap strictly smaller than four. For EJR+, we found disjoint committees for 189 out of 269 solved instances, and a maximum overlap of eight among all solved instances. Given our choice of $k = \lfloor \frac{m}{2} \rfloor$, the existence of disjoint proportional committees is quite remarkable, as it implies that one can partition the candidates into two halves in a way that both are representative.

4.2 What Makes a Candidate Important for Proportionality?

We analyze three candidate importance measures: the prevalence and power index, which aim to capture importance for proportionality, and the traditional approval score (see Section 2 for definitions).[9] In this section, we only consider instances with an EJR+-fraction $\leq 95\%$, resulting in 267 instances. We do this because when nearly all committees are proportional, there are no substantial differences in candidates' importance for proportionality, skewing the results.

Correlation Between Measures. We are interested in finding out (i) how closely our two new ways to measure candidates' importance for proportionality are related, and (ii) whether they differ from the established approach of assessing candidates' merit by approval score. To this end, we compute the Pearson correlation coefficient (PCC) between each pair of measures across all candidates within one instance. Regarding question (a), we find a very strong correlation for EJR+, with an average PCC value of 0.99. For JR, the average PCC is with 0.72 lower; however, this is also partly due to the fact that, in contrast as for EJR+, our dataset contains instances with JR-ratio very close to 1: for these cases, the prevalences of candidates do not vary significantly, and therefore, small sampling errors can have quite a large impact on the PCC. Regarding question (b), we find that the approval score is notably less correlated with our two proportionality-based measures: the average PCC between approval score and prevalence is only 0.22 for JR and 0.65 for EJR+. To make this observation more tangible, we show in Figure 6 the election from Warszawa, Wysokie-Okoecie (2017), which plots candidates' approval score (x-axis) against their prevalence (y-axis). We observe virtually no connection between approval scores and prevalence: candidates with the lowest and highest approval scores have similar prevalence, while a candidate with a medium approval score appears in all JR and EJR+ committees. While it is intuitive that such behavior can appear in the worst case, it is noteworthy that it also occurs in real-world instances. Practically, a candidate's strength and merit to be selected is often assessed by their approval score. However, we conclude that from the perspective of proportionality, this approach falls short and can be actively misleading.

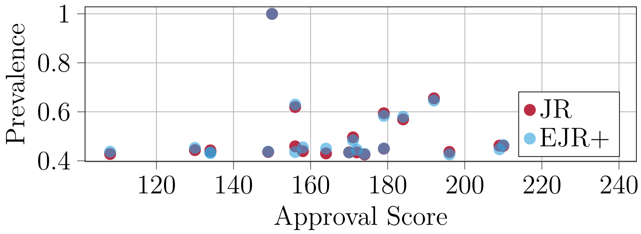


Figure 6: Each point corresponds to one candidate in the Warszawa, Wysokie-Okoecie (2017) election and one axiom (JR in red, EJR+ in blue). We plot the candidate's approval score against its axiomatic prevalence.

Lastly, we observe that proportional voting rules typically do not select the k candidates with the highest candidate importance for proportionality, see our full version (Boehmer, Glessen, and Peters 2025).

5 Conclusion

We conducted an algorithmic and experimental investigation into the impact of proportionality on feasible committees and candidate importance in approval-based multiwinner elections. We believe our findings can inform future discussions around proportionality in several ways: (i) Our observation that proportionality axioms can significantly restrict the outcome space on real-world voting data highlights their practical relevance and can help guide discussions on the adoption of proportional voting rules. (ii) Our experiments highlighting the influence of the committee size k can inform the design and interpretation of future experiments, emphasizing the need for greater care when selecting k. (iii) Our prevalence and power index measures offer a new perspective on candidate importance under proportionality and can contribute to the transparency of election outcomes.

Acknowledgments

This work was partially supported by the Singapore Ministry of Education under grant number MOE-T2EP20221-0001.

---

[8] We compute the maximum distance by solving the ILP formulation for DIFF-COMMITTEES (see our full version (Boehmer, Glessen, and Peters 2025)) using Gurobi. We set a time limit of 30 minutes for each instance, resulting in all instances being solved for JR, and 269 instances for EJR+.

[9] To compute the prevalence and power index of a candidate, we sampled 5000 JR committees and 5000 EJR+ committees using rejection sampling. For the prevalence, we take the fraction of sampled committees containing the candidate. For the power index, we count how many sampled committees containing the candidate stop to satisfy the respective axiom when the candidate is removed.

This research was (partially) funded by the HPI Research School on Foundations of AI (FAI).

References

Aziz, H.; Brill, M.; Conitzer, V.; Elkind, E.; Freeman, R.; and Walsh, T. 2017. Justified Representation in Approval-Based Committee Voting. *Social Choice and Welfare*, 48(2): 461–485.

Aziz, H.; Lu, X.; Suzuki, M.; Vollen, J.; and Walsh, T. 2023. Best-of-Both-Worlds Fairness in Committee Voting. In *Proceedings of the 19th International Conference on Web and Internet Economics (WINE)*, 676. Full version arXiv:2303.03642 [cs.GT].

Banzhaf, J. F. 1965. Weighted Voting Doesn't Work: A Mathematical Analysis. *Rutgers Law Review*, 19: 317–343.

Bardal, T.; Brill, M.; McCune, D.; and Peters, J. 2025. Proportional Representation in Practice: Quantifying Proportionality in Ordinal Elections. In *Proceedings of the 39th AAAI Conference on Artificial Intelligence (AAAI)*, 13581–13588.

Boehmer, N.; Brill, M.; Cevallos, A.; Gehrlein, J.; Sánchez-Fernández, L.; and Schmidt-Kraepelin, U. 2024a. Approval-Based Committee Voting in Practice: A Case Study of (Over-)Representation in the Polkadot Blockchain. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence (AAAI)*, 9519–9527.

Boehmer, N.; Faliszewski, P.; Janeczko, L.; Kaczmarczyk, A.; Lisowski, G.; Pierczyński, G.; Rey, S.; Stolicki, D.; Szufa, S.; and Wąs, T. 2024b. Guide to Numerical Experiments on Elections in Computational Social Choice. In *Proceedings of the 33rd International Joint Conference on Artificial Intelligence (IJCAI)*, 7962–7970.

Boehmer, N.; Glessen, L.; and Peters, J. 2025. Understanding the Impact of Proportionality in Approval-Based Multiwinner Elections. arXiv:2511.09479.

Bredereck, R.; Faliszewski, P.; Niedermeier, R.; and Kaczmarczyk, A. 2019. An Experimental View on Committees Providing Justified Representation. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, 109–115.

Brill, M.; and Peters, J. 2023. Robust and Verifiable Proportionality Axioms for Multiwinner Voting. In *Proceedings of the 24th ACM Conference on Economics and Computation (ACM-EC)*, 301. ACM Press. Full version arXiv:2302.01989 [cs.GT].

Cygan, M.; Fomin, F. V.; Kowalik, L.; Lokshtanov, D.; Marx, D.; Pilipczuk, M.; Pilipczuk, M.; and Saurabh, S. 2015. *Parameterized algorithms*, volume 5. Springer.

Dong, C.; Frank, F.; Peters, J.; and Suksompong, W. 2026. Reconfiguring Proportional Committees. In *Proceedings of the 40th AAAI Conference on Artificial Intelligence (AAAI)*. Forthcoming.

Elkind, E.; Faliszewski, P.; Igarashi, A.; Manurangsi, P.; Schmidt-Kraepelin, U.; and Suksompong, W. 2023. Justifying Groups in Multiwinner Approval Voting. *Theoretical Computer Science*, 969: 114039.

Elkind, E.; Faliszewski, P.; Igarashi, A.; Manurangsi, P.; Schmidt-Kraepelin, U.; and Suksompong, W. 2024. The Price of Justified Representation. *ACM Transactions on Economics and Computation*, 12(3): 11:1–11:27.

Faliszewski, P.; Fils, J.; Peters, D.; Pierczyński, G.; Skowron, P.; Stolicki, D.; Szufa, S.; and Talmon, N. 2023a. Participatory Budgeting: Data, Tools, and Analysis Participatory Budgeting: Data, Tools, and Analysis. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)*, 2667–2674. Full version: https://arxiv.org/pdf/2305.11035.

Faliszewski, P.; Lackner, M.; Sornat, K.; and Szufa, S. 2023b. An experimental comparison of multiwinner voting rules on approval elections. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)*, 2675–2683.

Halpern, D.; Kehne, G.; Procaccia, A. D.; Tucker-Foltz, J.; and Wüthrich, M. 2025. Representation with Incomplete Votes. *Theory and Decision*. Forthcoming.

Kehne, G.; Schmidt-Kraepelin, U.; and Sornat, K. 2025. Robust Committee Voting, or The Other Side of Representation. In *Proceedings of the 26th ACM Conference on Economics and Computation (ACM-EC)*, 1131 – 1151. ACM Press.

Lenstra, H. W. 1983. Integer programming with a fixed number of variables. *Mathematics of Operations Research*, 8: 538–548.

Peters, D.; Pierczyński, G.; and Skowron, P. 2021. Proportional Participatory Budgeting with Additive Utilities. In *Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS)*, 12726–12737.

Peters, D.; and Skowron, P. 2020. Proportionality and the Limits of Welfarism. In *Proceedings of the 21st ACM Conference on Economics and Computation (ACM-EC)*, 793–794. ACM Press.

Revel, M.; Milli, S.; Lu, T.; Watson-Daniels, J.; and Nickel, M. 2025. Representative Ranking for Deliberation in the Public Sphere. In *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, 51583–51613.

Szufa, S.; Faliszewski, P.; Janeczko, L.; Lackner, M.; Slinko, A.; Sornat, K.; and Talmon, N. 2022. How to Sample Approval Elections? In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*, 496–502.