Strategic Behavior, Fairness, and Social Optimality in Multi-Winner Elections under Uncertainty

Janik Bürgermeister
Technical University of Munich
Munich, Germany
janik.buergermeister@tum.de

Chris Dong
HPI, University of Potsdam
Potsdam, Germany
chrisshuyu.dong@hpi.de

Michael Lewicki
Technical University of Munich
Munich, Germany
michael.lewicki@tum.de

ABSTRACT

Multi-winner approval voting (MWAV) is widely studied in computational social choice, yet most analyses abstract away from the fact that voter preferences are non-dichotomous and that voters may behave strategically. We model all voters as strategic agents using deep reinforcement learning (DRL) and study equilibrium behavior under both complete and incomplete information. While agents learn to manipulate elections under complete information, we find that under uncertainty, they converge to simple cutoff strategies that are stable, fair, but often socially suboptimal. We further show that mild regulatory interventions can substantially improve welfare.

KEYWORDS

Multi-Winner Elections, Deep Reinforcement Learning, Bayesian Nash Equilibrium, Incentive Compatibility

ACM Reference Format:
Janik Bürgermeister, Chris Dong, and Michael Lewicki. 2026. Strategic Behavior, Fairness, and Social Optimality in Multi-Winner Elections under Uncertainty. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. https://doi.org/10.65109/GEVH7247

1 INTRODUCTION

Voting mechanisms are a central object of study in computational social choice and multi-agent systems, with applications ranging from political elections and participatory budgeting to algorithmic decision-making and recommender systems. A fundamental challenge in this area is strategic behavior: as shown by Gibbard [3], manipulation is unavoidable in general, and voters may benefit from misreporting their preferences.

This issue is also relevant in MWAV, where voters elect a committee by submitting approval ballots to a voting rule. Prominent rules include Chamberlin–Courant (CC), Proportional Approval Voting (PAV), and the Method of Equal Shares (MES), each offering different trade-offs between proportionality, welfare, and fairness. While these rules are widely studied, none are strategyproof, and most analyses assume that approval ballots truthfully reflect voter preferences. This assumption is problematic: preferences are rarely dichotomous, and even natural translations from ordinal preferences may create incentives for strategic misreporting.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/GEVH7247

We study how *rational agents should vote* in MWAV when underlying preferences are ordinal and privately known. Our approach uses DRL, which is well suited to environments in which agents interact repeatedly and adapt their behavior. Beyond its behavioral interpretation, DRL serves as a numerical method for approximating Bayesian Nash equilibria (BNE) in voting games that are analytically intractable.

Contributions. We analyze strategic behavior in MWAV under both complete and incomplete information. Under complete information, DRL agents reliably learn profitable manipulation, confirming the vulnerability of standard approval-based rules. Under incomplete information, agents converge to simple and stable *cutoff (Top-k) strategies*, approving their most-preferred candidates.

These equilibria are typically not socially optimal: agents approve fewer candidates than would maximize utilitarian welfare. To study this inefficiency systematically, we introduce a *reduced game* in which agents commit ex ante to a cutoff strategy. This yields a tractable normal-form game that enables direct equilibrium computation and principled evaluation of regulatory interventions. Across a wide range of environments, requiring voters to approve at least $\lfloor 0.4K \rfloor$ candidates improves welfare in about half of all cases, while only rarely decreasing it. Overall, our results connect learning-based analysis with classical equilibrium reasoning and provide actionable insights for the design of multi-winner voting systems.

2 RELATED WORK

Our work relates to research at the intersection of machine learning and social choice, including the learnability and design of voting rules [2, 7, 8], as well as applications of ML to empirical voting problems. Most closely related are studies on learning to manipulate elections [4] and on multi-agent learning in voting environments [1, 6]. These works primarily focus on single-winner settings or iterative voting dynamics. In contrast, we model *many strategic voters* interacting in one-shot multi-winner approval elections and use DRL to approximate equilibrium behavior at scales that are difficult to analyze theoretically.

3 MODEL

We model MWAV as a single-stage simultaneous game. Let $N = \{1, \ldots, n\}$ denote the voters and $C = \{c_1, \ldots, c_m\}$ the candidates. A committee $W \subseteq C$ satisfies $|W| = K$. Each voter submits an approval ballot $A_i \in \{0, 1\}^C$ and derives utility $U_i(W)$ from the elected committee.

Voter preferences are ordinal rankings over candidates, generated using impartial culture, Euclidean, or Mallows models. We assume *separable utilities* and normalize each voter's utilities to sum to one. Further, we discuss when separable voter preferences may occur: only assuming expected utility maximization, we provide axiomatic characterizations for separability. Cardinal utilities are derived from ordinal rankings using linear, convex, or concave mappings; for Euclidean preferences, we additionally use distance-based utilities. This allows us to test whether strategic behavior depends on how sharply voters distinguish between top-ranked and mid-ranked candidates.

We study both complete-information games, where preferences are publicly known, and Bayesian games, where preferences are drawn from commonly known distributions and private knowledge. The solution concept is Bayesian Nash equilibrium.

Voting rules and fairness. We study MES and sequential variants of CC (seq-CC), PAV (seq-PAV), which are computationally efficient approximations [5]. MES satisfies strong proportionality axioms (JR, PJR, EJR, and Priceability), seq-CC satisfies only JR, and seq-PAV satisfies none of these axioms. This variation allows us to examine how strategic incentives interact with formal fairness guarantees.

4 LEARNING FRAMEWORK

We formulate the voting game as a reinforcement learning problem. The state consists of utility information (full or private, depending on the information structure). Actions correspond to approval ballots, and rewards equal the utility derived from the elected committee. Agents are trained using independent deep RL learners with shared architecture but no explicit coordination or communication.

Our experiments scale to elections with up to 40 voters and 20 candidates. For each configuration, agents are trained until convergence, and learned policies are evaluated on large out-of-sample batches of preference profiles. Importantly, DRL is not used to *design* voting rules, but to approximate equilibrium play under fixed mechanisms. Thus, learned policies are interpreted as numerical approximations of best responses in large Bayesian games.

We evaluate learned behavior using four criteria: (i) approximate equilibrium stability (ϵ-BNE), (ii) utilitarian social welfare, (iii) truthfulness, defined as approval of a prefix of the preference ranking, and (iv) fairness, measured via standard proportionality axioms.

5 RESULTS

Complete information. In constructed instances that admit profitable deviations, DRL agents consistently learn to misreport their ballots, often by withholding approval from their most-preferred candidate. This demonstrates that DRL reliably uncovers manipulation strategies whenever such behavior is payoff-improving.

Incomplete information. Across all Bayesian settings, agents converge to *cutoff strategies*, approving their top k candidates. The learned strategies form ϵ-BNE with $\epsilon \leq 0.05$ and consistently outperform random voting. Importantly, agents do not converge to degenerate equilibria in which all candidates are approved or rejected. Instead, equilibrium behavior concentrates on cutoffs that vary systematically with the voting rule and committee size.

Equilibrium selection and structure. Although multiple cutoff equilibria may exist in principle, learning dynamics consistently select equilibria with relatively small to moderate cutoff values. Under seq-PAV and MES, agents approve more candidates than under seq-CC, which rewards minimal representation and strongly incentivizes sparse approval. These equilibria are symmetric and stable across random seeds, suggesting that they are focal under decentralized learning. This observation motivates our focus on symmetric cutoff strategies in the reduced game.

Welfare. While cutoff strategies are simple and stable, they are typically not welfare-optimal. In many environments, agents approve substantially fewer candidates than would maximize utilitarian welfare.

Fairness. Despite known theoretical counterexamples, all committees elected in our experiments satisfy JR, PJR, EJR, and Priceability and belong to the core. This suggests that, under commonly studied preference distributions, equilibrium behavior induced by learning does not undermine proportional fairness, even when agents act purely in their own interest.

Reduced game and regulation. Motivated by the empirical prevalence of cutoff strategies, we introduce a reduced game in which agents commit ex ante to a cutoff $k \in \{1, \dots, K\}$. This restriction reduces the action space from exponential to linear in K and yields a complete-information normal-form game that can be solved directly. Across 576 election experiments, requiring voters to approve at least $X = \lfloor 0.4K \rfloor$ strictly improves welfare in about half of the cases, while reducing welfare in only 3%. Lower thresholds are always safe but yield smaller gains. For seq-CC, regulation has little effect, as equilibrium behavior already involves minimal approval.

6 CONCLUSION

We demonstrate that deep multi-agent reinforcement learning offers a practical and informative framework for analyzing strategic behavior in multi-winner approval voting. Agents learn simple and stable cutoff strategies that are typically fair but socially suboptimal. By introducing a reduced game, we identify simple regulatory interventions that can substantially improve welfare with limited downside risk. Our results highlight the value of combining learning-based methods with equilibrium analysis to inform the design of robust, efficient, and strategically aware voting systems.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon Europe research and innovation programme (grant agreement No 101198689). This project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - GRK 2201/2 - Project Number 277991500 and BI 1057/9, as well as BR 2312/11-2. Most of this research was done when Chris Dong was a PhD student at the Technical University of Munich.

REFERENCES

[1] Stéphane Airiau, Umberto Grandi, and Filipo Studzinski Perotto. 2017. Learning Agents for Iterative Voting. In *Algorithmic Decision Theory (Lecture Notes in Computer Science, Vol. 10576)*, Jörg Rothe (Ed.). Springer, Cham, Switzerland, 139–152. https://doi.org/10.1007/978-3-319-67504-6_10

[2] Cem Anil and Xuchan Bao. 2021. Learning to Elect. https://doi.org/10.48550/arXiv.2108.02768

[3] Allan Gibbard. 1973. Manipulation of Voting Schemes: A General Result. *Econometrica* 41, 4 (1973), 587–601. http://www.jstor.org/stable/1914083

[4] Wesley H. Holliday, Alexander Kristoffersen, and Eric Pacuit. 2024. Learning to Manipulate under Limited Information. In *Proceedings of the 39th Annual AAAI Conference on Artificial Intelligence (AAAI-25)*. https://doi.org/10.48550/arXiv.2401.16412

[5] Martin Lackner and Piotr Skowron. 2023. *Multi-Winner Voting with Approval Preferences* (1 ed.). Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-031-09016-5

[6] Loujain Liekah and Umberto Grandi. 2019. Multi-Agents Reinforcement Learning In Iterative Voting. https://hal.science/hal-04277513v1/file/Layka.MARL_for_Strategic_Voting___Thesis.pdf

[7] Ariel D. Procaccia, Aviv Zohar, Yoni Peleg, and Jeffrey S. Rosenschein. 2009. The learnability of voting rules. *Artificial Intelligence* 173, 12-13 (2009), 1133–1149. https://doi.org/10.1016/j.artint.2009.03.003

[8] Lirong Xia. 2013. Designing Social Choice Mechanisms Using Machine Learning. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-Agent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 471–474.