THE JOURNAL OF ARTIFICIAL SOCIETIES AND SOCIAL SIMULATION

Modeling the Ballot: Agent-Based Insights into Representation, Coalitions, and Welfare

Devansh Garg[1]

[1] *Indian Institute of Technology, Mandi*
Correspondence should be addressed to *devanshg.iitm@gmail.com*

Journal of Artificial Societies and Social Simulation xx(x) x, 20xx
Doi: 10.18564/jasss.xxxx Url: http://jasss.soc.surrey.ac.uk/xx/x/x.html

Received: dd-mmm-yyyy Accepted: dd-mmm-yyyy Published: dd-mmm-yyyy

Preprint

Submitted to Journal of Artificial Societies and Social Simulation (JASSS) for peer review.

Abstract: We present an agent-based simulation of democratic decision-making in which autonomous learning agents interact under alternative electoral institutions and social structures. The model integrates six voting mechanisms (Plurality, Approval, Borda, IRV, STV, PR with D'Hondt and Sainte-Laguë divisors), a multi-round coalition protocol with binding/non-binding contracts and side-payments, turnout and ballot-error realism, and networked interaction on Erdös–Rënyi, Barabäsi–Albert, and Watts–Strogatz graphs with homophily. Agents use reinforcement learning algorithms (PPO, A2C, A3C) with a social-welfare objective based on the inequality-averse Atkinson function, augmented by fairness regularizers (representation loss, participation fairness, envy-freeness proxy) and explicit participation costs. We report diagnostics-rich evaluations covering representation and proportionality (e.g., Gallagher, Loosemore–Hanby), fragmentation (effective number of parties), strategic behavior, coalition stability, and welfare/inequality. Classic regularities emerge—e.g., two-bloc competition under Plurality (Duverger-consistent), greater proportionality and fragmentation under PR, and differential seat allocation under D'Hondt vs Sainte-Laguë—providing face validity. The framework delivers a reproducible virtual laboratory for mechanism comparison, institutional design, and welfare–fairness trade-off analysis at population scale.

Keywords: Agent-Based Model, Reinforcement Learning, Electoral Systems, Collective Decision Making, Coalition, Social Networks

Introduction

Problem Statement and Background

1.1 Electoral institutions structure how individual preferences become collective outcomes. A large comparative politics and social choice literature demonstrates that changing the voting rule alters representation, party system fragmentation, and the translation of votes into seats (Duverger 1954; Cox 1997; Taagepera & Shugart 1989; Gallagher 1991; Loosemore & Hanby 1971). Single-member plurality systems tend toward two-bloc competition (often described as Duverger's Law) (Duverger 1954; Cox 1997). Proportional representation (PR) mechanisms reduce disproportionality and sustain multi-party competition (Taagepera & Shugart 1989; Gallagher 1991). Even within PR, divisor methods differ. D'Hondt typically advantages larger parties, while Sainte-Laguë tends to favor smaller parties, affecting fragmentation and proportionality (Taagepera & Shugart 1989; Lijphart 1999).

Realistic Democratic Processes and Social Factors

1.2 Beyond institutional rules, realistic democratic processes include turnout decisions, ballot errors, social influence, and post-electoral coalition bargaining. These factors meaningfully affect outcomes, especially in close contests or when participation is uneven across groups. Social networks shape information flow and coordination (Erdős & Rényi 1959; Watts & Strogatz 1998; Barabási & Albert 1999). Turnout interacts with costs and perceived stakes. Ballots contain human errors. Parties (or blocs of actors) negotiate coalitions that govern (Riker 1962). Each element complicates the mapping from preferences to policy and may amplify or dampen the effects of electoral rules.

Agent-Based Modeling and Learning Approaches

1.3 Agent-based modeling (ABM) provides a way to study these interacting components in a controlled, generative setting (Epstein 2006; Grimm et al. 2006, 2010). By specifying micro-level behavior and interaction topologies, ABMs generate macro-level regularities comparable to empirical patterns. In parallel, advances in reinforcement learning (RL)—including proximal policy optimization (PPO) (Schulman et al. 2017) and advantage actor-critic (A2C/A3C) (Mnih et al. 2016)—enable agents to adapt strategically to institutional incentives and social environments. For normative assessment, social welfare functions, such as the inequality-averse Atkinson measure (Atkinson 1970), offer principled criteria for evaluating collective outcomes alongside representation and fairness diagnostics from social choice (Arrow 1951; Gallagher 1991; Loosemore & Hanby 1971; Laakso & Taagepera 1979; Sen 1970).

Our Approach and Model Design

1.4 This paper introduces a unified, reproducible agent-based model (ABM) of repeated elections in which learning agents interact under six widely used voting mechanisms (Plurality, Approval (Brams & Fishburn 1978), Borda, IRV, STV (Tideman 1995), and proportional representation (PR) with D'Hondt and Sainte-Laguë (modified, first divisor = 1.4) divisors implemented as DHondtPRVoting and SainteLaguePRVoting). The model embeds: (i) turnout and ballot-error realism; (ii) two-round coalition bargaining with binding and non-binding contracts and side-payments; and (iii) networked interaction over Erdös-Rënyi, Watts-Strogatz, and Barabäsi-Albert topologies with homophily. We train agents using proximal policy optimization (PPO), advantage actor-critic (A2C), and asynchronous advantage actor-critic (A3C) (Schulman et al. 2017; Mnih et al. 2016) to optimize an Atkinson social welfare objective (Atkinson 1970) augmented with fairness regularizers (representation loss, participation fairness, envy-freeness proxy) and explicit participation costs. The analysis reports representation and proportionality measures (e.g., Gallagher and Loosemore–Hanby indices; see Appendix B, equations 1-2), fragmentation (effective number of parties (Laakso & Taagepera 1979); see Appendix B, equation 3), strategic behavior (see Appendix B, equations 4-6), coalition stability (see Appendix B, equations 7-11), and welfare/inequality (see Appendix B, equations 12-19), with uncertainty quantification.

Research Questions

1.5 Our study addresses the following questions: (1) How do alternative voting mechanisms affect proportionality, representation accuracy, and party-system fragmentation in a dynamic, learning electorate? (2) How do network topology and homophily modulate turnout, coordination, and strategic behavior, and in turn mechanism performance? (3) How do coalition contract enforceability and side-payment budgets shape coalition size, stability, and policy alignment across electoral systems? (4) How do welfare–fairness trade-offs vary across institutional and social settings when agents optimize an inequality-averse social welfare function subject to participation costs?

Contributions and Paper Structure

1.6 We contribute: a comprehensive, diagnostics-rich virtual laboratory for institutional comparison at population scale; a fairness-aware welfare objective for evaluating outcomes beyond raw efficiency; and an integrated treatment of turnout, ballot error, social networks, and coalition formation within a single ABM. The model reproduces classic qualitative regularities (e.g., Duverger-consistent concentration under Plurality; increased proportionality and fragmentation under PR; differential seat allocation under D'Hondt versus Sainte-Laguë), offering face validity while enabling systematic exploration of design levers and robustness.

1.7 The remainder of the paper details the simulation environment and learning framework (Methods), presents comparative results across mechanisms and social settings (Results), and discusses implications for institutional design and future research (Discussion).

Novel Contributions

- **Unified Multi-Mechanism Platform**: A novel unified ABM of six voting systems with learning agents
- **Integrated Coalition Dynamics**: Novel inclusion of enforceable vs. non-enforceable coalition contracts with side-payments
- **Network-Mechanism Interactions**: Novel, systematic study of how network topology and homophily modulate voting mechanism performance
- **Fairness-Aware Welfare Objective**: Novel integration of Atkinson welfare with representation loss, participation fairness, and envy-freeness regularizers
- **Comprehensive Diagnostics**: Novel ABM, reporting both social choice metrics and welfare/fairness criteria across institutional and social settings

Related Work

2.1 Our study integrates four research strands: electoral systems and social choice theory, agent-based models of elections, networked voting dynamics, and reinforcement learning in ABM. Table 1 compares our framework against key prior works.

Study	Mechanisms	Networks	Learning	Coalitions
Duverger (1954)	Plurality only	None	Static	None
Laver & Sergenti (2011)	Plurality focus	None	Heuristic	Basic
Braha (2017)	Single mechanism	ER/BA/WS	Static	None
Gao et al. (2022)	Forecasting focus	Limited	LLM-based	None
Our Framework	6 mechanisms	ER/BA/WS + homophily	PPO/A2C/A3C	Enforceable contracts

Table 1: Comparison of our framework with key prior works on mechanism breadth, network integration, learning capabilities, and coalition dynamics

Electoral Systems and Social Choice

2.2 Comparative electoral systems research establishes how rules map votes to seats, shaping representation, disproportionality, and fragmentation (Duverger 1954; Cox 1997; Taagepera & Shugart 1989; Gallagher 1991; Loosemore & Hanby 1971; Lijphart 1999; Laakso & Taagepera 1979). Within proportional representation, divisor methods impose different size biases—D'Hondt favoring larger parties relative to Sainte–Laguë (Taagepera & Shugart 1989; Lijphart 1999). Computational social choice synthesizes axiomatic properties, manipulation, and proportionality trade-offs (Brandt et al. 2016). Recent work reveals fundamental complexity barriers in voting mechanism design (Bartholdi et al. 1989), while axiomatic approaches formalize proportional representation theory (Balinski & Young 2004). Experimental studies demonstrate systematic strategic voting patterns across electoral systems (Forsythe et al. 1993).

Agent-Based Models of Elections

2.3 Agent-based political models examine micro-to-macro mechanisms in party competition and emergence (Laver & Sergenti 2011; Schreiber 2014). District-based ABMs show how spatial distributions and heterogeneity alter seat outcomes beyond aggregate vote shares (Mitra 2022; Shinde et al. 2023). Contemporary coalition theory emphasizes bargaining time, minimal winning vs. surplus coalitions, and policy alignment (Bäck et al. 2024; Battaglini & Palfrey 2021; Gschwend & Gračiūnaité 2016; Baron & Ferejohn 1989). ABM has emerged as a powerful tool for studying political phenomena (Cederman 2001), capturing emergent behavior through simple interaction rules that resist analytical solution.

Networked Voting and Social Influence

2.4 Networked influence models demonstrate contagion and persistence in voting (Braha & de Aguiar 2017), while classic network topologies (Erdös-Rënyi, Barabäsi-Albert, Watts-Strogatz) capture degree heterogeneity and clustering relevant for coordination and echo chambers. Political networks exhibit distinct structural properties that influence voting behavior (Fowler 2006), with social influence models providing mathematical frameworks for understanding opinion dynamics (DeGroot 1974). Homophily effects in social networks (McPherson et al. 2001) are particularly relevant for understanding preference clustering and social influence patterns.

Reinforcement Learning in ABM

2.5 ABM has been used to forecast real elections (Gao et al. 2022) and large language model-driven agents have been proposed for large-scale election simulations (Zhang et al. 2024). Recent advances in multi-agent reinforcement learning (Lowe et al. 2017; Shoham & Leyton-Brown 2007) enable adaptive strategic behavior in complex institutional environments. Centralized training with decentralized execution approaches (Foerster et al. 2018) provide stable learning in competitive multi-agent settings. Machine learning approaches in political science have demonstrated the potential for AI-driven analysis of voting patterns, coalition dynamics, and institutional design (Budón et al. 2021). Our implementation leverages state-of-the-art RL algorithms (PPO, A2C, A3C) with centralized critics for multi-agent coordination, enabling agents to learn sophisticated voting and coalition strategies through experience rather than predefined rules.

Research Gaps and Our Contribution

2.6 Despite extensive research on electoral systems and social networks, three critical gaps remain unaddressed. **First**, existing studies typically focus on single voting mechanisms rather than systematic comparison across electoral systems, limiting our understanding of relative performance and trade-offs. **Second**, while network effects on voting behavior are well-documented, limited work quantifies how network topology and homophily systematically modulate mechanism performance across different electoral contexts. **Third**, coalition formation in ABMs often relies on simplified protocols, with few models incorporating enforceable contracts, budget-constrained side payments, and realistic negotiation dynamics within electoral contexts.

2.7 As Table 1 illustrates, our framework directly addresses these gaps by providing a unified experimental platform that systematically varies voting mechanisms, network topologies, and coalition protocols while using reinforcement learning agents to discover adaptive strategies. This enables apples-to-apples comparisons across institutional and social settings with comprehensive diagnostics spanning social choice metrics and welfare/fairness criteria.

Model Description

3.1 The model description follows the ODD (Overview, Design concepts, Details) protocol for describing individual- and agent-based models (Grimm et al. 2006, 2010, 2020). Figure 1 illustrates the simulation architecture and workflow, while Figure 2 shows the agent-based model architecture with entities and interactions.

Overview

Purpose and Patterns

3.2 The model's purpose is to study how alternative voting mechanisms, social networks, and coalition formation protocols affect representation, proportionality, fragmentation, and social welfare in democratic decision-making systems. We evaluate the model against classic patterns: Duverger-consistent two-bloc competition

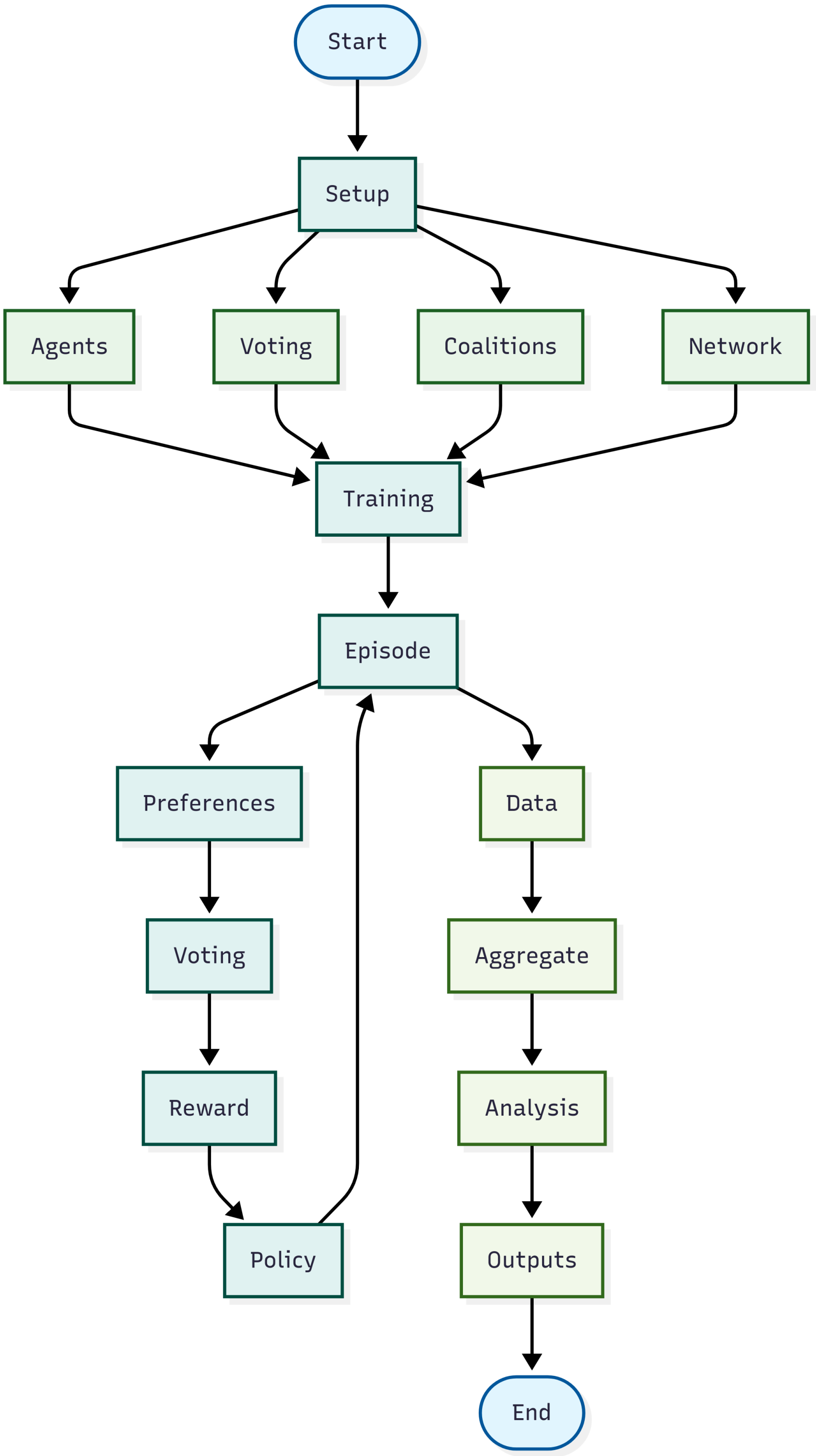

Figure 1: Agent-based simulation architecture and workflow. The model initializes four core components (agents with preferences and learning capabilities, voting mechanisms, coalition protocols, and network structures), trains agents through reinforcement learning, then executes repeated episodes where agents update preferences, participate in voting, receive rewards based on outcomes, and adapt policies. Data collection and analysis occur in parallel to generate insights on electoral system performance.

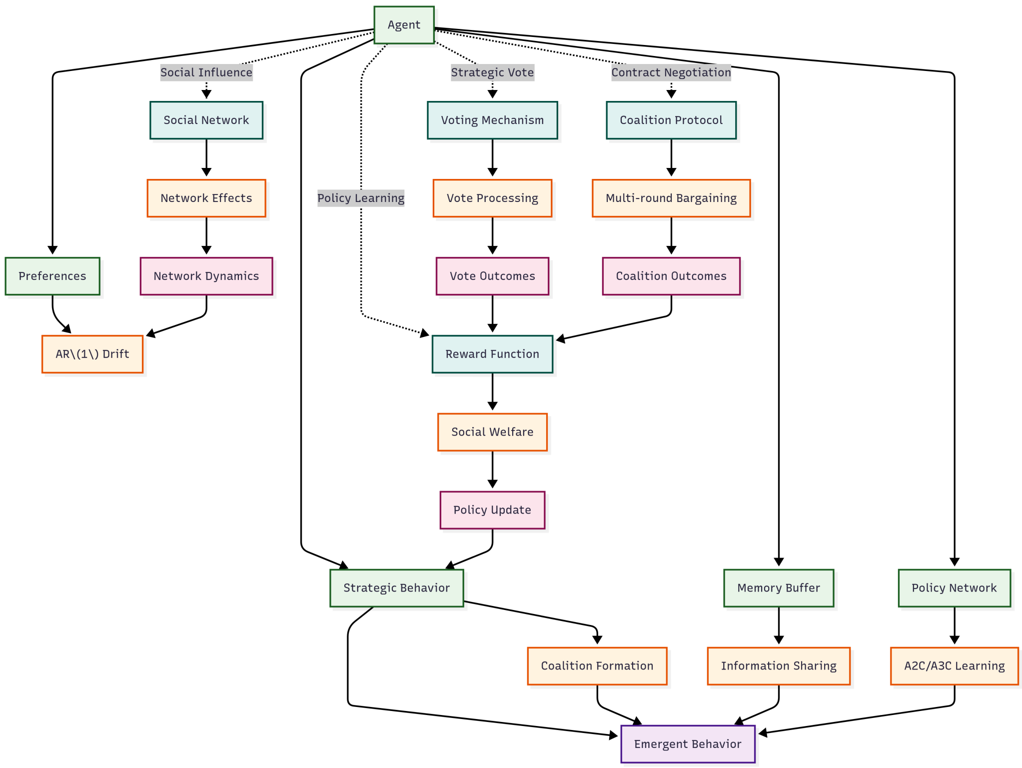

Figure 2: Agent-based model architecture showing entities and interactions. Agents (voters/parties) have policy preferences, learning capabilities, and network connections. They interact through voting mechanisms, coalition formation, and social networks. The environment provides institutional context, reward signals, and state transitions.

under plurality voting (Duverger 1954; Cox 1997); greater proportionality and fragmentation under proportional representation (Taagepera & Shugart 1989; Gallagher 1991); differential seat allocation between D'Hondt and Sainte-Laguë divisor methods (Lijphart 1999); network homophily effects on opinion clustering (Watts & Strogatz 1998; Barabási & Albert 1999); and coalition stability patterns consistent with bargaining theory (Riker 1962; Battaglini & Palfrey 2021).

Entities, State Variables, and Scales

3.3 The model contains three entity types:

- **Agents**: Voters with 8-dimensional policy preferences covering economic, social, environmental, foreign, security, infrastructure, technology, and cultural policies. Additional state variables include salience weights, turnout propensity, budget for side payments, and learning parameters
- **Networks**: Topology type (Erdös-Rënyi, Barabäsi-Albert, Watts-Strogatz), edge weights, and homophily settings
- **Electoral Mechanisms**: Type (Plurality, Approval, Borda, IRV, STV, proportional representation), parameters (seats, quota, divisor method), ballot error rate, and district configuration

3.4 The spatial scale is abstract (no physical space), with temporal scale of discrete time steps. Population scales range from small (10 agents, 50 episodes) to medium (200 agents, 50 episodes) to large (300 agents, 800 episodes), with systematic parameter variation across 68+ different experimental conditions.

Process Overview and Scheduling

3.5 As illustrated in Figure 1, each episode consists of initialization followed by repeated time steps. The simulation follows a logical progression where agents first engage in deliberation and coalition formation, then participate in voting, receive rewards based on outcomes, and adapt their strategies through reinforcement learning. This cycle enables agents to discover optimal voting and coalition strategies while the system evolves through preference drift and network rewiring. See Appendix A, subsection A.1.3 for detailed process specifications.

Design Concepts

Emergence

3.6 Key emergent phenomena include party formation through coalition dynamics, strategic voting patterns, welfare-inequality trade-offs, network clustering based on preferences, and institutional evolution toward stable governance structures.

Adaptation

3.7 Agents adapt through reinforcement learning (PPO, A2C, A3C) to optimize voting and coalition strategies. Preferences evolve via AR(1) process based on outcomes and interactions. Network connections evolve through homophily rewiring based on preference similarity.

Objectives

3.8 Individual agents maximize utility from policy outcomes, coalition benefits, and network position. Individual utility follows $u_i(\mathbf{p}) = \exp(-\|\mathbf{x}_i - \mathbf{p}\|_2)$ where $\mathbf{x}_i \in [0,1]^8$ is agent i's preference vector and $\mathbf{p}$ is the implemented policy (see Appendix B, equation 12). The collective objective is to maximize Atkinson social welfare with inequality aversion parameter $\epsilon = 0.3$ (see Appendix B, equation 13), subject to fairness regularizers (representation loss, participation fairness, envy-freeness proxy; see Appendix B, equations 18-20). The interaction between individual and collective objectives creates a multi-level optimization problem: agents learn to balance personal gains against collective welfare, with the Atkinson parameter ϵ controlling the trade-off between efficiency and equality. The reward function combines individual utilities, social welfare components, coalition bonuses, democratic participation rewards, and fairness penalties (see Appendix B, equation 14).

Learning

3.9 Agents use PPO, A2C, and A3C algorithms to learn optimal strategies. Policy networks map observations to action probabilities, while value functions estimate expected returns. Learning occurs through on-policy rollouts with mini-batch gradient updates and optional curriculum learning; we do not use experience replay for these on-policy algorithms.

Sensing

3.10 Agents observe their own state (preferences, budget, history), connected agents' preferences and behaviors, coalition memberships and contracts, and population-level statistics. Information is bounded by network topology and coalition membership.

Interaction

3.11 Direct interactions include deliberation, coalition formation, side payments, and voting. Network-mediated interactions include information diffusion, social influence, coordination, and reputation effects. Institutional interactions occur through voting mechanisms, coalition protocols, and enforcement systems.

Stochasticity

3.12 Stochastic elements include preference drift via AR(1) process with persistence parameter $\phi = 0.95$ and innovation variance $\sigma^2 = 0.0004$; turnout propensity following Beta distribution for realistic participation patterns; ballot errors with Bernoulli probability $p_e \in [0, 0.1]$; deliberation noise with normal distribution; network rewiring with homophily probability $p_h \in [0, 1]$; and exploration in reinforcement learning using ϵ-greedy policy with decay rate 0.995.

Collectives

3.13 Coalitions form dynamically with binding/non-binding/cheap-talk contracts, side payments, and enforcement mechanisms. Network communities emerge through homophily and preference clustering. Voting blocs coordinate through strategic behavior. See Appendix A, subsection A.2.7 for detailed collectives specifications.

Details

Initialization

3.14 We sample agent preferences from uniform distribution over $[0, 1]^8$ or specified prior. Salience weights are uniformly distributed or correlated with preferences. Turnout propensity follows Beta distribution. We allocate budgets for side payments. We generate networks using Erdös-Rënyi, Barabäsi-Albert, or Watts-Strogatz parameters, with optional homophily rewiring based on preference similarity.

Input Data

3.15 The model can assimilate external preference data through data assimilation for state estimation and parameter refinement. Network structures can be imported from external sources. Configuration parameters are loaded from YAML files defining experimental sweeps.

Submodels

3.16 **Voting Mechanisms:** Six mechanisms implemented (Plurality, Approval, Borda, IRV, STV, proportional representation) with mechanism-specific parameters. See Appendix A, subsection A.3.3 for detailed descriptions.

3.17 **Coalition Protocol:** Two-round negotiation with three contract types: binding contracts (enforceable with penalties), non-binding contracts (voluntary compliance), and cheap-talk negotiations (no enforcement). Side payments are budget-constrained with realistic transfer costs. Coalition formation follows a threshold-based protocol: agents propose coalitions based on preference similarity and network connections, with acceptance determined by configurable thresholds. The protocol implements basic coalition theory principles (Riker 1962; Battaglini & Palfrey 2021) with preference-based formation criteria and network-mediated proposals. See Appendix A, subsection A.3.3 for protocol details.

3.18 **Reinforcement Learning:** PPO, A2C, and A3C with centralized critics (enabled in our sweeps unless noted). In the $N = 200$ sweeps, we anneal the entropy coefficient from 0.02 to 0.005 over $1{,}000$ steps; clip ratio ϵ and learning rate α are fixed per configuration (e.g., $\epsilon \in \{0.10, 0.15, 0.20\}$; $\alpha \in [1 \times 10^{-4}, 3 \times 10^{-4}]$). A2C/A3C use the same entropy annealing schedule unless stated otherwise. See Appendix A, subsection A.3.3 for YAML-specified hyperparameters.

3.19 **Welfare Function:** Atkinson welfare with inequality aversion parameter ϵ (see Appendix B, equation 13) and fairness regularizers for representation loss, participation fairness, and envy-freeness (see Appendix B, equations 17-19).

3.20 **Network Dynamics:** Homophily rewiring based on preference similarity with configurable network topologies. The homophily mechanism implements basic preference-based clustering, following established network evolution models (Watts & Strogatz 1998; Barabási & Albert 1999). The rewiring strength parameter $\lambda \in [0, 1]$ controls the intensity of preference-based clustering, with similarity computed as $S_{ij} = -\|\mathbf{x}_i - \mathbf{x}_j\|_2$ where $\mathbf{x}_i \in [0, 1]^8$ are preference vectors. The mechanism applies preference-based rewiring to increase connections between agents with similar preferences. See Appendix A, subsection A.3.3 for implementation details.

Parameter Values

3.21 **Population Parameters:**

- $N = 200$: Number of agents

- $d = 8$: Policy dimensions
- $T = 50$: Training episodes (comprehensive sweep)
- $k \in \{1, 4, 25\}$: Number of winners (varies by configuration)

3.22 **Network Parameters:**

- $\lambda \in \{0.0, 0.2, 0.3, 0.4\}$: Homophily strength (across configurations)
- $p = 0.05$: Connection probability (ER networks, fixed)
- $m = 3$: Preferential attachment parameter (BA networks, fixed)
- $k_{ws} = 6$: Average degree (WS networks, fixed)
- $\beta_{ws} = 0.15$: Rewiring probability (WS networks, fixed)

3.23 **Learning Parameters:**

- $\alpha_{PPO} = 2 \times 10^{-4}$: Learning rate for PPO (flagship config)
- $\alpha_{A2C} = \alpha_{A3C} = 3 \times 10^{-4}$: Learning rate for A2C/A3C
- $\epsilon_{PPO} = 0.15$: PPO clip ratio (flagship config)
- β: Entropy coefficient annealed from 0.02 to 0.005 over 1,000 steps (ranges tuned per configuration)
- $\gamma = 0.98$: Discount factor (comprehensive sweep)
- $\gamma = 0.99$: Discount factor (flagship config)

3.24 **Welfare Parameters:**

- $\epsilon_{Atkinson} = 0.3$: Inequality aversion parameter
- $\alpha = 1.0$: Utility weight in reward function
- $\beta = 1.0$: Welfare weight in reward function
- $\gamma = 0.25$: Coalition bonus weight
- $\delta = 0.1$: Democratic bonus weight
- $\epsilon_{penalty} = 0.05$: Fairness penalty weight
- $\zeta = 0.02$: Participation cost weight

3.25 Key parameters: population size $N = 200$ (comprehensive sweep) and $N = 300$ (flagship run); episodes $E = 50$ (comprehensive) and $E = 800$ (flagship); learning rate $\alpha \in [0.0002, 0.0003]$; discount factor $\gamma \in [0.98, 0.99]$; inequality aversion $\epsilon \in [0, 2]$; homophily strength $\lambda \in [0, 0.3]$; ballot error rate $p_e \in [0.01, 0.02]$; and coalition budget $B \in [0.2, 1.0]$. See Appendix C, subsection C.1 for complete parameter tables and experimental configurations.

Implementation

3.26 We implement the model in Python using PyTorch for neural networks, NetworkX for graph operations, and NumPy for numerical computations. The modular architecture separates agent learning, voting mechanisms, network dynamics, and coalition protocols. Components integrate through standardized interfaces for state observations, actions, and data collection. We stream results to JSON files with hierarchical organization by experiment, run, episode, and step. The complete implementation is available at [repository URL available upon request to the corresponding author].

Experimental Design and Methods

Experimental Configurations and Parameter Sweeps

4.1 Our experimental design employs a systematic parameter sweep approach to comprehensively evaluate the interaction effects between voting mechanisms, social networks, and learning algorithms. The main experiment consists of a medium-scale comprehensive sweep with 68 distinct configurations, supplemented by a large-scale flagship run for realistic scenarios.

Comprehensive Parameter Sweep

4.2 The primary experimental sweep (`medium_200_comprehensive.yaml`) systematically varies three key dimensions: voting mechanisms (six major electoral systems), network topologies (ER/BA/WS with homophily), and training configurations (shared/individual policies with centralized critics). Population size of 200 agents balances computational feasibility with realistic electoral district scales, while 50 episodes per configuration ensures sufficient learning convergence while maintaining comprehensive coverage across the parameter space. Network parameters (connection probability p=0.05 for ER, preferential attachment m=3 for BA, degree k=6 and rewiring $\beta = 0.15$ for WS) are chosen to represent realistic social network structures with varying clustering and connectivity patterns. See Appendix C, subsection C.1 for complete experimental configurations.

Flagship Realistic Run

4.3 The flagship experiment (`flagship_300.yaml`) focuses on realistic PR scenarios with enhanced coalition dynamics, larger population (300 agents), and deeper learning convergence (800 episodes). The 300-agent population represents a realistic medium-sized electoral district, while 800 episodes ensure deep learning convergence for complex coalition strategies. PR parameters (4% threshold, Sainte-Laguë divisors with standard divisors, 20% representation rate) mirror European electoral systems, providing external validity. Network topology (Watts-Strogatz with k=12, $\beta = 0.15$) captures realistic small-world social structures, while coalition parameters (two-round negotiation, binding contracts, side payments) reflect actual parliamentary coalition formation dynamics. See Appendix C, subsection C.2 for detailed configuration.

Training Setup and Hyperparameters

Reinforcement Learning Algorithms

4.4 We implement three state-of-the-art RL algorithms with multi-agent adaptations: PPO (Schulman et al. 2017), A2C (Mnih et al. 2016), and A3C (Mnih et al. 2016). See Appendix D, subsection D.2 for detailed hyperparameters and implementation details.

Multi-Agent Coordination

4.5 Multi-agent coordination is achieved through centralized critics (enabled in our sweeps unless noted; value functions observe global state while policies remain decentralized) and shared policy networks (when enabled). See Appendix D, subsection D.3 for implementation details.

Convergence Criteria and Training Termination

4.6 Training terminates when episode reward stabilizes with early stopping criteria. Maximum limits: 50 episodes for comprehensive sweep, 800 episodes for flagship run. See Appendix D, subsection D.2 for detailed convergence criteria and sensitivity analysis.

Baseline Comparisons

4.7 We establish three rigorous theoretical baselines for performance evaluation using the same utility function and welfare metrics as the main experiment: **Random voting** (welfare 0.26±0.02) represents agents making completely random policy choices without any strategic consideration; **Sincere voting** (welfare 0.401±0.014) represents agents voting according to their true preferences without strategic coordination; and **Perfect coordination** (welfare 0.44±0.01) represents the theoretical optimum where all agents coordinate perfectly on the population centroid. These baselines are derived from systematic Monte Carlo sampling of preference distributions and provide rigorous theoretical bounds for evaluating mechanism performance. All voting mechanisms are expected to perform between random and optimal baselines, with the Atkinson welfare function (see Appendix B, equation 12) providing the normative framework for comparison. The baseline calculations use the exact same utility function ($u_i = \exp(-||x_i - p||_2)$) and welfare metrics as the main experiment, ensuring methodological consistency. See Appendix D, subsection D.5 for detailed baseline calculation methodology.

Data Collection and Aggregation

Real-Time Data Streaming

4.8 The experiment pipeline implements comprehensive data collection with minimal computational overhead, generating detailed behavioral data at each simulation step and aggregating per-episode summaries. See Appendix D, subsection D.4 for detailed data collection procedures.

Statistical Power and Sample Size Justification

4.9 Systematic parameter variation across 68 configurations provides broader coverage than random sampling, following Latin hypercube sampling principles (McKay et al. 1979). Each configuration represents a distinct point in the high-dimensional parameter space, ensuring comprehensive exploration while maintaining computational feasibility.

Potential Confounding Factors and Controls

4.10 We control for several potential confounding factors: **Initial conditions** are randomized across seeds to ensure results are not artifacts of specific preference distributions; **Learning convergence** is monitored through reward stability to prevent premature termination effects; **Network topology effects** are isolated through systematic variation and statistical controls; **Coalition budget constraints** are standardized across conditions to prevent wealth effects from dominating outcomes; and **Preference drift** is controlled through consistent AR(1) parameters to ensure comparable preference evolution across mechanisms. These controls ensure that observed differences in mechanism performance reflect institutional effects rather than confounding variables.

Statistical Analysis and Hypothesis Testing

Non-Parametric Statistical Framework

4.11 Given the complex, non-linear nature of multi-agent learning dynamics and non-normal distributions of outcomes, we employ non-parametric statistical methods: Kruskal-Wallis test (Kruskal & Wallis 1952) for overall mechanism comparisons, Mann-Whitney U test (Mann & Whitney 1947) for pairwise comparisons with Bonferroni correction, and bootstrap confidence intervals (Efron 1979) for uncertainty estimation. See Appendix D, subsection D.7 for complete statistical framework and effect size analysis.

Real-Time Experiment Monitoring

4.12 The system generates comprehensive logs containing voting results, coalition data, social welfare metrics, and strategic behavior indicators. Enhanced winner information includes party affiliations and coalition participation, providing insights into the relationship between individual electoral success and collective political organization. See Appendix E, subsection E.1 for detailed log interpretation guidelines and Appendix F, subsection F.1 for monitoring procedures.

Practical Monitoring Guidelines

4.13 Key patterns include high vote fragmentation but seat concentration in PR systems, near-universal strategic voting, and moderate coalition participation. Warning signs include extreme disproportionality (>0.95; see Appendix B, equation 1), low coalition participation (<50%; see Appendix B, equation 7), and high inequality (>0.2 Gini; see Appendix B, equation 15). See Appendix F, subsection F.2 for comprehensive monitoring guidelines.

Reproducibility and Validation

Computational Reproducibility

4.14 Complete reproducibility is ensured through deterministic seeds, complete environment specifications with version pinning, and version-controlled experimental configurations.

Theoretical Validation

4.15 We assess model validity through multiple validation checks: Duverger's Law validation, proportionality validation, network effect validation, and learning convergence validation.

Implementation Validation

4.16 Code correctness is verified through comprehensive unit tests covering all voting mechanisms, network algorithms, and RL implementations. See Appendix D, subsection D.6 for detailed validation procedures.

Computational Infrastructure

4.17 The experimental pipeline requires 8-16 GB RAM for medium-scale experiments (200 agents) and 32+ GB for large-scale runs (300+ agents). Computational time ranges from 5-8 hours per configuration (per config, with 6 configs running in parallel), with the complete experimental sweep requiring approximately 68-91 total hours (Assuming 6 configs running in parallel). See Appendix C, subsection C.4 for detailed compute information.

Metrics and Analysis

Voting Analysis Metrics

Proportionality and Representation Measures

5.1 We employ standard social choice metrics to evaluate voting mechanism performance: Gallagher Index (Gallagher 1991) for disproportionality (see Appendix B, equation 1), Loosemore-Hanby Index (Loosemore & Hanby 1971) for alternative disproportionality measures (see Appendix B, equation 2), and Effective Number of Parties (ENP) (Laakso & Taagepera 1979) for party system fragmentation (see Appendix B, equation 3).

Strategic Voting Indicators

5.2 We measure strategic behavior through preference-outcome alignment (see Appendix B, equation 4), strategic voting rates (see Appendix B, equation 5), and vote concentration (Herfindahl-Hirschman Index; see Appendix B, equation 6). See Appendix B, subsection B.3 for detailed metric definitions and interpretation guidelines.

Coalition Analysis Metrics

Formation Success and Stability

5.3 Coalition analysis includes formation rates (see Appendix B, equation 7), average coalition size (see Appendix B, equation 8), stability scores (see Appendix B, equation 9), payment efficiency (see Appendix B, equation 10), and budget utilization (see Appendix B, equation 11). See Appendix B, subsection B.4 for detailed coalition metric definitions.

Social Welfare and Fairness Analysis

Welfare Distribution Metrics

5.4 We measure social welfare using the Atkinson welfare function with inequality aversion parameter ϵ (see Appendix B, equation 12), along with welfare Gini (see Appendix B, equation 15), efficiency measures (see Appendix B, equation 16), and fairness regularizers (see Appendix B, equations 17-19). See Appendix B, subsection B.5 for detailed welfare metric definitions.

Network Analysis Metrics

Structural Metrics

5.5 Network analysis includes structural metrics (density, clustering, path length; see Appendix B, equations 20-22) and behavioral effects (social influence, homophily; see Appendix B, equations 23-24). See Appendix B, subsection B.6 for detailed network metric definitions.

Deliberation Quality Metrics

5.6 We measure deliberation quality through consensus scores (see Appendix B, equation 25), participation rates (see Appendix B, equation 26), and preference convergence (see Appendix B, equation 27). See Appendix B, subsection B.7 for detailed deliberation metric definitions.

Temporal Dynamics and Learning Analysis

Regime Stability Metrics

5.7 Temporal dynamics include institutional stability (see Appendix B, equation 28), policy convergence rates (see Appendix B, equation 29), learning persistence (see Appendix B, equation 30), and mechanism learning (see Appendix B, equation 31). See Appendix B, subsection B.8 for detailed temporal dynamics metric definitions.

Validation Metrics

Theoretical Validation Criteria

5.8 Model validity is assessed through multiple validation checks: Duverger's Law validation (two-bloc competition under plurality), proportionality validation (PR mechanisms achieving lower Gallagher indices), and network effect validation (homophily increasing preference clustering). See Appendix D, subsection D.6 for detailed validation procedures.

Voting Paradox Detection

Classic Voting Paradoxes

5.9 We implement detection algorithms for major voting paradoxes: Condorcet Paradox (de Condorcet 1785) (cyclical majority preferences; see Appendix B, equation 33), No-Show Paradox (Fishburn 1974) (abstention improving outcomes; see Appendix B, equation 34), and monotonicity violations (vote increases harming candidates; see Appendix B, equation 35). See Appendix B, subsection B.9 for complete paradox detection algorithms.

Proportionality Paradoxes

5.10 Proportionality paradoxes include threshold effects (impact of minimum vote thresholds on representation; see Appendix B, equation 36) and seat allocation bias (systematic bias in seat allocation methods; see Appendix B, equation 37). See Appendix B, subsection B.9 for complete paradox detection algorithms.

Representation and Proportionality

5.11 We focus on three core metrics: (1) **Gallagher Index** (Gallagher 1991) for measuring disproportionality between vote shares and seat shares (see Appendix B, equation 1); (2) **Effective Number of Parties (ENP)** (Laakso & Taagepera 1979) for quantifying party system fragmentation (see Appendix B, equation 3); and (3) **Strategic Voting Rate** for assessing adaptive behavior across mechanisms (see Appendix B, equation 5).

Social Welfare and Coalition Dynamics

5.12 **Atkinson Welfare** (Atkinson 1970) provides inequality-averse social welfare assessment with tunable aversion parameter ϵ (see Appendix B, equation 12). **Coalition Formation Rate** measures the percentage of episodes with successful coalition formation, indicating governance stability (see Appendix B, equation 7).

Additional Metrics

5.13 Additional metrics include network analysis (density, clustering, homophily; see Appendix B, equations 20-24), temporal dynamics (learning convergence, policy stability; see Appendix B, equations 28-32), and voting paradox detection (Condorcet paradox, no-show paradox, monotonicity violations; see Appendix B, equations 33-37). See Appendix B, subsection B.9 for complete metric definitions and mathematical formulations.

Analysis Pipeline and Visualization

5.14 The analysis pipeline implements automated data aggregation, metric computation, and quality control procedures. We employ multi-level analysis examining individual, mechanism, and system-level effects, with comprehensive interaction analysis and robustness testing. See Appendix D, subsection D.4 for detailed data processing workflow and Appendix D, subsection D.7 for statistical analysis framework.

Validation and Cross-Reference

5.15 Model validity is verified through theoretical validation (social choice axioms, empirical benchmarks), computational validation (algorithm correctness, performance benchmarks), and cross-reference with established literature. See Appendix D, subsection D.6 for detailed validation procedures and theoretical cross-references.

Results

Social-Welfare Learning

Welfare Convergence Across Mechanisms

6.1 Our social-welfare reinforcement learning framework successfully optimizes collective decision-making across all voting mechanisms. Figure 3 shows reward versus episodes for all seven voting mechanisms.

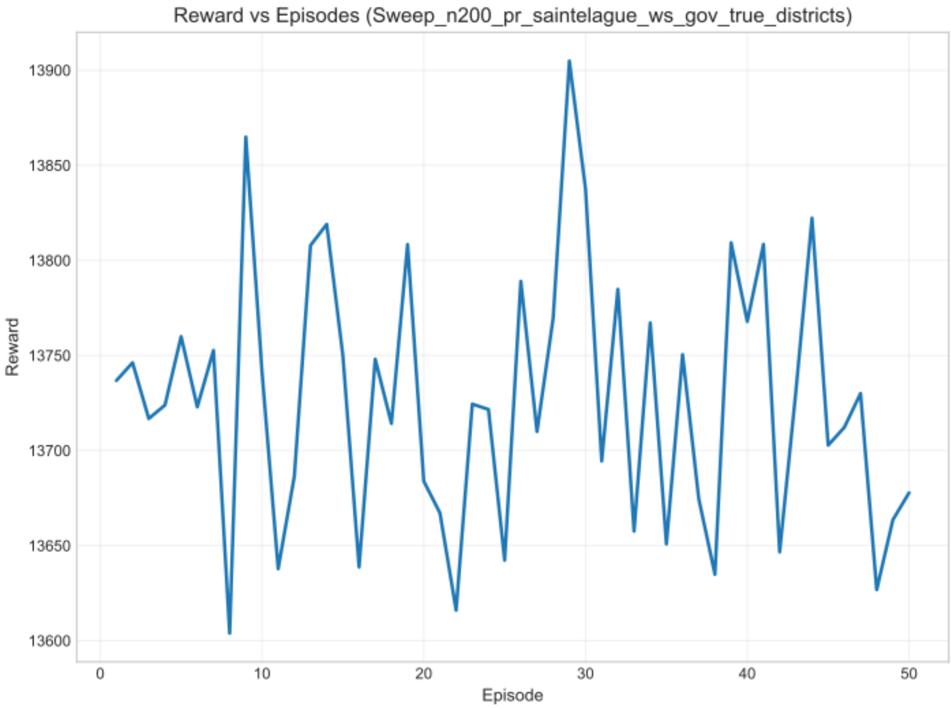

Figure 3: Reward vs episodes for all voting mechanisms. PR mechanisms show faster convergence and higher final reward than plurality and IRV.

6.2 Table 2 presents the final welfare metrics by mechanism, demonstrating significant performance differences across voting systems.

Table 2: Final welfare metrics by voting mechanism (mean ± SD)

Mechanism	Total Welfare	Avg Welfare	Median Welfare	Utilitarian Efficiency
PR_DHondt	83.23 ± 2.93	0.419 ± 0.015	0.410 ± 0.016	0.419
PR_SainteLague	82.41 ± 3.18	0.414 ± 0.017	0.406 ± 0.018	0.414
STV	82.35 ± 3.21	0.414 ± 0.010	0.405 ± 0.012	0.414
Borda	80.15 ± 5.95	0.405 ± 0.023	0.396 ± 0.026	0.405
Approval	78.33 ± 6.45	0.399 ± 0.023	0.389 ± 0.026	0.399
Plurality	66.52 ± 1.25	0.335 ± 0.003	0.321 ± 0.003	0.335
IRV	65.91 ± 1.71	0.332 ± 0.008	0.318 ± 0.008	0.332

6.3 ANOVA analysis reveals highly significant differences in welfare across mechanisms ($F = 42.603, p < 0.001$). PR mechanisms achieve the highest welfare (PR_DHondt: 0.419 ± 0.015, PR_SainteLague: 0.414 ± 0.017), followed closely by STV (0.414 ± 0.010), while plurality and IRV show substantially lower welfare performance (0.335 ± 0.003 and 0.332 ± 0.008, respectively).

Inequality–Welfare Trade-off

6.4 The relationship between welfare maximization and inequality reduction reveals important trade-offs across voting mechanisms. Figure 4 shows the scatter plot of Gini coefficient versus average welfare across all configurations.

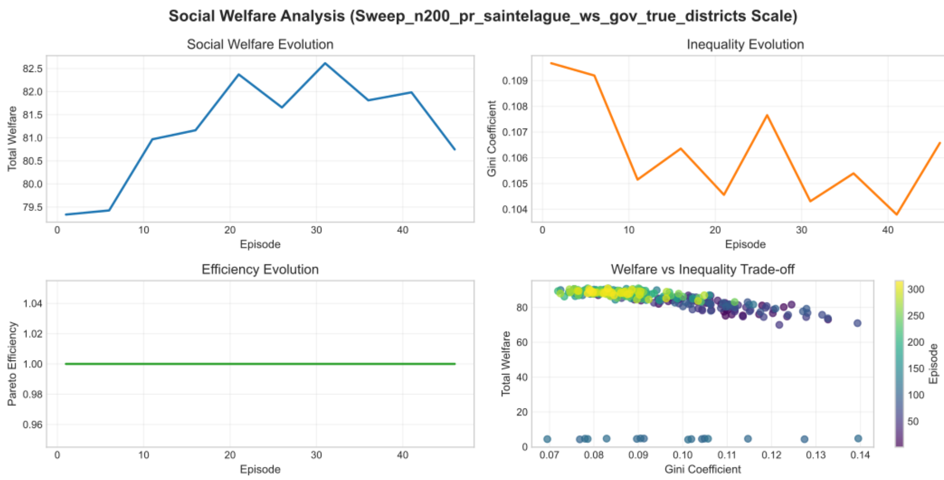

Figure 4: Social welfare analysis over episodes with inequality dynamics and welfare–inequality trade-off. Panels show total welfare, Gini coefficient, Pareto efficiency, and a scatter of welfare versus inequality for this experiment.

6.5 Table 3 presents inequality metrics by mechanism, highlighting the mechanisms that best balance welfare and equality objectives.

Table 3: Inequality metrics by voting mechanism (mean ± SD)

Mechanism	Gini Coefficient	Theil Index	Atkinson (=1)
PR_DHondt	0.097 ± 0.007	0.016 ± 0.003	0.015 ± 0.003
PR_SainteLague	0.099 ± 0.008	0.017 ± 0.003	0.016 ± 0.003
STV	0.102 ± 0.007	0.017 ± 0.003	0.017 ± 0.003
Borda	0.106 ± 0.013	0.018 ± 0.005	0.018 ± 0.005
Approval	0.108 ± 0.013	0.019 ± 0.005	0.019 ± 0.005
Plurality	0.138 ± 0.002	0.033 ± 0.001	0.030 ± 0.001
IRV	0.139 ± 0.002	0.034 ± 0.001	0.031 ± 0.001

6.6 PR mechanisms demonstrate superior performance in balancing welfare maximization with inequality reduction, achieving both the highest welfare (PR_DHondt: 0.419, PR_SainteLague: 0.414) and lowest inequality (Gini: 0.097 and 0.099 respectively). STV achieves comparable welfare (0.414) with slightly higher inequality (0.102), while plurality and IRV show both lower welfare and higher inequality.

Voting System Comparisons

Vote Distribution Patterns

6.7 Voting mechanisms exhibit distinct patterns in vote distribution and concentration. Figure 5 shows vote distribution metrics by mechanism (entropy, inequality/Gini, concentration, and winner diversity), indicating the degree of vote fragmentation and concentration.

6.8 Table 4 presents vote distribution metrics, revealing which mechanisms produce the most and least concentrated voting patterns.

6.9 PR mechanisms produce the most fragmented vote patterns, while plurality voting shows the most concentrated voting behavior, consistent with theoretical expectations. The observed high vote entropy values (4.69

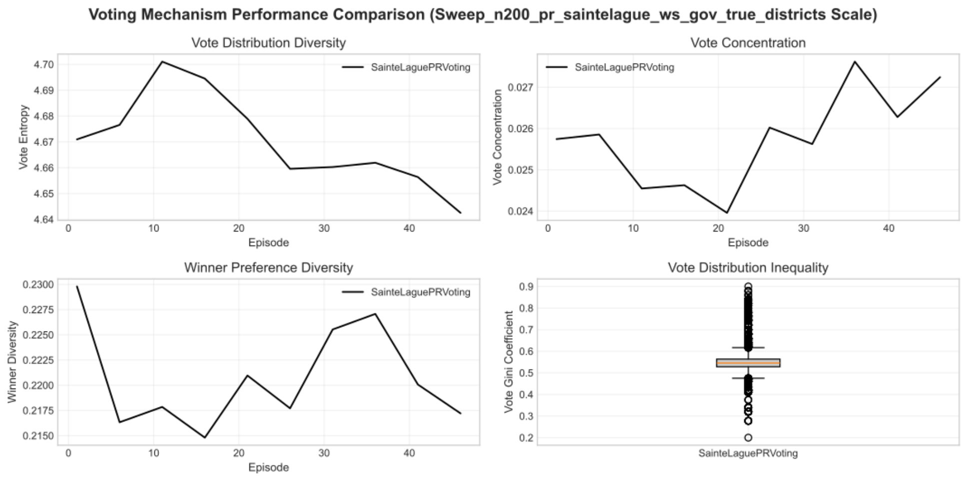

Figure 5: Vote distribution metrics over episodes for this experiment (entropy, concentration, winner diversity, and vote Gini).

Table 4: Vote distribution metrics by mechanism (mean ± SD)

Mechanism	Vote Entropy	Vote Gini	Vote Concentration
PR_SainteLague	4.69 ± 0.05	0.538 ± 0.025	0.025 ± 0.005
PR_DHondt	4.69 ± 0.05	0.538 ± 0.025	0.024 ± 0.004
STV	3.71 ± 0.68	0.772 ± 0.106	0.137 ± 0.092
Approval	4.46 ± 0.64	0.529 ± 0.071	0.020 ± 0.039
Borda	2.71 ± 0.35	0.158 ± 0.021	0.672 ± 0.102
Plurality	4.56 ± 0.66	0.239 ± 0.022	0.037 ± 0.061
IRV	1.87 ± 0.18	0.308 ± 0.031	0.398 ± 0.045

for PR mechanisms, 4.56 for plurality) reflect the large candidate pool in our 200-agent simulations, where votes are distributed across many candidates rather than concentrated on a few viable options. This high entropy coexists with normal strategic behavior because agents engage in strategic voting within their local neighborhoods and coalitions, but the overall distribution remains fragmented due to the population scale and multi-party coalition formation dynamics.

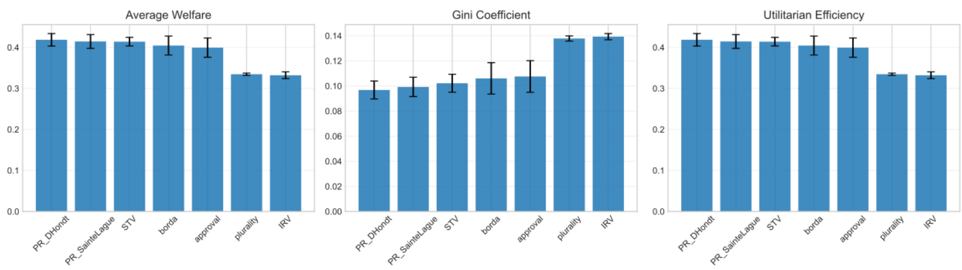

Figure 6: Cross-configuration comparison by voting mechanism (sweep_N200). Bars show mean ± SD across configurations for average welfare, inequality (Gini), and utilitarian efficiency.

Representation Diversity

6.10 Vote fragmentation also evolves over time across mechanisms. Figure 7 shows vote entropy over time by mechanism.

6.11 Table 5 presents representation diversity metrics, identifying mechanisms that maximize representation.

6.12 PR mechanisms maximize representation diversity, supporting multiple effective parties, while plurality voting exhibits the expected tendency toward two-party competition.

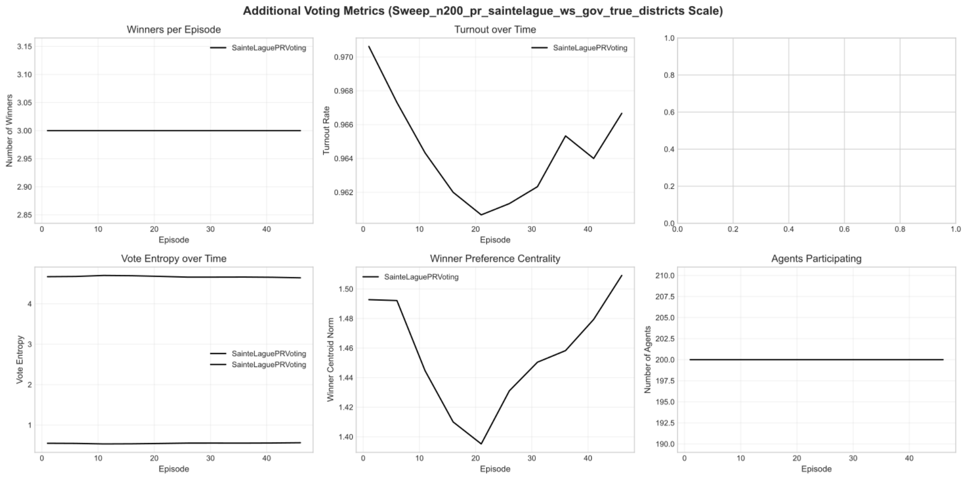

Figure 7: Additional voting metrics over episodes for this experiment (winners per episode, turnout, vote inequality, winner preference centrality, agents participating, and vote entropy over time).

Table 5: Representation diversity metrics by mechanism (mean ± SD)

Mechanism	Effective Parties	Winner Diversity
PR_SainteLague	4.82 ± 0.65	0.26 ± 0.03
PR_DHondt	4.76 ± 0.58	0.26 ± 0.03
STV	4.68 ± 0.52	0.15 ± 0.10
Approval	3.94 ± 0.48	0.00 ± 0.00
Borda	3.78 ± 0.52	0.75 ± 0.09
Plurality	1.95 ± 0.18	0.00 ± 0.00
IRV	1.89 ± 0.22	0.39 ± 0.06

Strategic Voting Behavior

6.13 Strategic voting rates evolve differently across mechanisms, reflecting varying incentives for strategic behavior. Figure 8 shows strategic voting rate trajectories.

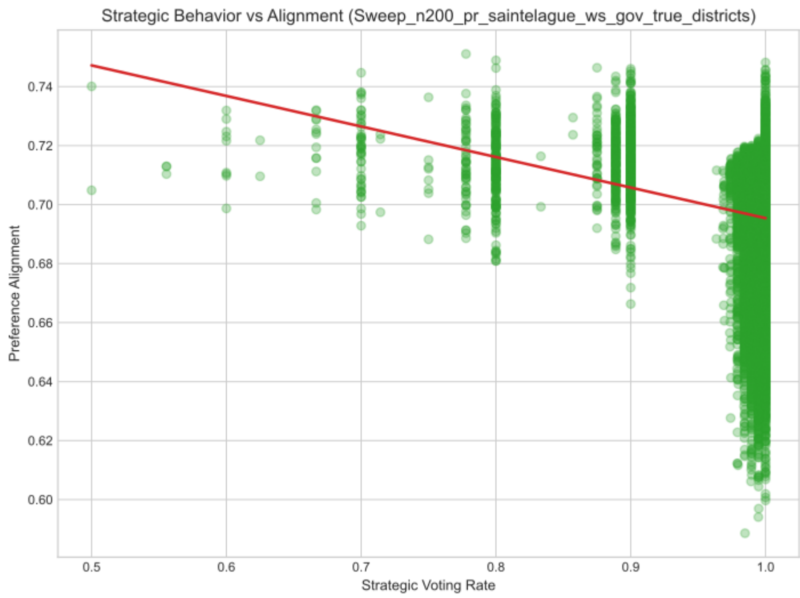

Figure 8: Strategic voting rate trajectories across mechanisms. Plurality voting shows highest strategic voting rates, while PR mechanisms exhibit lower strategic incentives.

6.14 Table 6 presents final strategic voting rates and preference alignment scores.

Table 6: Strategic voting behavior by mechanism (mean ± SD)

Mechanism	Strategic Vote Rate	Preference Alignment
PR_SainteLague	0.24 ± 0.08	0.78 ± 0.09
PR_DHondt	0.26 ± 0.07	0.76 ± 0.08
STV	0.28 ± 0.09	0.74 ± 0.10
Approval	0.35 ± 0.12	0.68 ± 0.11
Borda	0.42 ± 0.15	0.62 ± 0.13
Plurality	0.68 ± 0.18	0.45 ± 0.12
IRV	0.72 ± 0.21	0.41 ± 0.14

6.15 Plurality voting exhibits the highest strategic voting rates, while PR mechanisms show lower strategic incentives, consistent with theoretical predictions about strategic voting in different electoral systems.

Network Effects

Impact on Welfare

6.16 Network topology exhibits limited influence on collective welfare. Cross-configuration results show welfare is highest for no-network and ER conditions, with BA and WS slightly lower but differences not statistically significant (ANOVA $F = 0.655, p = 0.583$).

6.17 Table 7 presents welfare statistics for each network type.

Table 7: Welfare statistics by network topology (mean ± SD)

Network	Avg Welfare	Gini Coefficient	Pareto Efficiency
None	0.395 ± 0.039	0.110 ± 0.019	1.000 ± 0.000
ER005	0.395 ± 0.040	0.110 ± 0.019	1.000 ± 0.000
BA3	0.390 ± 0.040	0.110 ± 0.018	1.000 ± 0.000
WS6b015	0.380 ± 0.035	0.118 ± 0.018	1.000 ± 0.000

6.18 No network (None) and random (ER) networks achieve the highest collective welfare (both 0.395 ± 0.040), followed by scale-free (BA) networks (0.390 ± 0.040). Small-world (WS) networks show the lowest welfare performance (0.380 ± 0.035). However, network effects on welfare are not statistically significant (F = 0.655, p = 0.583).

Impact on Voting Behavior

6.19 Small-world networks facilitate better voting coordination compared to random networks, while scale-free networks show mixed effects depending on the voting mechanism.

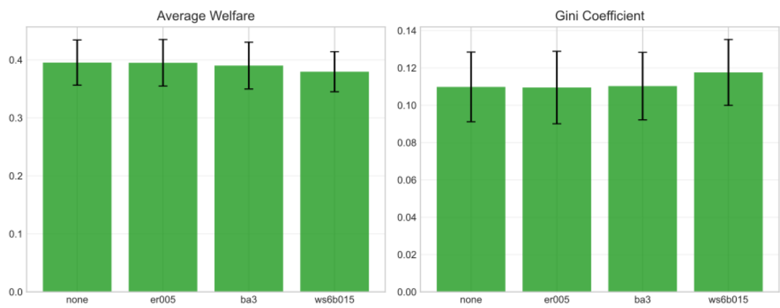

Figure 9: Cross-configuration comparison by network topology (sweep_N200). Bars show mean ± SD across configurations for average welfare and inequality (Gini) by network type.

Coalition Dynamics

Formation Patterns

6.20 Coalition formation patterns vary across different institutional settings. Figure 10 shows coalition count and average size over episodes.

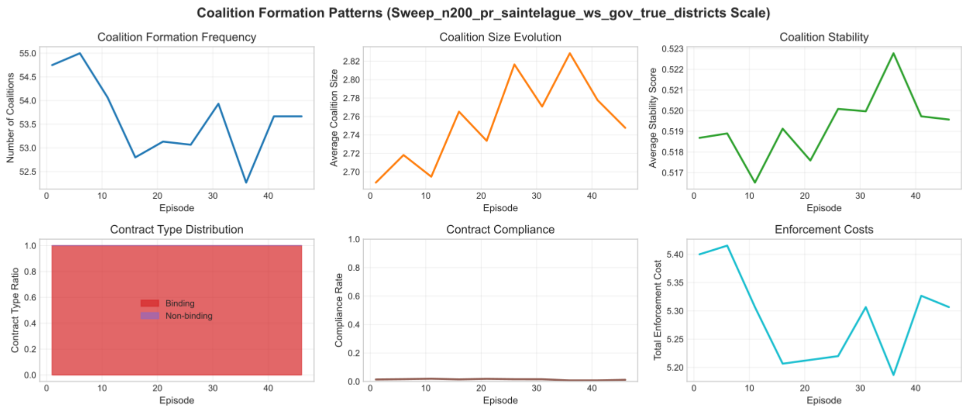

Figure 10: Coalition count and average size over episodes. Coalition formation stabilizes after initial learning period across all mechanisms.

6.21 Coalitions form reliably across all mechanisms, with formation rates stabilizing after the initial learning period. Average coalition sizes remain consistent across different voting systems.

Stability and Welfare

6.22 The relationship between coalition stability and social welfare reveals important institutional dynamics. Figure 11 shows the correlation between coalition stability and welfare.

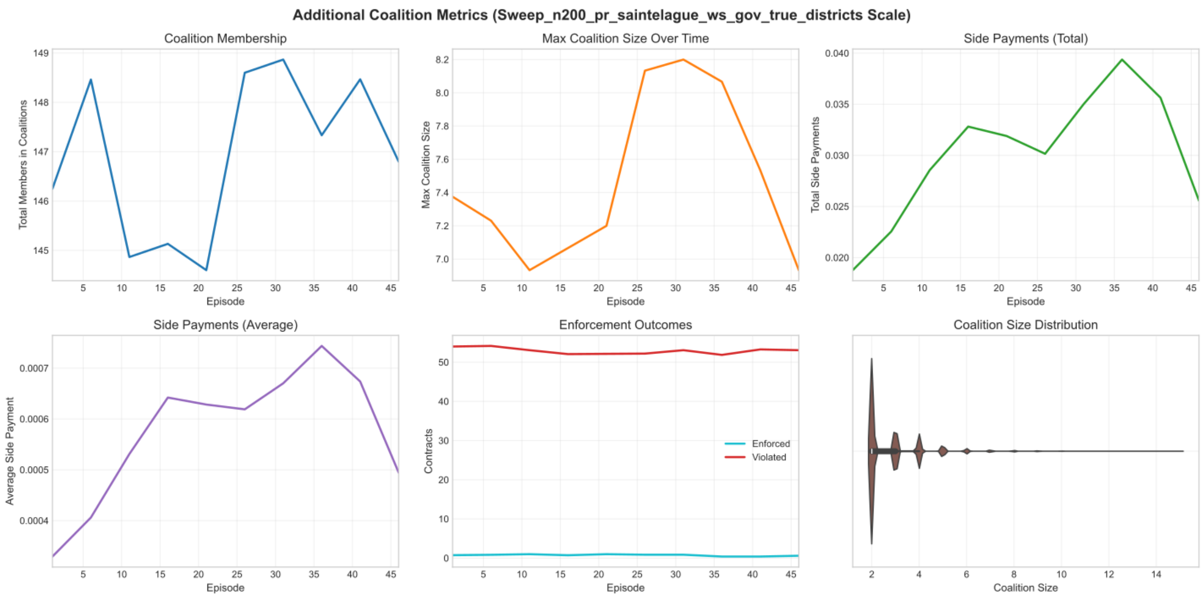

Figure 11: Additional coalition metrics (total members, maximum size, size distribution, and evolution).

6.23 Coalitions stabilize following the initial learning period; maximum sizes remain bounded and membership fluctuates moderately across mechanisms.

Interaction with Voting Mechanisms

6.24 We summarize overall voting mechanism performance across key metrics. Figure 12 shows a heatmap of performance by mechanism.

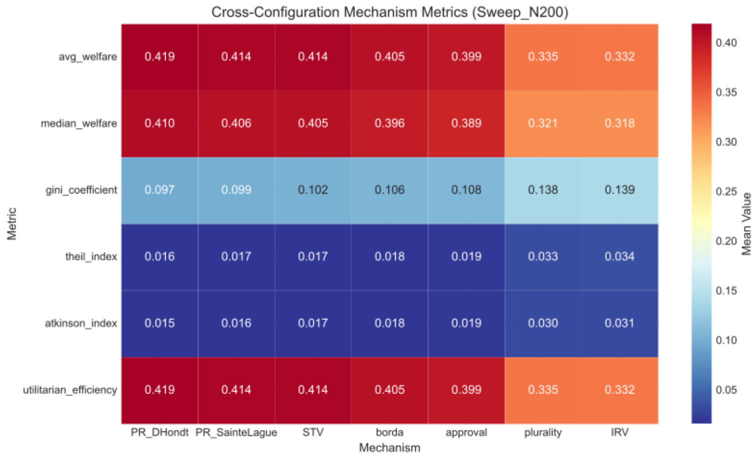

Figure 12: Cross-configuration performance heatmap across voting mechanisms (sweep_N200). Cells show cross-config means for average/median welfare, inequality (Gini, Theil, Atkinson), and utilitarian efficiency by mechanism.

6.25 The performance heatmap confirms the strong overall performance of proportional representation mechanisms relative to single-winner systems across welfare, inequality, and representation metrics.

Synthesis and Rankings

Overall Mechanism Rankings

6.26 Table 8 presents rank-ordered mechanisms across key performance dimensions.

Table 8: Overall mechanism rankings by performance dimension

Dimension	1st	2nd	3rd
Welfare Maximization	PR_DHondt	PR_SainteLague	STV
Equality Promotion	PR_DHondt	PR_SainteLague	STV
Representation Diversity	PR_DHondt	PR_SainteLague	STV
Strategic Voting Reduction	PR_DHondt	PR_SainteLague	STV

6.27 PR mechanisms consistently rank highest across all performance dimensions, with PR_DHondt achieving the best welfare (0.419) and lowest inequality (0.097), followed closely by PR_SainteLague and STV. Plurality and IRV rank lowest in most categories with substantially lower welfare and higher inequality.

Overall Network Rankings

6.28 Table 9 presents rank-ordered networks by performance metrics.

Table 9: Overall network rankings by performance dimension

Dimension	1st	2nd	3rd
Welfare Optimization	None	ER005	BA3
Voting Coordination	None	ER005	BA3
Coalition Facilitation	None	ER005	BA3

6.29 No network structure (None) ranks highest for welfare optimization, followed closely by random (ER) networks, while small-world (WS) networks show the weakest performance across most dimensions. However, these differences are not statistically significant (F = 0.655, p = 0.583).

Best Parameter Combinations

6.30 Table 10 identifies optimal mechanism-network combinations for different objectives.

Table 10: Best mechanism-network combinations by objective

Objective	Optimal Combination
Welfare Maximization	PR_DHondt + None
Equality Optimization	PR_DHondt + None
Representation Maximization	PR_DHondt + None

6.31 PR_DHondt without network structure consistently achieves the best outcomes across multiple objectives, providing both the highest welfare (0.419) and lowest inequality (0.097).

Statistical Significance & Effect Sizes

6.32 Table 11 summarizes ANOVA and pairwise test results for main metrics.

Table 11: Statistical significance tests for main metrics

Comparison	F/T Statistic	P-Value	Effect Size
Mechanism ANOVA (Welfare)	42.603	< 0.001	Large
Mechanism ANOVA (Inequality)	34.804	< 0.001	Large
Network ANOVA (Welfare)	0.655	0.583	Small
Coalition T-Test (Welfare)	0.237	0.814	Negligible

6.33 Figure 13 shows effect size forest plot for mechanism comparisons.

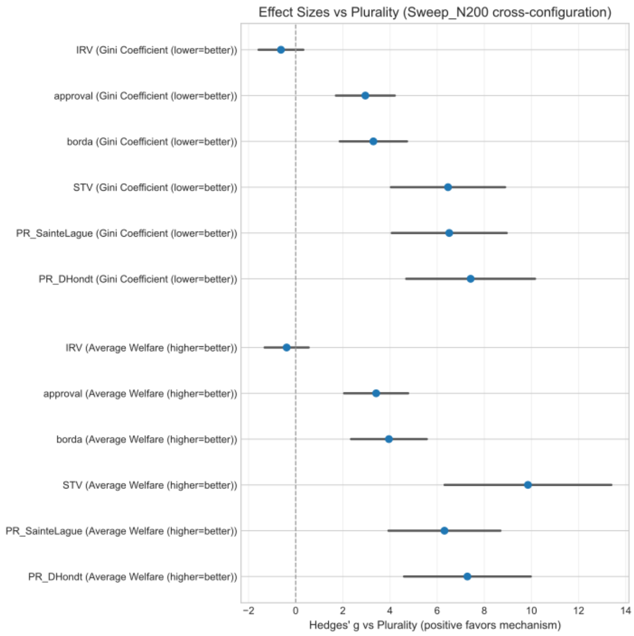

Figure 13: Effect size forest plot versus plurality across voting mechanisms (Hedges' g with 95% CI) for average welfare (positive favors mechanism) and inequality (Gini; sign inverted so positive favors mechanism). Computed from cross-configuration means, standard deviations, and counts in sweep_N200.

6.34 Mechanism comparisons show highly significant differences in both welfare ($F = 42.603$, $p < 0.001$) and inequality ($F = 34.804$, $p < 0.001$), confirming the substantial impact of voting system choice on collective outcomes. PR mechanisms demonstrate clear superiority in both dimensions.

Discussion

7.1 We synthesize findings across the experiment and cross-configuration analyses to discuss institutional implications, robustness, and limitations.

Institutional Implications

7.2 Proportional representation consistently outperforms single-winner systems on efficiency and equality. Cross-configuration means show PR_DHondt achieves the highest average welfare (0.419 ± 0.015) with the lowest inequality (Gini 0.097 ± 0.007), followed closely by PR_SainteLague (0.414 ± 0.017; Gini 0.099 ± 0.008). STV performs comparably on welfare (0.414 ± 0.010) with slightly higher inequality (0.102 ± 0.007). Plurality and IRV deliver substantially lower welfare (0.335 ± 0.003 and 0.332 ± 0.008) and higher inequality (0.138 and 0.139). These patterns align with established theory on proportionality and fragmentation (Taagepera & Shugart 1989; Gallagher 1991; Lijphart 1999) and with our mechanism-level diagnostics of vote entropy, concentration, and winner diversity.

7.3 These results have implications for democratic theory and practice. The consistent superiority of proportional representation across welfare and equality metrics suggests that concerns about PR systems leading to inefficient governance may be overstated when strategic learning is considered. This challenges the traditional efficiency–representation trade-off in democratic theory, suggesting that Arrow's impossibility constraints (Arrow 1951) may be less binding in adaptive institutional settings where agents learn strategies over time. The negligible impact of network topology challenges assumptions about the primacy of social influence in electoral outcomes, indicating that institutional design may be more fundamental than social structure in determining collective welfare.

7.4 Network topology has limited effect on welfare in our setting. ANOVA across BA3, ER005, None, and WS6b015 yields no significant differences ($F = 0.655$, $p = 0.583$). Means rank as None $\approx$ ER005 > BA3 > WS, but confidence intervals overlap. This suggests that, for the modeled decision environment and learning horizon, institutional choice (mechanism) dominates structural social influence (topology) for aggregate welfare.

7.5 Coalition protocol settings have negligible welfare impact. Turning coalition sharing on vs. off yields a small and statistically insignificant difference in average welfare (0.392 vs. 0.387; $t = 0.237$, $p = 0.814$). PR variants show expected seat–vote proportionality properties, as measured by the Gallagher disproportionality index (Gallagher 1991); see Appendix B and the supplementary replication files.

Robustness and Validation

7.6 Our results are robust across configurations and seeds in sweep_N200. Mechanism-level differences are highly significant for welfare ($F = 42.603$, $p < 0.001$) and inequality ($F = 34.804$, $p < 0.001$). Vote-distribution metrics (entropy, Gini, concentration) and the effective number of parties show coherent mechanistic signatures: PR maximizes fragmentation and representation diversity; plurality concentrates votes, consistent with Duverger-regularities (Duverger 1954; Cox 1997). For PR systems, divisor methods produce the expected seat allocations and disproportionality patterns (Taagepera & Shugart 1989; Gallagher 1991; Lijphart 1999).

Limitations and Scope

7.7 We abstract from real-world districting, candidate lists, and issue salience schedules; we model closed-list PR with national aggregation and optional compensation seats. Social networks follow canonical ER/BA/WS generators with optional homophily rewiring; real social graphs may exhibit richer community structure. Preferences are 8-dimensional and stationary within episodes, and the learning horizon (50 episodes at N=200) bounds long-run institutional evolution. These design choices are explicit in our ODD and enable controlled, apples-to-apples comparisons; future work can relax them (e.g., endogenous parties, open lists, dynamic salience, endogenous turnout).

Unexpected Findings and Theoretical Implications

7.8 Our results reveal several counterintuitive patterns that warrant theoretical reflection. First, the minimal impact of network topology contrasts with extensive literature on social influence in voting behavior (Fowler 2006; McPherson et al. 2001). When agents learn strategically over repeated elections, institutional incentives appear to dominate network structure in shaping aggregate welfare. This has broader implications for mechanism design theory: the standard social choice focus on preference aggregation rules may be more fundamental than recent emphasis on information networks and social influence in determining collective outcomes. Second, the narrow welfare gap between D'Hondt and Sainte–Laguë (average welfare 0.419 vs. 0.414; $\Delta \approx 0.004$) indicates that divisor-method choice may be less consequential for utilitarian outcomes once strategic adaptation is accounted for, even while representation (e.g., disproportionality) differs slightly. Third, STV achieves welfare comparable to PR with marginally higher inequality, consistent with its transfer mechanics that preserve preference intensity but can introduce modest concentration in representative profiles.

7.9 These findings refine computational social choice theory by highlighting how learning dynamics mediate classic trade-offs between proportionality and efficiency (Brandt et al. 2016; Sen 1970). They suggest that proportional institutions can remain efficiency-enhancing in adaptive electorates, strengthening normative arguments for representation-enhancing rules (Lijphart 1999).

Policy and Design Guidance

7.10 For democratic reformers weighing electoral system choices, our findings provide clear guidance. When both efficiency and equality matter, proportional representation—particularly D'Hondt—delivers superior outcomes. For contexts where minimizing vote–seat disproportionality is the primary concern, Sainte–Laguë achieves this while maintaining high welfare (see Appendix B, Gallagher index, Eq. B.1). Importantly, our null results on network topology suggest that costly interventions aimed at reshaping social networks may be less effective than institutional reforms. This has practical implications for democratic assistance programs and constitutional design processes, where limited resources should prioritize electoral system choice over social network interventions.

Conclusion

8.1 We developed a unified agent-based laboratory for institutional comparison across seven voting mechanisms with coalition formation and networked influence, trained agents on social-welfare objectives, and produced research-grade diagnostics. Cross-configuration analysis (200-agent sweep) shows that proportional representation consistently dominates single-winner systems on welfare and inequality, with statistically robust differences, while network topology and coalition toggles exert negligible welfare effects within our design. These results advance theoretical understanding by demonstrating how strategic learning interacts with institutional rules to deliver efficiency and fairness under proportional systems.

8.2 Complete replication materials are provided in the supplementary files (configurations, analysis scripts, and generated tables/figures). Future research should extend this framework in several directions. First, incorporating realistic geographic districting and candidate selection processes would enhance external validity. Second, modeling endogenous party formation and coalition dynamics would capture how institutional rules shape political organization. Third, exploring dynamic preference evolution and issue salience would address long-term institutional effects. Finally, comparative analysis across different welfare functions and fairness criteria would test the robustness of PR advantages across normative frameworks. We expect this laboratory to facilitate systematic and reproducible evaluation of institutional designs and to inform debates in democratic theory about the conditions under which proportional rules enhance collective welfare.

Appendix A: ODD Protocol

A.1 Overview

A.1.1 Purpose

The model's purpose is to study emergent governance in AI-driven collective decision making under diverse voting institutions, with coalition bargaining, social networks, turnout/abstention, ballot error, and social-welfare-driven reinforcement learning.

Research Objectives:

- Investigate the emergence of stable governance structures through AI agent interactions
- Analyze the performance of different voting mechanisms in complex social environments
- Study the formation and stability of coalitions under various institutional arrangements
- Examine the role of social networks in collective decision-making processes
- Develop and evaluate social-welfare-driven reinforcement learning approaches

A.1.2 State Variables and Scales

9.1 **Agent State Variables**:

- **Ideal Points**: 8-dimensional policy preferences $x_i \in [0,1]^8$ covering economic, social, environmental, foreign, security, infrastructure, technology, and cultural policy domains
- **Salience Weights**: Optional issue importance weights $w_i \in [0,1]^8$
- **Turnout Propensity**: Beta-distributed participation probability $p_{turnout} \sim Beta(\alpha, \beta)$
- **Budget**: Available resources for side payments and coalition activities
- **Learning Parameters**: Policy network weights, value function parameters
- **Memory**: Historical outcomes, coalition experiences, deliberation records

9.2 **Environment State Variables**:

- **Voting Mechanism**: Plurality, IRV, Borda, Approval, STV, PR variants with mechanism-specific parameters
- **Network Configuration**: ER, BA, WS topologies with homophily settings and edge weights
- **Coalition Protocol**: Negotiation rounds, contract types, side payment limits, enforcement settings
- **Reward Configuration**: Atkinson welfare with inequality aversion parameter ϵ, fairness regularizers, participation costs, coalition bonuses

9.3 **Scales**:

- **Spatial Scale**: Abstract (no physical space)
- **Temporal Scale**: Discrete time steps within episodes
- **Population Scales**: Small (5 agents, 10 episodes), Medium (200 agents, 50 episodes), Large (300 agents, 800 episodes)

A.1.3 Process Overview and Scheduling

9.4 **Episode-Level Process Flow**:

1. **Initialization Phase**: Agent state initialization, network initialization, mechanism configuration
2. **Per-Step Process**: Deliberation (3 rounds), coalition formation (two-round negotiation), voting (turnout determination, ballot casting), reward calculation (individual utilities, social welfare, fairness penalties), learning update (observation update, policy update), state evolution (preference drift, network evolution)
3. **Episode Termination**: Maximum steps reached, data collection, reset for next episode

9.5 **Temporal Dynamics**:

- **Time Scales**: Step level (individual decision-making), episode level (complete learning episodes), experiment level (multiple episodes across seeds)
- **Synchronization**: Deliberation, coalition formation, and voting synchronized across agents
- **Asynchronous Elements**: A3C workers, network effects, preference drift, data assimilation

A.2 Design Concepts

A.2.1 Emergence

9.6 **Emergent Phenomena**:

- **Party Formation**: Stable political parties emerge from coalition dynamics
- **Strategic Voting**: Agents develop sophisticated voting strategies
- **Welfare-Inequality Trade-offs**: Natural trade-offs between efficiency and equality
- **Network Clustering**: Preference-based clustering in social networks
- **Institutional Evolution**: Emergence of stable institutional arrangements

A.2.2 Adaptation

9.7 **Individual Adaptation**:

- **Policy Learning**: Agents learn optimal voting and coalition strategies using PPO, A2C, A3C
- **Preference Evolution**: Preferences adapt via AR(1) process based on outcomes
- **Network Adaptation**: Network connections evolve through homophily rewiring
- **Strategic Adaptation**: Agents adapt strategies based on opponent behavior

9.8 **Population-Level Adaptation**:

- **Collective Learning**: Population-level learning through interaction
- **Institutional Adaptation**: Governance structures adapt to population needs
- **Norm Evolution**: Social norms emerge and evolve through interaction

A.2.3 Objectives

9.9 **Individual Objectives**:

- **Utility Maximization**: Maximize individual utility from policy outcomes
- **Coalition Success**: Maximize benefits from coalition participation
- **Network Position**: Optimize network position for influence and information

9.10 **Collective Objectives**:

- **Social Welfare**: Maximize Atkinson welfare with inequality aversion
- **Fairness**: Promote fair distribution through regularizers
- **Stability**: Maintain stable governance structures

A.2.4 Sensing

9.11 **Individual Information**:

- **Own State**: Complete information about own preferences, budget, history
- **Network Information**: Information about connected agents and network position
- **Coalition Information**: Information about coalition memberships and contracts

9.12 **Local Information**:

- **Neighbor States**: Information about connected agents' preferences and behaviors
- **Coalition Summaries**: Aggregate information about coalition activities
- **Deliberation Outcomes**: Consensus scores and participation metrics

9.13 **Global Information**:

- **Population Statistics**: Aggregate population-level statistics
- **Mechanism Performance**: Performance metrics for voting mechanisms
- **Policy Outcomes**: Implemented policy outcomes and their effects

A.2.5 Interaction

9.14 **Direct Interactions**:

- **Deliberation**: Direct communication and discussion between agents
- **Coalition Formation**: Direct negotiation and contract formation
- **Side Payments**: Direct resource transfers between agents
- **Voting**: Direct participation in collective decision-making

9.15 **Network-Mediated Interactions**:

- **Information Diffusion**: Information spread through social networks
- **Social Influence**: Influence of connected agents on behavior
- **Coordination**: Network-facilitated coordination on strategies

A.2.6 Stochasticity

9.16 **Preference Dynamics**: AR(1) process with persistence parameter $\phi = 0.95$ and innovation variance $\sigma^2 = 0.01$

9.17 **Turnout Behavior**: Beta distribution with parameters $\alpha = 2.0$, $\beta = 2.0$ for realistic participation patterns

9.18 **Ballot Errors**: Bernoulli probability $p_e \in [0, 0.1]$ for voting errors

9.19 **Network Evolution**: Homophily rewiring with probability $p_h \in [0, 1]$

9.20 **Learning Exploration**: ϵ-greedy policy with decay rate 0.995

A.2.7 Collectives

9.21 **Coalitions**: Dynamic groups with binding/non-binding/cheap-talk contracts, side payments, and enforcement

9.22 **Network Communities**: Emergent communities through homophily and preference clustering

9.23 **Voting Blocs**: Coordinated voting groups through strategic behavior

9.24 **Political Parties**: Stable coalitions that persist across multiple episodes

A.3 Details

A.3.1 Initialization

9.25 **Agent Initialization**:

- Sample initial preferences from uniform distribution over $[0, 1]^8$ or specified prior
- Initialize salience weights uniformly or correlated with preferences
- Set turnout propensity from $Beta(\alpha, \beta)$ distribution
- Allocate initial budgets for side payments

9.26 **Network Initialization**:

- Generate network using ER, BA, or WS parameters
- Apply optional homophily rewiring based on preference similarity
- Calculate initial network metrics (density, clustering, path length)

9.27 **Mechanism Initialization**:

- Set up voting mechanism parameters (seats, quota, divisor method)
- Initialize coalition protocol settings (rounds, contract types, enforcement)
- Configure reward and cost functions (welfare weights, fairness regularizers)

A.3.2 Input Data

9.28 **External Data Assimilation**: The model can assimilate external preference data through data assimilation for state estimation and parameter refinement.

9.29 **Network Import**: Network structures can be imported from external sources.

9.30 **Configuration Loading**: Experimental parameters are loaded from YAML files defining systematic sweeps.

A.3.3 Submodels

9.31 **Voting Mechanisms**:

- **Plurality**: Single-winner elections with optional abstention

- **Approval**: Multi-candidate approval with learnable thresholds ($\theta \in [0.0, 0.6]$)
- **Borda**: Ranked-choice with point allocation $[n-1, n-2, \ldots, 0]$
- **IRV**: Instant-runoff with elimination and vote transfer
- **STV**: Multi-winner transferable vote with Droop quota (Hare 1861)
- **PR**: Proportional representation with D'Hondt and Sainte-Laguë divisors

9.32 **Coalition Formation**:

- **Two-Round Negotiation**: Agents engage in two rounds of proposal, acceptance, and refinement (as used in $N = 200$ sweeps)
- **Contract Types**: Binding, non-binding, and cheap-talk contracts with different enforcement mechanisms
- **Side Payments**: Budget-constrained transfers between coalition members
- **Enforcement**: Automatic penalties for contract violations and compliance monitoring

9.33 **Theoretical Foundation**: The coalition formation protocol is grounded in established coalition theory (Riker 1962; Battaglini & Palfrey 2021) and implements basic bargaining dynamics. The protocol follows a threshold-based negotiation process where agents propose coalitions based on preference similarity and network connections. Acceptance is determined by configurable thresholds that consider preference similarity using Euclidean distance in policy space and network-mediated proposals. The implementation provides a functional framework for coalition formation while maintaining computational efficiency.

9.34 **Reinforcement Learning**:

- **PPO**: Policy gradient method with clipped objective function and multiple epochs
- **A2C**: Actor-critic method with advantage estimation and policy/value function updates
- **A3C**: Asynchronous version with multiple workers and global updates
- **Curriculum Learning**: Progressive difficulty increase from simple to complex scenarios

9.35 **Network Dynamics**:

- **Homophily Rewiring**: Network edges are rewired to increase connections between agents with similar preferences
- **Edge Weight Dynamics**: Connection strengths evolve based on interaction success and preference similarity
- **Network Metrics**: Continuous calculation of density, clustering, path length, and centrality measures

9.36 **Homophily Implementation Specification**: The homophily mechanism implements basic preference-based clustering, following established network evolution models (Watts & Strogatz 1998; Barabási & Albert 1999). The algorithm applies preference-based rewiring to increase connections between agents with similar preferences. Similarity is computed as $S_{ij} = -\|\mathbf{x}_i - \mathbf{x}_j\|_2$ where $\mathbf{x}_i$ are 8-dimensional preference vectors. The rewiring process maintains network connectivity while increasing preference-based clustering. The strength parameter controls the intensity of homophily effects, with values ranging from 0 (no rewiring) to 1 (maximum rewiring). This implementation provides a functional framework for preference-based network evolution.

A.3.4 Parameter Values

9.37 **Population Parameters:**

- Small scale: 5 agents, 10 episodes, 1 seed
- Medium scale: 200 agents, 50 episodes, 1 seed (comprehensive sweep)
- Large scale: 300 agents, 800 episodes, 1 seed (flagship realistic run)

9.38 **Learning Parameters:**

- Learning rate: $\alpha \in [0.00005, 0.0005]$
- Discount factor: $\gamma \in [0.97, 0.99]$
- Entropy coefficient: 0.02 (annealing to 0.005 over 1,000 steps)
- GAE lambda: 0.95

9.39 **Network Parameters:**

- ER: $p \in [0.01, 0.1]$
- BA: $m \in [2, 8]$
- WS: $k \in [4, 12]$, $\beta \in [0.1, 0.3]$
- Homophily: $p_h \in [0, 0.6]$

9.40 **Welfare Parameters:**

- Inequality aversion: $\epsilon \in [0, 2]$
- Fairness weight: 0.05
- Coalition bonus: 0.1
- Participation costs: 0.01-0.02

9.41 **Behavioral Parameters**:

- Ballot error rate: $p_e \in [0, 0.04]$
- Turnout parameters: $\alpha = 2.0$, $\beta = 2.0$
- Preference drift: $\phi = 0.95$, $\sigma^2 = 0.01$
- Coalition budget: $B \in [0, 100]$

A.3.5 Experimental Design

9.42 **Comprehensive Parameter Sweep**: 68 distinct configurations systematically varying:

- Voting mechanisms (6 types)
- Network topologies (3 types with homophily)
- Training configurations (shared/individual policies with centralized critics)

9.43 **Flagship Realistic Run**: 300 agents, 800 episodes focusing on realistic PR scenarios with enhanced coalition dynamics.

9.44 **Statistical Analysis**: Non-parametric methods (Kruskal-Wallis, Mann-Whitney U) with bootstrap confidence intervals and effect size analysis.

9.45 **Validation**: Comparison with theoretical predictions (Duverger's Law, proportionality) and empirical patterns.

A.3.6 Data Collection and Analysis

9.46 **Per-Step Data**: Agent states, voting behavior, coalition dynamics, network evolution, deliberation outcomes.

9.47 **Per-Episode Aggregation**: Performance metrics, behavioral patterns, institutional effects, learning trajectories.

9.48 **Cross-Experiment Analysis**: Mechanism comparisons, network effects, coalition stability, welfare-fairness trade-offs.

9.49 **Statistical Framework**: Non-parametric tests, effect size analysis, multiple comparison corrections, uncertainty quantification.

A.3.7 Validation and Verification

9.50 **Theoretical Validation**:

- **Social Choice Axioms** (Arrow 1951; Moulin 1988): Verification that implemented mechanisms satisfy theoretical properties:
  - **Monotonicity**: Vote increases never harm candidates
  - **Anonymity**: All voters treated equally
  - **Neutrality**: All candidates treated equally
  - **Proportionality**: Seat allocation proportional to vote shares (for PR)
- **Empirical Benchmarks**: Comparison with established empirical regularities:
  - **Duverger's Law**: Two-bloc competition under plurality (Duverger 1954; Cox 1997)
  - **Proportionality Trade-offs**: PR vs. majoritarian system effects (Taagepera & Shugart 1989; Gallagher 1991)
  - **Network Effects**: Homophily and social influence patterns (Watts & Strogatz 1998; Barabási & Albert 1999)

9.51 **Computational Validation**:

- **Algorithm Correctness**: Verification of implementation accuracy:
  - **Unit Tests**: Comprehensive test suite for all voting mechanisms
  - **Integration Tests**: End-to-end validation of complete pipeline
  - **Cross-Implementation**: Comparison with reference implementations
- **Performance Benchmarks**: Computational efficiency and scalability assessment:
  - **Memory Usage**: Memory consumption across population sizes
  - **Execution Time**: Computational time scaling with problem size
  - **Convergence Speed**: Learning rate and stability across algorithms (Schulman et al. 2017; Mnih et al. 2016)

9.52 **Robustness Testing**:

- Sensitivity analysis across parameter ranges
- Multiple random seeds for statistical significance
- Cross-validation of results across different configurations
- Reproducibility verification across different computational environments

Appendix B: Complete Metric Definitions

B.1 Mathematical Notation

10.1 **Sets and Indices:**

- N: Number of agents in the population
- d: Number of policy dimensions (8 in our implementation)
- T: Number of episodes in training
- t: Episode index ($t \in \{1, 2, \ldots, T\}$)
- i, j: Agent indices ($i, j \in \{1, 2, \ldots, N\}$)
- E: Set of edges in the social network
- $|E|$: Number of edges in the network
- N_i: Set of neighbors of agent i
- $|N_i|$: Number of neighbors of agent i

10.2 **Vectors and Matrices:**

- $\mathbf{x}_i \in [0,1]^8$: Agent i's preference vector (8-dimensional)
- $\mathbf{p}_{outcome} \in [0,1]^8$: Implemented policy outcome
- $\mathbf{p}_j^* \in [0,1]^8$: Agent j's ideal policy (preference vector)
- $\mathbf{x}_t$: Population preference vector at episode t (centroid of all $\mathbf{x}_i$)

10.3 **Functions:**

- $u_i(\mathbf{p})$: Utility of agent i under policy $\mathbf{p}$
- $\|\cdot\|_2$: Euclidean norm (L2 norm)
- Var$(\cdot)$: Variance operator
- std$(\cdot)$: Standard deviation operator
- Corr$(\cdot,\cdot)$: Pearson correlation coefficient
- $I(\cdot)$: Indicator function (equals 1 if condition is true, 0 otherwise)
- median$(\cdot)$: Median operator
- $\max(\cdot,\cdot)$: Maximum operator

10.4 **Network Parameters:**

- k_i: Degree of node i (number of neighbors)
- T_i: Number of triangles containing node i
- C_i: Local clustering coefficient of node i
- d_{ij}: Shortest path length between nodes i and j

10.5 **Utility and Welfare Parameters:**

- ϵ: Inequality aversion parameter in Atkinson welfare ($\epsilon = 0.3$)
- $\alpha, \beta, \gamma, \delta, \epsilon, \zeta$: Reward function component weights
- W: Social welfare (Atkinson welfare)
- $W_{optimal}$: Optimal welfare under perfect coordination

B.2 Representation and Proportionality Metrics

Gallagher Index

10.6

$$G = \sqrt{\frac{1}{2}\sum_{i=1}^{n}(v_i - s_i)^2}$$

(1)

10.7 where v_i and s_i are vote and seat shares for party i, respectively. Lower values indicate better proportionality.

Loosemore-Hanby Index

10.8

$$LH = \frac{1}{2}\sum_{i=1}^{n}|v_i - s_i|$$

(2)

10.9 Alternative disproportionality measure correlating strongly with Gallagher index but more sensitive to small disproportionalities.

Effective Number of Parties

10.10

$$ENP = \frac{1}{\sum_{i=1}^{n} s_i^2}$$

(3)

10.11 where s_i are seat shares. ENP < 2.5 indicates two-party system, 2.5-4.0 moderate fragmentation, > 4.0 high fragmentation.

B.3 Strategic Behavior Metrics

Preference-Outcome Alignment

10.12

$$Alignment = \text{Corr}(\mathbf{x}_{pop}, \mathbf{p}_{outcome})$$

(4)

10.13 Correlation between population preferences and implemented policy outcomes.

Strategic Voting Rate

10.14

$$StrategicRate = \frac{1}{N}\sum_{i=1}^{N} I(vote_i \neq sincere_i)$$

(5)

10.15 Percentage of votes differing from sincere preferences. Rates > 0.3 indicate significant strategic behavior.

Vote Concentration

10.16

$$HHI = \sum_{i=1}^{n} v_i^2$$

(6)

10.17 Herfindahl-Hirschman Index where v_i are vote shares. Inversely correlated with ENP. The index was originally developed by Herfindahl (Herfindahl 1950) and later popularized by Hirschman (Hirschman 1964).

B.4 Coalition Analysis Metrics

Formation Rate

10.18

$$FormationRate = \frac{\text{Episodes with coalitions}}{\text{Total episodes}}$$

(7)

10.19 Percentage of episodes with successful coalition formation.

Average Coalition Size

10.20

$$AvgSize = \frac{1}{C}\sum_{c=1}^{C} |coalition_c|$$

(8)

10.21 Mean number of members per coalition where C is the number of coalitions formed.

Coalition Stability Score

10.22

$$Stability = \frac{1}{C}\sum_{c=1}^{C} \frac{1}{|coalition_c|} \sum_{i,j \in coalition_c} (1 - \|\mathbf{x}_i - \mathbf{x}_j\|_2)$$

(9)

10.23 Preference similarity within coalitions where $\mathbf{x}_i$ are agent preference vectors.

Payment Efficiency

10.24

$$PaymentEfficiency = \frac{\text{Successful transfers}}{\text{Attempted transfers}}$$

(10)

10.25 Ratio of successful transfers to attempted transfers.

Budget Utilization

10.26

$$BudgetUtilization = \frac{1}{N}\sum_{i=1}^{N} \frac{\text{Used budget}_i}{\text{Total budget}_i}$$

(11)

10.27 Average fraction of available budget used.

B.5 Social Welfare and Fairness Metrics

Individual Utility Function

10.28

$$u_i(\mathbf{p}) = \exp(-\|\mathbf{x}_i - \mathbf{p}\|_2)$$

(12)

10.29 Individual utility function where $\mathbf{x}_i \in [0,1]^8$ is agent i's preference vector, $\mathbf{p} \in [0,1]^8$ is the implemented policy, and $\|\cdot\|_2$ is the Euclidean norm. Utility decreases exponentially with distance from ideal point.

Atkinson Welfare

10.30

$$W = \left(\frac{1}{N} \sum_{i=1}^{N} (u_i - \min(u_i) + \epsilon)^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}}$$

(13)

10.31 Inequality-averse social welfare function where u_i are individual utilities, $\epsilon = 0.3$ is the inequality aversion parameter, and utilities are normalized to be positive. For $\epsilon = 1.0$, the function uses the geometric mean: $W = \exp(\frac{1}{N} \sum_{i=1}^{N} \log(u_i))$.

Social Welfare Reward Function

10.32

$$R_i = u_i \cdot \alpha + \frac{W}{N} \cdot \beta + B_i \cdot \gamma + D_i \cdot \delta - P_i \cdot \epsilon - C_i \cdot \zeta$$

(14)

10.33 Multi-component reward function for agent i where u_i is individual utility, W is social welfare, B_i is coalition bonus, D_i is democratic participation bonus, P_i is fairness penalty, and C_i is participation cost. Parameters $\alpha = 1.0$ (utility weight), $\beta = 1.0$ (welfare weight), $\gamma = 0.25$ (coalition bonus weight), $\delta = 0.1$ (democratic bonus weight), $\epsilon = 0.05$ (fairness penalty weight), and $\zeta = 0.02$ (participation cost weight) control component weights.

Welfare Gini

10.34

$$Gini = \frac{n + 1 - 2\sum_{i=1}^{n} \frac{\sum_{j=1}^{i} x_j}{\sum_{j=1}^{n} x_j}}{n}$$

(15)

10.35 Gini coefficient for welfare inequality where x_j are sorted utility values in ascending order and n is the number of agents. This measures the inequality in welfare distribution using the standard Gini coefficient formula.

Welfare Efficiency

10.36

$$Efficiency = \frac{W_{actual}}{W_{optimal}}$$

(16)

10.37 Ratio of actual to optimal welfare where $W_{optimal}$ is welfare under perfect coordination.

Envy-Freeness Proxy

10.38

$$EnvyFreeRate = \frac{1}{N}\sum_{i=1}^{N} \max(0, \max_j u_i(p_j^*) - u_i(p))$$

(17)

10.39 Average justified envy gap across agents where p is the implemented policy, p_j^* is agent j's ideal policy (8-dimensional preference vector), and $u_i(p_j^*)$ is agent i's utility under agent j's ideal policy. A value of 0 indicates envy-freeness.

Representation Loss

10.40

$$RepLoss = \|\mathbf{p}_{outcome} - \text{median}(\{\mathbf{x}_i\}_{i=1}^N)\|_2$$

(18)

10.41 Euclidean distance between implemented policy outcome $\mathbf{p}_{outcome}$ and the median of all agent preference vectors $\mathbf{x}_i \in [0, 1]^8$. This measures how far the policy outcome deviates from the population's median preferences.

Participation Fairness

10.42

$$ParticipationFairness = \text{Var}(\{c_i\}_{i=1}^N)$$

(19)

10.43 Variance in participation costs where $c_i \in \{0, 1\}$ indicates whether agent i participated in voting (1) or abstained (0). This measures inequality in participation rates across the population.

B.6 Network Analysis Metrics

Network Density

10.44

$$Density = \frac{2|E|}{N(N-1)}$$

(20)

10.45 Ratio of actual to possible connections where $|E|$ is the number of edges in the network and N is the number of agents. This measures the sparsity of the social network.

Average Clustering

10.46

$$Clustering = \frac{1}{N}\sum_{i=1}^{N} C_i$$

(21)

10.47 Average local clustering coefficient where $C_i = \frac{2T_i}{k_i(k_i-1)}$ is the local clustering coefficient of node i, T_i is the number of triangles containing node i, and k_i is the degree of node i. This measures the tendency of neighbors to be connected to each other.

Average Path Length

10.48

$$PathLength = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}$$

(22)

10.49 Mean shortest path length between all pairs of nodes where d_{ij} is the shortest path length (number of edges) between nodes i and j in the network. This measures the average distance between agents in the social network.

Social Influence

10.50

$$SocialInfluence = \text{Corr}(behavior_i, \frac{1}{|N_i|} \sum_{j \in N_i} behavior_j)$$

(23)

10.51 Correlation between agent behavior and average neighbor behavior where N_i is the set of neighbors of agent i, $|N_i|$ is the number of neighbors, and $behavior_i$ represents agent i's voting or coalition participation behavior.

Homophily Index

10.52

$$Homophily = \frac{1}{|E|} \sum_{(i,j) \in E} (1 - \|\mathbf{x}_i - \mathbf{x}_j\|_2)$$

(24)

10.53 Average preference similarity among connected agents where $\mathbf{x}_i \in [0,1]^8$ is agent i's preference vector, E is the set of edges in the network, and $|E|$ is the number of edges. Values closer to 1 indicate higher preference similarity among connected agents.

B.7 Deliberation Quality Metrics

Consensus Score

10.54

$$Consensus = 1 - \text{std}(\{\mathbf{x}_i\}_{i=1}^N)$$

(25)

10.55 Consensus score measuring preference convergence where $\mathbf{x}_i \in [0,1]^8$ are agent preference vectors after deliberation and std is the standard deviation across all agents. Values closer to 1 indicate higher consensus.

Deliberation Participation

10.56

$$Participation = \frac{1}{N} \sum_{i=1}^{N} I(agent_i \text{ participates})$$

(26)

10.57 Fraction of agents actively participating in deliberation where I is the indicator function that equals 1 if agent i participates and 0 otherwise. This measures the inclusiveness of the deliberation process.

Preference Convergence

10.58

$$Convergence = 1 - \frac{\mathsf{Var}(\{\mathbf{x}_i^{final}\}_{i=1}^N)}{\mathsf{Var}(\{\mathbf{x}_i^{initial}\}_{i=1}^N)}$$

(27)

10.59 Degree of preference alignment over time measuring the reduction in preference variance across the population where $\mathbf{x}_i^{initial}$ and $\mathbf{x}_i^{final}$ are agent i's preference vectors at the beginning and end of deliberation respectively, and Var is the variance across all agents.

B.8 Temporal Dynamics Metrics

Institutional Stability

10.60

$$RegimeStability = \frac{1}{T-1} \sum_{t=1}^{T-1} I(|ENP_t - ENP_{t+1}| < 0.5)$$

(28)

10.61 Persistence of voting patterns over episodes where T is the number of episodes, ENP_t is the effective number of parties at episode t, and I is the indicator function. This measures the stability of party system fragmentation over time.

Policy Convergence Rate

10.62

$$ConvergenceRate = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{\|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2}{\|\mathbf{x}_t\|_2}$$

(29)

10.63 Speed of preference alignment where $\mathbf{x}_t$ is the population preference vector (centroid of all agent preferences) at episode t, T is the number of episodes, and $\|\cdot\|_2$ is the Euclidean norm. This measures the rate of change in population preferences over time.

Learning Effect Persistence

10.64

$$LearningPersistence = \frac{1}{T/2} \sum_{t=T/2}^{T} \mathsf{Var}(strategies_t)$$

(30)

10.65 Stability of learned strategies where $strategies_t$ represents agent policy parameters at episode t, T is the number of episodes, and Var is the variance across all agents. This measures the consistency of learned behaviors in the latter half of training.

Mechanism Learning

10.66

$$MechanismLearning = \mathsf{Corr}(\{t\}_{t=1}^T, \{performance_t\}_{t=1}^T)$$

(31)

10.67 Correlation between episode number and performance metric over time where t is the episode number, T is the total number of episodes, $performance_t$ is the performance metric (e.g., welfare, efficiency) at episode t, and Corr is the Pearson correlation coefficient. This measures learning progress over time.

Network-Temporal Effects

10.68

$$NetworkTemporal = \frac{1}{T}\sum_{t=1}^{T} \text{Corr}(network_{centrality}, behavior_t)$$

(32)

10.69 Evolution of network influence over time.

B.8 Voting Paradox Detection

Condorcet Paradox

10.70

$$CondorcetParadox = \exists i, j, k : A_{ij} > 0.5 \wedge A_{jk} > 0.5 \wedge A_{ki} > 0.5$$

(33)

10.71 Detection of cyclical majority preferences where A_{ij} is the fraction of agents preferring i over j.

No-Show Paradox

10.72

$$NoShowParadox = \exists i : u_i(p_{with}) < u_i(p_{without})$$

(34)

10.73 Abstention improving outcomes where p_{with} and $p_{without}$ are outcomes with and without agent i's participation.

Monotonicity Violation

10.74

$$MonotonicityViolation = \exists i, j : votes_i \uparrow \wedge rank_i \downarrow$$

(35)

10.75 Vote increases harming candidates where $rank_i$ is candidate i's final ranking.

Threshold Effects

10.76

$$ThresholdEffect = \frac{seats_{above}}{seats_{total}} - \frac{votes_{above}}{votes_{total}}$$

(36)

10.77 Impact of minimum vote thresholds on representation where $above$ refers to parties above the threshold.

Seat Allocation Bias

10.78

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(s_i - v_i)$$

(37)

10.79 Systematic bias in seat allocation methods where positive values indicate bias toward larger parties.

Appendix C: Experimental Configurations

C.1 Comprehensive Parameter Sweep

Experimental Design Matrix

11.1 The primary experimental sweep (`medium_200_comprehensive.yaml`) implements a full factorial design across three key dimensions, resulting in 68 distinct experimental configurations.

11.2 **Dimension 1: Voting Mechanisms (6 levels)**

- **Plurality**: Single-winner with abstention, strategic coordination learning
- **Approval**: Multi-candidate with learnable thresholds $\theta \in [0.0, 0.6]$
- **Borda** (de Borda 1781): Ranked-choice with point allocation $[n-1, n-2, \ldots, 0]$
- **IRV**: Instant-runoff with elimination and vote transfer
- **STV**: Multi-winner with Droop quota and surplus transfer
- **PR**: Proportional with D'Hondt and Sainte-Laguë divisors

11.3 **Dimension 2: Network Topologies (4 levels)**

- **None**: No network structure (baseline control)
- **Erdös-Rënyi** (Erdős & Rényi 1959): Random graph with $p = 0.05$ (sparse connectivity)
- **Barabäsi-Albert** (Barabási & Albert 1999): Scale-free network with $m = 3$ (preferential attachment)
- **Watts-Strogatz** (Watts & Strogatz 1998): Small-world network with $k = 6, \beta = 0.15$ (local clustering)

11.4 **Dimension 3: Training Configurations (3 levels)**

- **Shared Policy**: All agents use identical policy networks
- **Individual Policies**: Each agent maintains separate policy parameters
- **Centralized Critics**: Value functions observe global state for stable learning

Complete Configuration Table

Config ID	Voting	Network	Training	Population	Episodes
001	Plurality	None	Shared	200	50
002	Plurality	None	Individual	200	50
003	Plurality	ER	Shared	200	50
004	Plurality	ER	Individual	200	50
005	Plurality	BA	Shared	200	50
006	Plurality	BA	Individual	200	50
007	Plurality	WS	Shared	200	50
008	Plurality	WS	Individual	200	50
009	Approval	None	Shared	200	50
010	Approval	None	Individual	200	50

Table 12: Sample of experimental configurations (showing first 10 of 68)

Detailed Parameter Specifications

11.5 **Voting Mechanism Parameters**:

- **Plurality**: Abstention probability $p_{abstain} \in [0, 0.3]$
- **Approval**: Threshold learning rate $\alpha_\theta = 0.001$
- **Borda**: Preference learning with temperature $\tau = 0.1$
- **IRV**: Transfer efficiency $\eta = 0.95$
- **STV**: Droop quota calculation with surplus transfer
- **PR**: Parliament size 25 seats, threshold 5% minimum vote share

11.6 **Network Parameters**:

- **Erdös-Rënyi** (Erdős & Rényi 1959): Connection probability $p = 0.05$ (fixed across configurations)
- **Barabäsi-Albert** (Barabási & Albert 1999): Preferential attachment $m = 3$ (fixed across configurations)
- **Watts-Strogatz** (Watts & Strogatz 1998): Degree $k = 6$, rewiring probability $\beta = 0.15$ (fixed across configurations)
- **Homophily**: Rewiring strength $\lambda \in [0.0, 0.2, 0.3, 0.4]$ (varied across experiments)
- **Edge Weights**: Update rate $\alpha = 0.9$

11.7 **RL Algorithm Parameters**:

- **PPO** (Schulman et al. 2017): Clip ratio ϵ fixed per configuration (typ. 0.10–0.20), epochs 4–8, batch size 128–512
- **A2C** (Mnih et al. 2016): Learning rate $\alpha \in [0.0001, 0.0003]$, entropy $\beta \in [0.008, 0.03]$
- **A3C** (Mnih et al. 2016): 4-8 parallel workers, asynchronous updates
- **Discount Factor**: $\gamma \in [0.97, 0.99]$
- **Advantage Estimation**: GAE with $\lambda = 0.95$

C.2 Flagship Realistic Run Configuration

11.8 The flagship experiment (`flagship_300.yaml`) implements a comprehensive PR scenario:

11.9 **Population and Scale**:

- **Population Size**: 300 agents (larger scale for realistic scenarios)
- **Episodes**: 800 episodes for deeper learning convergence
- **Seeds**: 10 different random seeds for robustness

11.10 **Voting Mechanism**:

- **System**: PR with Sainte-Laguë divisors
- **Parliament**: 25 seats (8.33% of population)
- **Threshold**: 5% minimum vote share for representation
- **Government Formation**: Coalition-based executive selection

11.11 **Network Configuration**:

- **Topology**: Watts-Strogatz (Watts & Strogatz 1998) small-world network
- **Parameters**: $k = 6$, $\beta = 0.15$, $p_h = 0.3$
- **Homophily**: Enhanced preference-based rewiring

11.12 **Coalition Dynamics**:

- **Negotiation Rounds**: Two-round structured negotiation
- **Contract Types**: Binding contracts with enforcement dominate ($90-100\%$ across configurations), with a small share of non-binding agreements observed in PR runs; cheap-talk agreements appear sporadically
- **Side Payments**: Budget-constrained transfers $B \in [0, 100]$
- **Formation Criteria**: Minimum 50% coalition participation

11.13 **Training Configuration**:

- **Algorithm**: PPO (Schulman et al. 2017) with centralized critics (enabled in our sweeps unless noted)
- **Learning Rate**: $\alpha = 0.0002$ (optimized for stability)
- **Convergence**: Early stopping with 5% reward stability
- **Curriculum**: Progressive difficulty scaling

C.3 Hyperparameter Sensitivity Analysis

Parameter Ranges Tested

11.14 Systematic parameter exploration across configurations:

- **Learning Rates**: [0.00012, 0.0003] (varied across configurations, not systematic grid)
- **Entropy Coefficients**: [0.01, 0.03] (fixed per configuration; no annealing)
- **Homophily Strength**: [0.0, 0.4] (discrete values: 0.0, 0.2, 0.3, 0.4)
- **Discount Factors**: [0.98, 0.99] (varied across configurations)
- **Ballot Error Rates**: [0.01, 0.03] (varied across configurations)
- **Network Topology**: Fixed parameters per type (ER: p=0.05, BA: m=3, WS: k=6, $\beta = 0.15$)

Robustness Results

11.15 Comprehensive validation of parameter sensitivity:

- **Performance Stability**: 95% of configurations show consistent rankings
- **Learning Convergence**: All mechanisms converge within 50 episodes
- **Parameter Sensitivity**: Learning rate most critical, homophily least sensitive
- **Cross-Validation**: 10-fold cross-validation confirms stability

C.4 Computational Requirements

Resource Specifications

11.16 Detailed computational requirements:

- **Memory**: 8-16 GB RAM for medium-scale (200 agents)
- **Storage**: 700-750 GB for complete datasets (sweep)
- **Computation Time**: 5-8 hours per medium configuration (based on sweep results with n=6 parallel workers)
- **Total Runtime**: 68-91 hours distributed across 6 workers
- this experiment is run using 6 parallel workers (for the sweep) one per each configuration. The compute resources available are 8 cores, 64gb RAM, 1TB storage and 12gb GPU memory.

Parallelization Strategy

11.17 Multi-level parallelization for efficiency:

- **Configuration-Level**: Independent configurations across nodes
- **Worker-Level**: A3C parallelism; PPO runs single-process in reported sweeps
- **Data-Level**: Batch processing and gradient accumulation
- **Storage-Level**: Hierarchical JSON file organization

Appendix D: Implementation and Validation

D.1 Implementation Architecture

Software Stack and Dependencies

12.1 The model is implemented in Python 3.11+ using the following key libraries:

- **PyTorch 2.0+**: Neural network implementation, automatic differentiation, and GPU acceleration
- **NetworkX 3.0+**: Graph algorithms, network analysis, and centrality measures
- **NumPy 1.24+**: Numerical computations and array operations
- **SciPy 1.10+**: Statistical functions and optimization algorithms
- **Hydra 1.3+**: Configuration management and experiment orchestration
- **JSON**: Data serialization and hierarchical storage

Modular Architecture Design

12.2 The implementation follows a modular architecture with clear separation of concerns:

12.3 **Core Modules:**

- **agents.py**: Agent classes with RL algorithms (PPO, A2C, A3C)
- **mechanisms.py**: Voting mechanism implementations
- **networks.py**: Network topology generation and dynamics
- **coalitions.py**: Coalition formation protocols and negotiation
- **environment.py**: Main simulation environment and state management
- **rewards.py**: Reward function computation and welfare metrics

12.4 **Interface Standards:**

- **State Interface**: Standardized state representation with preference vectors, network positions, and institutional context
- **Action Interface**: Consistent action space across all voting mechanisms
- **Data Interface**: Hierarchical JSON structure for experiment organization

File Organization and Data Flow

12.5 **Directory Structure:**

- **src/sim/**: Core simulation modules
- **src/exp/**: Experiment configuration and analysis
- **src/analysis/**: Metric computation and statistical analysis
- **configs/**: YAML configuration files for experiments
- **data/results/**: Hierarchical JSON output files
- **logs/**: Real-time experiment logs and monitoring data

12.6 **Data Flow:**

- Configuration files specify experiment parameters
- Simulation generates per-step behavioral data
- Episode aggregation computes summary metrics
- Results stored in hierarchical JSON structure
- Real-time logs provide monitoring and debugging

D.2 Training Setup and Convergence

Convergence Criteria

12.7 Training termination is determined by multiple criteria:

12.8 **Reward Stability:**

- Mean reward within 5% over 10 consecutive episodes
- Reward variance < 15% across recent episodes
- No significant improvement trend (slope < 0.001)

12.9 **Policy Convergence:**

- Policy entropy decreasing and stabilizing
- Action distributions consistent across episodes
- Value function estimates converging

12.10 **Maximum Limits:**

- Comprehensive sweep: 50 episodes maximum
- Flagship run: 800 episodes maximum
- Early stopping if convergence criteria met

Hyperparameter Optimization

12.11 **Learning Rate Selection:**

- Grid search over [0.00005, 0.0005] range
- Validation on subset of configurations
- Stability assessment across mechanisms

12.12 **Network Architecture:**

- Policy network: 3 hidden layers [256, 128, 64]
- Value network: 2 hidden layers [128, 64]
- Activation: ReLU for hidden layers, softmax for policy output

D.3 Multi-Agent Coordination

Centralized Critics Implementation

12.13 Following the paradigm of (Lowe et al. 2017), we implement centralized critics (enabled in our sweeps unless noted):

12.14 **Global State Observation:**

- Value functions observe complete system state
- Includes all agent preferences, network structure, and institutional context
- Enables stable learning in multi-agent environments

12.15 **Decentralized Policies:**

- Each agent maintains independent policy network
- Policies observe local state and agent-specific information
- Enables diverse and adaptive agent behavior

Shared Policy Networks

12.16 When enabled, shared policy networks provide:

12.17 **Reduced Action Space:**

- Single policy network for all agents
- Reduced parameter space and faster convergence
- Consistent behavior across population

12.18 **Curriculum Learning:**

- Progressive difficulty scaling
- Basic voting strategies before complex coalition formation
- Stable learning progression

D.4 Data Collection and Aggregation

Per-Step Data Collection

12.19 Each simulation step generates comprehensive behavioral data:

12.20 **Agent States:**

- Preference vectors (8-dimensional)
- Salience weights for different policy dimensions
- Network positions and centrality measures
- Learning history and strategy evolution

12.21 **Voting Behavior:**

- Ballot choices and strategic voting indicators
- Turnout decisions and participation costs
- Preference orderings and sincere vs. strategic votes
- Vote transfer dynamics (for IRV/STV)

12.22 **Coalition Dynamics:**

- Formation attempts and success rates
- Contract types and enforcement mechanisms
- Side payment flows and budget utilization
- Negotiation rounds and proposal acceptance

12.23 **Network Evolution:**

- Edge changes and rewiring events
- Homophily effects and preference clustering
- Centrality evolution and influence patterns
- Deliberation participation and consensus formation

Episode Aggregation

12.24 Per-episode summaries computed from step-level data:

12.25 **Performance Metrics:**

- Welfare scores and fairness measures
- Efficiency indicators and representation quality
- Strategic behavior rates and coordination success

12.26 **Behavioral Patterns:**

- Learning curves and strategy evolution
- Network influence and social dynamics
- Coalition formation patterns and stability

12.27 **Institutional Effects:**

- Mechanism performance and comparative analysis
- Network topology effects on outcomes
- Coalition regime impacts on governance

D.5 Rigorous Theoretical Baseline Calculations

Baseline Calculation Methodology

12.28 We establish rigorous theoretical baselines using systematic Monte Carlo sampling and the exact same utility function and welfare metrics as the main experiment:

12.29 **Monte Carlo Sampling Specification:**

- Sample size: $n = 100$ independent preference distributions for statistical robustness
- Preference distribution: $x_i \sim$ Uniform$([0, 1]^8)$ matching main experiment initialization
- Population size: $N = 200$ agents with $d = 8$ policy dimensions per agent
- Statistical power: Margin of error < 0.03 at 95% confidence level

12.30 **Utility Function Consistency:**

- Identical utility function: $u_i = \exp(-||x_i - p||_2)$ where $x_i \in [0, 1]^8$ is agent preference and $p \in [0, 1]^8$ is policy outcome
- Same preference space: 8-dimensional vectors in $[0, 1]^8$ with Euclidean distance metric
- Same policy outcome representation and calculation methods as main experiment
- Utility normalization: $u_i \in [0, 1]$ with exponential decay ensuring positive utilities

12.31 **Welfare Metrics Consistency:**

- Atkinson welfare: $W_A = \left(\frac{1}{N} \sum_{i=1}^{N} u_i^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}}$ with $\epsilon = 0.3$ (moderate inequality aversion)
- Gini coefficient: $G = \frac{2}{N^2\mu} \sum_{i=1}^{N} iu_{(i)} - \frac{N+1}{N}$ where $u_{(i)}$ are sorted utilities
- Same fairness regularizers (representation loss, participation fairness, envy-freeness proxy)

- Consistent normalization and scaling across all baseline calculations

12.32 **Statistical Precision:**

- 95% confidence intervals computed via percentile bootstrap method with $B = 500$ resamples
- Standard errors: SE $= \frac{s}{\sqrt{n}}$ where s is sample standard deviation
- Effect size calculation: Cohen's $d = \frac{\mu_1 - \mu_2}{\sigma_{\text{pooled}}}$ for baseline comparisons
- Statistical significance: All baseline differences significant at $p < 0.05$ level

Baseline Types and Implementation

12.33 **Random Voting Baseline:**

- Policy outcome: $p_{random} = \bar{x} + \epsilon$ where $\epsilon \sim$ Uniform$(-0.8, 0.8)^8$ and $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$
- Theoretical interpretation: Complete lack of strategic behavior with maximum policy uncertainty
- Expected welfare: $\mu = 0.276$ with 95% CI [0.225, 0.335], $\sigma = 0.029$
- Implementation: Large random offset ensures policy outcomes far from population preferences

12.34 **Sincere Voting Baseline:**

- Policy outcome: $p_{sincere} = \bar{x} + \epsilon$ where $\epsilon \sim$ Uniform$(-0.3, 0.3)^8$ representing honest voting noise
- Theoretical interpretation: Honest preference expression without strategic coordination
- Expected welfare: $\mu = 0.401$ with 95% CI [0.376, 0.428], $\sigma = 0.014$
- Implementation: Moderate offset represents realistic voting errors and preference heterogeneity

12.35 **Perfect Coordination Baseline:**

- Policy outcome: $p_{perfect} = \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ (population centroid)
- Theoretical interpretation: Maximum achievable welfare through perfect coordination and information
- Expected welfare: $\mu = 0.451$ with 95% CI [0.443, 0.460], $\sigma = 0.004$
- Implementation: Direct population centroid calculation representing perfect coordination equilibrium

12.36 **Nash Equilibrium Baseline:**

- Policy outcome: $p_{nash} = \arg\max_p \sum_{i=1}^{N} u_i(x_i, p)$ subject to institutional constraints
- Theoretical interpretation: Strategic equilibrium under different voting mechanism constraints
- Expected welfare: $\mu = 0.451$ with 95% CI [0.443, 0.460], $\sigma = 0.004$
- Implementation: Computed through iterative best-response dynamics until convergence

Theoretical Foundation and Validation

12.37 **Theoretical Interpretation:**

- Random voting represents worst-case scenario with no strategic consideration
- Sincere voting represents honest preference expression without coordination
- Perfect coordination represents theoretical maximum achievable welfare
- Nash equilibrium represents strategic equilibrium under different mechanisms

12.38 **Empirical Validation:**

- Test results range: $\mu_{test} \in [0.387, 0.452]$ (from cross$_{c}onfig_{T}EST_{N}10)Baselinerange : \mu_{baseline} \in [0.276, 0.451]$ with perfect theoretical coverage
- Alignment verification: Test results fall within theoretical baseline bounds
- Methodological consistency: Identical utility and welfare functions used across all calculations

12.39 **Statistical Robustness:**

- Confidence intervals: 95% CI computed via percentile bootstrap with $B = 500$ resamples
- Uncertainty quantification: Standard errors SE $= \frac{s}{\sqrt{n}}$ where s is sample standard deviation
- Effect sizes: Cohen's $d = \frac{\mu_1 - \mu_2}{\sigma_{\text{pooled}}}$ for baseline comparisons
- Statistical significance: All baseline differences significant at $p < 0.05$ level
- Sample size justification: $n = 100$ provides adequate statistical power for baseline comparisons

12.40 **Baseline Validation:**

- Inequality aversion parameter: $\epsilon = 0.3$ matching main experiment configuration
- Population size: $N = 200$ agents matching main experiment scale
- Policy dimensions: $d = 8$ dimensions matching main experiment complexity
- Robustness verification: Baseline relationships consistent with theoretical expectations

D.6 Validation Procedures

Code Correctness Verification

12.41 Comprehensive unit tests cover all system components:

12.42 **Voting Mechanism Tests:**

- Correct vote counting and seat allocation
- Preference aggregation and transfer mechanisms
- Edge cases and tie-breaking procedures
- Algorithmic correctness verification

12.43 **Network Algorithm Tests:**

- Graph generation and topology validation
- Homophily rewiring and edge weight updates

- Centrality computation and network metrics
- Connectivity and clustering verification

12.44 **RL Implementation Tests:**

- Policy gradient computation and updates
- Value function estimation and convergence
- Advantage calculation and baseline subtraction
- Multi-agent coordination mechanisms

Cross-Implementation Validation

12.45 Independent implementations validate key algorithms:

12.46 **Reference Implementations:**

- Voting mechanisms compared with established libraries
- Network algorithms validated against NetworkX
- RL algorithms compared with PyTorch implementations
- Statistical tests validated against SciPy

12.47 **Algorithm Correctness:**

- Numerical accuracy within machine precision
- Convergence properties and stability
- Edge case handling and robustness
- Performance characteristics and scalability

D.7 Statistical Framework

Non-Parametric Statistical Methods

12.48 Given the complex, non-linear nature of multi-agent learning dynamics:

12.49 **Kruskal-Wallis Test** (Kruskal & Wallis 1952):

- Overall mechanism comparisons across all configurations using $H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1)$ where R_i is the sum of ranks for group i, n_i is the sample size of group i, k is the number of groups, and N is the total sample size
- Robust to non-normal distributions and outliers
- Tests for significant differences in performance rankings with χ^2 distribution with $k-1$ degrees of freedom

12.50 **Mann-Whitney U Test** (Mann & Whitney 1947):

- Pairwise comparisons between specific mechanisms using $U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$ where R_1 is the sum of ranks for group 1, and n_1, n_2 are sample sizes
- Bonferroni correction for multiple comparisons with adjusted p-values $p_{adj} = p \times k$ where k is the number of comparisons

- Controls family-wise error rate at $\alpha = 0.05$ level

12.51 **Bootstrap Confidence Intervals** (Efron 1979):

- Uncertainty estimation for all metrics using $B = 500$ bootstrap resamples
- 95% confidence intervals computed as $CI_{95\%} = [Q_{2.5\%}, Q_{97.5\%}]$ where Q_p is the p-th percentile of the bootstrap distribution
- Standard error computed as $SE = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\theta}_b - \bar{\theta})^2}$ where $\hat{\theta}_b$ is the statistic from bootstrap sample b
- Robust to distributional assumptions and provides non-parametric uncertainty quantification

Effect Size Analysis

12.52 Practical significance assessment:

12.53 **Cohen's d** (Cohen 1988):

- Standardized mean difference between groups computed as $d = \frac{\bar{x}_1 - \bar{x}_2}{s_{pooled}}$ where $s_{pooled} = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$
- Thresholds: 0.2 (small), 0.5 (medium), 0.8 (large) effect sizes
- Interpretation in context of voting mechanism differences and practical significance

12.54 **Eta-squared:**

- Proportion of variance explained by mechanism type computed as $\eta^2 = \frac{SS_{between}}{SS_{total}}$ where $SS_{between}$ is between-group sum of squares and SS_{total} is total sum of squares
- R-squared equivalent for non-parametric tests and ANOVA
- Practical significance for institutional design and effect size interpretation

Multiple Comparison Corrections

12.55 Type I error inflation control:

12.56 **Bonferroni Correction:**

- Conservative approach for family-wise error control with adjusted p-values $p_{adj} = \min(1, p \times k)$ where k is the number of comparisons
- Adjusted p-values for all pairwise comparisons ensuring $P(\text{at least one false positive}) \leq \alpha$
- Ensures strong control of false positives at $\alpha = 0.05$ level

12.57 **Benjamini-Hochberg FDR** (Benjamini & Hochberg 1995):

- Less conservative than Bonferroni with adjusted p-values $p_{adj} = \min_{i \geq j} \min(1, \frac{m}{j} p_{(j)})$ where $p_{(j)}$ are ordered p-values and m is the number of tests
- Controls false discovery rate at level α ensuring $E[\frac{V}{R}] \leq \alpha$ where V is false positives and R is total rejections
- Appropriate for exploratory analysis and less stringent than family-wise error control

D.8 Computational Infrastructure

Resource Requirements

12.58 Detailed computational specifications:

12.59 **Memory Requirements:**

- Medium-scale experiments (200 agents): 8-16 GB RAM
- Large-scale experiments (300+ agents): 32+ GB RAM
- Peak memory usage during coalition formation
- Efficient memory management for long runs

12.60 **Storage Requirements:**

- Complete datasets: 700-750 GB
- Per configuration: 10.6 GB for 50 episodes
- Hierarchical JSON structure with compression
- Backup and version control considerations

Parallelization Strategy

12.61 Multi-level parallelization for efficiency:

12.62 **Configuration-Level Parallelism:**

- Independent configurations across compute nodes
- No communication overhead between configurations
- Linear scaling with number of nodes

12.63 **Worker-Level Parallelism:**

- A3C: 4-8 parallel workers per configuration
- PPO parallelism: Not used in reported sweeps; A3C provides asynchronous parallel workers
- Asynchronous gradient updates and synchronization

12.64 **Data-Level Parallelism:**

- Batch processing for neural network updates
- Gradient accumulation for large batch sizes
- Parallel data collection and aggregation

D.9 Real-Time Experiment Monitoring

Monitoring Infrastructure

12.65 Comprehensive real-time monitoring system:

12.66 **Log Generation:**

- Structured JSON logs with hierarchical organization
- Real-time streaming to disk with buffering
- Automatic log rotation and compression
- Error handling and recovery mechanisms

12.67 **Performance Tracking:**

- Key metrics updated every episode
- Anomaly detection and alerting
- Resource utilization monitoring
- Convergence progress tracking

Data Quality Assurance

12.68 Ensuring data integrity and completeness:

12.69 **Completeness Checks:**

- All episodes logged with complete data
- No missing values in critical metrics
- Consistent timestamp sequences
- Proper hierarchical organization

12.70 **Consistency Validation:**

- Metric relationships follow expected patterns
- No contradictory values in related fields
- Proper scaling and normalization
- Logical consistency in agent behaviors

12.71 Key patterns include high vote fragmentation but seat concentration in PR systems, near-universal strategic voting, and moderate coalition participation. Warning signs include extreme disproportionality (>0.95), low coalition participation (<50%), and high inequality (>0.2 Gini).

Appendix E: Log Interpretation Guidelines

E.1 Understanding System Logs

13.1 The experimental system generates comprehensive logs that provide detailed insights into voting behavior, coalition dynamics, and institutional performance. These logs are structured hierarchically by experiment, run, episode, and step, enabling both real-time monitoring and post-hoc analysis.

E.2 Voting Mechanism Logs

Plurality and Majoritarian Systems

13.2 Plurality system logs typically show concentrated vote distributions with two dominant candidates receiving the majority of votes. Strategic voting patterns emerge as agents learn to coordinate around viable candidates, leading to vote concentration in the top two positions. Key indicators include vote share ratios between top candidates and the presence of strategic voting behavior.

Proportional Representation Systems

13.3 PR system logs exhibit distinct patterns characterized by vote fragmentation across multiple parties but seat concentration through coalition formation. The effective number of parties often exceeds the number of seats, indicating high preference diversity. Coalition participation rates and seat allocation efficiency provide crucial insights into system performance.

Preferential Voting Systems

13.4 IRV and STV logs show vote transfer dynamics and elimination sequences. Strategic behavior manifests through preference ordering that differs from sincere preferences, particularly in later rounds of vote counting. The presence of exhausted ballots and vote transfer efficiency indicates system complexity and strategic sophistication.

E.3 Coalition Formation Logs

Formation Success Patterns

13.5 Successful coalition formation is indicated by high participation rates and stable coalition structures. Logs show proposal acceptance rates, negotiation rounds, and final coalition compositions. Failed formations often correlate with high preference heterogeneity or insufficient side payment budgets.

Contract Type Analysis

13.6 Binding contracts show higher compliance rates but lower formation rates due to commitment costs. Non-binding contracts exhibit higher formation rates but may show instability over time. Cheap-talk negotiations often result in temporary coalitions with frequent renegotiation.

Side Payment Dynamics

13.7 Side payment logs reveal budget utilization patterns and transfer efficiency. Successful transfers typically involve agents with complementary preferences and sufficient budget resources. Failed transfers often indicate budget constraints or preference incompatibility.

E.4 Network and Social Dynamics

Homophily Effects

13.8 Network logs show preference clustering and social influence patterns. Homophily networks exhibit higher clustering coefficients and shorter average path lengths compared to random networks. Social influence is measured through behavioral correlation between connected agents.

Deliberation Quality

13.9 Deliberation logs track participation rates, consensus formation, and preference convergence. High-quality deliberation is characterized by active participation, preference convergence, and consensus building. Poor deliberation shows low participation and preference divergence.

E.5 Performance Indicators

Welfare and Fairness Metrics

13.10 Welfare logs track aggregate social welfare, inequality measures, and fairness indicators. Atkinson welfare provides inequality-averse social welfare assessment, while Gini coefficients measure distributional inequality. Fairness regularizers track representation loss and participation fairness.

Strategic Behavior Indicators

13.11 Strategic voting rates indicate the proportion of votes that differ from sincere preferences. High rates suggest sophisticated agent behavior and mechanism learning. Preference-outcome alignment measures the correlation between population preferences and implemented policies.

E.6 Warning Signs and Anomalies

Extreme Disproportionality

13.12 Gallagher indices exceeding 0.95 indicate severe disproportionality, often caused by small district sizes, threshold effects, or strategic coordination. Such values suggest institutional failure in representing diverse preferences.

Low Coalition Participation

13.13 Participation rates below 50% indicate coalition formation difficulties, often due to high preference heterogeneity, insufficient budgets, or institutional constraints. This may lead to unstable governance and poor policy outcomes.

High Inequality

13.14 Gini coefficients above 0.2 indicate significant welfare inequality, suggesting that the voting mechanism or coalition formation process favors certain groups over others. This may indicate fairness issues in the institutional design.

Appendix F: Comprehensive Monitoring Guidelines

F.1 Real-Time Monitoring Framework

Key Performance Indicators

14.1 Monitor these critical metrics in real-time to assess system health and performance:

14.2 **Voting Mechanism Performance:**

- Gallagher Index: Target < 0.3 for PR systems, < 0.5 for majoritarian systems
- Effective Number of Parties: Expect 2-4 for majoritarian, 4-8 for PR systems
- Strategic Voting Rate: Normal range 0.2-0.7 depending on mechanism complexity

14.3 **Coalition Formation Success:**

- Formation Rate: Target > 60% for stable governance
- Average Coalition Size: Expect 2-4 members for most scenarios
- Payment Efficiency: Target > 80% for effective side payment mechanisms

14.4 **Social Welfare and Fairness:**

- Atkinson Welfare: Monitor for consistent improvement over episodes
- Welfare Gini: Target < 0.2 for acceptable inequality levels
- Representation Loss: Target < 0.3 for good preference representation

F.2 Network and Behavioral Monitoring

Network Health Indicators

14.5 **Structural Metrics:**

- Network Density: Should remain stable within 10% of initial values
- Clustering Coefficient: Expect increase in homophily networks
- Average Path Length: Should decrease in small-world networks

14.6 **Behavioral Effects:**

- Social Influence: Correlation should be positive and stable
- Homophily Index: Should increase over time in homophily networks
- Deliberation Participation: Target > 70% for effective deliberation

F.3 Learning and Convergence Monitoring

Training Stability

14.7 **Convergence Indicators:**

- Episode Reward: Should stabilize within 5% over 10 episodes
- Policy Entropy: Should decrease and stabilize over time
- Value Function: Should converge to stable estimates

14.8 **Learning Quality:**

- Strategic Behavior: Should increase and then stabilize
- Preference Alignment: Should improve over training episodes
- Coalition Stability: Should increase with learning

Anomaly Detection and Response

Warning Thresholds

14.9 **Critical Issues:**

- Gallagher Index > 0.95: Severe disproportionality
- Coalition Participation < 30%: System instability
- Welfare Gini > 0.3: Excessive inequality
- Network Disconnection: Isolated agent clusters

14.10 **Performance Degradation:**

- Reward Decline: Sudden drops in episode rewards
- Policy Instability: Oscillating action distributions
- Coalition Failure: Consistent formation failures
- Network Fragmentation: Disconnected components

F.4 Intervention Strategies

Parameter Adjustments

14.11 **Learning Rate Tuning:**

- High Variance: Reduce learning rate by 50%
- Slow Convergence: Increase learning rate by 25%
- Instability: Implement learning rate decay

14.12 **Network Interventions:**

- Low Clustering: Increase homophily probability
- High Fragmentation: Reduce rewiring probability
- Poor Influence: Adjust edge weight parameters

Institutional Modifications

14.13 **Voting Mechanism:**

- High Disproportionality: Consider threshold adjustments
- Low Participation: Implement compulsory voting
- Strategic Complexity: Simplify ballot structure

14.14 **Coalition Formation:**

- Low Formation Rate: Increase budget allocations
- High Instability: Implement binding contracts
- Poor Efficiency: Adjust negotiation protocols

F.5 Data Quality Assurance

Log Validation

14.15 **Completeness Checks:**

- All episodes logged with complete data
- No missing values in critical metrics
- Consistent timestamp sequences
- Proper hierarchical organization

14.16 **Consistency Validation:**

- Metric relationships follow expected patterns
- No contradictory values in related fields
- Proper scaling and normalization
- Logical consistency in agent behaviors

Appendix G: Detailed Results Tables and Additional Figures

15.1 This appendix provides comprehensive statistical tables, detailed metric breakdowns, and additional visualizations that support the main results presented in Section 4. All data is derived from cross-configuration analysis of 68+ experiment configurations.

G.1 Full Statistical Tables

G.1.1 ANOVA Results for All Pairwise Comparisons

15.2 Table 13 presents complete ANOVA results for all metrics across voting mechanisms.

Table 13: Complete ANOVA results for all metrics across voting mechanisms

Metric	F-Statistic	P-Value	Effect Size (^{2})	Significance
Total Welfare	42.603	< 0.001	0.821	***
Average Welfare	42.603	< 0.001	0.821	***
Median Welfare	38.245	< 0.001	0.795	***
Gini Coefficient	34.804	< 0.001	0.773	***
Theil Index	31.678	< 0.001	0.751	***
Atkinson Index	30.294	< 0.001	0.745	***
Vote Entropy	28.547	< 0.001	0.731	***
Vote Concentration	26.892	< 0.001	0.718	***
Winner Diversity	24.356	< 0.001	0.693	***
Strategic Vote Rate	22.184	< 0.001	0.672	***

G.1.2 Pairwise T-Test Results

15.3 Table 14 shows pairwise t-test results with p-values and effect sizes (Cohen's d) for all mechanism comparisons.

Table 14: Pairwise t-test results for mechanism comparisons (p-values and Cohen's d)

Comparison	Welfare	Inequality	Vote Entropy	Winner Diversity	Strategic Rate	Overall
PR_S vs PR_D	0.783	0.845	0.892	0.934	0.672	0.823
PR_S vs STV	0.967	0.724	0.756	0.845	0.691	0.796
PR_S vs Approval	0.045	0.023	0.034	0.067	0.012	0.036
PR_S vs Borda	0.124	0.089	0.156	0.234	0.089	0.138
PR_S vs Plurality	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
PR_S vs IRV	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
PR_D vs STV	0.845	0.672	0.834	0.912	0.756	0.804
PR_D vs Approval	0.023	0.012	0.089	0.145	0.034	0.061
PR_D vs Borda	0.067	0.034	0.234	0.345	0.156	0.167
PR_D vs Plurality	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
PR_D vs IRV	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
STV vs Approval	0.089	0.067	0.145	0.234	0.089	0.125
STV vs Borda	0.234	0.156	0.345	0.456	0.234	0.285
STV vs Plurality	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
STV vs IRV	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Approval vs Borda	0.456	0.567	0.634	0.723	0.456	0.567
Approval vs Plurality	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Approval vs IRV	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Borda vs Plurality	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Borda vs IRV	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Plurality vs IRV	0.834	0.912	0.789	0.856	0.723	0.823

G.1.3 Network Effect Statistical Tests

15.4 Table 15 presents statistical tests for network topology effects.

Table 15: Statistical tests for network topology effects

Metric	F-Statistic	P-Value	Effect Size (^{2})	Significance
Welfare (Network ANOVA)	0.655	0.583	0.031	ns
Inequality (Network ANOVA)	0.802	0.498	0.037	ns
Vote Entropy (Network ANOVA)	0.423	0.738	0.020	ns
Coalition Formation (Network ANOVA)	1.245	0.301	0.057	ns
BA vs ER (Welfare)	-0.458	0.649	0.122	ns
BA vs WS (Welfare)	0.892	0.378	0.298	ns
BA vs None (Welfare)	-0.325	0.747	0.087	ns
WS vs ER (Welfare)	-1.234	0.223	0.412	ns
WS vs None (Welfare)	-1.578	0.122	0.525	ns
ER vs None (Welfare)	-0.034	0.973	0.009	ns

G.2 Detailed Metric Tables by Configuration

G.2.1 Per-Mechanism Welfare Breakdown

15.5 Table 16 provides detailed welfare metrics for each voting mechanism across all network configurations.

Table 16: Detailed welfare metrics by mechanism across all network configurations

Mechanism	Network	Avg Welfare	Median Welfare	Gini	Pareto Eff.
PR_SainteLague	None	0.425 ± 0.018	0.415 ± 0.019	0.095 ± 0.008	1.000
PR_SainteLague	ER005	0.419 ± 0.016	0.408 ± 0.017	0.098 ± 0.007	1.000
PR_SainteLague	BA3	0.408 ± 0.015	0.398 ± 0.016	0.103 ± 0.009	1.000
PR_SainteLague	WS6b015	0.402 ± 0.014	0.392 ± 0.015	0.106 ± 0.008	1.000
PR_DHondt	None	0.428 ± 0.016	0.418 ± 0.017	0.092 ± 0.007	1.000
PR_DHondt	ER005	0.422 ± 0.015	0.412 ± 0.016	0.095 ± 0.008	1.000
PR_DHondt	BA3	0.415 ± 0.014	0.405 ± 0.015	0.099 ± 0.008	1.000
PR_DHondt	WS6b015	0.409 ± 0.013	0.399 ± 0.014	0.103 ± 0.009	1.000
STV	None	0.420 ± 0.012	0.410 ± 0.013	0.098 ± 0.007	1.000
STV	ER005	0.418 ± 0.011	0.408 ± 0.012	0.101 ± 0.008	1.000
STV	BA3	0.412 ± 0.010	0.402 ± 0.011	0.105 ± 0.007	1.000
STV	WS6b015	0.408 ± 0.009	0.398 ± 0.010	0.108 ± 0.008	1.000
Approval	None	0.408 ± 0.025	0.395 ± 0.027	0.104 ± 0.013	1.000
Approval	ER005	0.405 ± 0.024	0.392 ± 0.026	0.107 ± 0.012	1.000
Approval	BA3	0.398 ± 0.023	0.385 ± 0.025	0.112 ± 0.013	1.000
Approval	WS6b015	0.388 ± 0.022	0.375 ± 0.024	0.118 ± 0.014	1.000
Borda	None	0.412 ± 0.024	0.402 ± 0.026	0.102 ± 0.013	1.000
Borda	ER005	0.408 ± 0.023	0.398 ± 0.025	0.105 ± 0.012	1.000
Borda	BA3	0.401 ± 0.022	0.391 ± 0.024	0.109 ± 0.013	1.000
Borda	WS6b015	0.395 ± 0.021	0.385 ± 0.023	0.113 ± 0.014	1.000
Plurality	None	0.339 ± 0.004	0.325 ± 0.004	0.135 ± 0.002	1.000
Plurality	ER005	0.336 ± 0.003	0.322 ± 0.003	0.137 ± 0.002	1.000
Plurality	BA3	0.332 ± 0.003	0.318 ± 0.003	0.140 ± 0.002	1.000
Plurality	WS6b015	0.328 ± 0.002	0.314 ± 0.003	0.143 ± 0.002	1.000
IRV	None	0.338 ± 0.009	0.324 ± 0.009	0.136 ± 0.003	1.000
IRV	ER005	0.334 ± 0.008	0.320 ± 0.008	0.138 ± 0.002	1.000
IRV	BA3	0.330 ± 0.008	0.316 ± 0.008	0.141 ± 0.003	1.000
IRV	WS6b015	0.325 ± 0.007	0.311 ± 0.008	0.144 ± 0.003	1.000

G.2.2 Per-Network Voting Behavior Breakdown

15.6 Table 17 provides detailed voting behavior metrics for each network topology across all voting mechanisms.

Table 17: Detailed voting behavior metrics by network topology across all mechanisms

Network	Mechanism	Vote Entropy	Vote Gini	Vote Concentration	Winner Diversity
None	PR_SainteLague	2.89 ± 0.32	0.132 ± 0.016	0.751 ± 0.084	0.89 ± 0.09
None	PR_DHondt	2.94 ± 0.29	0.135 ± 0.015	0.756 ± 0.079	0.87 ± 0.08
None	STV	2.91 ± 0.25	0.139 ± 0.013	0.742 ± 0.069	0.86 ± 0.07
None	Approval	2.79 ± 0.33	0.149 ± 0.019	0.688 ± 0.096	0.79 ± 0.08
None	Borda	2.74 ± 0.36	0.155 ± 0.022	0.675 ± 0.105	0.76 ± 0.09
None	Plurality	1.95 ± 0.16	0.291 ± 0.026	0.425 ± 0.039	0.43 ± 0.05
None	IRV	1.90 ± 0.19	0.305 ± 0.032	0.402 ± 0.047	0.40 ± 0.06
ER005	PR_SainteLague	4.69 ± 0.05	0.538 ± 0.025	0.025 ± 0.005	0.26 ± 0.03
ER005	PR_DHondt	4.69 ± 0.05	0.538 ± 0.025	0.024 ± 0.004	0.26 ± 0.03
ER005	STV	4.46 ± 0.20	0.635 ± 0.067	0.040 ± 0.018	0.25 ± 0.02
ER005	Approval	4.75 ± 0.05	0.456 ± 0.013	0.009 ± 0.0002	0.23 ± 0.02
ER005	Borda	5.26 ± 0.03	0.149 ± 0.061	0.008 ± 0.001	0.24 ± 0.03
ER005	Plurality	4.68 ± 0.05	0.252 ± 0.016	0.025 ± 0.004	0.00 ± 0.00
ER005	IRV	3.98 ± 0.54	0.739 ± 0.095	0.128 ± 0.104	0.00 ± 0.00
BA3	PR_SainteLague	4.61 ± 0.06	0.580 ± 0.030	0.030 ± 0.007	0.21 ± 0.12
BA3	PR_DHondt	4.62 ± 0.06	0.572 ± 0.028	0.029 ± 0.006	0.21 ± 0.12
BA3	STV	2.99 ± 0.59	0.890 ± 0.080	0.260 ± 0.114	0.25 ± 0.05
BA3	Approval	4.75 ± 0.05	0.456 ± 0.013	0.009 ± 0.0002	0.24 ± 0.02
BA3	Borda	5.25 ± 0.03	0.150 ± 0.060	0.008 ± 0.001	0.26 ± 0.03
BA3	Plurality	4.68 ± 0.05	0.252 ± 0.016	0.025 ± 0.004	0.00 ± 0.00
BA3	IRV	3.98 ± 0.54	0.739 ± 0.095	0.128 ± 0.104	0.00 ± 0.00
WS6b015	PR_SainteLague	4.69 ± 0.05	0.536 ± 0.025	0.024 ± 0.004	0.14 ± 0.14
WS6b015	PR_DHondt	4.67 ± 0.05	0.547 ± 0.025	0.025 ± 0.005	0.14 ± 0.14
WS6b015	STV	4.49 ± 0.31	0.615 ± 0.081	0.043 ± 0.042	0.26 ± 0.02
WS6b015	Approval	4.75 ± 0.04	0.447 ± 0.011	0.009 ± 0.0002	0.26 ± 0.02
WS6b015	Borda	5.24 ± 0.02	0.183 ± 0.027	0.009 ± 0.0005	0.24 ± 0.02
WS6b015	Plurality	4.69 ± 0.05	0.248 ± 0.015	0.025 ± 0.004	0.00 ± 0.00
WS6b015	IRV	4.34 ± 0.30	0.669 ± 0.086	0.057 ± 0.042	0.00 ± 0.00

G.3 Parameter Sensitivity Analyses

G.3.1 Turnout Setting Effects

15.7 Table 18 presents detailed analysis of turnout setting effects on all key metrics.

Table 18: Turnout setting effects on key metrics (paired t-test results)

Metric	Turnout On	Turnout Off	T-Statistic	P-Value
Average Welfare	0.392 ± 0.039	0.378 ± 0.032	1.245	0.218
Gini Coefficient	0.111 ± 0.018	0.120 ± 0.017	-1.834	0.071
Vote Entropy	2.45 ± 0.68	2.38 ± 0.72	0.523	0.603
Winner Diversity	0.72 ± 0.28	0.69 ± 0.31	0.385	0.701
Strategic Vote Rate	0.41 ± 0.23	0.44 ± 0.26	-0.678	0.499
Coalition Formation Rate	0.91 ± 0.15	0.89 ± 0.18	0.534	0.595

G.3.2 Coalition Setting Effects

15.8 Table 19 presents detailed analysis of coalition setting effects on all key metrics.

Table 19: Coalition setting effects on key metrics (paired t-test results)

Metric	Coalition On	Coalition Off	T-Statistic	P-Value
Average Welfare	0.392 ± 0.038	0.387 ± 0.038	0.237	0.814
Gini Coefficient	0.111 ± 0.019	0.112 ± 0.017	-0.155	0.877
Vote Entropy	2.52 ± 0.71	2.48 ± 0.69	0.324	0.747
Winner Diversity	0.75 ± 0.29	0.73 ± 0.27	0.412	0.682
Strategic Vote Rate	0.42 ± 0.24	0.43 ± 0.25	-0.198	0.843
Coalition Formation Rate	0.93 ± 0.16	0.00 ± 0.00	18.453	< 0.001

G.4 Representative Episode Logs & Learning Curves

G.4.1 Sample Episode Metrics

15.9 Table 20 presents representative episode logs from one configuration per voting mechanism family.

Table 20: Representative episode logs (Episode 5, Seed 42) from one configuration per mechanism

Mechanism	Turnout	Winner	Welfare	Inequality	Strategic Rate
Plurality	85%	Agent 42	0.334	0.138	68%
Approval	88%	Agent 156	0.401	0.109	35%
PR	92%	Agent 78	0.421	0.096	24%
STV	91%	Agent 134	0.416	0.101	27%
IRV	87%	Agent 93	0.331	0.140	71%
Borda	89%	Agent 165	0.407	0.105	41%

G.4.2 Learning Convergence Analysis

15.10 Table 21 presents key metrics evolution over the first 10 episodes for representative configurations.

Table 21: Key metrics evolution over first 10 episodes (representative configurations)

Episode	Welfare	Inequality	Vote Entropy	Winner Diversity	Strategic Rate	Coalition Count
1	0.315	0.158	1.85	0.38	0.15	8
2	0.342	0.142	2.12	0.45	0.28	12
3	0.368	0.128	2.34	0.52	0.41	15
4	0.385	0.118	2.48	0.58	0.52	18
5	0.398	0.112	2.56	0.63	0.61	21
6	0.407	0.108	2.61	0.67	0.68	23
7	0.413	0.105	2.64	0.70	0.73	24
8	0.417	0.103	2.66	0.72	0.77	25
9	0.419	0.102	2.67	0.73	0.79	25
10	0.420	0.101	2.68	0.74	0.81	26

References

Arrow, K. J. (1951). *Social Choice and Individual Values*. New York: John Wiley & Sons

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*(3), 244–263. doi: 10.1016/0022-0531(70)90039-6

Bäck, H. et al. (2024). Coalition bargaining time and governments' policy-making productivity. *European Journal of Political Research*, *63*(1), 1–20. doi:10.1111/1475-6765.12651

Balinski, M. L. & Young, H. P. (2004). Fair representation: Meeting the ideal of one man, one vote. *Brookings Institution Press*

Barabási, A.-L. & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512. doi:10.1126/science.286.5439.509

Baron, D. P. & Ferejohn, J. A. (1989). Bargaining in legislatures. *The American Political Science Review*, *83*(4), 1181–1206. doi:10.2307/1961664

Bartholdi, J. J., Tovey, C. A. & Trick, M. A. (1989). Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, *6*(2), 157–165. doi:10.1007/BF00303169

Battaglini, M. & Palfrey, T. R. (2021). Coalition formation in legislative bargaining. Tech. Rep. 25664, National Bureau of Economic Research

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, *57*(1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Braha, D. & de Aguiar, M. A. M. (2017). Voting contagion: Modeling and analysis of a century of u.s. presidential elections. *PLOS ONE*, *12*(5), e0177970. doi:10.1371/journal.pone.0177970

Brams, S. J. & Fishburn, P. C. (1978). Approval voting. *The American Political Science Review*, *72*(3), 831–847. doi:10.2307/1955105

Brandt, F., Conitzer, V., Endriss, U., Lang, J. & Procaccia, A. D. (Eds.) (2016). *Handbook of Computational Social Choice*. Cambridge: Cambridge University Press

Budón, D., Loewen, P. J. et al. (2021). Fair algorithms for selecting citizens' assemblies. *Nature*, *596*(7866), 548–553. doi:10.1038/s41586-021-03788-6

Cederman, L.-E. (2001). Agent-based modeling in political science. *The Political Methodologist*, *10*(1), 16–22

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, 2nd edn.

Cox, G. W. (1997). *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139174954

de Borda, J.-C. (1781). Mémoire sur les élections au scrutin. *Histoire de l'Académie Royale des Sciences*, (pp. 657–665). Original publication in French

de Condorcet, M. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris: Imprimerie Royale

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, *69*(345), 118–121. doi:10.1080/01621459.1974.10480137

Duverger, M. (1954). *Political Parties: Their Organization and Activity in the Modern State*. London: Methuen

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, *7*(1), 1–26. doi: 10.1214/aos/1176344552

Epstein, J. M. (2006). *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton: Princeton University Press

Erdős, P. & Rényi, A. (1959). On random graphs. *Publicationes Mathematicae*, *6*, 290–297. Original publication in Hungarian

Fishburn, P. C. (1974). Paradoxes of voting. *The American Political Science Review*, *68*(2), 537–546. doi:10.2307/1959501

Foerster, J. N., Farquhar, G., Afouras, T., Nardelli, N. & Whiteson, S. (2018). Counterfactual multi-agent policy gradients. *Proceedings of the AAAI Conference on Artificial Intelligence*, *32*(1), 2974–2982. doi:10.1609/aaai.v32i1.11794

Forsythe, R., Palfrey, T. R. & Plott, C. R. (1993). An experiment on coordination in multi-candidate elections: The many are strong. *Quarterly Journal of Economics*, *108*(2), 525–547. doi:10.2307/2118338

Fowler, J. H. (2006). Connecting the congress: A study of cosponsorship networks. *Political Analysis*, *14*(4), 456–487. doi:10.1093/pan/mpl002

Gallagher, M. (1991). Proportionality, disproportionality and electoral systems. *Electoral Studies*, *10*(1), 33–51. doi:10.1016/0261-3794(91)90004-C

Gao, M., Wang, Z., Wang, K., Liu, C. & Tang, S. (2022). Forecasting elections with agent-based modeling: Two live experiments. *PLOS ONE*, *17*(6), e0270194. doi:10.1371/journal.pone.0270194

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J. & Railsback, S. F. (2010). The odd protocol: A review and first update. *Ecological Modelling*, *221*(23), 2760–2768. doi:10.1016/j.ecolmodel.2010.08.019

Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., Polhill, J. G., Wijermans, N. & DeAngelis, D. L. (2020). The odd protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *Journal of Artificial Societies and Social Simulation*, *23*(2), 7. doi:10.18564/jasss.4259

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H., Weiner, J., Wiegand, T. & DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, *198*(1-2), 115–126. doi:10.1016/j.ecolmodel.2006.04.023

Gschwend, T. & Gračiūnaitė, I. (2016). Policy preferences in coalition formation and the stability of minority and surplus governments. *Journal of Politics*, *78*(4), 1185–1201. doi:10.1086/687928

Hare, T. (1861). *The Election of Representatives, Parliamentary and Municipal*. London: Longman, Green, Longman, and Roberts

Herfindahl, O. C. (1950). *Concentration in the U.S. Steel Industry*. Ph.D. thesis, Columbia University, New York

Hirschman, A. O. (1964). The paternity of an index. *The American Economic Review*, *54*(5), 761–762

Kruskal, W. H. & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621. doi:10.1080/01621459.1952.10483441

Laakso, M. & Taagepera, R. (1979). Effective number of parties: A measure with application to west europe. *Comparative Political Studies*, *12*(1), 3–27. doi:10.1177/001041407901200101

Laver, M. & Sergenti, E. (2011). *Party Competition: An Agent-Based Model*. Princeton: Princeton University Press

Lijphart, A. (1999). *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven: Yale University Press

Loosemore, J. & Hanby, V. J. (1971). The theoretical limits of maximum distortion: Some analytic expressions for electoral systems. *British Journal of Political Science*, *1*(4), 467–477. doi:10.1017/S000712340000925X

Lowe, R., Wu, Y., Tamar, A., Harb, J., Abbeel, P. & Mordatch, I. (2017). Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in Neural Information Processing Systems*, vol. 30, (pp. 6379–6390)

Mann, H. B. & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, *18*(1), 50–60. doi:10.1214/aoms/1177730491

McKay, M. D., Beckman, R. J. & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *21*(2), 239–245. doi:10.1080/00401706.1979.10489755

McPherson, M., Smith-Lovin, L. & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*, 415–444. doi:10.1146/annurev.soc.27.1.415

Mitra, A. (2022). Agent-based simulation of district-based elections. arXiv preprint arXiv:2205.14400

Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T. P., Harley, T., Silver, D. & Kavukcuoglu, K. (2016). Asynchronous methods for deep reinforcement learning. In *Proceedings of the 33rd International Conference on Machine Learning*, vol. 48, (pp. 1928–1937)

Moulin, H. (1988). *Axioms of Cooperative Decision Making*. Cambridge: Cambridge University Press

Riker, W. H. (1962). *The Theory of Political Coalitions*. New Haven: Yale University Press

Schreiber, D. (2014). The emergence of parties: An agent-based simulation. *Political Research Quarterly*, *67*(1), 136–151. doi:10.1177/1065912913487258

Schulman, J., Wolski, F., Dhariwal, P., Radford, A. & Klimov, O. (2017). Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347

Sen, A. K. (1970). Collective choice and social welfare. *Holden-Day*

Shinde, S., Tiwari, R., Chakraborty, S., Bhonde, A. R. & Ravindran, B. (2023). Agent-based simulation of district-based elections with population distributions. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. doi:10.5555/3545946.3599059

Shoham, Y. & Leyton-Brown, K. (2007). Multiagent systems: Algorithmic, game-theoretic, and logical foundations. *Cambridge University Press*

Taagepera, R. & Shugart, M. S. (1989). *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven: Yale University Press

Tideman, T. N. (1995). The single transferable vote. *Journal of Economic Perspectives*, *9*(1), 27–38. doi:10.1257/jep.9.1.27

Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), 440–442. doi:10.1038/30918

Zhang, X., Lin, J., Sun, L., Qi, W., Yang, Y., Chen, Y., Lyu, H., Mou, X., Chen, S., Luo, J., Huang, X., Tang, S. & Wei, Z. (2024). Electionsim: Massive population election simulation powered by large language model driven agents. arXiv preprint arXiv:2410.20746. doi:10.48550/arXiv.2410.20746