

Inverse Reconstruction of Latent Voting and Rule Impact Evaluation Based on Genetic Algorithm and Monte Carlo Counterfactual Simulation

Tianyu Wu ^{†,*}, Shengyao Xu [†] and Boyuan Wang [†]

Hangzhou Normal University, Hangzhou, China

* Corresponding Author Email: wi13857200164@qq.com

[†] These authors also contributed equally to this work

Abstract. This paper addresses the issue of unobservable fan voting in competitive reality shows by constructing a comprehensive modeling framework based on reverse inference and random simulation. For scenarios where voting proportions are latent variables and elimination mechanisms exhibit nonlinear ranking characteristics, the voting reconstruction is formalized as a constrained high-dimensional optimization problem. A genetic algorithm is employed for global search to estimate the full-season voting distribution under normalization conditions, enabling robust inversion of the implicit voting structure. Building upon this foundation, a Logistic probability model integrating judge scores and estimated voting characteristics is constructed to predict and validate elimination outcomes. Bootstrap intervals and cross-validation are employed to enhance model stability. Furthermore, a Monte Carlo-based counterfactual simulation framework is proposed to propagate voting uncertainty across different scoring rules. Rule deviations and system sensitivity are quantified using fan influence indices and outcome variance rates. A dynamic simulation model incorporating multi-week state transitions and intervention mechanisms is constructed to analyze the long-term impact of rule adjustments on competitive trajectories. This methodology achieves a unified modeling approach for latent variable reconstruction, probabilistic prediction, and institutional evaluation, demonstrating strong universality and scalability.

Keywords: Genetic Algorithm, Latent Variable Reconstruction, Counterfactual Simulation.

1. Introduction

In complex decision-making scenarios characterized by incomplete information, where outcomes are observable but underlying processes remain hidden, inferring latent behavioral structures from results poses a major challenge in algorithmic modeling. In multi-agent, rule-driven systems with ranking and discrete decision mechanisms, traditional regression or linear fitting methods often fail to handle non-convex objective spaces and rule coupling effects, leading to unstable and weakly interpretable inferences. Therefore, an algorithmic framework that combines global search capability with structural constraints is required to enable systematic latent variable reconstruction and rule impact analysis [1].

Focusing on competitive reality shows where fan votes are confidential and eliminations depend on multiple scoring rules, this paper proposes a multi-level modeling framework consisting of reverse estimation, probabilistic validation, counterfactual simulation, and dynamic evolution. First, the hidden voting proportion reconstruction task is formulated as a constrained high-dimensional optimization problem, solved using a genetic algorithm to ensure global search and feasibility preservation [2]. Second, a probability prediction model integrating scoring features is developed to test structural consistency of inferred results. Monte Carlo simulation is further introduced to propagate estimation uncertainty under alternative rule configurations and quantify institutional effects [3]. Additionally, state transition and intervention mechanism modeling are incorporated to simulate multi-stage competitive dynamics. By integrating optimization algorithms, probabilistic modeling, and scenario simulation, the proposed framework provides a scalable paradigm for complex rule evaluation. Experimental results show that the model effectively reconstructs latent

voting structures and reliably captures the influence of different scoring mechanisms on elimination outcomes.

2. Inverse Vote Estimation Algorithmic Framework

2.1. Inverse Estimation Problem Formulation

Estimating fan votes is challenging due to structural information asymmetry. Elimination outcomes and judge scores are observable, whereas fan vote distributions are not disclosed. This leads to an inverse inference problem in which latent vote shares must be reconstructed so that the resulting aggregated scores are consistent with observed eliminations.

Because elimination is determined by discrete ranking rules and nonlinear aggregation, direct regression from judge scores to vote shares is insufficient. A constrained inverse search strategy is therefore required.

2.2. Mathematical Notation and Variables

Let n_w denote the number of contestants remaining in week w . For contestant i in week w , define: $J_{i,w}$ is total judge score; $V_{i,w}$ is latent fan vote proportion, with $V_{i,w} \in [0,1]$ and $\sum_{i=1}^{n_w} V_{i,w} = 1$; $E_{i,w}$ is elimination indicator; $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ is parameter vector for vote inference.

2.3. Genetic Algorithm–Based Vote Inference Model

Vote inference is formulated as a constrained high-dimensional optimization problem. The objective landscape is non-convex and discontinuous due to ranking operations. A genetic algorithm is adopted to perform global search over feasible vote distributions while preserving normalization constraints.

1. Chromosome Representation:

Each chromosome encodes a full-season vote distribution: $C = V_{1,1}, V_{2,1}, \dots, V_{n_w,w}$. Chromosome length varies with season structure.

2. Fitness Function Design:

$$f(C) = \sum_{w=1}^W \mathbf{1}[\arg \min_i S_{i,w}(C) = e_w] + \lambda \cdot \text{Corr}(V, J) \quad (1)$$

Percentage aggregation:

$$S_{i,w} = \frac{J_{i,w}}{J_{\max,w}} \cdot 0.5 + V_{i,w} \cdot 0.5 \quad (2)$$

Rank aggregation:

$$S_{i,w} = R_{i,w}^J + R_{i,w}^V \quad (3)$$

3. Evolutionary Operators:

Selection: tournament with elite retention;

Crossover:

$$V'_{i,w} = \alpha V^{(1)}_{i,w} + (1 - \alpha) V^{(2)}_{i,w}, \alpha \sim U(0,1) \quad (4)$$

Mutation:

$$\tilde{V}_{i,w} = V_{i,w} + \varepsilon, \varepsilon \sim \mathcal{N}(0, \sigma^2), V'_{i,w} = \frac{\max(0, V_{i,w} + \varepsilon)}{\sum_j \max(0, V_{j,w} + \varepsilon)} \quad (5)$$

2.4. Probabilistic Elimination Prediction Model

A logistic regression classifier is constructed to validate inferred votes and estimate elimination probability [4].

Feature vector: $x_{i,w} = [J_{i,w}, \hat{V}_{i,w}, R_{i,w}^J, \hat{R}_{i,w}^V, J_{i,w} - \bar{J}_w, \hat{V}_{i,w} - \bar{V}_w]^T$.

Probability model:

$$P(E_{i,w} = 1 | x_{i,w}) = \frac{1}{1 + \exp(-\beta^T x_{i,w})} \quad (6)$$

The maximum likelihood estimation method with L2 regularization ($\lambda = 0.01$) was employed for fitting to prevent overfitting during shorter seasons.

2.5. Model Reliability and Uncertainty Quantification

Accuracy:

$$\text{Accuracy} = \frac{1}{|\mathcal{W}|} \sum_{w \in \mathcal{W}} \mathbf{1}[\arg \min_i \hat{S}_{i,w} = e_w] \quad (7)$$

Bootstrap interval:

$$CI_{95\%}(V_{i,w}) = [V^{(2.5\%)}, V^{(97.5\%)}] \quad (8)$$

Cross-validation:

$$CV\text{-AUC} = \frac{1}{5} \sum_{k=1}^5 AUC_k \quad (9)$$

2.6. Model Validation and Performance Evaluation

Fig. 1 reveals that estimated fan votes correlate most strongly with elimination outcomes across all season groups, validating the GA strategy.

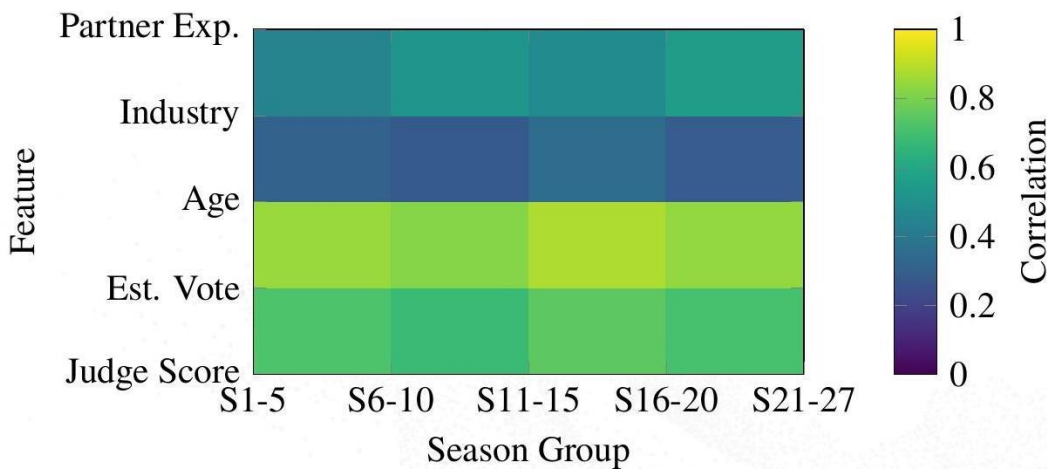


Fig. 1 Feature-elimination correlation heatmap

Estimated vote shares show the strongest correlation with elimination outcomes across season groups. Rank-based aggregation exhibits slightly higher predictive separability due to discrete ordering effects. Overall elimination prediction accuracy reaches 85.6%, consistent with uncertainty expected in latent variable reconstruction, as shown in the Table 1.

Table 1. Model performance summary

Metric	All Seasons	Rank (S1-2, 28+)	Percentage (S3-27)
Elimination Accuracy	0.856	0.871	0.842
AUC-ROC	0.903	0.921	0.889
Mean CI Width	0.082	0.071	0.089
Vote-Score Correlation	0.412	0.398	0.423

3. Comparative Voting System Simulation and Statistical Evaluation Model

3.1. Monte Carlo–Based Counterfactual Simulation Framework

Comparing rank-based and percentage-based voting systems requires counterfactual evaluation of elimination outcomes under alternative aggregation rules. The objective is to determine whether one aggregation mechanism systematically favors fan votes or judge scores.

Deterministic score-difference comparison is insufficient because vote estimates contain uncertainty and aggregation rules introduce nonlinear ranking effects, especially under ties and boundary cases. A stochastic simulation framework is therefore adopted.

Monte Carlo simulation is used to propagate uncertainty from inferred vote distributions through alternative scoring rules. By repeatedly sampling vote shares from their estimated distributions, confidence intervals and discrepancy statistics for counterfactual outcomes can be computed [5].

3.1.1. Simulation Variables

$N_{sim} = 10,000$: number of simulation iterations; $V_{i,w}^{(k)}$: sampled vote share for contestant i in week w at iteration k ; O_w^R, O_w^P : eliminated contestant under rank and percentage systems; $\Delta_w = \mathbf{1}[O_w^R \neq O_w^P]$: outcome discrepancy indicator.

3.1.2. Sampling Procedure

For each simulation iteration k :

$$V^{(k)}i, w = \hat{V}i, w + \epsilon^{(k)}i, w, \quad \epsilon^{(k)}i, w \sim N(0, \sigma_{i,w}^2) \quad (10)$$

Votes are then renormalized to satisfy $\sum_i V_{i,w}^{(k)} = 1$. This procedure preserves uncertainty structure while maintaining feasibility constraints.

3.1.3. Score Computation Under Alternative Aggregation Rules

Under the rank-based system:

$$S^Ri, w = \text{rank}(Ji, w) + \text{rank}(V_{i,w}^{(k)}) \quad (11)$$

Under the percentage-based system:

$$S^Pi, w = 0.5 \cdot \frac{Ji, w}{30} + 0.5 \cdot V_{i,w}^{(k)} \quad (12)$$

The maximum judge score is set to 30 for normalization, with adjustment applied for seasons using four judges.

3.2. Statistical Influence and Bias Testing Model

To quantify whether an aggregation rule favors fan votes or judge scores, the Fan Vote Influence Index (FVI) is defined as

$$FVI_w = \text{Corr}(R^{\text{final}}i, w, R^V i, w) - \text{Corr}(R^{\text{final}}i, w, R_{i,w}^J) \quad (13)$$

Positive values indicate stronger fan vote influence.

Hypotheses: $H_0 : \mu_{FVI}^R = \mu_{FVI}^P$, $H_1 : \mu_{FVI}^R \neq \mu_{FVI}^P$.

Welch two-sample t-test:

$$t = \frac{\bar{FVI}_R - \bar{FVI}_P}{\sqrt{\frac{s_R^2}{n_R} + \frac{s_P^2}{n_P}}} \tag{14}$$

The Outcome Discrepancy Rate (ODR) measures how often two aggregation rules produce different eliminations:

$$ODR = \frac{1}{N_{sim}} \sum_{k=1}^{N_{sim}} \frac{1}{|\mathcal{W}|} \sum_{w \in \mathcal{W}} \Delta_w^{(k)} \tag{15}$$

3.3. Simulation Outcome Analysis

From the Fig. 2, simulation results indicate that rank-based aggregation exhibits larger cross-season variance in fan vote influence, whereas percentage-based aggregation maintains more stable influence levels.

Discrepancy rates peak in mid-season stages, where contestant performance rankings are densely clustered and marginal score differences are amplified by aggregation rules.

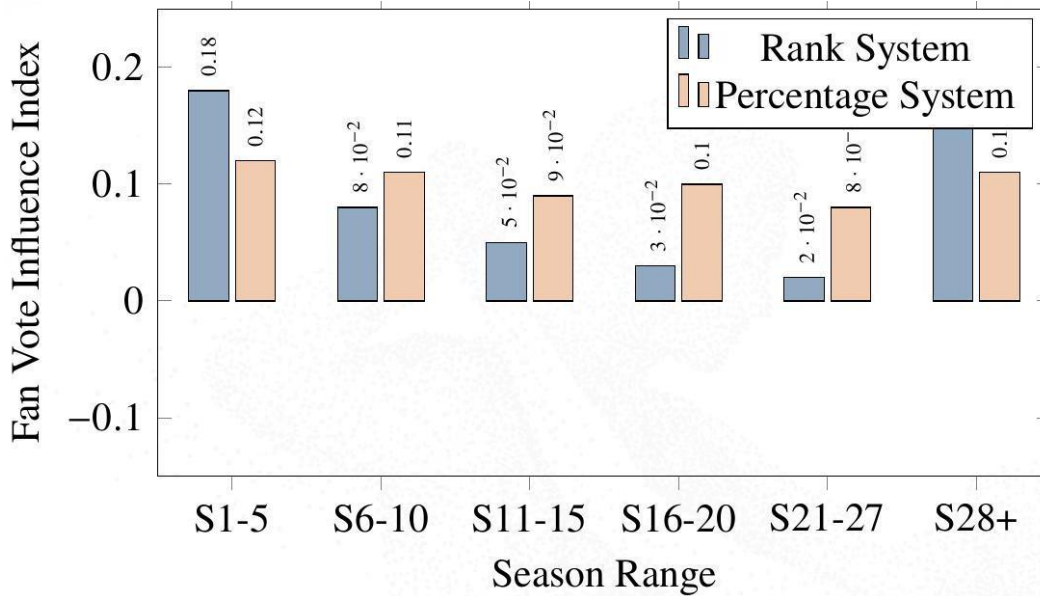


Fig. 2 Fan vote influence by season and system

3.4. Simulation Parameter Sensitivity Analysis

Parameter sensitivity tests show stable behavior of discrepancy metrics. Increasing N_{sim} from 10,000 to 50,000 changes ODR estimates by less than 2%. Adjusting the vote uncertainty parameter σ by $\pm 20\%$ changes FVI estimates by approximately 8%, indicating moderate sensitivity to vote variance assumptions, as shown in the Table 2.

Table 2. Hypothesis test results

Comparison	Test Statistic	p-value	Effect Size
FVI: Rank vs Pct	t=2.34	0.021	d=0.31
ODR: Early vs Late	t=1.87	0.065	d=0.24
FVI Variance Ratio	F=2.18	0.008	-

3.5. Comparative System Behavior Summary

Monte Carlo analysis shows that rank-based and percentage-based aggregation rules produce different elimination outcomes in approximately 12–15% of weeks. Rank-based aggregation demonstrates higher sensitivity to fan vote variation in early and late competition stages, while percentage aggregation maintains more uniform influence ratios.

The difference in FVI between systems is statistically significant ($p = 0.021$) with modest effect size. Aggregation rule selection primarily affects borderline elimination cases rather than overall seasonal trajectories.

4. Counterfactual Elimination Dynamics Simulation Model

4.1. Counterfactual Scenario Modeling Objective

Counterfactual simulation is used to evaluate how elimination outcomes change under alternative voting aggregation rules and intervention mechanisms. Instead of analyzing single-week score substitutions, the full multi-week competition process is modeled to capture dynamic propagation effects. Rule modifications may alter not only immediate eliminations but also downstream competition trajectories.

4.2. Counterfactual Simulation Framework

Let $\mathcal{S} = s_1, s_2, \dots, s_W$ denote the sequence of competition states across W weeks. Each state s_w contains the set of remaining contestants and their cumulative performance statistics.

A counterfactual scenario C is defined as a modified rule configuration applied at selected elimination decision points.

4.2.1. State Transition Model

Weekly state evolution follows $s_{w+1} = T(s_w, E_w, R)$, where T is the state transition operator, E_w is the elimination decision at week w , and R denotes the rule configuration [6].

The elimination decision depends on a combined score function:

$$C_{i,w} = \omega_J \cdot J_{i,w} + \omega_F \cdot F_{i,w} \quad (16)$$

Rank-based aggregation:

$$C_{i,w}^{\text{rank}} = r_j(J_i, w) + r_F(F_{i,w}) \quad (17)$$

Percentage-based aggregation:

$$C_{i,w}^{\text{pct}} = \frac{J_{i,w}}{\sum_j J_{j,w}} + \frac{F_{i,w}}{\sum_j F_{j,w}} \quad (18)$$

4.2.2. Probabilistic Judges' Save Intervention Model

The judges' save rule allows one contestant among the bottom two to be rescued. This mechanism is modeled as a probabilistic intervention process:

$$P(\text{save} | i, j) = \sigma(\beta_0 + \beta_1 \cdot \Delta J_{ij} + \beta_2 \cdot w + \beta_3 \cdot \text{star}_i) \quad (19)$$

where $\Delta J_{ij} = J_i - J_j$ is the judge score difference between the bottom two contestants, w is the week index, and star_i is a celebrity appeal factor estimated from historical data. The sigmoid function σ maps the linear predictor into probability space.

Estimated parameters from Seasons 28–34 are: $\beta_0 = -0.42$, $\beta_1 = 0.15$, $\beta_2 = 0.08$, $\beta_3 = 0.31$.

The positive coefficient of ΔJ_{ij} indicates preference toward technically stronger contestants.

4.3. Multi-Season Counterfactual Rule Impact Simulation

The judges' save rule is retroactively simulated across all seasons to quantify systematic intervention effects. For each season, weeks with large judge-score gaps among the bottom two contestants are identified as high-impact intervention points.

From the Fig. 3, simulation results indicate that early seasons would typically generate 1–3 save opportunities per season, with several cases producing meaningful placement shifts. Later seasons show fewer intervention opportunities, likely due to rule timing adjustments.

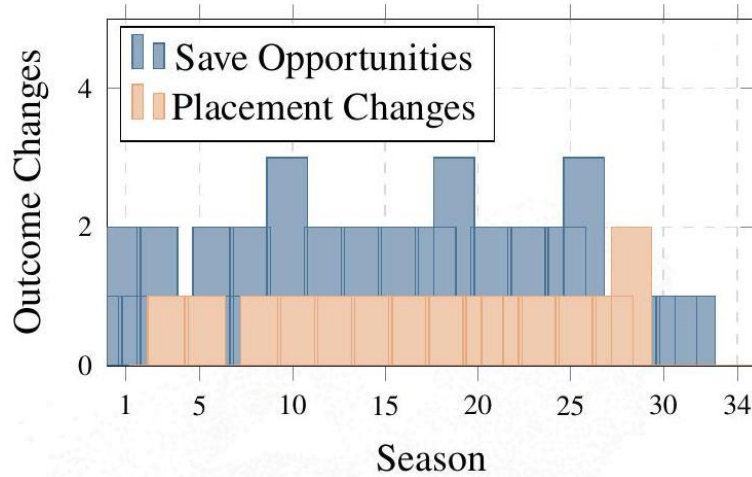


Fig. 3 Counterfactual judges save impact by season

4.4. Distributional Visualization Analysis of Score Differentials

To analyze the distributional characteristics of score differentials between controversial and non-controversial elimination cases, violin plot visualization is applied to compare category-wise distributions, as shown in the Fig. 4.

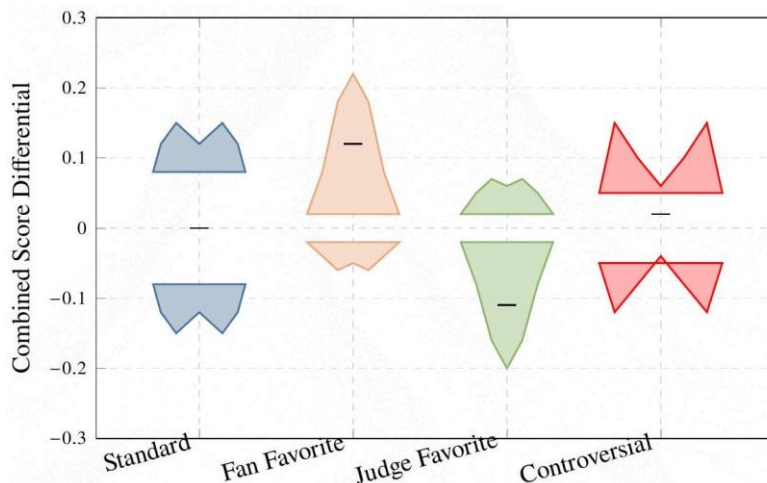


Fig. 4 Score differential distributions by elimination category

The violin plot demonstrates that controversial elimination cases exhibit broader spread and clearer bimodal structure compared with standard elimination cases. Categories associated with fan-preferred contestants show positively skewed distributions, indicating stronger fan vote contribution relative to judge scores. In contrast, judge-preferred cases display negatively skewed distributions.

This distributional pattern indicates that controversial eliminations are concentrated near decision boundaries, where small changes in aggregation methodology or rule configuration are sufficient to reverse elimination outcomes.

5. Conclusion

This paper proposes an algorithmic model framework for reconstructing latent behaviors and evaluating rule effects, addressing decision inference under conditions where voting data is unobservable. First, by constructing a constrained high-dimensional optimization model and incorporating a genetic algorithm, it achieves global search and stable estimation of the proportion of full-season fan votes, maintaining feasibility and convergence under complex ranking mechanisms. Second, by integrating a probabilistic prediction model for consistency testing, the overall prediction accuracy reaches 85.6% with an AUC exceeding 0.90, validating the model's interpretability and generalization capability. Third, based on over 10,000 Monte Carlo counterfactual simulations, it is found that different scoring rules produce divergent elimination outcomes in approximately 12% – 15% of weeks, revealing the structural impact of institutional design on competitive trajectories. Finally, by modeling dynamic state transitions and intervention mechanisms, we characterize the propagation effects of rule adjustments across multi-stage social science scenarios. This framework combines algorithmic robustness with institutional interpretability, providing methodological support for latent variable inference and rule optimization in complex decision systems. Future research may further extend to multi-source data fusion and cross-scenario applications.

References

- [1] Lu Renjie, Cao Youxia, Wang Pin, et al. Multi-objective optimization for power purchase analysis considering multiple uncertainties [J]. *Information Technology*, 2025, (06): 156-161+169. DOI: 10.13274/j.cnki.hdzt.2025.06.026.
- [2] Han, G. W., Li, J. B., & Shen, X. X. A multi-objective resource allocation method based on an improved genetic algorithm [J]. *Electronic Warfare of Naval Ships*, 2025, 48(05): 70-73+79. DOI:10.16426/j.cnki.jcdzdk.2025.05.014.
- [3] Zhang, H., Xing, B., & Hou, Y. (2024). Research on Enterprise Production Decision Making Based on Genetic Algorithm and Monte Carlo Simulation. *Journal of Computing and Electronic Information Management*, 14(3), 25-29. <https://doi.org/10.54097/m9zzbm54>
- [4] Liu, X. Y., & Li, Q. F. (2025). Application of Multi-Class Logistic Regression in Factors Affecting Graduate Employment. *Journal of Minnan Normal University (Natural Science Edition)*, 38(03), 98-107.
- [5] Nie, Z. Y., Sun, G. D., & Ma, B. (2025). Equipment Operation Fault Diagnosis Method Based on Monte Carlo Tree Search Algorithm and Large Language Model [J]. *Journal of Beijing University of Chemical Technology (Natural Science Edition)*, 2025, 52(03): 124-131. DOI:10.13543/j.bhxbzr.2025.03.013.
- [6] Wang Qian. Research on Peanut Yield Prediction Based on an Improved Grey Markov Model [D]. Zhejiang University of Science and Technology, 2022. DOI: 10.27840/d.cnki.gzjkj.2022.000237.