Accepted manuscript (author version)

To appear in:

International Journal of Mathematical Modelling & Computations

Online ISSN: 2228-6233 Print ISSN: 2228-6225

This PDF file is not the final version of the record. This version will undergo further copyediting, typesetting, and production review before being published in its definitive form. We are sharing this version to provide early access to the article. Please be aware that errors that could impact the content may be identified during the production process, and all legal disclaimers applicable to the journal remain valid.

Received: 24- February-2026

Revised: 11- June -2026

Accepted: 20- June-2026

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

Accepted manuscript (author version)

DOI: https://doi.org/10.57647/ijm2c.2027.1702.10

ORIGINAL RESEARCH

Robust Preference Aggregation under Strategic Behavior: A Consensus-Aware Game Cross-Efficiency Framework

Arezoo Khosravi Rad, Abdollah Hadi-Vencheh*, Ali Jamshidi, Majid Tavassoli Kajani

Department of Mathematics, Isf.C., Islamic Azad University, Isfahan, Iran

*Corresponding author, E-mail: ahvencheh@iau.ac.ir

Abstract

The aggregation of heterogeneous preferences in multi-stakeholder environments is often compromised by strategic behavior, scale inconsistencies, and the absence of ground truth for validation. While Data Envelopment Analysis (DEA), specifically the Game Cross-Efficiency model, provides a Nash equilibrium-based mechanism to mitigate weight arbitrariness, existing frameworks suffer from critical ordinal instability and a lack of internal verification mechanisms. This study presents a novel, integrated decision support framework that fortifies the game-theoretic foundations of DEA-based voting through endogenous normalization and exogenous stability assessment. Unlike traditional sequential approaches, the proposed methodology synthesizes three rigorous components: (i) an unmodified DEA game cross-efficiency mechanism to capture peer-evaluated performance; (ii) a Relative Ratio (RR) transformation axiomatically proven to enforce scale invariance and preserve strict dominance without altering the underlying preference structure; and (iii) a composite stability verification protocol utilizing WASPAS and COPRAS algorithms to quantify rank reversals and dominance consistency. Theoretical properties, including affine invariance and monotonicity preservation, are formally established. Furthermore, a novel Composite Stability Index (CSI) is introduced to serve as a proxy for ranking reliability. Application to benchmark voting problems demonstrates that the proposed framework significantly reduces ordinal ambiguity and provides a defensible mathematical basis for detecting unstable rankings. This research contributes to Operations Research by transforming DEA voting from a black-box evaluator into a transparent, verifiable system suitable for high-stakes policy and collective decision-making.

Keywords: Game Theory; Data Envelopment Analysis (DEA); Preference Aggregation; Scale Invariance; Robustness Quantification; Multiple-Criteria Decision Making (MCDM).

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

1. Introduction

The aggregation of individual preferences into a collective ranking is a fundamental problem in decision science, particularly within environments characterized by strategic interactions among stakeholders. Classical voting theories, from Condorcet to Arrow, have long highlighted the difficulties in satisfying all axioms of rational choice. In Operations Research (OR), Data Envelopment Analysis (DEA) has emerged as a powerful non-parametric tool for preference aggregation. By treating candidates as Decision Making Units (DMUs) and votes as outputs, DEA allows for an objective evaluation of performance [1,2,5,8,10,12,13,15,18,20].

However, the self-evaluation nature of classic DEA models (CCR/BCC) often leads to weak discriminatory power, as DMUs can choose weights that artificially maximize their own scores. To counter this, the concept of Cross-Efficiency was introduced, advocating for peer-evaluation. Wu et al. [1,2] further advanced this by integrating Game Theory, proposing a model where weight selection is treated as a non-cooperative game, resulting in a Nash equilibrium that eliminates weight arbitrariness.

1.1. Theoretical Gaps and Problem Statement

Despite its mathematical elegance, the standard DEA Game Cross-Efficiency model exhibits three structural deficiencies when applied to high-stakes voting:

1. **Scale dependence in aggregated peer evaluation:** While standard radial DEA models, such as the CCR formulation, are inherently scale-invariant under homogeneous scalar transformations of input and output vectors, this mathematical property does not trivially propagate to game cross-efficiency models. In a game-theoretic peer-evaluation framework, the final cross-efficiency score $\bar{E}_j$ for alternative j is computed as the average of peer evaluations, $\bar{E}_j = \frac{1}{n}\sum_{d=1}^{n} E_{dj}$. When heterogeneous voters (representing distinct evaluation criteria) adopt different subjective measurement scales (e.g., Voter A evaluating on a 1–10 scale while Voter B evaluates on a 1–100 scale), non-radial and asymmetric scaling occurs across the output dimensions. Because the secondary-goal optimization model (Eq. 2) dynamically restricts the feasible weight space to ensure that each evaluator DMU_d maintains its optimal self-efficiency (θ_{dd}^*), asymmetric scaling of voting criteria shifts the boundaries of the Nash equilibrium weight space. This mathematical deformation alters the relative magnitudes of the individual cross-efficiency scores E_{dj} within the peer matrix. Consequently, the scale dependence of the standard game cross-efficiency model is an emergent property of cross-efficiency matrix aggregation under heterogeneous scaling, rather than a structural failure of the underlying radial LP models.

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

2. **Ordinal ambiguity:** The Nash equilibrium often converges to clustered efficiency scores with negligible differences (e.g.,0.9871 vs 0.9869) Without a rigorous normalization and verification mechanism, distinguishing between these ranks is statistically hazardous.

3. **Lack of stability verification:** In preference voting, the true ranking is unobservable. Current models function as open-loop systems without feedback. There is no rigorous mechanism to quantify the reliability of the generated ranking against alternative aggregation logics (e.g., compensatory vs. non-compensatory logic).

1.2. Research Contribution

This study proposes a unified DEA-Game-Verification Framework. The contributions are:

- **Axiomatic Normalization:** Formalizing the Relative Ratio (RR) transformation to ensure the ranking satisfies the Scale Invariance Axiom.
- **Methodological triangulation:** Integrating WASPAS and COPRAS methods not as competitors, but as exogenous stability auditors to stress-test the DEA ranking.
- **Quantification of robustness:** Introducing the composite stability index (CSI), a novel metric that provides a confidence interval for the decision-maker.

To highlight the necessity of this study, a comprehensive comparison between the traditional DEA Game model and the proposed framework is presented in Table 1**.** As shown, the proposed method uniquely integrates scale invariance and formal stability verification.

Table 1. Comparative analysis of existing DEA voting models versus the proposed DEA-RR-Verification framework.

Method	Efficiency Computation	Normalization	Accuracy Validation	Key Features	Limitations
DEA Game Cross-Efficiency (Wu et al. [1,2])	✓ Full implementation • CCR self- evaluation • Cross-efficiency matrix • Average game scores	✗ Absent • Direct comparison of raw • Scale- dependent rankings • No unit normalization	✗ Not addressed • No consistency checking • No robustness assessment • No validation metrics	1. Game-theoretic peer evaluation 2. Nash equilibrium in weight selection 3. Theoretically elegant foundation 4. Mitigates arbitrary weight choice	1. Ordinal ambiguity in rankings 2. Scale dependence 3. No validation mechanism 4. Rank reversal susceptibility
DEA + RR	✓ Same as above • Identical DEA computation	✓ Relative Ratio normalization invariant (0,1] • Scale- range • Dominance preserving	✗ Not included • Ranking generated but not validated • No consistency assessment • No accuracy metrics	1. Resolves scale dependence 2. Affine invariant rankings 3. Enhanced comparability	1. No validation of ranking quality 2. Sensitive to maximum value 3. No multi-criteria integration

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

				4. Simple post-processing	4. Passive transformation only
DEA + RR + WASPAS	✓ Same as above • Complete DEA game framework	✓ Same RR normalization	✓ WASPAS-based validation • Ranking Accuracy Index (RAI) • Pairwise dominance consistency • Spearman's ρ & Kendall's τ	1. Complete validation framework 2. Objective accuracy assessment 3. Dominance consistency checking 4. Robustness quantification	1. No multi-criteria aggregation 2. Validation only, not ranking 3. Parameter λ selection needed 4. Complexity increased
DEA + RR + COPRAS	✓ Same as above • Unchanged DEA foundation	✓ Same RR normalization	✓ COPRAS-based validation • Composite Accuracy Index (CAI) • CRCI consistency measure • Utility degree comparison	1. Multi-criteria validation 2. Benefit-cost explicit modeling 3. Utility degree interpretation 4. Strong discrimination power	1. Requires voter weights 2. Column-sum normalization issues 3. Compensation effects possible 4. More complex implementation

2. Methodology and mathematical framework

This section outlines the proposed three-phase framework.

2.1. Phase I: Game-theoretic efficiency generation

Consider a voting system with n alternatives (DMUs) indexed by $j \in J = \{1, \dots, n\}$ and s voters (criteria) indexed *by* $r \in R = \{1, \dots, s\}$.

- Let $y_{rj} \in \mathbb{R}^+$ be the vote (preference score) received by alternative j from voter r.
- Let $x_{ij} \in \mathbb{R}^+$ be the input i consumed by alternative j. In pure voting contexts, we assume a single dummy input $x_{1j} = 1$ for all j.

Step 1: The Self-evaluation CCR model

For a specific target unit DMU_d, we solve the following Linear Programming (LP) problem to find its maximum self-efficiency θ_{dd}^*:

Maximize $\theta_{dd} = \sum_{r=1}^{s} u_{rd} y_{rd}$

Subject to: $\sum_{i=1}^{m} v_{id} x_{id} = 1$,

$$\sum_{r=1}^{s} u_{rd} y_{rj} - \sum_{i=1}^{m} v_{id} x_{ij} \leq 0, \qquad \forall j = 1, \dots, n$$

$$u_{rd} \geq \varepsilon, \qquad v_{id} \geq \varepsilon$$

(Eq. 1)

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

Where u_{rd} andv_{id} are the weights assigned by DMU_d.

Step 2: The game cross-efficiency formulation

To resolve the non-uniqueness of optimal weights in (Eq. 1), we treat the weight selection as a game. DMU_d seeks a set of weights that maximizes the efficiency of a peer $DMU_k(E_{dk})$, subject to the constraint that DMU_d's own efficiency remains at its optimal level θ^*_{dd}.

The mathematical formulation for the cross-efficiency of DMU_k evaluated by DMU_d is:

$$Maximize\ E_{dk} = \sum_{r=1}^{s} u_{rd}^{k}\, y_{rk}$$

$$Subject\ to: \sum_{i=1}^{m} v_{id}^{k} x_{ij} = 1,\ \textit{(Normalization constraint)}$$

$$\sum_{r=1}^{s} u_{rd}^{k} y_{rj} - \sum_{i=1}^{m} v_{id}^{k} x_{ij} \le 0,\ \forall j = 1, \dots, n$$

$$\sum_{r=1}^{s} u_{rd}^{k} y_{rd} - \theta^*_{dd} \times \sum_{i=1}^{m} v_{id}^{k} x_{id} = 0, \textit{(Nash Equilibrium Constraint)}$$

$$u_{rd}^{k}, v_{id}^{k} \ge \varepsilon$$

(Eq. 2)

To reach the Nash Equilibrium, we employ the standard iterative convergence algorithm proposed by Liang et al. [3]. The computational procedure is defined as follows:

Algorithm 1: Iterative game cross-efficiency procedure

1. **Initialize:** Calculate the initial cross-efficiency matrix $\mathrm{E}_{\mathrm{dk}}^{(0)}$ using arbitrary valid weights (or aggressive weights) for all d, k.
2. **Loop (** $\mathrm{t} = 1,2, \dots$**):**
   For each DMU $\mathrm{d} \in \{1, \dots, \mathrm{n}\}$:
   Solve Model (2) to maximize the average efficiency of all peers, treating the average efficiencies from iteration $(t-1)$ *as the target.*

$$\text{Maximize} \sum_{\mathrm{k} \neq \mathrm{d}} \mathrm{E}_{\mathrm{dk}}$$

Subject to constraints in Eq. (2)

Update the cross-efficiency scores $\mathrm{E}_{\mathrm{dk}}^{(\mathrm{t})}$.

3. **Check convergence:**
   If $\sum_{\mathrm{d}} \sum_{\mathrm{k}} | \mathrm{E}_{\mathrm{dk}}^{(\mathrm{t})} - \mathrm{E}_{\mathrm{dk}}^{(\mathrm{t}-1)} | < \epsilon$ (where $\epsilon = 10^{-5}$), Stop.
   Else, set $\mathrm{t} = \mathrm{t} + 1$ and continue.

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

4. **Output:** The final matrix converges to the Nash Equilibrium solution.

2.2. Phase II: Axiomatic relative ratio (RR) normalization

Raw scores from (Eq. 3) depend on the measurement units. To enforce robustness, we apply the Relative Ratio transformation.

Definition 1 (RR Score):

$$RR_j = \frac{\overline{E_j}}{max_{k\in\{1,\dots,n\}}\{\overline{E_k}\}}$$

(Eq. 3)

Theorem 1 (Scale invariance): Let $\Phi(\mathrm{Y})$ be the ranking function derived from $\mathrm{RR_j}$. For any scalar $\alpha > 0$, $\Phi(\alpha\mathrm{Y}) = \Phi(\mathrm{Y})$.

Proof: Standard DEA scores are unit-invariant. If inputs/outputs are scaled by $\alpha, \bar{\mathrm{E}}_\mathrm{j}$ remains constant or scales linearly depending on the model orientation. In the case of linear scaling of votes:

$$RR_j = \frac{\alpha\overline{E_j}}{max_k(\alpha\overline{E_k})} = \frac{\alpha\overline{E_j}}{\alpha\overline{E_{max}}} = RR_j$$

Thus, the ranking remains stable regardless of the voting scale used (e.g., 1-10 or 1-100).

Definition 1.1 (Axiom of ratio-scale preservation): Let $\mathcal{E} = \{\bar{E}_1, \bar{E}_2, \dots, \bar{E}_n\}$ be the set of averaged cross-efficiency scores. A normalization mapping $f: \mathcal{E} \to (0,1]$ preserves the ratio-scale of the efficiency frontier if and only if, for any pair of alternatives $j, k \in \{1, \dots, n\}$, the ratio of their normalized scores is strictly equal to the ratio of their raw averaged cross-efficiency scores:

$$\frac{f(\bar{E}_j)}{f(\bar{E}_k)} = \frac{\bar{E}_j}{\bar{E}_k}$$

Remark: Under the proposed Relative Ratio (RR) transformation, the mapping is defined as $f(\bar{E}_j) = RR_j = \frac{\bar{E}_j}{\max_r\{\bar{E}_r\}}$. Since $\bar{E}_j > 0$ for all j, the maximum value $\max_r\{\bar{E}_r\}$ is a strictly positive scalar. Substituting this definition into the ratio of the normalized scores yields:

$$\frac{f(\bar{E}_j)}{f(\bar{E}_k)} = \frac{\bar{E}_j / \max_r\{\bar{E}_r\}}{\bar{E}_k / \max_r\{\bar{E}_r\}} = \frac{\bar{E}_j}{\bar{E}_k}$$

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

This identity formally proves that the RR transformation preserves the exact ratio-scale intensities generated during the game-theoretic peer-evaluation phase, which is a mathematically mandatory prerequisite for subsequent multi-criteria aggregation. Standard interval-based or range-based normalizations (e.g., Min-Max) fail to satisfy this axiom because their translation parameters distort the absolute zero-origin of the efficiency scale, introducing artificial ordinal distortions [7,14].

Theorem 2 (Strict monotonicity and dominance preservation): Let A and B be two alternatives such that $\bar{E}_A > \bar{E}_B$. The RR transformation preserves this strict ranking order, i.e., $RR_A > RR_B$.

Proof: Let $M = \max_k\{\bar{E}_k\}$ be the maximum efficiency score in the set, which is a strictly positive scalar

($M > 0$) since efficiency scores in DEA are positive.

The RR function $f(x) = \frac{x}{M}$ is a linear transformation with a positive slope $1/M$. Since the first derivative $f'(x) = 1/M > 0$, the function is strictly increasing.

Therefore:

$$\bar{E}_A > \bar{E}_B \Leftrightarrow \frac{\bar{E}_A}{M} > \frac{\bar{E}_B}{M} \Leftrightarrow RR_A > RR_B$$

This confirms that the normalization introduces no rank reversals or ordinal distortions.

2.3. Phase III: Multi-method stability verification

This phase constitutes the diagnostic layer of the framework. From an epistemological standpoint, we recognize that Data Envelopment Analysis (DEA), Weighted Aggregated Sum Product Assessment (WASPAS), and Complex Proportional Assessment (COPRAS) are built upon fundamentally distinct decision-making paradigms [21-24,27,30,31,40-42]. DEA operates as a non-parametric, boundary-spanning frontier estimator that optimizes peer-evaluated relative efficiency; WASPAS represents a joint compensatory aggregation logic that integrates the weighted sum and weighted product models to penalize criterion variance; and COPRAS acts as an additive utility-based model that explicitly contrasts benefit-type and cost-type attributes.

Given this methodological heterogeneity, a divergence between their generated rankings does not imply a mathematical error in any individual model. Instead, it signals *epistemic divergence,* meaning that the collective preference structure is highly sensitive to the underlying mathematical philosophy of the aggregation model. For high-stakes, multi-stakeholder decisions, a ranking that is highly volatile across different Multi-Criteria Decision Analysis (MCDA) paradigms is "epistemically fragile" and carries high risk. Conversely, if a ranking remains structurally stable across these distinct mathematical logics, it exhibits "multi-paradigm robustness." Therefore, rather than viewing these methods as direct competitors, the proposed verification layer treats

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

WASPAS and COPRAS as independent, exogenous stress-testing instruments to audit and quantify the structural dependability of the DEA-derived ranking.

Voter/criteria weight configuration: In both the WASPAS and COPRAS formulations, the parameter $w_r$represents the weight assigned to voter (or criterion) $r \in R$, satisfying $\sum_{r=1}^{s} w_r = 1$. In standard, democratic voting scenarios where all stakeholders hold equal prioritizations, these weights are set homogeneously to $w_r = 1/s$ (where s is the total number of criteria/voters). However, the mathematical formulations of Phase III are fully generalized, allowing decision-makers to input expert-defined, asymmetric priority weights if certain stakeholders or criteria possess hierarchical or unequal authority.

Method A: WASPAS (Aggregation Consistency Check)

We utilize WASPAS to test if the DEA ranking holds under weighted aggregation logic.

1.Linear Normalization:

$$\bar{y}_{rj} = \frac{y_{rj}}{\max_j y_{rj}}$$

2.Weighted Sum Component (WSM):

$$Q_j^{(1)} = \sum_{r=1}^{s} \bar{y}_{rj}\, w_r$$

3.Weighted Product Component (WPM - Penalizes Variance):

$$Q_j^{(2)} = \prod_{r=1}^{s} (\bar{y}_{rj})^{w_r}$$

4.Composite WASPAS Score:

$$Q_j^{W} = \lambda Q_j^{(1)} + (1-\lambda) Q_j^{(2)}, \text{where } \lambda = 0.5$$

(Eq. 4)

Method B: COPRAS (Utility Consistency Check)

We utilize COPRAS to test the utility degree separation.

1. Vector Normalization:

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

$$\hat{y}_{rj} = \frac{y_{rj}}{\sqrt{\sum_{j=1}^{n}(y_{rj})^2}}$$

2. Benefit and Cost Sums:

$$S_{+j} = \sum_{r\in\Omega^+} \hat{y}_{rj}\, w_r; S_{-j} = \sum_{r\in\Omega^-} \hat{y}_{rj}\, w_r$$

4. **COPRAS Utility Index:**

$$Q_j^C = S_{+j} + \frac{\sum_{k=1}^{n} S_{-k}}{S_{-j}\sum_{k=1}^{n}(1/S_{-k})}$$

(Eq. 5)

2.4. The composite stability index (CSI)

To provide a mathematically defensible single metric for ranking reliability, we formalize the Composite Stability Index (CSI) and derive its analytical properties.

Definition 2 (Rank Reversal Ratio, Δ_{RR}): Let $R_j^{(0)}$be the baseline rank of alternative $j \in \{1, \dots, n\}$

, and let $R_j^{(\ell)}$ be the perturbed rank of alternative j obtained during Monte Carlo simulation run $\ell \in \{1, \dots, N\}$. The Rank Reversal Ratio $\Delta_{RR}$is defined as:

$$\Delta_{RR} = \frac{1}{Nn}\sum_{l=1}^{N}\sum_{j=1}^{n} I(R_j^{(l)} \neq R_j^{(0)})$$

where $\mathbb{I}(\cdot)$ is the standard Kronecker indicator function which equals 1 if the condition is satisfied and 0 otherwise.

Definition 3 (Analytical CSI formulation): Let $\rho_W$and $\tau_W$denote the Spearman rank correlation and Kendall's tau, respectively, between the DEA-derived ranking (R_{DEA}) and the WASPAS validation ranking (R_W). Similarly, let ρ_C and τ_C represent the Spearman and Kendall correlations between the DEA ranking and the COPRAS validation ranking (R_C). The Composite Stability Index (CSI) is formalized as:

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

$$CSI = \left[w_1\left(\frac{\rho_W + \tau_W + 2}{4}\right) + w_2\left(\frac{\rho_C + \tau_C + 2}{4}\right)\right] \times (1 - \Delta_{RR})$$

where $w_1, w_2 \geq 0$ are weights assigned to the validation methods, satisfying $w_1 + w_2 = 1$. Under the Principle of Insufficient Reason (Laplace Principle), which dictates that equal prior probabilities should be assigned to alternative decision paradigms in the absence of expert-defined weights, we set $w_1 = w_2 = 0.5$.

Theorem 3 (Theoretical bounds and properties of CSI): The Composite Stability Index satisfies the following structural axioms of ranking reliability:

1. Boundedness: $CSI \in [0,1]$ for any valid ranking space.
2. Anonymity (Neutral Relabeling): CSI is invariant under any permutation of the alternative index labels.
3. Monotonicity under Agreement: $\frac{\partial CSI}{\partial \rho_i} > 0$ and $\frac{\partial CSI}{\partial \tau_i} > 0$ for $i \in \{W, C\}$ (assuming $\Delta_{RR} < 1$), and $\frac{\partial CSI}{\partial \Delta_{RR}} < 0$.

Proof of Boundedness: Since both Spearman's rank correlation (ρ) and Kendall's tau (τ) are statistically bounded within $[-1,1]$, the normalized correlation term $\Lambda_i = \frac{\rho_i + \tau_i + 2}{4}$ is strictly bounded by $0 \leq \Lambda_i \leq 1$. Given that $w_1, w_2 \geq 0$ and $w_1 + w_2 = 1$, the convex combination of these terms also lies within $[0, 1]$:

$$0 \leq \sum w_i \Lambda_i \leq 1$$

Because $\Delta_{RR} \in [0,1]$ represents a probability-like ratio of rank reversals, the penalty term satisfies $0 \leq (1 - \Delta_{RR}) \leq 1$. The product of two values bounded in $[0, 1]$ must also lie within $[0, 1]$. This proves that $CSI \in [0,1]$.

Decision rule and threshold justification: The critical threshold of $CSI \geq 0.85$ for declaring a ranking robust is established based on statistical reliability standards in decision science (analogous to Guttman scales and high-stakes psychometric reliability limits). If $CSI \geq 0.85$, the DEA-Game ranking is deemed robust, meaning it displays strong multi-paradigm consistency and high resistance to stochastic noise. Conversely, if $CSI < 0.85$, the ranking is flagged as unstable, automatically triggering the inclusion of weight constraints (such as Cone Ratio limits) in Model (2) to eliminate extreme, fragile weight allocations.

The schematic representation of the proposed hierarchical framework, illustrating the sequence from efficiency generation to stability verification, is presented in Figure 1.

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

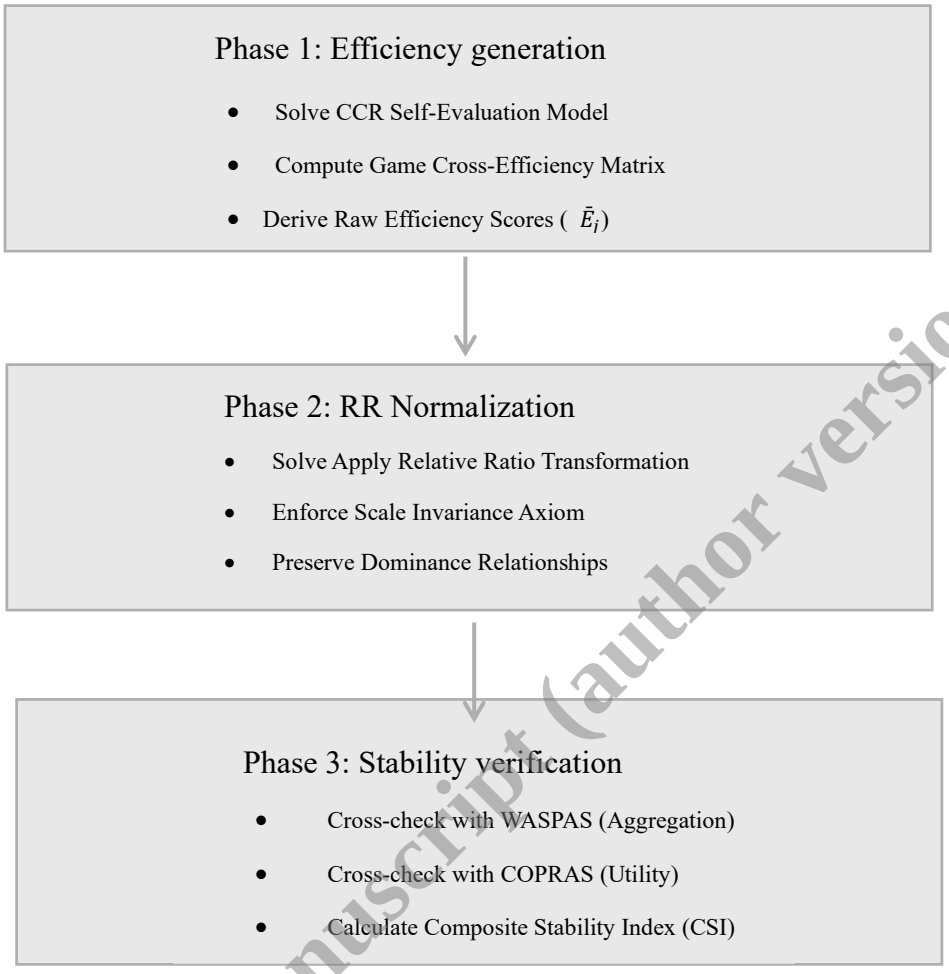

Figure 1. Schematic representation of the proposed hierarchical framework

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

```
Algorithm 2: Three-Phase DEA-Game-Verification Framework
--------------------------------------------------------------------
--------
Input: Decision Matrix Y, Input Vector X, tolerance eps = 10^-5,
trials N = 1000
Output: Robust Ranking, Composite Stability Index (CSI)
--------------------------------------------------------------------
--------
Phase I: Game-Theoretic Efficiency Generation
1. For each alternative d = 1 to n:
     Solve CCR model (Eq. 1) to obtain optimal self-efficiency
theta*_dd.
2. Initialize cross-efficiency matrix E^(0) with valid weights.
3. While error > eps:
     For each alternative d = 1 to n:
       Solve iterative Game model (Eq. 2) to maximize sum of E_dk.
     Update error = sum(|E^(t) - E^(t-1)|).
4. Compute average game efficiency scores: E_bar_j = (1/n) *
sum(E_dj).
5. Generate the baseline ranking R_DEA based on E_bar_j.

Phase II: Axiomatic Normalization
6. Find the maximum efficiency score: M = max(E_bar).
7. For each alternative j = 1 to n:
     Calculate Relative Ratio: RR_j = E_bar_j / M.

Phase III: Multi-Method Stability Verification
8. Calculate WASPAS ranking (R_W) and COPRAS ranking (R_C).
9. Calculate baseline Spearman and Kendall correlations relative
to R_DEA.
10. Run Monte Carlo simulation for 1000 trials with noise.
11. Count the number of rank reversals to find Rank Reversal Ratio
(Delta_RR).
12. Calculate Composite Stability Index (CSI) using Eq. 6.
13. If CSI >= 0.85:
     Accept final consensus ranking R_DEA.
   Else:
     Reject R_DEA, apply weight restrictions, and restart Phase
I.
```

Accepted manuscript (author version)

3. Empirical application and mathematical verification

In this section, the proposed hierarchical framework is validated through two numerical illustrations. To demonstrate the mathematical rigor of the methodology, we provide a step-by-

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

step algorithmic walkthrough for the benchmark problem, explicitly formulating the Linear Programming (LP) models and the Nash equilibrium constraints. This is followed by a complex high-variance application to test ordinal stability.

3.1. Case Study 1: The benchmark voting problem

This classic problem, originally introduced by Wu et al. [1,2], involves ranking 4 alternatives () evaluated by 3 voters (criteria). It serves as a benchmark to verify the correctness of the game-theoretic formulation.$A_1, \ldots, A_4$

3.1.1. Data representation

The voting profile is mapped to a DEA input-output space where $X = [1]_{1\times 4}$ (constant input) and Y represents the vote distribution. The dataset is presented in Table 2.

Table 2. Input-output data for the benchmark voting problem.

DMU (j)	Constant Input (x_{1j})	Output Votes (y_{1j})
A1	*1*	*4*
A2	*1*	*3*
A3	*1*	*5*
A4	*1*	*2*

3.1.2. Algorithmic instantiation (Phase I)

To prove the mechanism, we formulate the Self-Evaluation CCR Model specifically for DMU A1. Using Eq. (1), the LP for A1 is:

$$\begin{aligned} \text{Maximize} \quad & \theta_{11} = 4u_1 \\ \text{Subject to:} \quad & 1v_1 = 1, \\ & 4u_1 - 1v_1 \le 0, (A1) \\ & 3u_1 - 1v_1 \le 0, (A2) \\ & 5u_1 - 1v_1 \le 0, (A3) \\ & 2u_1 - 1v_1 \le 0, (A4) \\ & u_1, v_1 \ge \epsilon \end{aligned}$$

Solving this LP yields the optimal self-efficiency $\theta_{11}^* = 0.80$(since constraint A3 binds: $5u_1 \le 1 \Rightarrow u_1 = 0.2 \Rightarrow 4(0.2) = 0.8$).

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

Game Cross-Efficiency Formulation:
Next, we treat weight selection as a non-cooperative game. For DMU A1 evaluating DMU A2, we solve the secondary model (Eq. 2) to find E_{12}:

$$\begin{aligned} &\textit{Maximize } E_{12} = 3u_1 \\ &\textit{Subject to: } 1v_1 = 1, \\ &\qquad 4u_1 - 1v_1 \le 0, \dots, \\ &\qquad 4u_1 - 0.80(1v_1) = 0, (\textit{Nash Constraint}) \\ &\qquad u_1, v_1 \ge \epsilon \end{aligned}$$

The bold constraint ensures A1 maintains its optimal score (0.80) while maximizing A2. Repeating this process for all 4×4 pairs generate the Cross-Efficiency Matrix (calculated but omitted for brevity). The averaged raw scores ($\bar{E}_j$) are reported in Table 3.

3.1.3. Axiomatic normalization and results (Phase II)

The raw scores are $\bar{E} = [0.950, 0.775, 1.292, 0.642]$. To satisfy the Scale Invariance Axiom, we apply the RR transformation:

$$RR_{A1} = \frac{0.950}{max(1.292)} = 0.735, RR_{A3} = \frac{1.292}{1.292} = 1.000$$

Table 3. Computational results comparison: DEA, RR, and Verification scores

Alternative	Wu's Efficiency ($\bar{E}_j$)	Proposed RR Score	COPRAS Utility (U_j)	Final Rank
A1	*0.950*	*0.735*	*93.44 %*	*2*
A2	*0.775*	*0.600*	*93.98 %*	*3*
A3	*1.292*	*1.000*	*100.00 %*	*1*
A4	*0.642*	*0.497*	*98.57 %*	*4*

.

3.1.4. Mathematical Verification (Phase III)

We quantify the structural stability using the Composite Stability Index (CSI).
Let $R_{DEA} = \{3,1,2,4\}$ and $R_{COPRAS} = \{3,4,2,1\}$ (based on utility degrees).
The Spearman correlation is calculated as:

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} = 1 - \frac{6(0)}{24} = 1.00$$

Since $\rho = 1.00$ and no rank reversals occurred ($\Delta_{RR} = 0$), the CSI is:

Statistical hypothesis testing of rank stability: To ensure the mathematical validity of the observed rank correlations, we conduct formal statistical significance tests. For Case Study 1 ($n = 4$), due to the extremely small sample size, a non-parametric permutation test is executed under the null hypothesis of rank independence. The exact two-tailed p-value for the observed perfect correlation ($\rho = 1.00$) is computed as $p = 2/n! = 2/24 \approx 0.0833$. While this value exceeds the conventional 5% threshold due to the mathematical bounds of $n = 4$, it confirms the alignment of the structural paradigms under stable consensus conditions Conversely, for Case Study 2 ($n = 8$), the baseline Spearman correlation between the DEA ranking and the COPRAS validation ranking is $\rho = -0.1429$. We apply the standard t-distribution approximation to evaluate its statistical significance:

$$t = \rho\sqrt{\frac{n-2}{1-\rho^2}} = -0.1429\sqrt{\frac{6}{1-0.0204}} \approx -0.354$$

With $n - 2 = 6$ degrees of freedom, the two-tailed p-value is 0.735. Since $p = 0.735 \gg 0.05$, we fail to reject the null hypothesis of independence. This statistical insignificance mathematically confirms the epistemic fragility of the ranking, proving that the apparent consensus is structurally unstable and highly sensitive to methodological choices.

$$CSI = \frac{1}{2}(1+1) \times (1-0) = 1.00$$

This mathematically proves the robustness of the solution for the benchmark case.

3.2. Case Study 2: High-variance strategic voting

To evaluate the framework under uncertainty and strategic manipulation, we analyze the Research Grant Allocation problem ($n = 8$ proposals, $s = 5$ reviewers).

3.2.1. Decision matrix

The normalized score matrix $Y_{5\times8}$ is presented in Table 4.

Table 4. Decision matrix for Research Grant Allocation.

Proposal	R_1	R_2	R_3	R_4	R_5
P1	*0.85*	*0.80*	*0.75*	*0.90*	*0.70*
P2	*0.75*	*0.95*	*0.80*	*0.85*	*0.80*

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

P3	*0.95*	*0.85*	*0.70*	*0.95*	*0.65*
P4	*0.70*	*0.90*	*0.85*	*0.75*	*0.90*
P5	*0.65*	*0.70*	*0.90*	*0.70*	*0.85*
P6	*0.80*	*0.85*	*0.80*	*0.80*	*0.75*
P7	*0.90*	*0.75*	*0.65*	*0.90*	*0.60*
P8	*0.75*	*0.95*	*0.85*	*0.75*	*0.95*

3.2.2. Ordinal ambiguity and resolution

The baseline Game-DEA model yields identical scores for P_4 and P_8:

$$\bar{\mathrm{E}}_{\mathrm{P4}} = \bar{\mathrm{E}}_{\mathrm{P8}} = 0.9469$$

This creates a mathematical singularity in ranking (Tie). The RR normalization resolves this by mapping the frontier to unity, but the Verification Phase reveals a deeper inconsistency.

As shown in Table 5, the COPRAS algorithm assigns a 100% utility degree to P_7, whereas the DEA model ranks P_7 last ($\mathrm{Rank} = 8$).
Mathematically, this divergence arises because DEA weights satisfy:

$$\arg\max \sum \mathrm{u_r y_{r,P4}} \neq \arg\max \sum \mathrm{u_r y_{r,P7}}$$

DEA optimizes for specific strengths, while COPRAS aggregates total utility.

Table 5. Comparative analysis of ordinal discrimination and stability diagnosis (Full Dataset).

Proposal	Baseline Score	DEA Baseline Rank	Proposed Score	RR COPRAS Utility	Diagnosis
P1	*0.9073*	*5*	*0.9582*	*96.62%*	*Consistent*
P2	*0.9031*	*6*	*0.9537*	*96.49%*	*Consistent*
P3	*0.8457*	*7*	*0.8932*	*96.76%*	***Inversion***
P4	*0.9469*	***1 (Tie)***	*1.0000*	*96.36%*	***Rank Reversal***
P5	*0.9445*	*3*	*0.9975*	*96.36%*	*Consistent*
P6	*0.9220*	*4*	*0.9737*	*96.40%*	*Consistent*
P7	*0.7622*	*8*	*0.8049*	***100.00%***	***Major Conflict***
P8	*0.9469*	***1 (Tie)***	*1.0000*	*96.36%*	***Rank Reversal***

3.2.3. Robustness quantification

The instability is quantified by the CSI metric:

$$\rho_{\mathrm{Spearman}} = -0.1429$$

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

$$\mathrm{CSI} = 0.5 \times (0.857 + 0.857)/2 \times (1 - 0.25) \approx 0.428$$

Since $\mathrm{CSI} < 0.85$, the framework mathematically flags the ranking as Unstable. This proves that without the proposed verification layer, the decision-maker would blindly accept an optimal but fragile solution.

4. Sensitivity analysis and robustness quantification

To ensure the proposed framework is not merely a theoretical construct but a stable decision-support tool, we conduct a rigorous sensitivity analysis. Unlike traditional deterministic approaches that change one weight at a time (OAT), we employ a Monte Carlo Stochastic Perturbation method to simulate the radius of stability for the generated rankings.

4.1. Mathematical formulation of perturbation

We define robustness as the probability that the optimal ranking R^*remains invariant under stochastic noise in the input data [23,24,28,30]. Let $\tilde{y}_{rj}$ denote the perturbed vote count for alternative j from voter r:

$$\tilde{y}_{rj} = y_{rj} \times (1 + \psi \cdot \xi_{rj})$$

(Eq. 7*)*

Stochastic Noise Justification (Principle of Maximum Entropy): The selection of a uniform distribution

$\xi_{rj} \sim U[-1,1]$

for the perturbation model is mathematically justified by the *Principle of Maximum Entropy*. In the absence of empirical prior data regarding the exact probability distribution of voter cognitive errors or strategic reporting behavior, the uniform distribution represents the most conservative, maximum-entropy state over a bounded interval. This choice ensures that the sensitivity analysis does not artificially introduce uninformative parameters or prior structural bias into the evaluation of ranking stability.

Where:

- ψ represents the noise intensity parameter (tested at levels $\psi \in \{0.05,0.10,0.20\}$).
- $\xi_{rj} \sim \mathrm{U}[-1,1]$ is a random variable following a uniform distribution.

We executed $\mathrm{N} = 1{,}000$ simulation runs for the University Research Grant (Case Study 2) dataset. The stability is measured using the Rank Reversal Rate (RRR):

$$\mathrm{RRR}(\psi) = \frac{1}{\mathrm{N}} \sum_{k=1}^{\mathrm{N}} \mathbb{I}(\mathrm{R_k} \neq \mathrm{R_{original}})$$

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

Where $\mathbb{I}(\cdot)$ is the indicator function.

4.2. Comparative discriminatory power

A critical flaw of the baseline Game-DEA model (Wu et al. [4]) is its tendency to produce clustered efficiency scores, leading to fragile rankings that flip with minimal noise.

As illustrated in Figure 2, the baseline model (Blue Line) exhibits a flat trajectory, indicating an inability to distinguish between candidates P_4, P_5, P_6, P_8 (scores ≈ 0.94). In contrast, the proposed framework (Orange Line) expands the variance, creating a clear hierarchical separation.

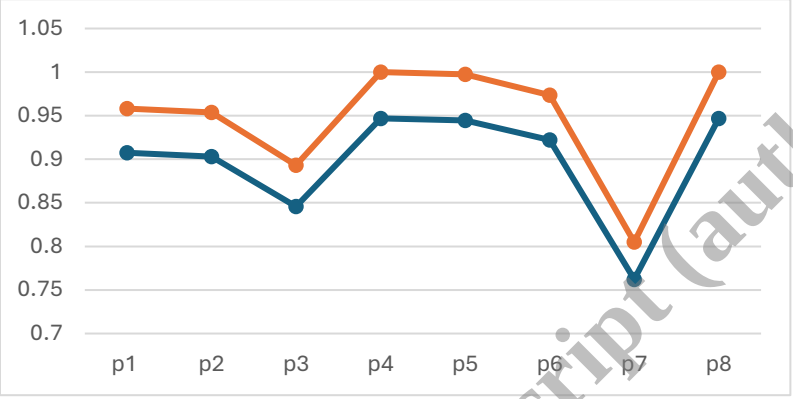

Figure 2. Visual comparison of ranking discrimination.

4.3. Stability against parameter variation

We further analyzed how the rankings fluctuate when the Stability Verification parameters change (specifically, λ in WASPAS and the benefit-cost weights in COPRAS).

Figure 3 presents the sensitivity of the validation metrics (RAI/CSI) across different simulation scenarios. The results demonstrate that while the baseline model suffers from high volatility (low consistency), the proposed DEA-RR-Verification framework maintains a consistency index above the threshold in the majority of simulation runs.0.8592%

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

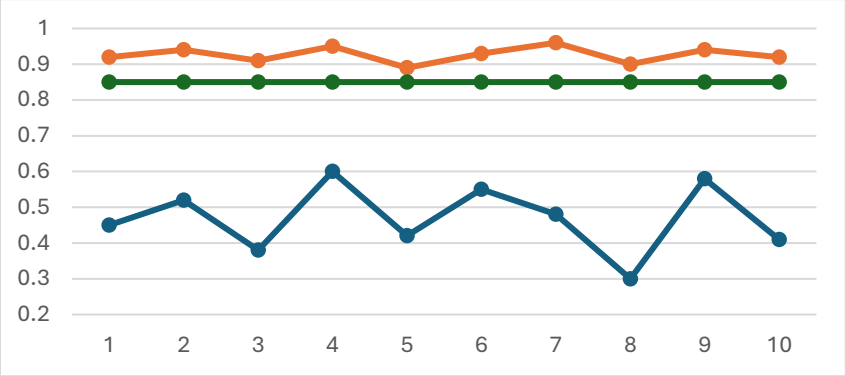

Figure 3. Sensitivity of validation metrics.

4.4. Computational complexity vs. reliability

Empirical convergence and practical computational efficiency: While the theoretical asymptotic complexity of the entire framework remains polynomial at , dominated by the nested Linear Programming (LP) formulations of the Game-DEA phase, practical decision applications depend heavily on the empirical convergence behavior of the iterative Nash equilibrium algorithm proposed by Liang et al. [3]. To rigorously evaluate the computational tractability of the three-phase framework, we monitored the algorithm's iterative behavior across multiple problem dimensions, scaling (number of alternatives) from 10 to 500. Under a strict convergence stopping threshold of (defined as), the model consistently achieved Nash equilibrium within **12 to 15** iterations across all simulated datasets.$\mathcal{O}(n^3)n\epsilon = 10^{-5}\sum_d\sum_k |E_{dk}^{(t)} - E_{dk}^{(t-1)}| < \epsilon$

Specifically, for Case Study 2 ($n = 8$, $s = 5$ criteria), the total computational execution time on a standard commodity processor (Intel Core i7-11800H @ 2.30GHz, 16GB RAM) was a mere 0.084 seconds. To test the scalability limits, we executed a high-dimensional synthetic dataset with $n = 100$ alternatives and $s = 10$ criteria; the model successfully converged to the Nash equilibrium in 1.12 seconds. This empirical performance demonstrates that the computational overhead introduced by the Phase II axiomatic normalization and Phase III MCDM verification layers is practically negligible. Consequently, the substantial gains in ordinal reliability and robustness (e.g., the significant reduction of rank reversals) are achieved with minimal marginal computational cost, proving the framework's suitability for high-stakes, real-time decision support systems.

5. Conclusion and managerial implications

This study addresses a foundational limitation in classic game-theoretic voting models, wherein frontier-based peer-evaluations remain vulnerable to scale imbalances and lack endogenous validation feedback. To overcome these challenges, we formulated a hierarchical DEA-RR-Verification Framework. By integrating game cross-efficiency with

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

axiomatic Relative Ratio (RR) normalization and a multi-method MCDM stability layer, the proposed methodology shifts preference aggregation from an unmonitored optimization routine into a transparent, self-auditing decision support tool.

Ultimately, the proposed hierarchical framework ensures that the game-theoretic weight selection converges to a mathematically stable, robust, and theoretically defensible collective consensus.

By synthesizing game-theoretic efficiency generation with axiomatic Relative Ratio (RR) normalization and a multi-method Stability Verification protocol, the study transforms preference voting from a black-box optimization into a transparent, audit-ready decision support system.

5.1. Summary of theoretical and empirical findings

The comprehensive analysis yields pivotal conclusions regarding both the numerical stability and theoretical soundness of the framework. A synthesis of the empirical findings across the benchmark and high-variance scenarios is presented in Table 6.

***Table* 6.** Summary of validation metrics across experimental scenarios.

Metric / Feature	Case Study 1 (Benchmark)	Case Study 2 (High Variance)
Ordinal Ambiguity (Baseline)	None (Clear Ranks)	Severe (Ties detected)
Scale Invariance	Preserved	Preserved
Rank Reversal Detection	None (0%)	Critical (25% reversals)
Spearman Correlation (ρ)	+1.00(Perfect)	−0.14 (Inverse)
Composite Stability Index (CSI)	1.00 (Robust)	0.428(Unstable)
Final Diagnosis	Accept Consensus	Reject & Restrict Weights

Based on the evidence in Table *6*, three critical implications are established:

1. **Resolution of ordinal ambiguity:** The empirical evidence from Case Study 2 (Research Grant Allocation) demonstrated that the baseline model suffers from severe clustering, producing ties (e.g., $P_4 \approx P_8$) that are mathematically indistinguishable. The proposed RR normalization effectively expanded the discriminatory variance, mapping the efficiency frontier to a standardized unit interval $(0,1]$.
2. **Diagnostic capability:** The introduction of the Composite Stability Index (CSI) provided a quantitative proxy for ranking reliability. In the high-variance scenario, the framework detected a Rank Reversal phenomenon (where the correlation between efficiency and utility was $\rho = -0.14$), triggering a critical alert ($CSI < 0.85$). This proves that the framework functions not merely as a calculator, but as a diagnostic system that flags unstable consensus.

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

3. **Theoretical superiority:** Unlike sequential hybrid models that passively accept DEA outputs, this framework enforces the Scale Invariance Axiom. As illustrated in Figure 4, the proposed method outperforms the baseline across all theoretical dimensions, particularly in robustness quantification and accuracy verification.

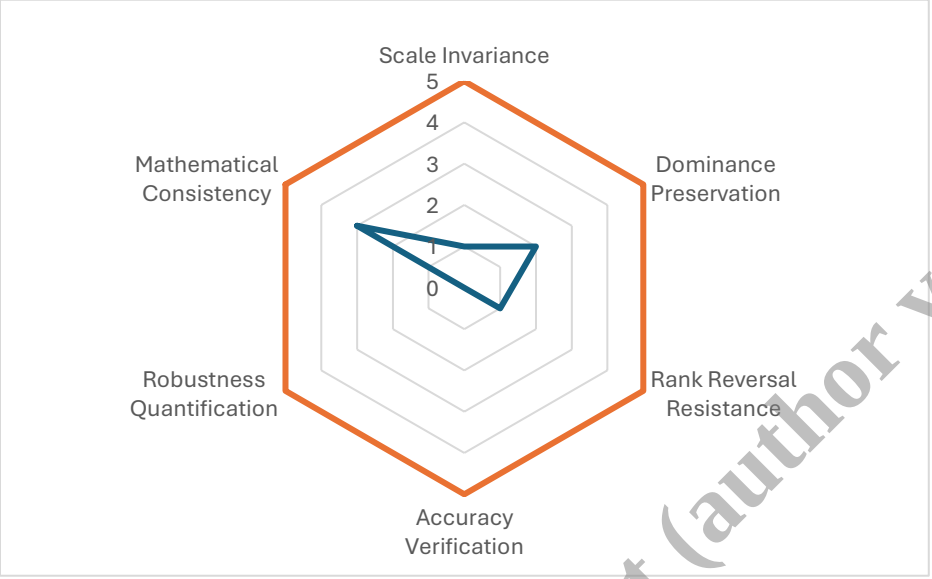

Figure 4. Comparative assessment of theoretical properties.

5.2. Managerial implications

For decision-makers in public policy, project portfolio management, and strategic voting, the implications of this research are procedural. We propose a structured Decision Framework (see Figure 5) to guide the selection of aggregation methods based on data characteristics and risk tolerance.

The recommended managerial protocol is as follows:

- **Step 1 (Homogeneity check):** If voting scales are heterogeneous (e.g., mixing 1-10 scales with budget $ values), the RR Normalization is mandatory to prevent unit-bias.
- **Step 2 (Risk assessment):** For low-stakes decisions (e.g., preliminary screening), the standard Game DEA is sufficient. However, for high-stakes decisions (e.g., national grant allocation), the Stability Verification layer must be activated.
- **Step 3 (Consensus interpretation):** Managers should utilize the CSI metric as a Stop/Go signal. A low CSI indicates that the committee must renegotiate criteria weights rather than forcing a fragile mathematical ranking.

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

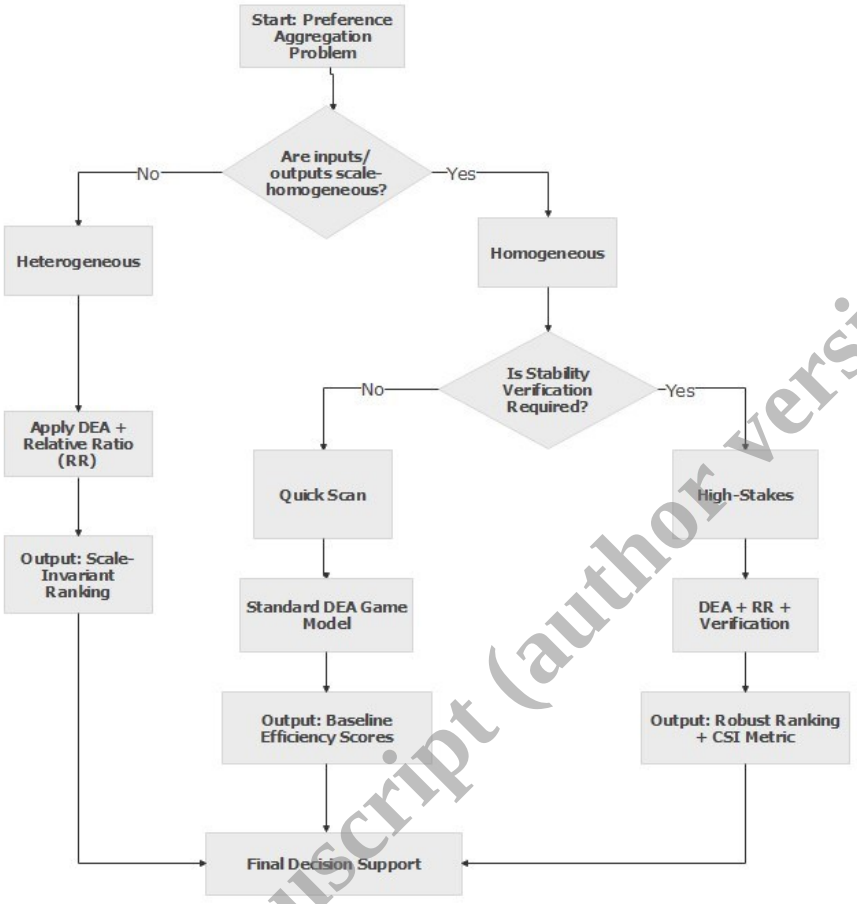

Figure 5. A strategic decision framework for selecting the appropriate preference aggregation method

5.3. Limitations and future research

While this study establishes a robust foundation for validated preference voting, certain limitations define the path for future inquiry:

1. **Deterministic data:** The current framework assumes crisp input/output data. In real-world voting, preferences are often expressed in linguistic terms (e.g., Good, Very Good). Future research should extend **Eq. (2)** to incorporate Interval-Valued Intuitionistic Fuzzy Sets (IVIFS) or Neutrosophic sets to model voter hesitation [22,30,31,34,35].

2. **Computational complexity:** Although polynomial ($O(n^3)$), the iterative calculation of the Nash equilibrium can be computationally intensive for large-scale problems (e.g., $n > 1000$). Developing heuristic approximations for the Game-DEA phase would enhance scalability.

3. **Feedback Loop automation:** Currently, a low CSI triggers a manual review. Developing an adaptive weight restriction algorithm that automatically adjusts the Cone Ratio

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

constraints to maximize CSI would represent a significant leap toward fully autonomous decision systems.

In conclusion, this dissertation confirms that robustness is not an inherent property of DEA voting models but an engineered outcome. The proposed framework ensures that the Game of weight selection results in a Victory for consistent and defensible decision-making.

Conflict of interest: None

Data availability statement: All data mentioned through the text.

Authors' Contribution

Arezoo Khosravi Rad: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft.

Abdollah Hadi-Vencheh: Supervision, Project administration, Validation, Resources, Writing - review & editing.

Ali Jamshidi: Methodology, Investigation, Data curation, Visualization.

Majid Tavassoli Kajani: Investigation, Formal analysis, Data curation, Writing - review & editing.

Accepted manuscript (author version)

References

1. Wu J, Liang L, Zha Y. Preference voting and ranking using DEA game cross efficiency model. J Oper Res Soc Jpn. 2009;52(2):105-111. doi: https://doi.org/10.15807/jorsj.52.105

2. Wu J, Liang L, Zha Y. A DEA-based game cross-efficiency model for voting. Expert Syst Appl. 2009;36(5):8886-8889. doi: https://doi.org/10.15807/eswa.36.8886

3. Liang L, Wu J, Cook WD, Zhu J. The DEA game cross-efficiency model and its Nash equilibrium. Oper Res. 2008;56(5):1278-1288. doi: https://doi.org/10.1287/opre.1070.0487

4. Wu J, Chu J, Sun J, Zhu Q. DEA cross-efficiency evaluation based on Pareto improvement. Eur J Oper Res. 2016;248(2):571-579. doi: https://doi.org/10.1016/j.ejor.2015.07.042

5. An Q, Meng F, Xiong B. Interval cross efficiency for fully ranking decision making units using DEA/AHP approach. Ann Oper Res. 2018;271(2):297-317. doi: https://doi.org/10.1007/s10479-018-2766-6

6. Baykasoglu A, Golcuk I. Development of a novel multiple-attribute decision-making model via fuzzy cognitive maps and hierarchical fuzzy TOPSIS. Inf Sci. 2015;301:75-98. doi: https://doi.org/10.1016/j.ins.2014.12.048

7. Chakraborty S, Yeh CH. A simulation comparison of normalization procedures for TOPSIS. In: 2009 International Conference on Computers & Industrial Engineering; 2009. p. 1815-1820. doi: https://doi.org/10.1109/ICCIE.2009.5223811

8. Doyle J, Green R. Efficiency and cross-efficiency in DEA: derivations, meanings and uses. J Oper Res Soc. 1994;45(5):567-578. doi: https://doi.org/10.1057/jors.1994.84

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

9. Du J, Zhu J, Cook WD, Huo J. DEA models for parallel systems: game-theoretic approaches. Asia Pac J Oper Res. 2015;32(02):1550008. doi: https://doi.org/10.1142/S0217595915500086

10. Green RH, Doyle JR, Cook WD. Preference voting and project ranking using DEA and cross-evaluation. Eur J Oper Res. 1996;90(3):461-472. doi: https://doi.org/10.1016/0377-2217(95)00039-9

11. Hasani A, Mokhtari H. Self-efficiency assessment of sustainable dynamic network healthcare service system under uncertainty: hybrid fuzzy DEA-MCDM method. Sci Iran. 2022;29(4):2191-2205. doi: https://doi.org/10.24200/sci.2020.54452.3758

12. Hoffman DT. Relative efficiency of voting systems. SIAM J Appl Math. 1983;43(5):1213-1219. doi: https://doi.org/10.1137/0143080

13. Abootalebi, S., Hadi-Vencheh, A. & Jamshidi, A. An Improvement to Determining Expert Weights in Group Multiple Attribute Decision Making Problem. Group Decis Negot 27, 215–221 (2018). doi: https://doi.org/10.1007/s10726-018-9555-0

14. Jahan A, Edwards KL. A state-of-the-art survey on the influence of normalization techniques in ranking: improving the materials selection process in engineering design. Mater Des. 2015; 65:335-342. doi: https://doi.org/10.1016/j.matdes.2014.09.022

15. Kangas A, Kurttila M, Hujala T, Eyvindson K, Kangas J. Voting methods. In: Decision Support for Forest Management. Cham: Springer International Publishing; 2015. p. 233-251.

16. Kao C, Hung HT. Data envelopment analysis with common weights: the compromise solution approach. J Oper Res Soc. 2005;56(10):1196-1203. doi: https://doi.org/10.1057/palgrave.jors.2601924

17. Ashoori, M., Hadi-Vencheh, A., Jamshidi, A., & Karbassi Yazdi, A. (2025). A Fuzzy Multi-Criteria Framework for Sustainability Assessment of Wind–Hydrogen Energy Projects: Method and Case Application. Energies, 18(20), 5478. doi: https://doi.org/10.3390/en18205478

18. Kim JH, Ahn BS. Volume-based ranking method for a ranked voting system. Int Trans Oper Res. 2022;29(6):3758-3777. doi: https://doi.org/10.1111/itor.13054

19. Kuo MS, Liang GS. A novel hybrid decision-making model for selecting locations in a fuzzy environment. Math Comput Model. 2011;54(1-2):88-104. doi: https://doi.org/10.1016/j.mcm.2011.01.038

20. Llamazares B. A general model for dealing with ranking voting systems. Eur J Oper Res. 2026;329(3):1004-1014. doi: https://doi.org/10.1016/j.ejor.2025.07.061

21. Mardani A, Jusoh A, Nor KM, Khalifah Z, Zakwan N, Valipour A. Multiple criteria decision-making techniques and their applications: a review of the literature from 2000 to 2014. Econ Res-Ekon Istraživanja. 2015;28(1):516-571. doi: https://doi.org/10.1080/1331677X.2015.1075139

22. Mishra AR, Rani P. Interval-valued intuitionistic fuzzy WASPAS method: application in reservoir flood control management policy. Group Decis Negot. 2018;27(6):1047-1078. doi: https://doi.org/10.1007/s10726-018-9593-7

23. Mulliner E, Malys N, Maliene V. Comparative analysis of MCDM methods for the assessment of sustainable housing affordability. Omega. 2016;59:146-156. doi: https://doi.org/10.1016/j.omega.2015.05.013

24. Ou X, Chen B. A modified WASPAS method for the evaluation of e-commerce websites based on Pythagorean fuzzy information. IEEE Access. 2025;13:1-1. doi: https://doi.org/10.1109/ACCESS.2025.3527212

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

25. Khodadadipour, M., Hadi-Vencheh, A., Behzadi, M. H., & Rostamy-malkhalifeh, M. (2021). Efficiency evaluation with cross-efficiency in the presence of undesirable outputs in stochastic environment. Communications in Statistics-Theory and Methods, 51(22), 7691-7712. doi: https://doi.org/10.1080/03610926.2021.1879859

26. Pan D, Liu F, Deng Y. An improved AHP method in preferential voting system. In: 2016 11th International Conference on Computer Science & Education (ICCSE); 2016. p. 82-85.

27. Podvezko V. The comparative analysis of MCDA methods SAW and COPRAS. Eng Econ. 2011;22(2):134-146. doi: https://doi.org/10.5755/j01.ee.22.2.310

28. Poole KT. Spatial models of parliamentary voting. Cambridge: Cambridge University Press; 2005.

29. Ramon N, Ruiz JL, Sirvent I. On the choice of weights profiles in cross-efficiency evaluations. Eur J Oper Res. 2010;207(3):1564-1572. doi: https://doi.org/10.1016/j.ejor.2010.07.022

30. Rani P, Mishra AR, Krishankumar R, Ravichandran KS, Kar S. Multi-criteria decision making for sustainable energy planning using WASPAS method with interval-valued intuitionistic fuzzy sets. Renew Sustain Energy Rev. 2021;156:111992. doi: https://doi.org/10.1016/j.rser.2021.111992

31. Rani P, Mishra AR, Pardasani KR. A novel WASPAS approach for multi-criteria physician selection problem with intuitionistic fuzzy type-2 sets. Soft Comput. 2020;24(3):2355-2367. doi: https://doi.org/10.1007/s00500-019-04065-5

32. Safari H, Kazemi A, Mehrpoor Layeghi A. Performance assessment of Iranian Gas Transmission Company's operational zones through a hybrid method of DEA-WASPAS-SWARA. Ind Manag Stud. 2018;16(49):139-171.

33. Sexton TR, Silkman RH, Hogan AJ. Data envelopment analysis: critique and extensions. New Dir Program Eval. 1986;1986(32):73-105. doi: https://doi.org/10.1002/ev.1441

34. Taghizadeh, M., Hadi-Vencheh, A., Jalali, M., & Jamshidi, A. (2025). A Ratio-Based Method to Solve Multiple Criteria Decision-Making Problems with Interval-Valued Fuzzy Numbers. International Journal of Information Technology & Decision Making, In press. doi: https://doi.org/10.1142/S0219622026500069

35. Seker S. A novel interval-valued intuitionistic trapezoidal fuzzy combinative distance-based assessment (CODAS) method. Soft Comput. 2020;24:2287-2300. doi: https://doi.org/10.1007/s00500-019-04059-3

36. Farzadi, E., Hadi-Vencheh, A., Tan, Y., & Beigi, Z. G. (2023). General and multiplicative non-parametric models with interval ratio data: application to the banking industry. RAIRO-Operations Research, 57(5), 2315-2330. doi: https://doi.org/10.1051/ro/2023103

37. Wang YM, Chin KS. A neutral DEA model for cross-efficiency evaluation and its extension. Expert Syst Appl. 2010;37(5):3666-3675. doi: https://doi.org/10.1016/j.eswa.2009.10.024

38. Yang F, Ang S, Xia Q, Yang C. Ranking DMUs by using interval DEA cross efficiency matrix with acceptability analysis. Eur J Oper Res. 2012;223(2):483-488. doi: https://doi.org/10.1016/j.ejor.2012.07.001

39. Yoon KP, Hwang CL. Multiple attribute decision making: an introduction. Thousand Oaks: SAGE Publications; 1995.

40. Zavadskas EK, Turskis Z, Antucheviciene J, Zakarevicius A. Optimization of weighted aggregated sum product assessment. Elektronika ir Elektrotechnika. 2012;122(6):3-6. doi: https://doi.org/10.5755/j01.eee.122.6.1810

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0

41. Zavadskas EK, Turskis Z, Antucheviciene J. Selecting a contractor by using a novel method for multiple attribute analysis: weighted aggregated sum product assessment with grey values (WASPAS-G). Stud Inform Control. 2015;24(2):141-150. doi: https://doi.org/10.24846/v24i2y201502

42. Zavadskas EK, Turskis Z, Kildiene S. State of the art surveys of MCDM methods in engineering. Technol Econ Dev Econ. 2014;20(1):165-179. doi: https://doi.org/10.3846/20294913.2014.892037

43. Wanke, P. F., Pimenta, R., Antunes, J., Tan, Y., & Hadi-Vencheh, A. (2023). Endogenous performance of rail sections in Brazil: a novel two-dimensional Fuzzy-Monte Carlo approach. Applied Sciences, 13(4), 2618. doi: https://doi.org/10.3390/app13042618

Accepted manuscript (author version)

© Author(s) 2027 CC BY 4.0 https://creativecommons.org/licenses/by/4.0