Article

Not peer-reviewed version

Meritocratic Prize Distribution via Nash-Antifragil Mechanism

Roberto Riva *

Posted Date: 4 July 2025

doi: 10.20944/preprints202507.0359.v1

Keywords: game theory; nash equilibrium; mreitocracy; antifragile systems; performance-based rewards; incentive mechanism; awards and prizes

 doi:10.20944/preprints202507.0359.v1

Article

Meritocratic Prize Distribution via Nash-Antifragile Mechanism

Roberto Riva

Affliation; orloch314@gmail.com

Abstract

This paper introduces a novel incentive distribution mechanism rooted in game theory, leveraging a non-collusive Nash equilibrium and the concept of antifragility. Participants are assessed on multiple objective criteria by a single evaluator, and only those who exceed a minimum score threshold become eligible for a share of a fixed prize pool. The rewards are allocated proportionally among these top performers, creating strong incentives for excellence while naturally discouraging underperformance. The mechanism is simple to implement, resistent to manipulation, and becomes more rewarding in the presence of poor performers. We present a full mathematical formalization, analyze key theoretical properties including meritocracy and Pareto-efficiency, and support our findings with a simulation on real-world-like data. The mechanism has also been successfully applied in a private company in Paraguay over a two-year period to reward approximately 70 employees.

Keywords: game theory; nash equilibrium; mreitocracy; antifragile systems; performance-based rewards; incentive mechanism; awards and prizes

---

1. Introduction

In organizational contexts where individual results are objectively measurable, distributing economic incentives represents a central issue in terms of fairness, efficiency, and transparency.

This work introduces a prize allocation mechanism grounded in game theory, particularly inspired by the Nash equilibrium [1] and antifragility [2]. The system allocates rewards only to those who exceed a fixed performance threshold, calculated using a consistent evaluation method. This results in a natural meritocratic effect, penalizing low performance and amplifying the rewards for excellence.

2. Mathematical Model

Let n be the number of participants and v_i the performance score of participant i.
Let s be a predefined minimum score threshold.
Define the set of eligible participants:

$$E = i \mid v_i > s \tag{1}$$

Let P be the total prize pool. Then, for each $i \in E$, their reward r_i is:

$$r_i = \frac{v_i}{\sum_{j \in E} v_j} \cdot P \tag{2}$$

To prevent division by zero, we assume:

$$\sum_{j \in E} v_j > 0 \tag{3}$$

This is guaranteed if at least one $v_i > s$.

 doi:10.20944/preprints202507.0359.v1

If no participants exceed the threshold, then $E = \emptyset$, and no rewards are distributed. A practical fallback rule is to either distribute P equally among all participants, or defer the prize to the next period.

To improve robustness and break ties, a small perturbation ϵ can be added to each v_i:

$$v_i^{\epsilon} = v_i + \epsilon_i, \quad \epsilon_i \sim \mathcal{U}(0, \delta) \tag{4}$$

where $\delta > 0$ is a small noise level.

2.1. Score Aggregation

Each participant i is evaluated on a set of m performance criteria. Let $s_i^{(k)}$ denote the score of participant i on criterion k, and w_k be the weight assigned to criterion k, with:

$$w_k \geq 0, \quad \sum_{k=1}^{m} w_k = 1$$

The final performance score v_i is computed as a weighted sum:

$$v_i = \sum_{k=1}^{m} w_k \cdot s_i^{(k)} \tag{5}$$

This formulation allows the model to flexibly reflect the relative importance of each criterion while maintaining linearity and transparency.

3. Properties of the Mechanism

3.1. Nash Equilibrium

Since participants do not control each other's scores, and evaluation is exogenous, they cannot directly influence eligibility beyond their own effort. Any attempt to reduce performance decreases their own chance of being rewarded. Thus, increasing one's score is a dominant strategy as long as it raises it above s [3].

3.2. Formal Proof of Nash Equilibrium Existence

Theorem 1. *Under the following assumptions:*

1. *Performance scores v_i are determined by individual effort $e_i \in [0, M]$ through continuous functions $v_i = g_i(e_i)$ with g_i strictly increasing and concave.*
2. *Effort cost $c_i(e_i)$ is continuous, strictly increasing and convex.*
3. *The threshold s is exogenously fixed.*

the game admits a pure-strategy Nash equilibrium.

Proof. We prove existence using the Debreu-Glicksberg-Fan theorem [4], which requires:
1. Compact and convex strategy space: For each player i, the effort space $[0, M]$ is:

- *Compact* (closed and bounded in $\mathbb{R}$)
- *Convex* (any linear combination $\lambda e_i + (1 - \lambda)e_i' \in [0, M]$ for $\lambda \in [0, 1]$)

2. Upper semi-continuous payoff functions: The payoff $u_i(e_i, \mathbf{e}_{-i}) = r_i(e_i, \mathbf{e}_{-i}) - c_i(e_i)$ has discontinuities only when $v_i = s$ or $v_j = s$ for some j. These critical points form a set of measure zero, and at discontinuity points we have:

$$\limsup_{(e_i', \mathbf{e}_{-i}') \to (e_i, \mathbf{e}_{-i})} u_i(e_i', \mathbf{e}_{-i}') \leq u_i(e_i, \mathbf{e}_{-i})$$

3. Quasi-concavity in own strategy: For fixed $\mathbf{e}_{-i}$, $u_i(e_i, \mathbf{e}_{-i})$ is quasi-concave in e_i. Consider two cases:

 doi:10.20944/preprints202507.0359.v1

- **Case 1:** $g_i(e_i) < s$
  Then $u_i = -c_i(e_i)$. Since $-c_i$ is concave (and thus quasi-concave), u_i is quasi-concave.
- **Case 2:** $g_i(e_i) \geq s$
  Let $V' = \sum_{j \neq i, j \in E} g_j(e_j)$. The reward component is:

$$r_i = P \cdot \frac{g_i(e_i)}{V' + g_i(e_i)}$$

The function $\frac{g_i(e_i)}{V' + g_i(e_i)}$ is concave in e_i because:

$$\frac{\partial^2}{\partial e_i^2}\left(\frac{g_i}{V' + g_i}\right) = \underbrace{\frac{g_i''}{(V' + g_i)}}_{\leq 0} - \underbrace{\frac{2g_i'^2(V' + g_i) + 2g_i'^2 g_i}{(V' + g_i)^3}}_{\leq 0} \leq 0$$

since g_i is concave ($g_i'' \leq 0$) and $g_i' \geq 0$. The cost term $-c_i(e_i)$ is concave. The sum of concave functions is concave, hence quasi-concave.

Since all conditions are satisfied, a pure-strategy Nash equilibrium exists by the Debreu-Glicksberg-Fan theorem [4]. □

3.3. Strategic Transparency

The mechanism's score computation relies on a weighted aggregation of multiple performance criteria:

$$v_i = \sum_{k=1}^{m} w_k \cdot s_i^{(k)}$$

Even if the weighting scheme $\{w_k\}_{k=1}^m$ is made fully transparent and known to all participants, the mechanism remains robust against strategic underperformance. Since scores are assigned externally and eligibility depends solely on surpassing a fixed threshold not relative ranking participants have no incentive to lower their effort in any component.

Attempting to manipulate the outcome by neglecting low-weighted criteria or selectively focusing on high-weighted ones does not yield a reward-maximizing strategy if it results in a lower total score. Therefore, maximizing individual performance across all evaluated dimensions remains the dominant and stable behavior, regardless of participants' knowledge of the scoring formula.

3.4. Antifragility

The proposed reward mechanism exhibits antifragile behavior in the sense introduced by Taleb [2]: it benefits from disorder, volatility, and underperformance by others.

Formally, consider the set of eligible participants E, such that $i \in E$ if $v_i > s$. The reward allocated to participant $i \in E$ is:

$$r_i = P \cdot \frac{v_i}{\sum_{j \in E} v_j}$$

Holding v_i fixed, if other participants $j \neq i$ experience a decrease in their scores or fall below the eligibility threshold s, then $\sum_{j \in E} v_j$ decreases. As a result, r_i increases, since the denominator shrinks.

Proposition 1. *For any eligible participant i, the partial derivative of r_i with respect to the score v_j of another eligible participant $j \neq i$ is negative:*

$$\frac{\partial r_i}{\partial v_j} = -P \cdot \frac{v_i}{\left(\sum_{k \in E} v_k\right)^2} < 0$$

 doi:10.20944/preprints202507.0359.v1

Proof. Immediate from the expression of r_i, as the numerator is independent of v_j, and the denominator increases with v_j, thus reducing r_i. □

This structure implies that the system amplifies reward concentration when volatility occurs: the worse others perform, the more concentrated the reward becomes for those who remain above threshold.

Unlike fragile systems—where underperformance degrades collective outcomes—or robust ones—where performance fluctuations are neutralized—this mechanism actively favors volatility among non-top performers. This dynamic creates a self-reinforcing incentive for continuous improvement: outperforming others not only increases one's absolute score but also redistributes unclaimed rewards.

Thus, the mechanism is not only stable under variation but benefits from it, satisfying the definition of antifragility.

Clarification on Pareto Efficiency and Antifragility

It is important to distinguish two distinct levels of analysis.

First, the mechanism is *Pareto efficient within* the set of eligible participants E. Once eligibility is fixed, the proportional allocation implies that any reallocation of the prize pool among members of E would necessarily reduce the reward of at least one participant without increasing the reward of another.

Second, the *antifragile* nature of the mechanism concerns system-level dynamics: when fewer participants exceed the minimum threshold s, the remaining eligible participants receive larger shares of the prize pool. This creates an inherent resistance to volatility and underperformance among participants, which strategically benefits those who maintain or increase their performance, without violating Pareto efficiency within E.

3.5. Pareto Efficiency

Given a fixed prize pool P and a proportional distribution among eligible participants E, any reallocation of rewards among members of E that changes the current distribution necessarily decreases the reward of at least one participant without benefiting another, unless the composition of E itself changes. Hence, the allocation is Pareto efficient within the subset E. [5].

4. Numerical Simulation of the Reward Mechanism

We consider a simple scenario with five participants evaluated according to their performance scores v_i. The minimum eligibility threshold is fixed at $s = 50$, and the total prize pool is $P = 100$ monetary units.

Table 1. Performance scores and eligibility.

Participant	Score v_i	Eligible ($v_i > s$)
Alice	80	Yes
Bob	70	Yes
Carla	45	No
Diego	55	Yes
Eva	30	No

Only participants who exceed the threshold form the eligible set $E = \{\text{Alice}, \text{Bob}, \text{Diego}\}$. The total score among eligible participants is:

$$\sum_{j \in E} v_j = 80 + 70 + 55 = 205$$

The reward for each eligible participant is calculated as:

 doi:10.20944/preprints202507.0359.v1

$$r_i = \frac{v_i}{\sum_{j \in E} v_j} \cdot P$$

Table 2. Reward allocation based on proportional scoring.

Participant	Reward Formula	Reward r_i
Alice	$\frac{80}{205} \cdot 100$	39.02
Bob	$\frac{70}{205} \cdot 100$	34.15
Diego	$\frac{55}{205} \cdot 100$	26.83
Carla	— (ineligible)	0.00
Eva	— (ineligible)	0.00

4.1. Observations

The mechanism rewards participants proportionally within the eligible group. When some participants fail to reach the threshold, the others receive a higher reward. This illustrates the antifragile property of the system: the failure of some benefits the rest without violating fairness or efficiency among those who qualify.

4.2. Implementation

A fully functional Excel implementation of the reward mechanism, used to develop and test the ideas presented in this paper, is publicly available at the following GitHub repository: https://github.com/Orloch314/Nash-Antifragile. The spreadsheet allows users to input individual scores, adjust the threshold and prize pool parameters, and automatically computes the proportional rewards. It is designed for immediate use in organizational contexts and can serve both as a demonstration tool and as a practical instrument for incentive distribution.

5. Conclusions

We have presented a robust and meritocratic mechanism for the distribution of fixed rewards in competitive environments where objective performance scores are available. The mechanism integrates Nash-stable incentives with antifragile reward dynamics, offering both theoretical soundness and practical applicability.

Its simplicity, transparency, and resistance to collusion make it suitable for a wide range of operational contexts. The model has been applied in practice within a call center environment, where it showed potential to support motivation and fairness, although no formal data collection was conducted.

Future work will focus on empirical validation and possible extensions of the model, including dynamic prize pools, adaptive thresholds, and multi-period evaluation schemes.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

References

1. Nash, J. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences* **1950**, *36*, 48–49.
2. Taleb, N.N. *Antifragile: Things That Gain from Disorder*; Random House, 2012.
3. Myerson, R.B. *Game theory: Analysis of conflict*; Harvard University Press, 1991.
4. Debreu, G. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences* **1952**, *38*, 886–893.
5. Roth, A.E. Game Theory as a Tool for Market Design. In *The New Palgrave Dictionary of Economics*, 2 ed.; Durlauf, S.N.; Blume, L.E., Eds.; Palgrave Macmillan, 2008. https://doi.org/10.1057/9780230226203.2912.

 doi:10.20944/preprints202507.0359.v1