Eur. J. Math. Anal. 6 (2026) 19
doi: 10.28924/ada/ma.6.19

Safety Critical Thresholds and Stability Analysis of Collective Cooperation with Heterogeneous Contributions

Maju Wu

School of Mathematics, Tianjin University, Tianjin, 300354, China
wmj_1251@tju.edu.cn

Abstract. This paper addresses the problem of individual defection (voluntary exit) in collective cooperation. Based on an upper-bounded S-shaped contribution-payoff function, we construct a Homogeneous Equal-Contribution Model and a One-Strong-n-Weak Heterogeneous Model with a high-contribution agent. We strictly define and derive three core safety critical thresholds for collective cooperation: the failure threshold of equal allocation, the failure threshold of proportional allocation, and the full stability threshold. Through rigorous algebraic derivation and functional property analysis, we provide explicit closed-form solutions and solving methods for all thresholds, prove the self-consistency of the model, and reveal the mechanisms by which individual heterogeneity, group size, contribution intensity, and task attributes affect the stability of collective cooperation. This paper provides a complete and unified mathematical theoretical framework for the stability analysis of collective cooperation.

1. Introduction

The breakdown of collective cooperation often arises when an individual's payoff from the group is lower than their payoff from working independently, leading to voluntary exit. This core contradiction is the endogenous conflict between individual rationality and collective rationality, which has long been a central topic in cooperative game theory and social economics [1, 3]. Whether it is the dissolution of corporate teams, the breakdown of project alliances, or free-riding in public goods provision, the underlying logic is the failure to satisfy individual participation constraints [10].

Existing literature has made systematic contributions to this field: Fichtenberger H, Rey A [2] and Magallán A, Carreras F. [7] analyzed the core stability, strong stability, and equilibrium properties of coalition structures and established the game-theoretic foundation of coalition formation; Lin Y H, et al. [9] explored the underlying logic and evolutionary path of dynamic coalition formation and characterized the time-varying changes of coalition structures; Weikard H P. [15] and Yang S B, Tan Z F, Zhao R, et al. [16] focused on the incentive effects of fair

Received: 9 Apr 2026.
Key words and phrases. Collective cooperation stability; Heterogeneous contribution; Defection critical threshold; S-shaped payoff function; Allocation rules.

allocation rules and the design of optimal sharing rules, and analyzed the impact of different allocation mechanisms on the efficiency of coalition cooperation.

However, two key gaps remain in existing research. First, few studies have rigorously mathematically characterized the underlying critical condition under which full non-defection is guaranteed regardless of the allocation rule adopted, failing to define the absolute safety boundary for the survival of collective cooperation. Second, existing studies generally default to the assumption that "higher total group contribution leads to stronger cooperation stability", ignoring the realistic scenario that excessive total contribution, under the constraint of payoff upper bound, will lead to individual independent payoff exceeding group allocation payoff, thus triggering cooperation breakdown.

This paper takes rigorous mathematical derivation as the core, referring to the classical framework of mathematical modeling of social group behavior [12, 20, 21]. We first present the basic model and core definitions of collective cooperation, and rigorously prove the core properties of the payoff function. Second, we derive the general form and necessary and sufficient conditions for the three types of safety critical thresholds, and rigorously demonstrate the order relationship between the thresholds. Subsequently, we analyze the critical conditions, explicit solutions, and stability properties of the Homogeneous Equal-Contribution Model and the One-Strong-n-Weak Heterogeneous Model respectively, and fully derive the erosion mechanism of heterogeneity on cooperation stability. Finally, we present conclusions and policy implications. All thresholds in this paper are upper-bound thresholds: the allocation rule is valid and cooperation is stable when total contribution is below the threshold, while the rule fails and cooperation breaks down when total contribution exceeds the threshold, which fully fits the realistic economic phenomenon of "excessive investment leading to cooperation collapse".

2. Model Construction and Core Definitions

2.1. Basic Notation and Assumptions.

Let the set of individuals in the collective be $N = \{1, 2, \ldots, n\}$, where $n \geq 2$ (a collective includes at least 2 individuals). All individuals are fully rational decision-makers, with the sole objective of maximizing their own payoff, and have perfect information to accurately calculate their independent payoff and group allocation payoff [29]. The research ethics and privacy norms for individual decision-making refer to Hilmar N. [5].

The core notation of this paper is defined as follows:

- Let $x_i \geq 0$ be the effective contribution of individual i, and the total effective contribution of the collective is $X = \sum_{i=1}^{n} x_i$.
- The feasible allocation set Φ is defined as the set of all allocation schemes satisfying budget balance and non-negativity constraints:

$$\Phi = \left\{ (\varphi_1, \varphi_2, \ldots, \varphi_n) \,\middle|\, \sum_{i=1}^{n} \varphi_i = Y(X),\ \varphi_i \geq 0,\ \forall i \in N \right\}$$

where φ_i is the payoff allocated to individual i from the total group payoff. The allocation must satisfy two basic axioms: (1) Budget Balance Axiom: total payoff is fully allocated with no surplus; (2) Individual Rationality Axiom: individual allocated payoff is non-negative with no forced transfer payments. The axiomatic system of allocation rules refers to Bich P, Laraki R. [14] and Öztürk M. [17].

- Efficiency parameter $k > 0$, which characterizes the individualization attribute and co-operation dependence of the task: the larger k is, the faster the independent payoff grows with small individual contribution, the more suitable the task is for independent work, and the weaker the scale effect of cooperation; the smaller k is, the slower the independent payoff grows, the more the task relies on collective cooperation, and the stronger the scale effect of cooperation.
- Maximum payoff upper bound $A > 0$, which characterizes the payoff ceiling of the industry or scenario, a positive constant, meaning that the payoff cannot exceed this upper bound regardless of the individual or collective contribution. The distribution properties and economic implications of bounded payoffs refer to Oancea B, Pirjol D, Andrei T. [18].

2.2. **Contribution-Payoff Function and Core Properties.** This paper adopts an upper-bounded S-shaped contribution-payoff function, which conforms to the universal payoff law of "first increasing marginal returns, then decreasing marginal returns, and finally converging to the payoff upper bound" in reality, and is widely used in social dynamics and cooperative game modeling [11]. The specific form is as follows:

- Individual independent payoff function:

$$y(x) = \frac{Ak^2x^2}{1+k^2x^2}$$

- Total collective cooperation payoff function:

$$Y(X) = \frac{Ak^2X^2}{1+k^2X^2}$$

This function form ensures that individual independent work and collective cooperation follow exactly the same payoff law, with only the difference in contribution scale, eliminating the interference of function form differences on the results and ensuring the self-consistency of the model.

Lemma 2.1 (Core Properties of the Payoff Function). *The payoff function $y(x)$ satisfies the following four core properties:*

(1) *Non-negativity: For all $x \geq 0$, $y(x) \geq 0$ and $y(0) = 0$;*

(2) *Strict Monotonicity: $y(x)$ is strictly monotonically increasing on $[0, +\infty)$, i.e., for any $0 \leq x_1 < x_2$, $y(x_1) < y(x_2)$;*

(3) *Concavity and Convexity: $y(x)$ is strictly convex on $\left[0, \frac{1}{\sqrt{3}k}\right]$ and strictly concave on $\left[\frac{1}{\sqrt{3}k}, +\infty\right)$, with the inflection point at $x = \frac{1}{\sqrt{3}k}$;*

(4) *Boundedness:* $\lim_{x\to+\infty} y(x) = A$, *i.e., the payoff has an absolute upper bound and cannot exceed* A*.*

Proof. (1) Non-negativity: Since $A > 0, k > 0, x \geq 0$, the numerator $Ak^2x^2 \geq 0$ and the denominator $1 + k^2x^2 > 0$, so $y(x) \geq 0$. Substituting $x = 0$ gives $y(0) = 0$.

(2) Strict Monotonicity: Take the first derivative of $y(x)$:

$$y'(x) = \frac{d}{dx}\left(\frac{Ak^2x^2}{1+k^2x^2}\right) = \frac{2Ak^2x}{(1+k^2x^2)^2}$$

For all $x > 0$, $y'(x) > 0$, so $y(x)$ is strictly monotonically increasing on $[0, +\infty)$.

(3) Concavity and Convexity: Take the second derivative of $y(x)$:

$$y''(x) = \frac{d}{dx}\left(\frac{2Ak^2x}{(1+k^2x^2)^2}\right) = \frac{2Ak^2(1-3k^2x^2)}{(1+k^2x^2)^3}$$

The denominator $(1 + k^2x^2)^3 > 0$ always holds, so the sign of the second derivative is determined by the numerator $1 - 3k^2x^2$: - When $x < \frac{1}{\sqrt{3}k}$, $y''(x) > 0$, the function is strictly convex; - When $x > \frac{1}{\sqrt{3}k}$, $y''(x) < 0$, the function is strictly concave; - When $x = \frac{1}{\sqrt{3}k}$, $y''(x) = 0$, which is the inflection point of the function.

(4) Boundedness: Calculate the limit of $y(x)$ at infinity:

$$\lim_{x\to+\infty} y(x) = \lim_{x\to+\infty} \frac{Ak^2x^2}{1+k^2x^2} = A$$

The proof is completed. The proof is based on the classical monotonicity, concavity-convexity judgment theorems and limit operation rules in mathematical analysis [26, 28]. □

Lemma 2.2 (Core Properties of the Unit Contribution Payoff Function)**.** *Define the unit contribution payoff function as* $r(x) = \frac{y(x)}{x}$ $(x > 0)$*, which satisfies the following properties:*

(1) *Explicit Form:* $r(x) = \frac{Ak^2x}{1+k^2x^2}$*;*

(2) *Unimodality:* $r(x)$ *is strictly monotonically increasing on* $\left(0, \frac{1}{k}\right]$ *and strictly monotonically decreasing on* $\left[\frac{1}{k}, +\infty\right)$*, reaching the maximum value* $r_{max} = \frac{Ak}{2}$ *at* $x = \frac{1}{k}$*;*

(3) *Limit Properties:* $\lim_{x\to0^+} r(x) = 0$ *and* $\lim_{x\to+\infty} r(x) = 0$*.*

Proof. The proof is completed by direct algebraic simplification and first-order derivative analysis, consistent with the method in Lemma 1.1. □

This lemma is the core foundation for the derivation of the proportional allocation failure threshold, and clarifies the unimodal characteristic of unit contribution payoff, providing a rigorous basis for subsequent inequality derivation.

2.3. Strict Mathematical Definitions of Core Concepts.

All core definitions in this paper are based on the individual rationality axiom, and are strictly aligned with the participation constraint and coalition stability definitions in cooperative game theory [2, 7].

Definition 2.1 (Defection Condition). *For individual i, if for $\forall\varphi \in \Phi$ (all feasible allocation schemes), we have*

$$\varphi_i < y(x_i)$$

then individual i is said to inevitably defect.

Its essential economic implication is: no matter how the total group payoff is allocated to this individual, the maximum payoff they can obtain in the group is lower than their payoff from working independently. Defection is the strictly dominant strategy for this individual, and no allocation rule can retain them.

If there exists at least one feasible allocation scheme $\varphi \in \Phi$ such that

$$\varphi_i \geq y(x_i)$$

then individual i is said to have no defection motivation and satisfies the individual participation constraint.

Definition 2.2 (Safety State (Full Non-Defection Condition)). *If there exists a feasible allocation scheme $\varphi \in \Phi$ such that for $\forall i \in N$ (all individuals), we have*

$$\varphi_i \geq y(x_i)$$

then the collective is in a safety state; otherwise, it is in an unsafe state.

The essence of an unsafe state is: at least one individual inevitably defects, the collective cooperation has endogenous breakdown risk, and no allocation rule can achieve full stable participation.

Definition 2.3 (Failure Threshold of Allocation Rule). *For a given allocation rule*

$$\varphi(X) = (\varphi_1(X), \varphi_2(X), \ldots, \varphi_n(X))$$

if for $\forall i \in N$, we have $\varphi_i(X) \geq y(x_i)$, then the rule is valid and can achieve full non-defection; otherwise, the rule is invalid.

All thresholds in this paper are upper-bound thresholds: under a fixed contribution proportion structure, the allocation rule is valid when the total contribution X is lower than the threshold, and the rule fails when the total contribution X exceeds the threshold. The failure threshold is the boundary total contribution X where the rule changes from valid to invalid. The definition of strategy-proof allocation rules refers to Rodríguez-Álvarez C. [8].

3. General Form of the Three Types of Safety Critical Thresholds

Based on the above definitions, this section rigorously derives the necessary and sufficient conditions, explicit expressions, and existence proofs of the three core safety critical thresholds, and strictly demonstrates the order relationship between the three types of thresholds. This paper fixes the contribution proportion structure, i.e., the relative proportion of individual contributions remains unchanged, and the total contribution X is the only variable. This setting ensures that the relative proportion of individual independent payoffs is fixed, eliminating the interference of contribution structure changes on the thresholds.

A unified core logic runs through this paper: the failure of equal allocation, the failure of proportional allocation, and collective unsafety are all essentially homologous: there exists at least one individual whose payoff from independent work is strictly greater than the payoff they can obtain under the collective allocation rule. As long as this condition holds, defection is the strictly dominant strategy for the individual, and collective cooperation has endogenous breakdown risk. The three types of thresholds are only specific manifestations of this core logic under different allocation rules.

3.1. **Failure Threshold of Equal Allocation.** Equal allocation is the most basic allocation rule, where all individuals equally share the total group payoff regardless of individual contribution differences, and is widely used in flat, low-heterogeneity collective cooperation scenarios [25].

Definition 3.1 (Equal Allocation Rule)**.** *The equal allocation rule is defined as: for* $\forall i \in N$*, the payoff allocated to individual* i *is*

$$\varphi_i^{eq} = \frac{Y(X)}{n}$$

i.e., the total group payoff is equally distributed to all individuals, and each individual receives exactly the same payoff.

Theorem 3.1 (Failure Threshold of Equal Allocation)**.** *The necessary and sufficient condition for the equal allocation rule to be valid is*

$$Y(X) \geq n \cdot \max_{i\in N} y(x_i)$$

The failure threshold total contribution X_{eq} *of equal allocation satisfies*

$$Y(X_{eq}) = n \cdot \max_{i\in N} y(x_i)$$

When the total contribution $X > X_{eq}$*, the equal allocation rule is invalid; when* $X \leq X_{eq}$*, the equal allocation rule is valid.*

Proof. Necessity: If the equal allocation rule is valid, then for $\forall i \in N$, we have

$$\frac{Y(X)}{n} \geq y(x_i)$$

This inequality holds for all individuals, so it must hold for the individual with the maximum independent payoff, i.e.,

$$\frac{Y(X)}{n} \geq \max_{i\in N} y(x_i)$$

Multiplying both sides by n gives the necessary condition.

Sufficiency: If $Y(X) \geq n \cdot \max_{i\in N} y(x_i)$, then for $\forall i \in N$, we have

$$y(x_i) \leq \max_{i\in N} y(x_i) \leq \frac{Y(X)}{n} = \varphi_i^{\mathrm{eq}}$$

i.e., the equal allocation rule satisfies the participation constraints of all individuals, and the rule is valid.

The failure threshold is the boundary state between valid and invalid, i.e., the state where the inequality takes the equal sign, corresponding to the critical total contribution X_{eq}. The proof is completed. □

Essential Interpretation: The core reason for the failure of equal allocation is that the individual with the highest contribution and maximum independent payoff has an independent payoff exceeding the equal allocation payoff. That is, there exists an individual i such that $y(x_i) > \frac{Y(X)}{n}$, this individual will inevitably defect, and the equal allocation rule fails, which fully fits the unified core logic of this paper.

3.2. **Failure Threshold of Proportional Allocation.** Proportional allocation (distribution according to work) is the most widely used allocation rule in reality, where individual payoff is proportional to their own contribution, reflecting the core logic of "more work, more pay" [23,24].

Definition 3.2 (Proportional Allocation Rule)**.** *The proportional allocation rule based on contribution proportion is defined as: for $\forall i \in N$, the payoff allocated to individual i is*

$$\varphi_i^{prop} = \frac{x_i}{X} Y(X)$$

i.e., individual payoff is proportional to the ratio of their own contribution to the total contribution, the higher the contribution, the more payoff is allocated.

Theorem 3.2 (Failure Threshold of Proportional Allocation)**.** *The necessary and sufficient condition for the proportional allocation rule to be valid is*

$$\frac{Y(X)}{X} \geq \max_{\substack{i \in N \\ x_i > 0}} \frac{y(x_i)}{x_i}$$

The failure threshold total contribution X_{prop} of proportional allocation satisfies

$$\frac{Y(X_{prop})}{X_{prop}} = \max_{\substack{i \in N \\ x_i > 0}} \frac{y(x_i)}{x_i}$$

When the total contribution $X > X_{prop}$, the proportional allocation rule is invalid; when $X \leq X_{prop}$, the proportional allocation rule is valid.

Proof. Necessity: If the proportional allocation rule is valid, then for all individuals $i \in N$ with $x_i > 0$, we have

$$\frac{x_i}{X} Y(X) \geq y(x_i)$$

Dividing both sides by x_i ($x_i > 0$, the direction of the inequality remains unchanged), we get

$$\frac{Y(X)}{X} \geq \frac{y(x_i)}{x_i}$$

For individuals with $x_i = 0$, $\varphi_i^{\text{prop}} = 0 \geq y(0) = 0$, the inequality holds automatically without additional constraints.

This inequality holds for all individuals with $x_i > 0$, so it must hold for the individual with the highest unit contribution payoff, i.e.,

$$\frac{Y(X)}{X} \geq \max_{\substack{i \in N \\ x_i > 0}} \frac{y(x_i)}{x_i}$$

The necessity is proved.

Sufficiency: If $\frac{Y(X)}{X} \geq \max_{\substack{i \in N \\ x_i > 0}} \frac{y(x_i)}{x_i}$, then for all individuals $i \in N$ with $x_i > 0$, we have

$$\frac{y(x_i)}{x_i} \leq \max_{\substack{i \in N \\ x_i > 0}} \frac{y(x_i)}{x_i} \leq \frac{Y(X)}{X}$$

Multiplying both sides by x_i, we get

$$y(x_i) \leq \frac{x_i}{X} Y(X) = \varphi_i^{\text{prop}}$$

For individuals with $x_i = 0$, the inequality holds automatically. Therefore, the proportional allocation rule satisfies the participation constraints of all individuals, and the rule is valid. The sufficiency is proved.

The failure threshold is the boundary state between valid and invalid, i.e., the state where the inequality takes the equal sign, corresponding to the critical total contribution X_{prop}. The proof is completed. □

Essential Interpretation: The core reason for the failure of proportional allocation is that there exists at least one individual whose unit contribution payoff from independent work exceeds the overall unit contribution payoff of the collective. That is, there exists an individual i such that $\frac{y(x_i)}{x_i} > \frac{Y(X)}{X}$, at this time, even with distribution according to work, the payoff allocated to this individual is still lower than the independent payoff, they will inevitably defect, and the proportional allocation rule fails, which fully fits the unified core logic of this paper.

3.3. **Full Stability Critical Threshold.** The full stability critical threshold is the underlying boundary for the survival of collective cooperation, which characterizes the minimum requirement for "full non-defection regardless of the allocation method adopted", and is a necessary condition for the stability of collective cooperation.

Theorem 3.3 (Full Stability Critical Threshold)**.** *The necessary and sufficient condition for the collective to be in a safety state is*

$$Y(X) \geq \sum_{i=1}^{n} y(x_i)$$

The full stability critical total contribution X_{safe} of the collective satisfies

$$Y(X_{safe}) = \sum_{i=1}^{n} y(x_i)$$

When the total contribution $X > X_{safe}$, the collective is in an unsafe state; when $X \leq X_{safe}$, the collective is in a safety state.

Proof. Sufficiency: If $Y(X) \geq \sum_{i=1}^{n} y(x_i)$, construct the allocation scheme:

$$\varphi_i = y(x_i) + \frac{Y(X) - \sum_{i=1}^{n} y(x_i)}{n}, \quad \forall i \in N$$

First, verify feasibility:

$$\sum_{i=1}^{n} \varphi_i = \sum_{i=1}^{n} y(x_i) + \left(Y(X) - \sum_{i=1}^{n} y(x_i) \right) = Y(X)$$

which satisfies the budget balance axiom; at the same time, since $Y(X) - \sum_{i=1}^{n} y(x_i) \geq 0$, we have $\varphi_i \geq y(x_i) \geq 0$, which satisfies the non-negativity constraint. This allocation scheme belongs to the feasible allocation set Φ.

Second, verify the participation constraint: for $\forall i \in N$, we have $\varphi_i \geq y(x_i)$, so the collective is in a safety state. The sufficiency is proved.

Necessity: If the collective is in a safety state, then there exists a feasible allocation scheme $\varphi \in \Phi$ such that for $\forall i \in N$, we have $\varphi_i \geq y(x_i)$. Summing both sides of the inequality, we get

$$Y(X) = \sum_{i=1}^{n} \varphi_i \geq \sum_{i=1}^{n} y(x_i)$$

The necessity is proved.

The full stability threshold is the boundary state between safety and unsafety, i.e., the state where the inequality takes the equal sign, corresponding to the critical total contribution X_{safe}. The proof is completed. □

Essential Interpretation: The core reason for collective unsafety is that the total group payoff is insufficient to cover the sum of the opportunity costs of independent work of all individuals. At this time, even if the total group payoff is accurately allocated completely according to the individual independent payoff, it still cannot meet the participation constraints of all individuals. There must be at least one individual whose independent payoff exceeds any feasible allocation, and they will inevitably defect, leading to the breakdown of collective cooperation, which fully fits the unified core logic of this paper. This theorem reveals the underlying boundary of the survival of collective cooperation, which is fully consistent with the participation constraint conditions for heterogeneous agent coalition formation in Castaneda Alvarez D C. [6].

3.4. **Order Relationship of Critical Thresholds.** This section rigorously demonstrates the order relationship of the three types of upper-bound critical thresholds, and clarifies the stringency and valid interval of different allocation rules.

Theorem 3.4 (Strict Order of Upper-Bound Critical Thresholds)**.** *Under a fixed contribution proportion structure, the three types of upper-bound critical thresholds satisfy the strict order relationship:*

$$X_{eq} \leq X_{prop} \leq X_{safe}$$

Its economic implication is: the equal allocation rule has the narrowest valid interval and the strictest requirements; the proportional allocation rule is the second; the full stability condition has the widest valid interval and the loosest requirements.

Proof. We prove the order relationship in two strict steps, with the core logic that "the stricter the rule, the narrower the valid interval, and the smaller the upper-bound critical threshold".

Step 1: Prove $X_{\text{eq}} \leq X_{\text{prop}}$, i.e., "Valid Equal Allocation $\implies$ Valid Proportional Allocation" If equal allocation is valid, by Theorem 2.1, we have

$$Y(X) \geq n \cdot \max_{i \in N} y(x_i)$$

Let the individual with the maximum independent payoff be s, i.e., $y(x_s) = \max_{i \in N} y(x_i)$. By Lemma 1.1, $y(x)$ is strictly monotonically increasing, so $x_s = \max_{i \in N} x_i$, i.e., the individual with the highest contribution is also the individual with the maximum independent payoff.

By Lemma 1.2, the unit contribution payoff function $r(x)$ is strictly increasing when $x \leq \frac{1}{k}$ and strictly decreasing when $x \geq \frac{1}{k}$, so the maximum value $\max_{i \in N} \frac{y(x_i)}{x_i}$ must correspond to the individual s with the highest contribution, i.e.,

$$\max_{i \in N} \frac{y(x_i)}{x_i} = \frac{y(x_s)}{x_s}$$

Valid equal allocation implies:

$$\frac{Y(X)}{X} \geq \frac{ny(x_s)}{X}$$

The total contribution $X = \sum_{i=1}^{n} x_i \leq nx_s$ (since x_s is the maximum contribution), so $\frac{n}{X} \geq \frac{1}{x_s}$, therefore

$$\frac{Y(X)}{X} \geq \frac{ny(x_s)}{X} \geq \frac{y(x_s)}{x_s} = \max_{i \in N} \frac{y(x_i)}{x_i}$$

By Theorem 2.2, this is exactly the necessary and sufficient condition for valid proportional allocation. Therefore, valid equal allocation necessarily implies valid proportional allocation, indicating that equal allocation has stricter requirements and a smaller upper-bound critical threshold, i.e., $X_{\text{eq}} \leq X_{\text{prop}}$.

Step 2: Prove $X_{\text{prop}} \leq X_{\text{safe}}$, i.e., "Valid Proportional Allocation $\implies$ Collective Safety" If proportional allocation is valid, by Theorem 2.2, we have

$$\frac{Y(X)}{X} \geq \max_{i \in N} \frac{y(x_i)}{x_i}$$

Therefore, for $\forall i \in N$, we have

$$\frac{y(x_i)}{x_i} \leq \frac{Y(X)}{X}$$

Multiplying both sides by x_i, we get

$$y(x_i) \leq \frac{x_i}{X} Y(X)$$

Summing over all individuals, we get

$$\sum_{i=1}^{n} y(x_i) \leq \frac{Y(X)}{X} \sum_{i=1}^{n} x_i = Y(X)$$

By Theorem 2.3, this is exactly the necessary and sufficient condition for collective safety. Therefore, valid proportional allocation necessarily implies collective safety, indicating that proportional allocation has stricter requirements and a smaller upper-bound critical threshold, i.e., $X_{\text{prop}} \leq X_{\text{safe}}$.

Combining the two steps of proof, we obtain the strict order of the three types of critical thresholds:

$$X_{\text{eq}} \leq X_{\text{prop}} \leq X_{\text{safe}}$$

The proof is completed. □

Corollary 3.1 (Coincidence of Critical Thresholds under Homogeneous Contribution)**.** *When individual contributions are homogeneous, i.e., $x_i = x, \forall i \in N$, the three types of critical thresholds are completely coincident, i.e.,*

$$X_{eq} = X_{prop} = X_{safe}$$

Proof. When $x_i = x$, $\max_{i\in N} y(x_i) = y(x)$, $\max_{i\in N} \frac{y(x_i)}{x_i} = \frac{y(x)}{x}$, $\sum_{i=1}^{n} y(x_i) = ny(x)$. Substituting into the three threshold conditions, we find that all three thresholds satisfy $Y(X) = ny(x)$, so they are completely equal. The proof is completed. □

This corollary shows that when individual contributions are completely homogeneous, the valid intervals of the three rules are completely coincident: either all are valid or all are invalid, with no intermediate state.

Corollary 3.2 (Stability Domain Division under Heterogeneous Contribution)**.** *Based on the strict order of critical thresholds, according to the size of the total actual contribution X, the stability of collective cooperation can be divided into four non-overlapping domains:*

(1) *Full Stability Domain: $X \leq X_{eq}$, both equal allocation and proportional allocation are valid, the collective is in a safety state, there are infinitely many feasible allocation rules to achieve full non-defection, and the collective stability is the strongest;*
(2) *Proportional Stability Domain: $X_{eq} < X \leq X_{prop}$, equal allocation is invalid, but proportional allocation is valid, the collective is still in a safety state, and full non-defection can be achieved through distribution according to work;*
(3) *Limited Stability Domain: $X_{prop} < X \leq X_{safe}$, both equal allocation and proportional allocation are invalid, but the collective is still in a safety state, and collective stability can only be maintained through differentiated allocation rules tilted towards high-contribution individuals;*
(4) *Full Failure Domain: $X > X_{safe}$, the collective is in an unsafe state, there is no feasible allocation rule to achieve full non-defection, individuals will inevitably defect, and collective cooperation will inevitably break down.*

This corollary fully characterizes the stable state of collective cooperation under different total contribution levels, and clarifies the applicable boundaries of different allocation rules.

The impact mechanism of individual heterogeneity on the degree of allocation inequality and cooperation stability refers to Victoria-Feser M P. [19].

4. Homogeneous Equal-Contribution Model

This section analyzes the Homogeneous Equal-Contribution Model where all individuals have exactly the same contribution, derives the explicit closed-form solutions, properties and stability domain division of the critical thresholds, and provides a benchmark reference for the heterogeneous model. The research on homogeneous individual alliances refers to Wakita D, Sobu Y, Kaneko N, et al. [4].

4.1. **Model Setting.** Let the total number of individuals in the collective be n $(n \geq 2)$, and the effective contribution of all individuals is exactly the same, i.e.,

$$x_i = x, \quad \forall i \in N$$

The total effective contribution of the collective is

$$X = \sum_{i=1}^{n} x_i = nx$$

All other notation, payoff functions, and core definitions are consistent with Section 1, with no additional assumptions.

4.2. **Explicit Solution of Critical Thresholds.** By Corollary 2.1, the three types of critical thresholds are completely coincident under the Homogeneous Equal-Contribution Model, so we only need to solve the full stability threshold to obtain the explicit solutions of all thresholds.

By Theorem 2.3, the safety critical condition is

$$Y(X) = \sum_{i=1}^{n} y(x_i)$$

Substituting the payoff function and $X = nx$, we get

$$\frac{Ak^2(nx)^2}{1 + k^2(nx)^2} = n \cdot \frac{Ak^2x^2}{1 + k^2x^2}$$

Both sides of the equation are positive, and the common constant Ak^2x^2 can be canceled out $(x > 0)$, giving

$$\frac{n^2}{1 + k^2n^2x^2} = \frac{n}{1 + k^2x^2}$$

Dividing both sides by n, we simplify to

$$\frac{n}{1 + k^2n^2x^2} = \frac{1}{1 + k^2x^2}$$

Cross-multiplying and expanding (all terms are positive, the direction of the inequality remains unchanged):

$$n(1 + k^2x^2) = 1 + k^2n^2x^2$$

Expanding the left side:

$$n + nk^2x^2 = 1 + n^2k^2x^2$$

Moving terms, we move the terms containing x^2 to the right and the constant terms to the left:

$$n - 1 = n^2k^2x^2 - nk^2x^2$$

Extracting the common factor nk^2x^2 on the right side:

$$n - 1 = nk^2x^2(n - 1)$$

Since $n \geq 2$, $n - 1 > 0$, we can cancel out $n - 1$ on both sides, getting

$$1 = nk^2x^2$$

Rearranging gives the explicit solution of the individual critical contribution:

$$x^* = \frac{1}{\sqrt{n}k}$$

The corresponding total contribution critical threshold is:

$$X^* = nx^* = \frac{\sqrt{n}}{k}$$

Combined with Corollary 2.1, the three types of critical thresholds are completely coincident under the Homogeneous Equal-Contribution Model, i.e.,

$$X_{\text{safe}} = X_{\text{prop}} = X_{\text{eq}} = X^* = \frac{\sqrt{n}}{k}$$

The algebraic equation solving and simplification in this section are based on the basic theory of polynomial operations [27].

4.3. **Property Analysis of Critical Thresholds.** Based on the above explicit solutions, this section derives the core properties of the critical thresholds under the Homogeneous Equal-Contribution Model, and reveals the impact mechanism of group size and efficiency parameters on the stability of collective cooperation.

Corollary 4.1 (Invariance of Critical Thresholds in Equal-Contribution Model)**.** *Under the Homogeneous Equal-Contribution Model, the three types of safety critical thresholds are completely coincident, there is a unique explicit closed-form solution for the individual critical contribution and total contribution critical threshold, and the critical thresholds are only determined by the group size n and efficiency parameter k, independent of the payoff upper bound A.*

Proof. From the derivation results in Section 3.2, the individual critical contribution is

$$x^* = \frac{1}{\sqrt{n}k}$$

The total contribution critical threshold is

$$X^* = \frac{\sqrt{n}}{k}$$

In the process of solving the critical thresholds, the payoff upper bound A is completely canceled out on both sides of the equation, so the critical thresholds are independent of A; at the same time, by Corollary 2.1, the three types of critical thresholds are completely coincident under homogeneous contribution. The invariance is proved. □

This corollary shows that in a homogeneous equal-contribution collective, the height of the payoff ceiling does not affect the critical threshold of cooperation stability, which is only determined by the cooperation dependence and group size.

Corollary 4.2 (Scale Effect of Group Size)**.** *Under the Homogeneous Equal-Contribution Model, the total contribution critical threshold $X^*(n)$ is strictly monotonically increasing with the increase of group size n, and satisfies the limit property:*

$$\lim_{n\to+\infty} X^*(n) = +\infty$$

That is, the larger the group size, the higher the total contribution critical threshold, the wider the valid interval, and the easier it is to maintain the stability of collective cooperation.

Proof. Rewrite the total contribution critical threshold as a function of n:

$$X^*(n) = \frac{\sqrt{n}}{k}$$

Take the first derivative with respect to n ($n \geq 2$):

$$\frac{dX^*}{dn} = \frac{1}{2k\sqrt{n}} > 0$$

Therefore, $X^*(n)$ is strictly monotonically increasing on $n \geq 2$.

Solve the limit:

$$\lim_{n\to+\infty} X^*(n) = \lim_{n\to+\infty} \frac{\sqrt{n}}{k} = +\infty$$

The proof is completed. □

This corollary reveals the scale advantage of homogeneous collectives: the larger the number of people, the wider the valid interval of collective cooperation, the higher total contribution it can withstand without cooperation breakdown, and the stronger the stability.

Corollary 4.3 (Threshold Effect of Efficiency Parameter)**.** *Under the Homogeneous Equal-Contribution Model, the critical contribution is in a strict inverse proportional relationship with the efficiency parameter k, i.e.,*

$$x^* \propto \frac{1}{k}, \quad X^* \propto \frac{1}{k}$$

The larger the efficiency parameter k, the more suitable the task is for independent work, the lower the critical threshold of collective cooperation, the narrower the valid interval, and the easier it is to fail; the smaller k is, the more the task relies on collective cooperation, the higher the critical threshold, the wider the valid interval, and the more stable the cooperation.

Proof. From the explicit solution of the critical thresholds, $x^* = \frac{1}{\sqrt{n}} \cdot \frac{1}{k}$, $X^* = \sqrt{n} \cdot \frac{1}{k}$, where $\frac{1}{\sqrt{n}}$ and $\sqrt{n}$ are positive constants only related to n, so x^* and X^* are in a strict inverse proportional relationship with k.

Combined with the definition of the efficiency parameter k in Section 1.1: the larger k is, the faster the independent payoff grows, the more suitable the task is for independent work, the weaker the scale effect of cooperation, at this time the critical threshold is lower and the valid interval is narrower; on the contrary, the smaller k is, the slower the independent payoff grows, the more the task relies on collective cooperation, the stronger the scale effect of cooperation, the higher the critical threshold and the wider the valid interval. The proof is completed. □

Corollary 4.4 (Stability Domain Division of Homogeneous Equal-Contribution Model)**.** *For a homogeneous equal-contribution collective, let the total actual contribution be X, then:*

(1) *Full Stability Domain: $X \leq X^*$, both equal allocation and proportional allocation are valid, the collective is in a safety state, there are infinitely many feasible allocation rules to achieve full non-defection, and the collective stability is the strongest;*
(2) *Full Failure Domain: $X > X^*$, the collective is in an unsafe state, there is no feasible allocation rule to achieve full non-defection, all allocation rules are invalid, individuals will inevitably defect, and collective cooperation will inevitably break down.*

Proof. By Corollary 2.1, the three types of critical thresholds are completely coincident under homogeneous equal-contribution, so $X \leq X^*$ is equivalent to satisfying the necessary and sufficient conditions for full safety, valid proportional allocation, and valid equal allocation at the same time, at this time all three types of rules are valid and in the full stability domain; when $X > X^*$, the necessary and sufficient condition for full safety is not satisfied, by Definition 1.2, the collective is in an unsafe state, there is no feasible allocation rule to achieve full non-defection, so it is in the full failure domain. The proof is completed. □

This corollary completes the full characterization of the stability boundary of the Homogeneous Equal-Contribution Model, and clarifies the strict boundary between the stability and failure of collective cooperation.

5. One-Strong-n-Weak Heterogeneous Model

This section focuses on the more common scenario of individual heterogeneity in reality, and constructs the One-Strong-n-Weak Heterogeneous Model: there is 1 high-contribution ability strong agent and n low-contribution ability weak agents in the collective. We rigorously derive the explicit expressions, order relationship and stability properties of the three types of safety critical thresholds under this model, and reveal the impact mechanism of individual heterogeneity on the stability of collective cooperation. The research on heterogeneous individual coalition formation refers to Castaneda Alvarez D C. [6].

5.1. **Heterogeneous Model Setting.** Let the total number of individuals in the collective be $N = n + 1$ ($n \geq 1$, i.e., the collective includes at least 1 strong agent and 1 weak agent), where:

(1) Strong Agent (denoted as individual s): effective contribution is $x_s > 0$, the high-contribution subject in the collective;

(2) Weak Agent Set (denoted as $L = \{1, 2, \ldots, n\}$): a total of n homogeneous weak agents, each with effective contribution $x_l > 0$, and satisfying $x_s > x_l$, i.e., the contribution of the strong agent is significantly higher than that of the weak agents;

(3) The total contribution of the collective is:

$$X = x_s + nx_l$$

(4) All other notation, payoff functions, and core definitions are consistent with Section 1, with no additional assumptions.

To simplify the subsequent derivation, we introduce the heterogeneity coefficient λ, defined as the ratio of the contribution of the strong agent to that of the weak agent:

$$\lambda = \frac{x_s}{x_l}, \quad \lambda > 1$$

The larger λ is, the higher the degree of heterogeneity of individual contributions; when $\lambda \to 1$, the model degenerates into the Homogeneous Equal-Contribution Model in Section 3.

5.2. **Explicit Solution of the Three Types of Safety Critical Thresholds.** Based on the above settings, combined with the general necessary and sufficient conditions of the three types of critical thresholds in Section 2, we derive the explicit expressions, existence and uniqueness proofs of the three types of critical thresholds under the heterogeneous model respectively.

5.2.1. *Full Stability Critical Threshold.* By Theorem 2.3, the necessary and sufficient condition for the collective to be in a safety state is:

$$Y(X) \geq y(x_s) + ny(x_l)$$

Substituting the payoff function and the heterogeneity coefficient $\lambda = \frac{x_s}{x_l}$ (i.e., $x_s = \lambda x_l$), the total contribution $X = \lambda x_l + nx_l = x_l(\lambda + n)$, we expand to get:

$$\frac{Ak^2X^2}{1 + k^2X^2} \geq \frac{Ak^2x_s^2}{1 + k^2x_s^2} + n \cdot \frac{Ak^2x_l^2}{1 + k^2x_l^2}$$

Both sides of the equation are positive, cancel out the common constant Ak^2, substitute $x_s = \lambda x_l$ and $X = x_l(\lambda + n)$, we get:

$$\frac{x_l^2(\lambda + n)^2}{1 + k^2x_l^2(\lambda + n)^2} \geq \frac{\lambda^2x_l^2}{1 + k^2\lambda^2x_l^2} + \frac{nx_l^2}{1 + k^2x_l^2}$$

Dividing both sides by x_l^2 ($x_l > 0$), let the dimensionless variable $t = k^2x_l^2$ ($t > 0$), then $k^2x_l^2 = t$, after substitution and canceling out the common terms, the equation is simplified to:

$$\frac{(\lambda + n)^2}{\frac{1}{t} + (\lambda + n)^2} \geq \frac{\lambda^2}{\frac{1}{t} + \lambda^2} + \frac{n}{\frac{1}{t} + 1}$$

Let $z = \frac{1}{t} > 0$, further simplified to an algebraic inequality about z:

$$\frac{(\lambda+n)^2}{z+(\lambda+n)^2} \geq \frac{\lambda^2}{z+\lambda^2} + \frac{n}{z+1} \tag{1}$$

Theorem 5.1 (Full Stability Critical Threshold of Heterogeneous Model)**.** *For the One-Strong-n-Weak Heterogeneous Model, there exists a unique positive real solution z^* satisfying the equality condition of Equation (1), corresponding to the critical contribution of the weak agent:*

$$x_l^* = \frac{1}{k\sqrt{z^*}}$$

The critical contribution of the strong agent is:

$$x_s^* = \frac{\lambda}{k\sqrt{z^*}}$$

The total contribution full stability critical threshold is:

$$X_{safe}^* = \frac{\lambda+n}{k\sqrt{z^*}}$$

And X_{safe}^* is strictly monotonically increasing with the increase of the heterogeneity coefficient λ, i.e., the higher the heterogeneity of individual contributions, the higher the critical threshold of collective safety and the wider the valid interval.

Proof. Let the difference between the two sides of Inequality (1) be the function $G(z)$:

$$G(z) = \frac{(\lambda+n)^2}{z+(\lambda+n)^2} - \frac{\lambda^2}{z+\lambda^2} - \frac{n}{z+1}, \quad z > 0$$

We prove the existence, uniqueness and monotonicity of the solution in three steps.

Step 1: Prove $G(z)$ is strictly monotonically decreasing Take the first derivative of $G(z)$:

$$G'(z) = -\frac{(\lambda+n)^2}{[z+(\lambda+n)^2]^2} + \frac{\lambda^2}{(z+\lambda^2)^2} + \frac{n}{(z+1)^2}$$

For any $a > b > 0$, the function $h_a(z) = \frac{a^2}{(z+a^2)^2}$ is strictly decreasing with the increase of a, i.e., the larger a is, the smaller $h_a(z)$ is. Since $\lambda + n > \lambda > 1$, we have

$$\frac{(\lambda+n)^2}{[z+(\lambda+n)^2]^2} > \frac{\lambda^2}{(z+\lambda^2)^2}, \quad \frac{(\lambda+n)^2}{[z+(\lambda+n)^2]^2} > \frac{n}{(z+1)^2}$$

Therefore

$$G'(z) = -\left(\frac{(\lambda+n)^2}{[z+(\lambda+n)^2]^2} - \frac{\lambda^2}{(z+\lambda^2)^2} - \frac{n}{(z+1)^2}\right) < 0$$

i.e., $G(z)$ is strictly monotonically decreasing on $z > 0$.

Step 2: Prove the existence and uniqueness of the solution Calculate the limit properties of $G(z)$: - When $z \to 0^+$:

$$G(0^+) = \frac{(\lambda+n)^2}{(\lambda+n)^2} - \frac{\lambda^2}{\lambda^2} - \frac{n}{1} = 1 - 1 - n = -n < 0$$

- When $z \to +\infty$:

$$G(+\infty) = \lim_{z\to+\infty}\left(\frac{(\lambda+n)^2}{z} - \frac{\lambda^2}{z} - \frac{n}{z}\right) = 0^+$$

Since $G(z)$ is strictly monotonically decreasing, and $G(0^+) < 0$, $G(+\infty) \to 0^+$, according to the Intermediate Value Theorem for continuous functions, there exists a unique positive real number z^* such that $G(z^*) = 0$, i.e., the equality of Equation (1) holds, corresponding to the unique critical contribution solution. The existence and uniqueness are proved.

Step 3: Prove monotonicity When the heterogeneity coefficient λ increases, the $G(z)$ curve shifts downward as a whole, so that the solution z^* of $G(z) = 0$ decreases, therefore

$$X^*_{\text{safe}} = \frac{\lambda + n}{k\sqrt{z^*}}$$

is strictly increasing with the increase of λ. The monotonicity is proved.

The proof is completed. $\square$

5.2.2. *Proportional Allocation Failure Threshold.* By Theorem 2.2, the necessary and sufficient condition for the proportional allocation rule to be valid is:

$$\frac{Y(X)}{X} \geq \max_{\substack{i \in N \\ x_i > 0}} \frac{y(x_i)}{x_i}$$

In the One-Strong-n-Weak Model, the strong agent's contribution $x_s > x_l$, combined with Lemma 1.2, the unit contribution payoff function $r(x)$ is strictly increasing when $x \leq \frac{1}{k}$, so $\max_{\substack{i \in N \\ x_i > 0}} \frac{y(x_i)}{x_i} = \frac{y(x_s)}{x_s}$, the necessary and sufficient condition is simplified to:

$$\frac{Y(X)}{X} \geq \frac{y(x_s)}{x_s}$$

Substituting the payoff function, we expand to get:

$$\frac{1}{X} \cdot \frac{Ak^2X^2}{1 + k^2X^2} \geq \frac{1}{x_s} \cdot \frac{Ak^2x_s^2}{1 + k^2x_s^2}$$

Cancel out the common constant Ak^2, simplify to get:

$$\frac{X}{1 + k^2X^2} \geq \frac{x_s}{1 + k^2x_s^2}$$

Cross-multiplying and rearranging (all terms are positive, the direction of the inequality remains unchanged):

$$X(1 + k^2x_s^2) \geq x_s(1 + k^2X^2)$$

Expanding both sides:

$$X + k^2Xx_s^2 \geq x_s + k^2x_sX^2$$

Moving terms and grouping to extract common factors:

$$(X - x_s) - k^2x_sX(X - x_s) \geq 0$$

Extract the common factor $(X - x_s)$ again, we obtain the core inequality:

$$(X - x_s)(1 - k^2x_sX) \geq 0 \tag{2}$$

Combined with the model setting, the total contribution $X = x_s + nx_l > x_s$ always holds, so $X - x_s > 0$. Therefore, Inequality (2) is equivalent to:

$$1 - k^2 x_s X \geq 0 \implies k^2 x_s X \leq 1$$

Theorem 5.2 (Proportional Allocation Failure Threshold of Heterogeneous Model)**.** *Under the One-Strong-n-Weak Heterogeneous Model, the necessary and sufficient condition for the proportional allocation rule to be valid is*

$$k^2 x_s X \leq 1$$

The corresponding total contribution failure threshold is:

$$X^*_{prop} = \frac{1}{k^2 x_s}$$

The failure threshold of proportional allocation is only determined by the strong agent's contribution x_s and the efficiency parameter k, and is independent of the weak agents' contribution and the total group size.

Proof. The above derivation has completed the proof of the necessary and sufficient condition. The failure threshold is the boundary state between valid and invalid, i.e., the state where the inequality takes the equal sign: $k^2 x_s X = 1$. Solving for X gives the explicit threshold $X^*_{\text{prop}} = \frac{1}{k^2 x_s}$. The expression only contains x_s and k, and is independent of x_l and n, so the validity of proportional allocation is solely determined by the strong agent's contribution level and the task's efficiency parameter. The proof is completed. □

This theorem reveals the core constraint of proportional allocation in heterogeneous collectives: the higher the strong agent's contribution level, the lower the failure threshold of proportional allocation, the narrower the valid interval, and the easier the proportional allocation rule fails.

5.2.3. *Equal Allocation Failure Threshold.* By Theorem 2.1, the necessary and sufficient condition for the equal allocation rule to be valid is:

$$Y(X) \geq (n+1) \cdot \max_{i \in N} y(x_i)$$

In the One-Strong-n-Weak Model, $y(x)$ is strictly monotonically increasing, and $x_s > x_l$, so $\max_{i \in N} y(x_i) = y(x_s)$. The necessary and sufficient condition is simplified to:

$$Y(X) \geq (n+1) y(x_s)$$

Substituting the payoff function, we expand to get:

$$\frac{Ak^2 X^2}{1 + k^2 X^2} \geq (n+1) \cdot \frac{Ak^2 x_s^2}{1 + k^2 x_s^2}$$

Cancel out the common positive constant Ak^2, cross-multiply and rearrange (all terms are positive, the direction of the inequality remains unchanged):

$$X^2 (1 + k^2 x_s^2) \geq (n+1) x_s^2 (1 + k^2 X^2)$$

Expand both sides of the inequality:

$$X^2 + k^2 x_s^2 X^2 \geq (n+1)x_s^2 + (n+1)k^2 x_s^2 X^2$$

Move terms with X^2 to the left-hand side and constant terms to the right-hand side:

$$X^2 - (n+1)k^2 x_s^2 X^2 + k^2 x_s^2 X^2 \geq (n+1)x_s^2$$

Combine like terms:

$$X^2 \left[1 - nk^2 x_s^2\right] \geq (n+1)x_s^2 \tag{3}$$

Theorem 5.3 (Equal Allocation Failure Threshold of Heterogeneous Model)**.** *Under the One-Strong-n-Weak Heterogeneous Model, the precondition for the equal allocation rule to have a valid interval is*

$$1 - nk^2 x_s^2 > 0 \implies x_s < \frac{1}{k\sqrt{n}}$$

When this precondition is satisfied, the total contribution failure threshold of equal allocation is:

$$X_{eq}^* = \frac{x_s\sqrt{n+1}}{\sqrt{1 - nk^2 x_s^2}}$$

If $x_s \geq \frac{1}{k\sqrt{n}}$, the equal allocation rule is permanently invalid, and there exists no total contribution X that makes equal allocation valid.

Proof. The right-hand side of Inequality (3) $(n+1)x_s^2 > 0$ always holds. Therefore, for the inequality to hold, the left-hand side coefficient $1 - nk^2 x_s^2$ must be strictly positive; otherwise, the left-hand side is non-positive and the inequality is permanently invalid. This gives the precondition for the existence of a valid interval.

When the precondition is satisfied, divide both sides of Inequality (3) by $1 - nk^2 x_s^2$ (positive, so the direction of the inequality remains unchanged):

$$X^2 \geq \frac{(n+1)x_s^2}{1 - nk^2 x_s^2}$$

Take the square root of both sides ($X > 0$) to obtain the explicit threshold:

$$X_{\text{eq}}^* = \sqrt{\frac{(n+1)x_s^2}{1 - nk^2 x_s^2}} = \frac{x_s\sqrt{n+1}}{\sqrt{1 - nk^2 x_s^2}}$$

If $x_s \geq \frac{1}{k\sqrt{n}}$, the left-hand side coefficient of Inequality (3) is non-positive, and the inequality never holds, so equal allocation is permanently invalid. The proof is completed. $\square$

This theorem reveals the strong constraint of equal allocation in heterogeneous collectives: when the strong agent's contribution exceeds the critical value $\frac{1}{k\sqrt{n}}$, no matter how high the total group contribution is, the equal allocation rule will inevitably fail and cannot satisfy the strong agent's participation constraint.

5.3. **Order of Critical Thresholds and Stability Analysis in Heterogeneous Model.** Based on the explicit solutions of the three types of critical thresholds above, this section derives the order relationship of the thresholds, and completes the stability domain division and stability property analysis of the heterogeneous model.

Theorem 5.4 (Order of Critical Thresholds in Heterogeneous Model)**.** *For the One-Strong-n-Weak Heterogeneous Model, when the heterogeneity coefficient $\lambda > 1$, the three types of critical thresholds satisfy the strict order relationship:*

$$X^*_{eq} < X^*_{prop} < X^*_{safe}$$

Proof. We prove the strict order in two rigorous steps, consistent with the unified logic of Theorem 2.4.

Step 1: Prove $X^*_{\text{eq}} < X^*_{\text{prop}}$ The precondition for equal allocation to be valid is $x_s < \frac{1}{k\sqrt{n}}$, i.e., $nk^2x_s^2 < 1$. By Theorem 4.2, $X^*_{\text{prop}} = \frac{1}{k^2x_s}$. Substituting the precondition gives:

$$X^*_{\text{prop}} = \frac{1}{k^2x_s} > \frac{\sqrt{n}}{k}$$

By Theorem 4.3, $X^*_{\text{eq}} = \frac{x_s\sqrt{n+1}}{\sqrt{1-nk^2x_s^2}}$. Let $u = kx_s$, then $u < \frac{1}{\sqrt{n}}$, and the threshold is rewritten as:

$$X^*_{\text{eq}} = \frac{u\sqrt{n+1}}{k\sqrt{1-nu^2}}$$

Taking the first derivative with respect to u shows that X^*_{eq} is strictly increasing with u. For all $u < \frac{1}{\sqrt{n}}$, we can verify that:

$$\frac{u\sqrt{n+1}}{\sqrt{1-nu^2}} < \frac{1}{u}$$

Cross-multiplying and squaring both sides (all terms are positive) gives:

$$u^4(n+1) < 1 - nu^2 \implies (n+1)u^4 + nu^2 - 1 < 0$$

Let $v = u^2 < \frac{1}{n}$, the quadratic function $f(v) = (n+1)v^2 + nv - 1$ is strictly increasing for $v > 0$, with $f(0) = -1 < 0$ and $f(\frac{1}{n}) = \frac{n+1}{n^2} > 0$. Therefore, for all valid u in the equal allocation valid interval, the inequality holds, i.e., $X^*_{\text{eq}} < X^*_{\text{prop}}$.

Step 2: Prove $X^*_{\text{prop}} < X^*_{\text{safe}}$ By the proof of Theorem 2.4, valid proportional allocation necessarily implies collective safety. This means that the maximum total contribution that satisfies proportional allocation validity is strictly smaller than the maximum total contribution that satisfies collective safety, i.e., $X^*_{\text{prop}} < X^*_{\text{safe}}$.

Combining the two steps, we obtain the strict order:

$$X^*_{\text{eq}} < X^*_{\text{prop}} < X^*_{\text{safe}}$$

The proof is completed. □

Corollary 5.1 (Four-Stage Stability Domain Division of Heterogeneous Model)**.** *Based on the order relationship of critical thresholds, according to the size of the total actual contribution X, the stability of collective cooperation is divided into four non-overlapping domains:*

(1) *Full Stability Domain: $X \leq X^*_{eq}$. Equal allocation and proportional allocation are both valid, the collective is in a safety state, there are infinitely many feasible allocation rules to achieve full non-defection, and the collective stability is the strongest.*
(2) *Proportional Stability Domain: $X^*_{eq} < X \leq X^*_{prop}$. Equal allocation is invalid, but proportional allocation is valid, the collective is still in a safety state, and full non-defection can be achieved through distribution according to work.*
(3) *Limited Stability Domain: $X^*_{prop} < X \leq X^*_{safe}$. Equal allocation and proportional allocation are both invalid, but the collective is still in a safety state. Collective stability can only be maintained through differentiated allocation rules tilted towards the strong agent.*
(4) *Full Failure Domain: $X > X^*_{safe}$. The collective is in an unsafe state, there is no feasible allocation rule to achieve full non-defection, individuals will inevitably defect, and collective cooperation will inevitably break down.*

Proof. By the order of critical thresholds in Theorem 4.4 and the definition of each threshold: - When $X \leq X^*_{\text{eq}}$, all three validity conditions are satisfied, so all allocation rules are valid and the collective is in the full stability domain. - When $X^*_{\text{eq}} < X \leq X^*_{\text{prop}}$, the equal allocation validity condition is violated, but the proportional allocation and safety conditions are satisfied, so the collective is in the proportional stability domain. - When $X^*_{\text{prop}} < X \leq X^*_{\text{safe}}$, the equal and proportional allocation validity conditions are violated, but the safety condition is satisfied. Only allocation rules tilted towards the strong agent can satisfy all participation constraints, so the collective is in the limited stability domain. - When $X > X^*_{\text{safe}}$, the safety condition is violated, so the collective is in the full failure domain.

The proof is completed. □

Corollary 5.2 (Stability Erosion Effect of Heterogeneity)**.** *The stability of collective cooperation decreases continuously with the increase of individual contribution heterogeneity, with the following specific manifestations:*

(1) *The larger the heterogeneity coefficient λ, the higher the full stability threshold X^*_{safe}, the wider the valid interval of collective safety, but the narrower the valid intervals of proportional allocation and equal allocation.*
(2) *The larger the heterogeneity coefficient λ, the lower the failure thresholds of proportional allocation and equal allocation, the easier the two allocation rules fail, the wider the limited stability domain, and the harder it is to reach the full stability domain.*
(3) *When the heterogeneity coefficient λ exceeds the critical value $\lambda_{max} = \frac{1}{kx_l\sqrt{n}}$, the equal allocation rule is permanently invalid, the full stability domain disappears, and the collective can only be in the proportional stability domain at most.*

Proof. 1. By Theorem 4.1, X^*_{safe} is strictly monotonically increasing with λ, so higher heterogeneity leads to a higher safety threshold. 2. By Theorem 4.2, $X^*_{\text{prop}} = \frac{1}{k^2 x_s} = \frac{1}{k^2 \lambda x_l}$, which is strictly decreasing with λ. By Theorem 4.3, X^*_{eq} decreases as $x_s = \lambda x_l$ increases, and when $x_s \geq \frac{1}{k\sqrt{n}}$ (i.e., $\lambda \geq \frac{1}{k x_l \sqrt{n}}$), equal allocation is permanently invalid. 3. Combined with the order of critical thresholds, the larger λ is, the lower X^*_{eq} and X^*_{prop} are, the narrower the full stability domain, and the wider the limited stability domain.

The proof is completed. □

This corollary fully reveals the mechanism by which individual heterogeneity erodes the stability of collective cooperation: the larger the contribution gap, the easier the fair and equal allocation rules fail. The collective can only maintain stability through allocation rules tilted towards the strong agent, and when heterogeneity exceeds the critical value, the equal allocation rule will be permanently invalid.

6. Conclusions and Implications

This paper constructs a Homogeneous Equal-Contribution Model and a One-Strong-n-Weak Heterogeneous Model for collective cooperation based on an upper-bounded S-shaped contribution-payoff function. We strictly define and derive the explicit expressions and properties of three core safety critical thresholds, and fully characterize the stability domains and failure boundaries of collective cooperation. The core conclusions are as follows:

1. **Unified Essence of Collective Cooperation Failure** The failure of equal allocation, the failure of proportional allocation, and collective unsafety are all essentially homologous: there exists at least one individual whose payoff from independent work is strictly greater than the payoff they can obtain under the collective allocation mechanism. As long as this condition holds, defection is the strictly dominant strategy for the individual, and collective cooperation has endogenous breakdown risk. The three types of critical thresholds are only specific manifestations of this core logic under different allocation rules.

2. **Upper-Bound Threshold Structure and Stability Domain Division** The stability of collective cooperation does not increase monotonically with total contribution. Instead, there is a clear upper-bound critical threshold: the allocation rule is valid and cooperation is stable when total contribution is below the threshold, while the rule fails and cooperation breaks down when total contribution exceeds the threshold. The three critical thresholds satisfy the strict order $X_{\text{eq}} \leq X_{\text{prop}} \leq X_{\text{safe}}$, corresponding to four stages: full stability domain, proportional stability domain, limited stability domain, and full failure domain, which clarify the applicable boundaries of different allocation rules.

3. **Scale Advantage of Homogeneous Collectives** In homogeneous equal-contribution collectives, the three types of safety critical thresholds are completely coincident. The larger the group size, the higher the total contribution critical threshold, the wider the valid interval, and the more stable the cooperation. The smaller the efficiency parameter k, the more the task relies

on collective cooperation, the higher the critical threshold, and the stronger the scale effect of cooperation.

4. **Stability Erosion Effect of Heterogeneity** Individual contribution heterogeneity significantly changes the order and magnitude of critical thresholds, breaking the coincidence of the three types of thresholds. The higher the strong agent's contribution and the stronger the heterogeneity, the lower the failure thresholds of proportional allocation and equal allocation, the narrower the valid intervals, and the easier the two rules fail. When heterogeneity exceeds the critical value, the equal allocation rule will be permanently invalid, and the collective can only maintain limited stability through differentiated allocation rules tilted towards the strong agent.

5. **Selection Logic of Allocation Rules** The equal allocation rule is only valid in collectives with low heterogeneity and low total contribution, and has the lowest tolerance for heterogeneity. The proportional allocation rule has a moderate tolerance for heterogeneity, and its validity is solely determined by the strong agent's contribution level, making it the optimal choice for moderately heterogeneous collectives. Differentiated allocation rules tilted towards high-contribution individuals are the only feasible way to maintain stability in collectives with high heterogeneity and high total contribution.

This paper constructs a complete mathematical framework for the stability analysis of collective cooperation through rigorous algebraic derivation and functional property analysis at the pure theoretical level, and reveals the mechanisms by which individual heterogeneity, group size, and task attributes affect the stability of collective cooperation. It provides a pure theoretical reference for realistic scenarios such as enterprise team management, alliance cooperation mechanism design, and public goods provision [13]. The practice of income distribution policy refers to Stiglitz J E [22]. Future research can be further extended to multi-type heterogeneous agent models, dynamic contribution adjustment models, and incomplete information scenarios to improve the theoretical system of collective cooperation stability.

AI Contribution Statement: This paper uses artificial intelligence tools for typesetting formatting, language polishing, and auxiliary translation of the theoretical framework. All core mathematical models, theorem derivation, logical construction, and economic interpretation are completed by the author. AI tools do not participate in the creation of the core research content and theoretical contributions of this paper.

Funding Statement: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability Statement: This research is purely based on theoretical mathematical modeling, and no empirical data or specific dataset is used in the study.

Acknowledgements: The author would like to thank the anonymous reviewers for their constructive comments on the theoretical framework and mathematical derivation of this paper.

Competing interests: The author declares that there is no conflict of interest regarding the publication of this paper.

References

[1] Individual vs. collective rationality, Encyclopedia of Law and Economics, Springer, 2025.

[2] H. Fichtenberger, A. Rey, Testing stability properties in graphical hedonic games, Auton. Agent Multi-Agent Syst. 35 (2021), 26.

[3] K.I. Shimomura, Individual Rationality and Collective Optimality, Ph.D. thesis, University of Rochester, 1994.

[4] D. Wakita, Y. Sobu, N. Kaneko, et al., A mathematical model suggests collectivity and inconstancy enhance the efficiency of neuronal migration in the adult brain, PLoS Comput. Biol. 21 (2025), e1013105.

[5] N. Hilmar, Anonymity confidentiality and invasions of privacy - responsibility of researcher, Amer. J. Public Health 58 (1968), 324–330.

[6] D.C. Castaneda Alvarez, Coalition Formation with Heterogeneous Agents, Ph.D. thesis, 2021.

[7] A. Magaña, F. Carreras, Coalition formation and stability, Group Decis. Negot. 27 (2018), 817–837.

[8] C. Rodríguez-Álvarez, Strategy-proof coalition formation, Int. J. Game Theory 38 (2009), 323–343.

[9] Y.H. Lin, et al., On the logic of dynamic coalition formation, Proc. Int. Conf. Intell. Syst. Knowl. Eng. 2008 (2008), 850–855.

[10] C.Q. Wang, et al., Modeling the social dilemma of involution on a square lattice, Chaos Solitons Fractals 158 (2022), 112092.

[11] M.L. Bertoni, M. Delitala, On the qualitative analysis of the solutions of a mathematical model of social dynamics, Appl. Math. Lett. 19 (2006), 1109–1112.

[12] S. Eshghi, C.R. Williams, G.B. Colombo, et al., Mathematical models for social group behavior, Proc. IEEE Smartworld 2017 (2017), 1–8.

[13] R. Zuo, C. He, J. Wu, et al., Simulating the impact of social resource shortages on involution competition: involution, sit-up, and lying-flat strategies, Comput. Math. Organiz. Theory 31 (2025), 27–42.

[14] P. Bich, R. Laraki, On the existence of approximate equilibria and sharing rule solutions in discontinuous games, Theor. Econ. 12 (2017), 79–108.

[15] H.P. Weikard, Cartel stability under an optimal sharing rule, Manch. Sch. 77 (2009), 575–593.

[16] S. Yang, Z. Tan, R. Zhao, et al., Operation optimization and income distribution model of park integrated energy system with power-to-gas technology and energy storage, J. Clean. Prod. 247 (2020), 119090.

[17] M. Öztürk, Distribution in Economic Systems and the Share Economy, Ph.D. thesis, Ïnönü University, 2013.

[18] B. Oancea, D. Pirjol, T. Andrei, A Pareto upper tail for capital income distribution, Physica A 492 (2018), 403–417.

[19] M.P. Victoria-Feser, Robust methods for the analysis of income distribution, inequality and poverty, Int. Stat. Rev. 68 (2000), 277–293.

[20] M.L. Bertotti, G. Modanese, Mathematical models for socio-economic problems, Mathematical Models and Methods for Planet Earth, Springer INdAM Series, Vol. 6, Springer, 2014, 123–134.

[21] G.A. Pushkarev, E.Y. Vorob'eva, V.A. Sokolov, On one modern approach in economic-mathematical modeling, Adv. Econ. Bus. Manag. Res. 47 (2018), 577–580.

[22] J.E. Stiglitz, Income Distribution and Redistribution Policies in the US Economy (1940–1980), China Social Sciences Press, Beijing, 1989.

[23] L.I. Abalkin, Distribution According to Work in the Socialist Economy of the USSR, People's Publishing House, Beijing, 1983.

[24] G.Y. Yu, On Socialist Distribution According to Work, People's Publishing House, Beijing, 1979.

[25] E.G. Luo, Collation and Annotations on the Celestial Field System, Zhonghua Book Company, Beijing, 1984.

[26] Department of Mathematics, East China Normal University, Mathematical Analysis (Vol. 1), 4th ed., Higher Education Press, Beijing, 2010.

[27] Department of Mathematics, East China Normal University, Mathematical Analysis (Vol. 2), 4th ed., Higher Education Press, Beijing, 2010.

[28] J.X. Chen, C.H. Yu, L. Jin, Mathematical Analysis, 2nd ed., Higher Education Press, Beijing, 2004.

[29] H.Y. Gao, Western Economics (Microeconomics), 7th ed., Renmin University of China Press, Beijing, 2018.