Model Extension: Author-Journal Game with Competition and Capacity Constraints

December 16, 2025

1 Motivation and Problem Statement

In our previous evolutionary game-theoretic model of the author-journal interaction, we observed that the equilibrium strategies were largely insensitive to the population size N. This phenomenon arises because, in the original formulation, the payoff of a focal author depends solely on their own strategy and the journal's acceptance policy, but is independent of the actions of other authors. Consequently, there is no *congestion effect* or *competition* for limited resources.

To address this and model a more realistic academic publishing environment (e.g., conferences with fixed acceptance rates or journals with limited pages), we introduce a capacity threshold K. This mechanism creates a zero-sum element: as the number of submissions increases, the probability of acceptance for a passed paper decreases, potentially disincentivizing opportunistic strategies like *Always Submit* (AS).

2 The Threshold Competition Model

The updated interaction proceeds in the following stages:

1. **Submission:** N authors choose a strategy $s \in \{OG, AS, \dots\}$.

2. **Peer Review (Perception):** All submissions undergo peer review characterized by false negative rate ϵ and false positive rate λ. Papers that pass peer review enter a "Perception Pool" of size M.

3. **Competition (Threshold Cut):** The journal has a maximum capacity K.

   - If $M \leq K$, all papers in the pool are accepted.
   - If $M > K$, the journal randomly selects K papers from the pool to accept. The remaining $M - K$ papers are rejected despite passing review.

Let α be the probability an author produces a good paper. The probability that a single author's submission passes the peer review stage, denoted as $P_{pass}(s)$, depends on their chosen strategy s:

$$P_{pass}(OG) = \alpha(1 - \epsilon) \tag{1}$$
$$P_{pass}(AS) = \alpha(1 - \epsilon) + (1 - \alpha)\lambda \tag{2}$$

Here, OG authors only submit good papers, while AS authors submit both good and bad papers, benefiting from the false positive rate λ.

3 Derivation of the Competition Factor γ

The key addition to the payoff function is the competition factor γ, which represents the probability that a paper is finally accepted *given that it has already passed peer review*.

Let us consider a *focal author* who has passed peer review. Let M_{-i} be the random variable representing the number of *other* authors (out of $N - 1$) whose papers also passed peer review.

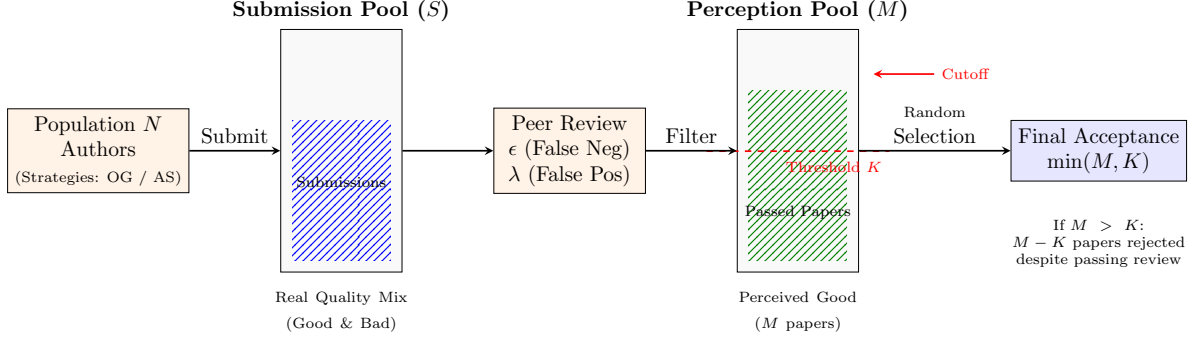The total number of passed papers is therefore $M = 1 + M_{-i}$.

Figure 1: **Schematic representation of the Threshold Competition Model.** The process unfolds in three stages: (1) Submission into a pool S; (2) Peer review filtering into a perception pool M; and (3) A hard capacity constraint K (red dashed line). If $M > K$, a random cutoff occurs, rejecting surplus papers even if they passed review. This introduces the competition factor γ.

The competition factor for the focal author is defined as:

$$\gamma_{\text{focal}} = \min\left(1, \frac{K}{1 + M_{-i}}\right) = \begin{cases} 1 & \text{if } 1 + M_{-i} \leq K \\ \frac{K}{1+M_{-i}} & \text{if } 1 + M_{-i} > K \end{cases} \tag{3}$$

3.1 Expected Competition Factor

Since M_{-i} is a random variable, authors base their decisions on the expected value $E[\gamma_{\text{focal}}]$.

Assume the population consists of N_{OG} authors playing OG and N_{AS} authors playing AS (where $N_{OG} + N_{AS} = N - 1$, excluding the focal author). The variable M_{-i} is the sum of two independent binomial distributions:

$$M_{-i} = X_{OG} + X_{AS} \tag{4}$$

where $X_{OG} \sim B(N_{OG}, P_{pass}(OG))$ and $X_{AS} \sim B(N_{AS}, P_{pass}(AS))$.

The expected competition factor is given by:

$$E[\gamma_{\text{focal}}] = \sum_{m=0}^{N-1} \Pr(M_{-i} = m) \cdot \min\left(1, \frac{K}{1+m}\right) \tag{5}$$

4 Payoff after competition

In this section, we rigorously define the payoff functions for both authors and journals under the Threshold Competition Model. The introduction of the capacity constraint K and the endogenous competition factor γ fundamentally alters the strategic landscape.

We first define the key variables used in the payoff derivation:

4.1 Journal

The journal's payoff is determined by the quality of the accepted papers minus the cost of reviewing. The crucial addition in this model is that the competition factor γ is endogenous to the journal's strategy. The journal maximizes its utility by choosing a strategy $S_J \in \{OG, AA, OB, AR\}$.

The competition factor γ is defined as the ratio of capacity K to the size of the perception pool M, capped at 1:

$$\gamma(S_J) = \min\left(1, \frac{K}{M(S_J)}\right) \tag{6}$$

Since the screening policy (determined by effective ϵ and λ) changes the size of M, γ varies across strategies. Let S_{good} and S_{bad} be the total number of good and bad papers submitted by the authors, such that $S = S_{good} + S_{bad}$.

Symbol	Description
N	Total number of authors
K	Journal capacity (maximum number of accepted papers)
S	Total number of submissions (Submission Pool)
M	Number of papers passing peer review (Perception Pool)
γ	Competition factor, probability of acceptance after passing review
α	Probability that an author produces a good paper
ϵ	False negative rate (probability of rejecting a good paper)
λ	False positive rate (probability of accepting a bad paper)
r	Reward for an accepted paper (Author)
c	Cost of submission (Author)
B	Benefit for accepting a good paper (Journal)
D	Penalty for accepting a bad paper (Journal)
$C(S)$	Cost of reviewing S submissions

Table 1: Key notation and definitions.

Only Good (OG) – Selective Strategy The journal applies standard peer review to filter papers, aiming to accept good papers and reject bad ones.

- **Screening parameters:** Standard ϵ and λ.

- **Perception Pool (M_{OG}):** Contains accepted good papers and falsely accepted bad papers.

$$M_{OG} = S_{good}(1 - \epsilon) + S_{bad}\lambda \tag{7}$$

- **Competition Factor (γ_{OG}):**

$$\gamma_{OG} = \min\left(1, \frac{K}{S_{good}(1 - \epsilon) + S_{bad}\lambda}\right) \tag{8}$$

- **Payoff ($U_J(OG)$):** The journal gains B from good papers and loses D from bad papers, scaled by γ_{OG}.

$$U_J(OG) = \gamma_{OG}\left[B \cdot S_{good}(1 - \epsilon) - D \cdot S_{bad}\lambda\right] - C(S) \tag{9}$$

Always Accept (AA) – Permissive Strategy The journal abandons screening, effectively setting $\epsilon = 0$ and $\lambda = 1$.

- **Perception Pool (M_{AA}):** All submissions enter the pool.

$$M_{AA} = S = S_{good} + S_{bad} \tag{10}$$

- **Competition Factor (γ_{AA}):** This strategy maximizes congestion.

$$\gamma_{AA} = \min\left(1, \frac{K}{S}\right) \tag{11}$$

- **Payoff ($U_J(AA)$):**

$$U_J(AA) = \gamma_{AA}\left[B \cdot S_{good} - D \cdot S_{bad}\right] - C(S) \tag{12}$$

Only Bad (OB) – Irrational Strategy The journal adopts an inverse selection policy, accepting papers that are perceived as "bad" by the peer review process. This strategy effectively targets papers that fail the standard quality threshold.

- **Perception Pool (M_{OB}):** The pool consists of good papers that were falsely rejected (false negatives) and bad papers that were correctly identified as bad (true negatives).

$$M_{OB} = S_{good}\epsilon + S_{bad}(1 - \lambda) \tag{13}$$

- **Competition Factor (γ_{OB}):**

$$\gamma_{OB} = \min\left(1, \frac{K}{S_{good}\epsilon + S_{bad}(1-\lambda)}\right) \tag{14}$$

- **Payoff ($U_J(OB)$):** Although the journal intends to select bad papers, it may accidentally accept good papers due to review errors (ϵ). However, it also accepts the vast majority of bad papers.

$$U_J(OB) = \gamma_{OB}\left[B \cdot S_{good}\epsilon - D \cdot S_{bad}(1-\lambda)\right] - C(S) \tag{15}$$

Note that despite the potential small gain from $B \cdot S_{good}\epsilon$, the term $-D \cdot S_{bad}(1-\lambda)$ typically dominates because $D > B$ and $(1-\lambda) \gg \epsilon$ in realistic settings. Thus, $U_J(OB)$ remains strictly dominated by strategies that do not actively seek penalties.

Always Reject (AR) – Baseline Strategy The journal rejects all submissions. The perception pool is empty ($M = 0$), and no papers are published.

- **Payoff ($U_J(AR)$):**

$$U_J(AR) = 0 \tag{16}$$

4.2   Author

An author chooses a strategy $S_A \in \{OG, OB, AS, NS\}$. The author's expected utility depends on the probability of producing a good paper (α), the cost of submission (c), the reward (r), and the probability of acceptance. The probability of acceptance is a two-step process: passing the peer review (denoted as P_{pass}) and surviving the competition cut (denoted by γ).

The expected payoff is generally given by:

$$U_A = P(\text{Good}) \cdot P(\text{Accept}|\text{Good}) \cdot r + P(\text{Bad}) \cdot P(\text{Accept}|\text{Bad}) \cdot r - \text{Total Cost} \tag{17}$$

Assuming the journal plays the standard *Only Good* strategy (implying the competition factor is γ_{OG}), the payoffs for the four author strategies are derived as follows:

Only Good (OG) The author submits only if the paper is good.

- Probability of submission: α

- Probability of passing review: $1 - \epsilon$

- **Payoff:**

$$U_A(OG) = \alpha\left[(1-\epsilon)\gamma_{OG}r - c\right] \tag{18}$$

Only Bad (OB) The author submits only if the paper is bad, attempting to exploit the false positive rate.

- Probability of submission: $1 - \alpha$

- Probability of passing review: λ

- **Payoff:**

$$U_A(OB) = (1-\alpha)\left[\lambda\gamma_{OG}r - c\right] \tag{19}$$

Always Submit (AS) The author submits regardless of quality.

- Probability of submission: 1

- Probability of passing review: The paper passes if it is good and correctly identified ($1-\epsilon$), or if it is bad and falsely accepted (λ).

- **Payoff:**

$$U_A(AS) = \left[\alpha(1-\epsilon) + (1-\alpha)\lambda\right]\gamma_{OG}r - c \tag{20}$$

Never Submit (NS)   The author does not submit any papers.

- **Payoff:**

$$U_A(NS) = 0 \tag{21}$$

The assumption that these constraints hold is strictly necessary to model a functional and sustainable scientific community rather than a degenerate system. Specifically, the condition $c > \lambda r$ ensures that the ecosystem is not overwhelmed by costless speculative spamming; the lower bound on γ posits that the journal has not reached a state of total collapse where even high-quality contributions are discouraged; and the reputation constraint on D distinguishes legitimate, quality-seeking venues from predatory publishers. By restricting our analysis to this valid parameter space, we filter out trivial breakdown scenarios to focus exclusively on the non-trivial evolutionary dynamics: how resource scarcity (K) and congestion precipitate a tragedy of the commons, and how the interplay between institutional screening and authorial strategy can resolve it.

We incorporate $E[\gamma_{\text{focal}}]$ into the utility functions. Let r be the reward for acceptance and c be the cost of submission.

An OG author submits with probability α. If they submit, they pay cost c. They pass review with probability $1 - \epsilon$. If they pass, they are accepted with probability $E[\gamma_{\text{focal}}]$.

$$U_A(OG) = \alpha \left[(1 - \epsilon) \cdot E[\gamma_{\text{focal}}] \cdot r - c \right] \tag{22}$$

Note: Assuming cost is paid per submission regardless of acceptance.

An AS author always submits (probability 1), paying cost c. They pass review with probability $P_{pass}(AS)$. If they pass, they are accepted with probability $E[\gamma_{\text{focal}}]$.

$$U_A(AS) = \left[(\alpha(1 - \epsilon) + (1 - \alpha)\lambda) \cdot E[\gamma_{\text{focal}}] \cdot r \right] - c \tag{23}$$

The journal's utility depends on the quality of the final accepted papers and the cost of reviewing the total volume of submissions.

Let B be the benefit for accepting a good paper and D be the penalty for accepting a bad paper. The total submission volume S is given by:

$$E[S] = N_{OG} \cdot \alpha + N_{AS} \cdot 1 \tag{24}$$

The "Perception Pool" M contains a mix of good and bad papers that passed review. The expected number of good papers (M_G) and bad papers (M_B) in the pool are:

$$E[M_G] = (N_{OG} + N_{AS}) \cdot \alpha(1 - \epsilon) \tag{25}$$
$$E[M_B] = N_{AS} \cdot (1 - \alpha)\lambda \tag{26}$$

Note: Both OG and AS authors contribute to M_G, but only AS authors contribute to M_B.

Since the journal selects K papers randomly from M (if $M > K$), the final accepted papers maintain the same quality ratio as the pool. The expected acceptance rate for any paper in the pool is approximated by the system-wide competition factor $\bar{\gamma} = E[\min(1, K/M)]$.

The journal's expected payoff is:

$$U_J = B \cdot (E[M_G] \cdot \bar{\gamma}) - D \cdot (E[M_B] \cdot \bar{\gamma}) - C(E[S]) \tag{27}$$

Where $C(S)$ is the cost function (e.g., linear or convex with respect to total submissions). This formulation highlights that while K limits the absolute number of bad papers accepted (by reducing $\bar{\gamma}$), it does not improve the *ratio* of good to bad papers, which is still determined by ϵ and λ.

5   Conditional Dominated strategies

To determine the evolutionary stability of the system, we perform a pairwise comparison of payoffs to identify strictly dominated strategies that can be eliminated from the strategy space. A strategy is considered strictly dominated if its expected utility is lower than that of another strategy under all admissible parameter configurations. Conversely, a strategy is conditionally viable if its dominance depends on specific inequalities involving the competition factor γ and the penalty structure.

We first analyze the author's strategy space. The strategy *Never Submit* (NS) serves as the baseline with a fixed payoff of $U_A(NS) = 0$. Consider the *Only Bad* (OB) strategy, where an author submits only

bad papers to exploit the false positive rate λ. The expected payoff is $U_A(OB) = (1-\alpha)[\lambda\gamma r - c]$. Given the fundamental cost constraint $c > \lambda r$ (assuming submission costs exceed the expected reward of a lucky acceptance in a vacuum) and the fact that $\gamma \leq 1$, the term $[\lambda\gamma r - c]$ is strictly negative. Consequently, $U_A(OB) < U_A(NS)$ holds universally, rendering OB a strictly dominated strategy. In contrast, the *Only Good* (OG) and *Always Submit* (AS) strategies are conditionally viable. $U_A(OG)$ remains positive as long as the competition factor satisfies $\gamma > c/[r(1-\epsilon)]$. Similarly, AS remains viable only when the dilution of acceptance probability γ does not render the submission of bad papers prohibitively costly. Thus, the author's effective strategy space reduces to $\{OG, AS, NS\}$.

Next, we examine the journal's strategy space, using *Always Reject* (AR) as the baseline with $U_J(AR) = 0$. The *Only Bad* (OB) strategy, representing an irrational inverse selection policy, yields a payoff of $U_J(OB) = \gamma_{OB}[BS_{good}\epsilon - DS_{bad}(1-\lambda)] - C(S)$. Although there is a non-zero probability ϵ of accepting a good paper by mistake, the operational parameters of a functional peer review system satisfy $(1-\lambda) \gg \epsilon$ and $D > B$. The penalty term from correctly identifying bad papers overwhelmingly dominates the accidental benefit term, ensuring $U_J(OB) < 0$ for any non-zero submission volume. Thus, Journal OB is strictly dominated by AR. The *Always Accept* (AA) strategy is conditionally dominated. Its payoff $U_J(AA)$ becomes negative if the penalty for bad papers satisfies the reputation constraint $D > B(S_{good}/S_{bad})$. Under strict reputation mechanisms, AA is eliminated, leaving *Only Good* (OG) as the sole dominant strategy for the journal, provided the screening costs do not exceed the utility of publishing good research.

The complete classification of strategy dominance is summarized in Table 2.

Strategy	Payoff Function	Status	Condition / Reasoning
Author Strategies (Baseline: $U_A(NS) = 0$)			
Only Bad (OB)	$(1-\alpha)[\lambda\gamma r - c]$	**Strictly Dominated**	Dominated by NS since $c > \lambda r$ and $\gamma \leq 1$.
Only Good (OG)	$\alpha[(1-\epsilon)\gamma r - c]$	Conditionally Viable	Viable if $\gamma > \frac{c}{r(1-\epsilon)}$.
Always Submit (AS)	$[\alpha(1-\epsilon) + (1-\alpha)\lambda]\gamma r - c$	Conditionally Viable	Dominated by OG if congestion is high ($\gamma \to 0$); viable if $\gamma \to 1$ and costs are low.
Never Submit (NS)	0	Survival	Optimal when competition is prohibitive.
Journal Strategies (Baseline: $U_J(AR) = 0$)			
Only Bad (OB)	$\gamma_{OB}[BS_{good}\epsilon - DS_{bad}(1-\lambda)] - C(S)$	**Strictly Dominated**	Dominated by AR. Penalties (D) outweigh accidental benefits ($B\epsilon$).
Always Accept (AA)	$\gamma_{AA}[BS_{good} - DS_{bad}] - C(S)$	Conditionally Dominated	Dominated by AR if $D > B\frac{S_{good}}{S_{bad}}$.
Only Good (OG)	$\gamma_{OG}[BS_{good}(1-\epsilon) - DS_{bad}\lambda] - C(S)$	**Dominant**	Strictly yields highest utility under standard parameters.
Always Reject (AR)	0	Survival	Optimal only if system collapses ($U_J(OG) < 0$).

Table 2: Pairwise comparison and dominance analysis of Author and Journal strategies under the Threshold Competition Model.

To establish the dynamic congestion model characterized by bistability (three equilibria), the game is governed by the following system of inequalities. Note that Condition (1) allows for conditional profitability of bad papers, which is essential for the survival of the "Always Submit" strategy.

$$\begin{cases} \textbf{(1) Profitability Condition:} & \lambda r > c \\ \textbf{(2) Reputation Constraint:} & D > \frac{\alpha}{1-\alpha}B \\ \textbf{(3) Viability Constraint:} & \gamma > \frac{c}{r(1-\epsilon)} \end{cases} \tag{28}$$

Where c is the submission cost, r is the reward for authors, D is the damage to the journal, B is the benefit to the journal, and λr represents the potential return of a bad paper if accepted.

1. Elimination of Author Strategies

The strategy **{Never Submit}**, defined as $\sigma_A = (NS|G, NS|B)$, is strictly dominated for Good (G) types due to the **Viability Constraint**.

$$U_A(NS \mid G) = 0$$
$$U_A(S \mid G) = \gamma \cdot r \cdot (1 - \epsilon) - c$$

From inequality (3), we have $\gamma > \frac{c}{r(1-\epsilon)}$, implying $U_A(S \mid G) > 0$. Rational authors holding good papers will always submit.

The strategy **{Only Bad}**, defined as $\sigma_A = (NS|G, S|B)$, is eliminated due to the violation of **Monotonicity**, rather than cost. While Inequality (1) $\lambda r > c$ implies that submitting a bad paper *can* be profitable (if the Journal accepts), submitting a good paper is strictly *more* profitable because the reward r is strictly greater than the discounted reward λr (assuming $\lambda < 1$).

$$U_A(S \mid G) \approx r - c$$
$$U_A(S \mid B) \approx \lambda r - c$$

Since $r > \lambda r$, we have $U_A(S \mid G) > U_A(S \mid B)$. If a rational author chooses to submit a Bad paper $(S|B)$, they must strictly prefer submitting a Good paper $(S|G)$. Therefore, the strategy $(NS|G, S|B)$ is strictly dominated by **{Always Submit}** $(S|G, S|B)$.

*Note: Unlike the simplified static model, the strategy **{Always Submit}** is NOT dominated here, as Inequality (1) allows $U_A(S|B) > 0$ under conditions of high congestion (low screening).*

2. Elimination of Journal Strategies

The strategy **{Always Accept}**, defined as accepting all submissions, is eliminated due to the **Reputation Constraint**. The expected utility is:

$$E[U_J(\text{Always Accept})] = \alpha B - (1 - \alpha)D$$

From inequality (2), $(1 - \alpha)D > \alpha B$, which implies $E[U_J] < 0$. The damage from bad papers outweighs the benefit of good ones in a blind acceptance scenario, forcing the journal to screen or reject.

The strategy **{Only Bad}**, defined as $\sigma_J = (R|G, A|B)$, is strictly dominated by rationality constraints:

$$U_J(R \mid G) = 0 < B = U_J(A \mid G)$$
$$U_J(A \mid B) = -D < 0 = U_J(R \mid B)$$

This strategy minimizes utility in all states and is discarded.

The strategy **{Reject All}**, defined as $\sigma_J = (R|G, R|B)$, is dominated by the conditional strategy {Only Good}. Since $U_J(A \mid G) = B > 0$, the Journal strictly prefers accepting Good papers over rejecting them. Thus, conditional acceptance dominates blanket rejection.

Conclusion: The resulting reduced game consists of the strategies:

- **Author:** $\mathcal{S}_A = \{\text{Only Good}, \text{Always Submit}\}$

- **Journal:** $\mathcal{S}_J = \{\text{Only Good}\}$

This strategy set allows for the evolutionary dynamics where the Author switches between "Only Good" and "Always Submit" based on the Journal's congestion-dependent screening effectiveness.

6 Replicator Dynamics and Bifurcation Analysis

6.1 Reduction to a Two-Strategy System and Congestion Model

Following the elimination of strictly dominated strategies (specifically {Only Bad} and {Never Submit}), the game is reduced to a dynamic contest between two Author strategies, while the Journal plays a fixed strategy of *Only Good*. We define the population state $x \in [0, 1]$ as follows:

- x: The proportion of Authors playing **{Only Good}**.

- $1 - x$: The proportion of Authors playing **{Always Submit}**.

Since the Journal's strategy is fixed, the evolutionary dynamics are driven entirely by the payoff difference between the Author's two remaining strategies. To allow for the survival of the *Always Submit* strategy and the emergence of complex dynamics, we assume the **Profitability Condition** ($\lambda B > c$) holds. This implies that submitting a "Bad" paper is profitable *if* the journal fails to screen it.

6.1.1 Physical Definitions and Variables

We introduce a finite resource constraint on the Journal to model the **Congestion Effect**. The parameters are defined as follows:

- N: Total population of authors.

- K: Journal processing capacity (total review effort available).

- C_J: Cost per review. Thus, K/C_J represents the maximum number of papers the journal can effectively screen.

- α: The prior probability that an author holds a "Good" paper.

- B: Benefit (Reward) of acceptance.

- C: Cost of submission.

- D: Reputational damage penalty for a rejected "Bad" paper.

- λ: The discount factor for a "Bad" paper (Payoff = λB).

- ϵ: A friction term to prevent singularities.

6.2 Derivation of the Polynomial Dynamics

To analyze the stability of the system, we express the rate of change $\frac{dx}{dt}$ as a function of the submission volume and screening effectiveness.

6.2.1 Volume and Screening Functions

The total submission volume $V(x)$ depends on the strategy distribution. "Only Good" authors submit with probability α, while "Always Submit" authors submit with probability 1:

$$V(x) = N[\alpha x + (1 - x)] = N[1 - (1 - \alpha)x] \tag{29}$$

Note that $V(x)$ is linear and decreasing in x; as honesty increases, junk submissions decrease.

The screening probability $\sigma(x)$ is endogenous and depends on the ratio of capacity to volume:

$$\sigma(x) = \frac{K}{C_J \cdot V(x) + \epsilon} \tag{30}$$

This inverse relationship creates the nonlinearity required for bistability.

6.2.2 Utility Functions and Replicator Equation

The evolutionary dynamics depend on the expected utility of submitting a "Bad" paper, $E[U_{Bad}]$.

- If screened (probability σ): Payoff is $-C - D$.

- If unscreened (probability $1 - \sigma$): Payoff is $\lambda B - C$.

The expected utility is:

$$E[U_{Bad}](x) = \sigma(x)(-C - D) + (1 - \sigma(x))(\lambda B - C)$$
$$= \sigma(x)(-D - \lambda B) + (\lambda B - C) \tag{31}$$

Substituting this into the standard replicator equation $\frac{dx}{dt} = -x(1 - x)(1 - \alpha)E[U_{Bad}]$, and clearing the denominator ($C_J V(x) + \epsilon$), we derive the polynomial $P(x)$ that characterizes the sign of the dynamics:

$$P(x) = \underbrace{x(1 - x)}_{\text{Deg 2}} \cdot \underbrace{\left[K(-D - \lambda B) + (\lambda B - C)(C_J N(1 - (1 - \alpha)x) + \epsilon) \right]}_{\text{Deg 1}} \tag{32}$$

6.2.3 Degree Analysis and Root Configuration

The polynomial $P(x)$ is of **Degree 3 (Cubic)**. This is confirmed by the product of the quadratic replicator term $x(1-x)$ and the linear volume term nested within the utility function. Consequently, the system admits up to three real roots in $[0, 1]$:

1. $x = 0$: **Full Corruption** (Stable if congestion makes cheating profitable).

2. $x = 1$: **Ideal Honesty** (Stable if low volume makes screening strict).

3. $x = x^*$: An internal **Tipping Point** (Unstable).

The analytical form of the internal root is derived as:

$$x^* = \frac{1}{1-\alpha} \left[1 + \frac{\epsilon}{C_J N} - \frac{K(D + \lambda B)}{C_J N(\lambda B - C)} \right] \tag{33}$$

6.3 Numerical Simulation and Parameter Analysis

6.3.1 Case 1: The Failure of Congestion ($K = 500$)

In the initial simulation with $N = 1000$ and $K = 500$, the journal capacity was sufficient to screen all papers even at maximum volume (N). Mathematically, this resulted in $x^* < 0$. Under these conditions, the effective screening rate $\sigma \approx 1$ rendered cheating globally unprofitable ($U_{Bad} < 0$), collapsing the game to a single stable equilibrium at $x = 1$.

6.3.2 Case 2: The Emergence of Bistability ($K = 100$)

By reducing capacity to $K = 100$, we induced the congestion effect. At $x = 0$, high volume overwhelms the journal (σ is low), satisfying the profitability condition $\lambda B > C$. At $x = 1$, volume is low, σ is high, and cheating is punished. This trade-off creates the internal root x^*.
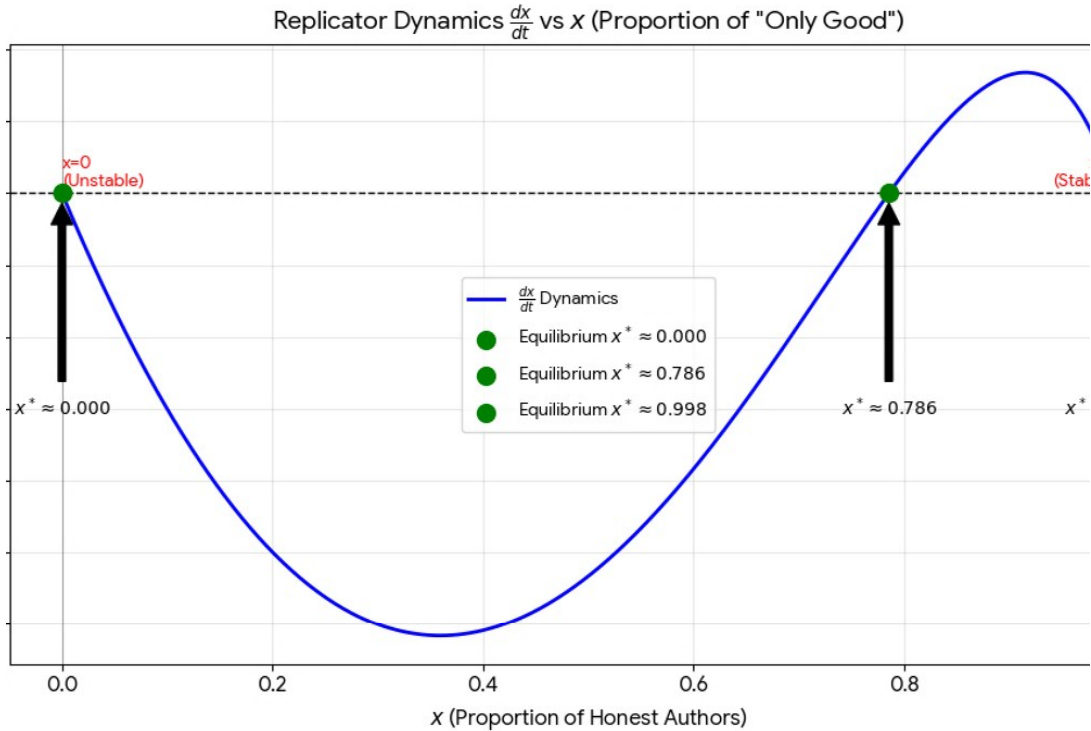

Figure 2: Replicator Dynamics $\frac{dx}{dt}$ vs x with $K = 100$. The system exhibits bistability. The internal root $x^* \approx 0.786$ acts as an unstable repeller. If the initial proportion of honest authors $x < x^*$, the system collapses to full corruption ($x = 0$). If $x > x^*$, it converges to full honesty ($x = 1$).

Figure 2 visualizes this dynamic. The existence of x^* in $(0, 1)$ confirms that the "Always Submit" strategy is viable only when the journal is congested.

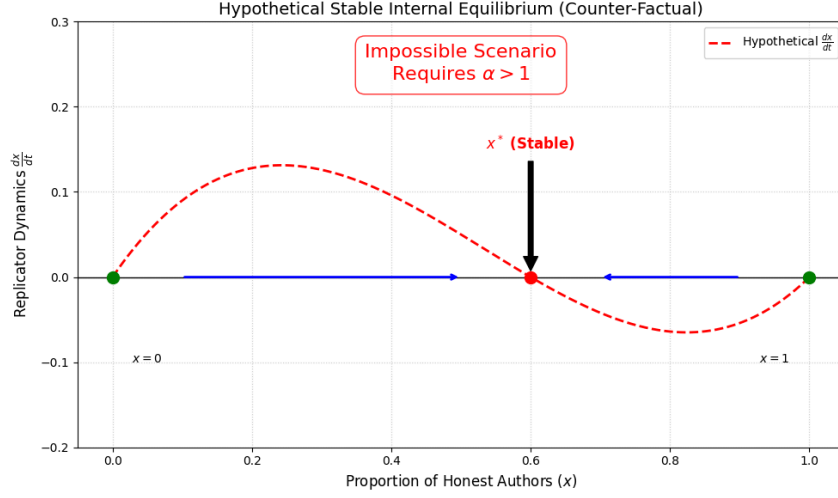### 6.3.3 Analytical Proof of Equilibrium Instability via Monotonicity

Figure 3: Hypothetical representation of a stable internal equilibrium (Counter-Factual). This dynamic profile, where $\dot{x}$ transitions from positive to negative at x^*, would require the strictly decreasing volume function $V(x)$ to be increasing. Mathematically, this implies $\alpha > 1$, which violates the physical definition of probability. Thus, a stable mixed strategy is structurally impossible in this model.

To rigorously characterize the stability of the internal equilibrium x^*, we analyze the first derivative of the evolutionary velocity field with respect to the population state x. The sign of the replicator dynamic equation $\dot{x}$ is determined strictly by the payoff differential between the "Only Good" and "Always Submit" strategies. Since the payoff for honest authors is strictly positive under the viability constraint, the system's trajectory is governed by the sign of the net disadvantage of the "Always Submit" strategy. We define the stability indicator function $\Psi(x)$ as the negative expected utility of submitting a "Bad" paper:

$$\Psi(x) = -E[U_{Bad}](x) = \sigma(x)(D + \lambda B) - (\lambda B - C) \tag{34}$$

The stability of any fixed point x^* where $\Psi(x^*) = 0$ depends on the slope $\Psi'(x^*)$. A positive slope implies that the dynamics transition from negative to positive as x increases, identifying x^* as an unstable repeller. We first observe that the total submission volume $V(x)$ is linearly decreasing in x, given the physical constraint $\alpha < 1$:

$$\frac{dV(x)}{dx} = \frac{d}{dx}\left[N(1 - (1-\alpha)x)\right] = -N(1-\alpha) < 0 \tag{35}$$

Since the screening probability $\sigma(x)$ is inversely proportional to the volume $V(x)$, and the indicator function $\Psi(x)$ is linear in $\sigma(x)$, we apply the chain rule to determine the monotonicity of the system:

$$\frac{d\Psi}{dx} = \underbrace{\frac{\partial\Psi}{\partial\sigma}}_{(D+\lambda B)>0} \cdot \underbrace{\frac{d\sigma}{dV}}_{-\frac{K}{C_J V^2}<0} \cdot \underbrace{\frac{dV}{dx}}_{-N(1-\alpha)<0} \tag{36}$$

The product of two negative terms and one positive term yields a strictly positive result: $\Psi'(x) > 0$ for all $x \in [0, 1]$. This strict monotonicity proves that the function $\Psi(x)$ can cross zero at most once, and must do so from below. Consequently, the replicator dynamics $\dot{x}$ must change sign from negative to positive at x^*. This mathematically precludes the existence of a stable internal equilibrium (which would require $\Psi'(x) < 0$) and confirms that the system exhibits global bistability, where x^* acts as a basin boundary separating the regions of convergence to full corruption ($x = 0$) and full honesty ($x = 1$).

6.3.4 Impact of Population (N) and Capacity (K)

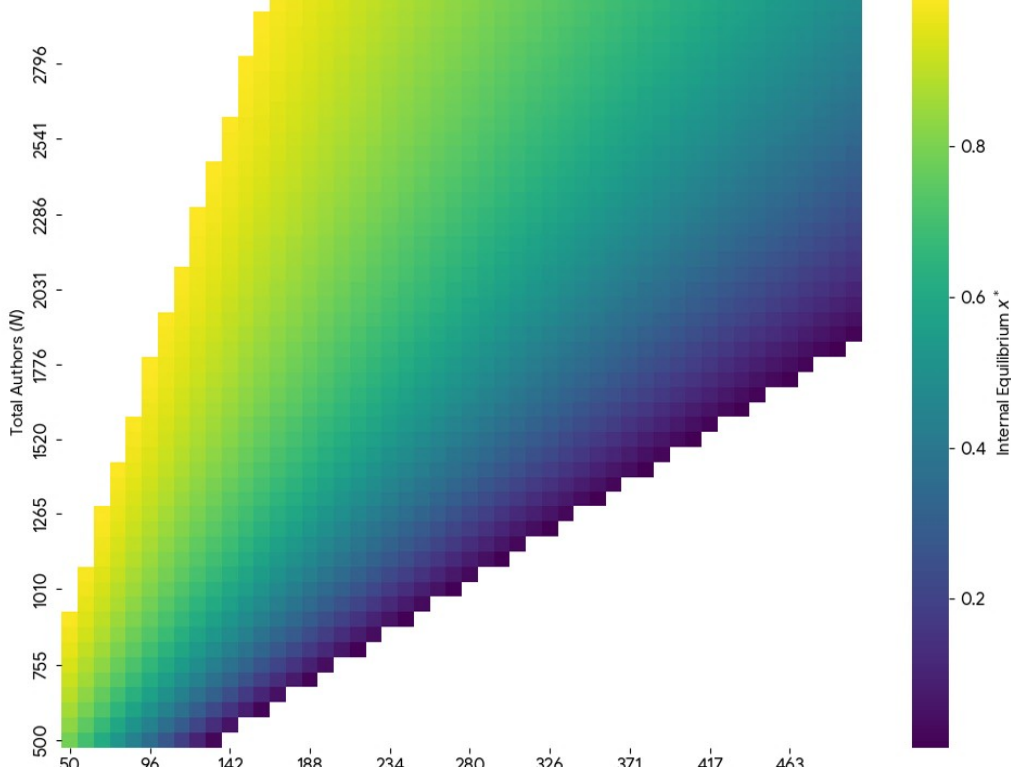We performed a parameter sweep to analyze how the tipping point x^* shifts with N and K.

Figure 4: Heatmap of the internal equilibrium x^* relative to Total Authors (N) and Journal Capacity (K). The color gradient represents the position of the tipping point.

Analysis of Figure 4:

- **White Region (Right):** Represents parameters where K is high relative to N. Here, the inequality $c > \lambda B$ effectively holds due to strict screening. No internal root exists, and honesty ($x = 1$) is the only equilibrium.

- **Colored Region (Left):** Represents the "Congestion Zone" where $\lambda B > c$ is activated by low screening rates. As N increases (moving up), the region of corruption expands (darker colors), indicating that a larger critical mass of honest authors is required to stabilize the system.

7 Parameter Sensitivity and Bifurcation Analysis

In this section, we explore how the internal equilibrium x^* shifts in response to variations in the journal capacity K, which is the primary determinant of the congestion effect in our model. We analyze the conditions under which the system transitions between the three identified regimes: Global Corruption ($x^* \to 1$), Bistability ($0 < x^* < 1$), and Global Honesty ($x^* \to 0$).

Unless otherwise stated, the numerical analysis assumes the default parameter set specified in Table 3. Note that the default capacity is set to $K = 100$, which corresponds to the bistable (congested) scenario observed in the previous analysis.

7.1 Impact of Journal Capacity K

The journal capacity K acts as the critical constraint on the screening effectiveness. We identify two critical threshold values for K that define the boundaries of the stability regimes. These thresholds are derived analytically by solving for the capacity required to make the expected utility of a bad paper zero ($E[U_{Bad}] = 0$) at the boundary states of the system ($x = 0$ and $x = 1$).

Table 3: Default Parameters for Numerical Analysis

Symbol	Description	Value
N	Total population of authors	1000
K	Journal Capacity	100
α	Probability of producing a good paper	0.2
ϵ	False negative rate	0.01
λ	Return rate for bad papers (False positive)	0.8
B	Benefit of acceptance	100
D	Penalty for accepted bad paper	50
C	Cost of submission	10
C_J	Unit cost of review	0.5

The required screening probability σ^* to neutralize the incentive for cheating is derived from the utility condition:

$$E[U_{Bad}] = \sigma^*(-D - \lambda B) + (\lambda B - C) = 0 \implies \sigma^* = \frac{\lambda B - C}{D + \lambda B} \tag{37}$$

Using the default parameters, we obtain $\sigma^* \approx 0.5385$. The corresponding capacity K is related to the submission volume $V(x)$ by $K = \sigma^*(C_J V(x) + \epsilon)$.

1. **Threshold for Global Honesty (K_{high})**: This occurs when the journal can maintain effective screening ($\sigma \geq \sigma^*$) even under the maximum submission load at full corruption ($x = 0, V(0) = N$).

$$K_{high} = \sigma^*(C_J N + \epsilon) \approx 269.2 \tag{38}$$

If $K > K_{high}$, the system converges to full honesty regardless of the initial state.

2. **Threshold for Global Corruption (K_{low})**: This occurs when the journal fails to maintain effective screening even under the minimum submission load at full honesty ($x = 1, V(1) = \alpha N$).

$$K_{low} = \sigma^*(C_J \alpha N + \epsilon) \approx 53.9 \tag{39}$$

If $K < K_{low}$, the system collapses to full corruption regardless of the initial state.

The dynamic behavior of the system under these three regimes is illustrated in Figure 5.

It is worth noting that the dynamic curve for the high capacity scenario (Green line in Figure 5) exhibits a non-smooth "kink" or discontinuity in its derivative. This phenomenon arises due to the saturation of the screening probability. At low values of x (high congestion), the screening probability is limited by capacity: $\sigma(x) = K/(C_J V(x) + \epsilon) < 1$. However, as the proportion of honest authors x increases, the submission volume $V(x)$ decreases significantly. Once x exceeds a certain threshold, the capacity K becomes sufficient to review every submission thoroughly, causing the calculated $\sigma(x)$ to reach its physical upper bound of 1.

$$\sigma(x) = \min\left(1, \frac{K}{C_J V(x) + \epsilon}\right) \tag{40}$$

The kink represents the transition point where the system switches from a congestion-constrained state to an uncongested state where $\sigma(x)$ is fixed at 1. Beyond this point, the dynamics are driven solely by the replicator term $x(1 - x)$, resulting in a change in the curvature of $\dot{x}$.

7.2 Impact of Total Population N

While the journal capacity K represents the supply side of the screening resources, the total population N determines the demand intensity. In this section, we analyze how the internal equilibrium x^* shifts as the population size varies, assuming a fixed capacity (default $K = 100$).

The relationship between population size and the tipping point is governed by the congestion mechanism. Analytically, the internal equilibrium x^* is the state where the incentive to cheat is exactly neutralized ($E[U_{Bad}] = 0$). This requires a specific screening probability $\sigma^* = \frac{\lambda B - C}{D + \lambda B}$. By substituting
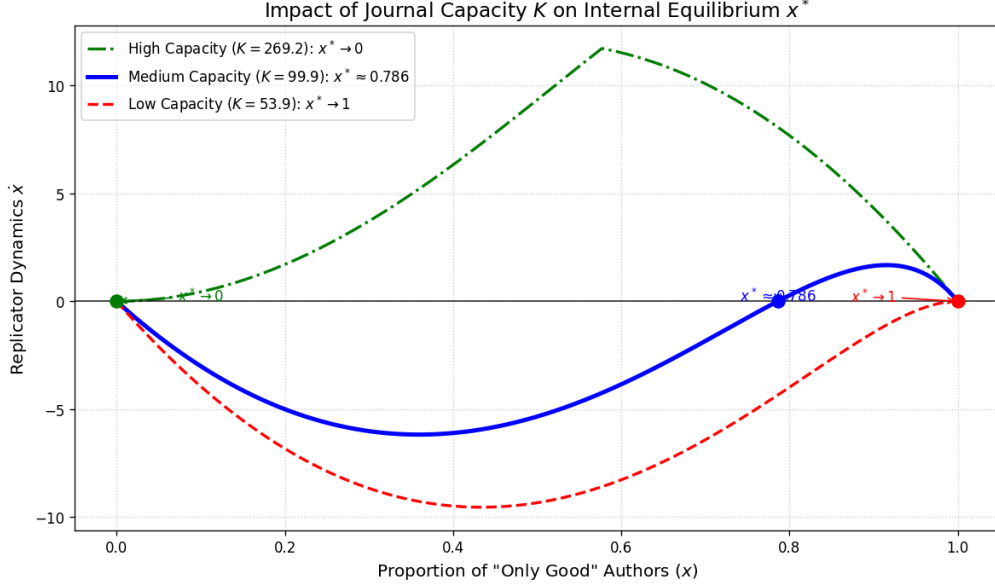
Figure 5: Phase portrait of the replicator dynamics $\dot{x}$ for three distinct capacity regimes. The red dashed line ($K \approx 53.9$) represents a low-capacity scenario leading to global corruption. The blue solid line ($K \approx 99.9$) represents the bistable regime with an internal unstable equilibrium at $x^* \approx 0.786$. The green dash-dotted line ($K \approx 269.2$) represents a high-capacity scenario ensuring global honesty.

the definition of σ into the volume function $V(x) = N[1 - (1 - \alpha)x]$, we can solve for x^* as a function of N:

$$x^*(N) = \frac{1}{1-\alpha} \left(1 - \frac{1}{N \cdot C_J} \left[\frac{K}{\sigma^*} - \epsilon \right] \right) \tag{41}$$

This equation reveals a non-linear relationship between population size and system stability, as illustrated in Figure 6. We identify three distinct regimes:

1. **Global Honesty Regime** ($x^* < 0$): When N is small (left side of the curve), the calculated tipping point x^* is negative. Mathematically, this implies that for any physically valid state $x \in [0, 1]$, the system lies to the "right" of the unstable repeller. Consequently, the evolutionary dynamics always drive the system towards $x = 1$. Physically, the population is small enough that the journal can effectively screen all submissions even in the worst-case scenario.

2. **Global Corruption Regime** ($x^* > 1$): When N is extremely large (right side of the curve), x^* exceeds 1. The entire physical range $[0, 1]$ lies to the "left" of the repeller, causing the system to collapse towards $x = 0$. Here, the sheer volume of submissions overwhelms the fixed capacity, making screening ineffective regardless of author behavior.

3. **Bistability Regime** ($0 \leq x^* \leq 1$): For intermediate populations (highlighted in yellow), the unstable equilibrium x^* lies within the physical range. This phenomenon, known as **Bistability**, indicates that the long-term outcome of the system is not determined solely by the parameters, but is *path-dependent*.

   - If the initial proportion of honest authors is high ($x_0 > x^*$), the system generates enough "self-discipline" (low submission volume) to maintain effective screening, converging to Honesty ($x = 1$).

   - If the initial proportion is low ($x_0 < x^*$), the high volume of junk submissions dilutes screening capacity, making cheating profitable and driving the system to Corruption ($x = 0$).

13

Bistability represents a "trap": a journal that has lost its reputation (low x) cannot recover simply by maintaining the same resources that would sustain a healthy journal. Recovery would require a massive reduction in N or increase in K to shift x^* back to zero.
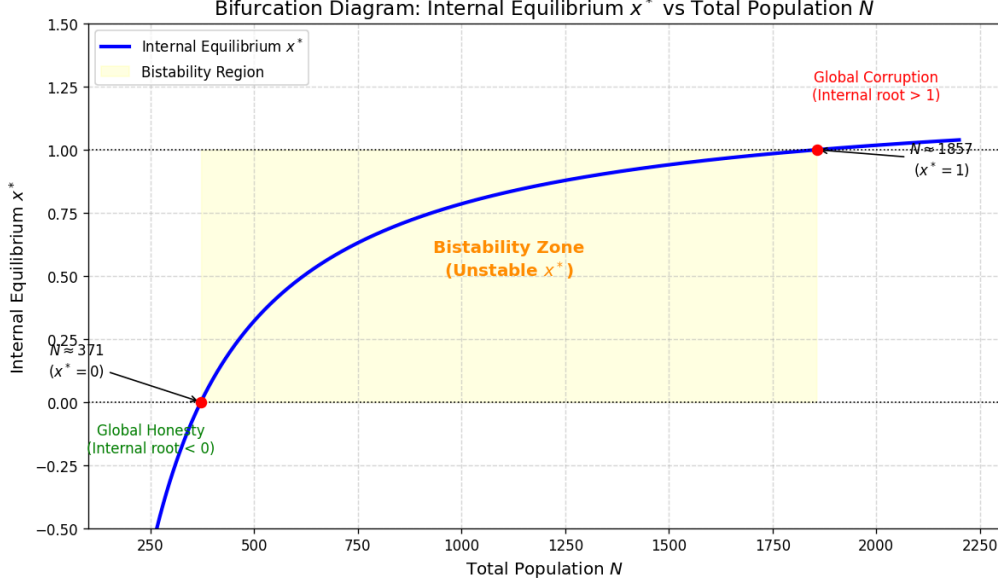

Figure 6: Bifurcation diagram showing the dependence of the internal equilibrium x^* on the total population N. The yellow shaded region indicates the zone of **Bistability**, where the system's fate depends on initial conditions. The critical thresholds mark the transitions to Global Honesty ($N \lesssim 371$) and Global Corruption ($N \gtrsim 1857$).

7.3 Impact of Author Productivity α

Finally, we investigate the influence of the author's productivity parameter α, which represents the probability that a submission is genuinely "Good". While one might intuitively expect higher productivity to benefit the system, our analysis reveals a counter-intuitive dynamic under capacity constraints. Holding the population $N = 1000$ and capacity $K = 100$ constant, the stability of the system is determined by whether the submission volume exceeds the critical threshold V_{crit} required to maintain effective screening. The relationship between the internal equilibrium x^* and α is derived by equating the actual submission volume to this critical limit:

$$V(x^*) = N[1 - (1 - \alpha)x^*] = V_{crit} \tag{42}$$

where the critical volume is determined by the screening parameters as $V_{crit} = \frac{1}{C_J}(\frac{K}{\sigma^*} - \epsilon) \approx 371.4$. Solving this equation for x^*, we obtain a hyperbolic relationship that characterizes the tipping point:

$$x^*(\alpha) = \frac{1 - \frac{V_{crit}}{N}}{1 - \alpha} \tag{43}$$

The resulting bifurcation diagram, presented in Figure 7, illustrates how the stability of the academic ecosystem is sensitive to the capabilities of its participants.

A striking feature of this relationship is that the equilibrium curve originates at a non-zero baseline of $x^* \approx 0.63$ when $\alpha = 0$. This intercept, which we term the system's "Base Load" of Honesty, reveals the fundamental structural constraints of the journal. In a scenario where honest authors produce no good papers ($\alpha = 0$), they effectively adopt a "Silence" strategy, submitting nothing. Meanwhile, opportunistic authors continue to saturate the journal with bad papers. To prevent the fixed capacity $K = 100$ from

being overwhelmed, the number of active submitters must be strictly limited. Mathematically, the proportion of authors who must remain silent (honest) is governed by the ratio of critical capacity to total population:

$$x^*_{min} = x^*(0) = 1 - \frac{N^{max}_{AS}}{N} \approx 1 - \frac{371.4}{1000} \approx 0.629 \tag{44}$$

This implies that even in the absence of valid scientific output, a significant majority of the population must abstain from spamming for the peer-review system to function.

As the productivity α increases, honest authors begin to contribute valid submissions to the pool, leading to a phenomenon we identify as the "Productivity Paradox". Paradoxically, the legitimate volume generated by productive honest authors adds to the overall congestion. To maintain the total volume at the sustainable level V_{crit}, the system must further compensate by reducing the number of opportunistic spammers. This necessitates an even higher proportion of honest authors to displace the noise with quality. Once α exceeds a critical threshold (≈ 0.37), the required honest fraction x^* exceeds unity:

$$x^*(\alpha) > 1 \quad \text{for} \quad \alpha > 1 - \frac{V_{crit}}{N} \tag{45}$$

At this point, the combined volume of high-quality submissions and opportunistic spam overwhelms the journal regardless of the strategy distribution, rendering the bistable equilibrium physically impossible and forcing the system into a state of Global Corruption.
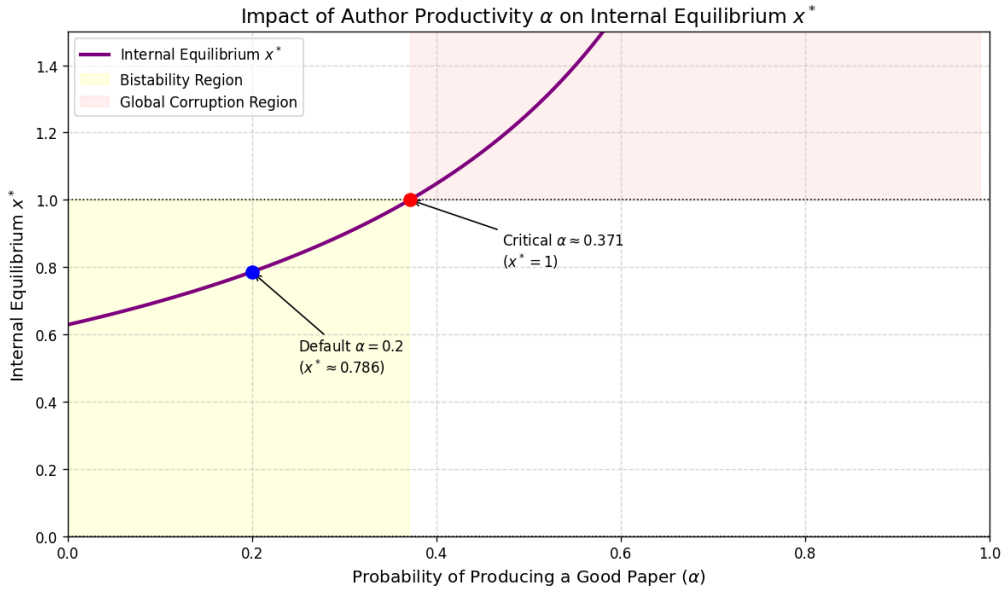

Figure 7: Dependence of the internal equilibrium x^* on author productivity α. The curve starts at a non-zero baseline of $x^* \approx 0.63$ at $\alpha = 0$. As productivity increases, the required proportion of honest authors rises until it exceeds physical bounds ($x^* > 1$) at $\alpha \approx 0.37$, triggering a transition from Bistability (Yellow) to Global Corruption (Red).

7.4 Impact of Screening Friction ϵ

The parameter ϵ in the screening function $\sigma(x) = \frac{K}{C_J V(x) + \epsilon}$ represents the operational friction or fixed overhead of the peer-review process that is independent of the submission volume. This could include administrative bottlenecks, platform latency, or the minimum time required to process a paper regardless of its quality. We analyze the sensitivity of the internal equilibrium x^* to ϵ by substituting the variable ϵ into the equilibrium condition $V(x^*) = \frac{1}{C_J}(\frac{K}{\sigma^*} - \epsilon)$.

Solving for x^*, we observe a linear relationship:

$$x^*(\epsilon) = \frac{1}{1-\alpha}\left[1 - \frac{1}{NC_J}\left(\frac{K}{\sigma^*} - \epsilon\right)\right] \tag{46}$$

Taking the derivative with respect to ϵ, we find:

$$\frac{dx^*}{d\epsilon} = \frac{1}{(1-\alpha)NC_J} > 0 \tag{47}$$

This positive derivative indicates that an increase in system friction strictly increases the threshold for stability. The dynamic response is illustrated in Figure 8.
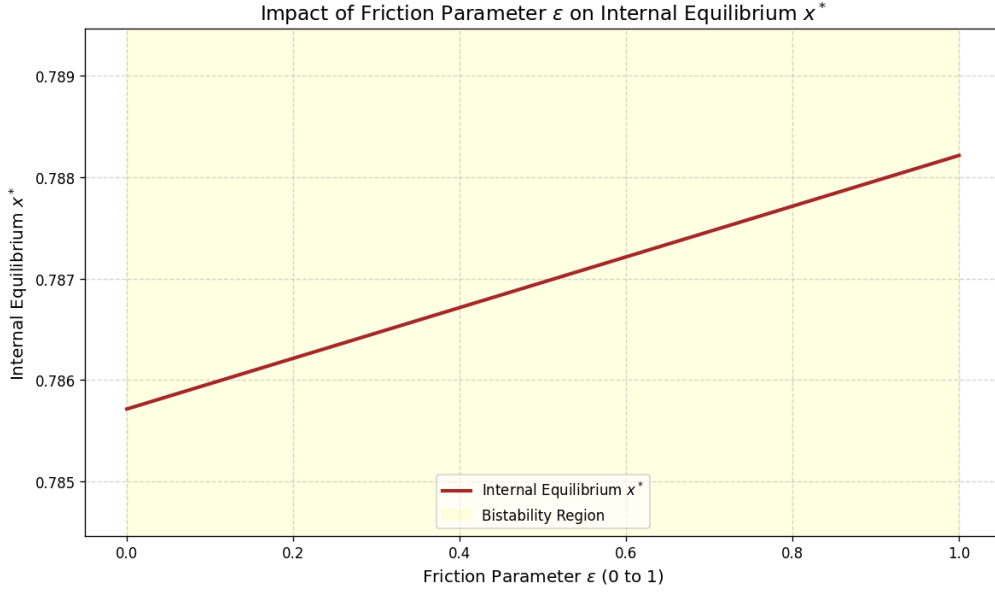

Figure 8: Linear dependence of the internal equilibrium x^* on the screening friction ϵ. An increase in friction consumes journal resources, necessitating a higher proportion of honest authors (higher x^*) to reduce volume and maintain screening effectiveness.

Physical Interpretation of Extremes

1. **Ideal Frictionless State ($\epsilon \to 0$):** In this theoretical limit, the journal suffers no fixed overhead. All capacity K is dedicated strictly to the variable cost of reviewing submissions. This scenario yields the minimum possible value for x^*, representing the most robust state of the system. Even here, however, stability is not guaranteed if N is sufficiently large, as the variable cost $C_J V(x)$ alone can saturate capacity.

2. **High Friction State ($\epsilon \to 1$):** When ϵ is large, a significant portion of the journal's "effective capacity" is consumed by overhead before any actual review takes place. To satisfy the critical screening condition $\sigma = \sigma^*$, the term $C_J V(x)$ must be reduced to compensate for the large ϵ. Physically, this means the community must drastically reduce its submission volume to accommodate the system's inefficiency. Since only honest authors reduce volume (by withholding bad papers), the system demands a much higher proportion of honest participants (higher x^*) to survive. If ϵ becomes too large, x^* may exceed 1, pushing the system into inevitable Global Corruption solely due to administrative inefficiency.

7.5 Impact of False Positive Reward λ

Finally, we investigate the role of λ, the false positive rate, which serves as a proxy for the return rate on bad papers. Unlike capacity parameters (K, ϵ), λ directly alters the payoff structure of the game. Our analysis highlights that the influence of λ cannot be evaluated in isolation; rather, it must be synthesized with the benefit B and submission cost C to determine the *Profitability Condition* of the "Always Submit" strategy.

The system's behavior is fundamentally governed by the net expected profit of a successful speculative submission, given by $\pi_{bad} = \lambda B - C$. We identify three distinct phases of stability based on this net profit, as illustrated in Figure 9:

1. **Phase I: Unconditional Honesty ($\lambda B \leq C$):** In the region where $\lambda \leq C/B$ (approximated as $\lambda \leq 0.1$ under default parameters), the expected reward from a bad paper, even if accepted, is insufficient to cover the submission cost. In this regime, the strategy "Always Submit" is strictly dominated by "Only Good" or "Never Submit" regardless of the journal's screening effectiveness. The internal equilibrium x^* is mathematically negative, implying that the system converges to Global Honesty unconditionally.

2. **Phase II: Bistability and the Rising Threshold ($0 \leq x^* \leq 1$):** Once λ crosses the profitability threshold, cheating becomes economically viable, and the system enters the Bistable regime. We observe a monotonic increase in x^* as λ grows. The mechanism driving this trend is the *Incentive-Congestion Trade-off*. As λ increases, the incentive to cheat rises. To neutralize this higher incentive ($E[U_{Bad}] = 0$), the journal must impose a strictly higher screening probability σ^*:

$$\sigma^*(\lambda) = \frac{\lambda B - C}{D + \lambda B} \tag{48}$$

Since $\frac{\partial \sigma^*}{\partial \lambda} > 0$, a higher λ demands stricter screening. Under fixed capacity K, increasing screening efficiency requires a reduction in submission volume $V(x)$. Consequently, the population must become increasingly honest (higher x^*) to sufficiently lower the volume and "afford" the required screening intensity.

3. **Phase III: Global Corruption ($x^* > 1$):** When λ exceeds a critical upper bound ($\lambda \approx 0.6$), the temptation to cheat becomes so high that the required screening probability σ^* demands a submission volume lower than physically possible (even if $x = 1$). Mathematically, x^* crosses unity. At this point, even a fully honest population generates too much volume to sustain the screening intensity needed to deter the high-reward cheating, causing the system to collapse into Global Corruption.

7.6 Impact of Acceptance Benefit B

We further examine the influence of the reward for acceptance, denoted by B. This parameter represents the career advancement, funding, or prestige associated with a publication. The relationship between B and the internal equilibrium x^* is governed by the profitability condition of the "Always Submit" strategy. Analytically, the critical screening probability σ^* required to neutralize the incentive to cheat is given by:

$$\sigma^*(B) = \frac{\lambda B - C}{D + \lambda B} \tag{49}$$

Substituting this into the equilibrium solution for x^*, we obtain the response curve illustrated in Figure 10.

For low values of B, specifically when $\lambda B \leq C$, the expected payoff of a speculative submission is negative even in the absence of screening ($\sigma = 0$). In this **Unprofitable Zone**, rational opportunistic authors voluntarily abstain from submitting bad papers because the cost of submission C outweighs the probability-weighted benefit. Mathematically, the equilibrium condition requires $\sigma^* \leq 0$, which is physically realized as unconditional stability ($x^* \to -\infty$). Under default parameters, this threshold occurs at $B \leq 12.5$.

A critical question arises regarding the system's behavior as the reward B becomes arbitrarily large. As shown in Figure **??**, x^* does not diverge to infinity but instead converges asymptotically to a value slightly greater than 1.
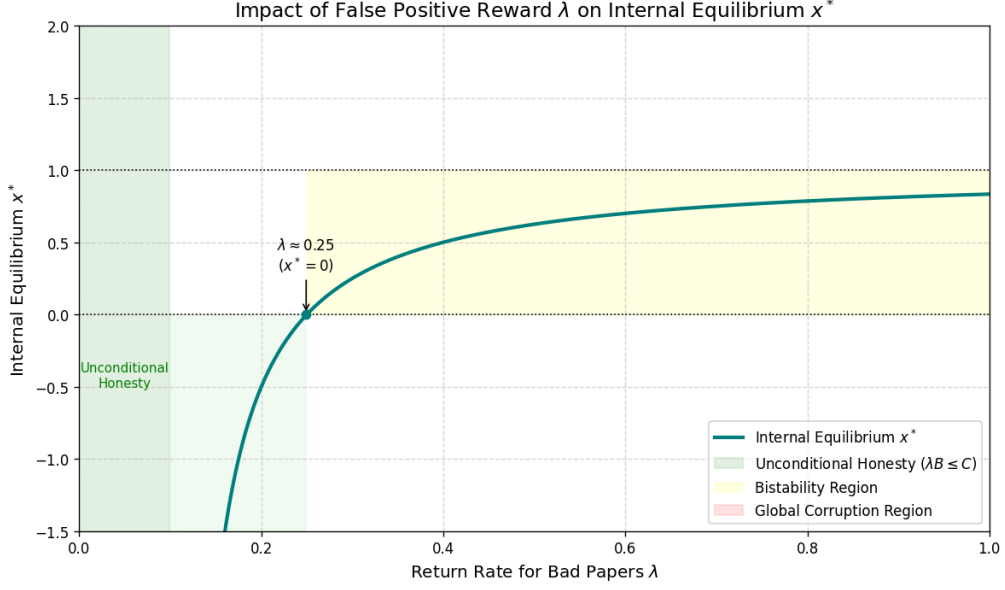
Figure 9: Impact of the false positive return rate λ on the internal equilibrium x^*. The system exhibits Unconditional Honesty (Green) when cheating is unprofitable ($\lambda B \leq C$). As λ increases, the profitable region requires an increasing proportion of honest authors to maintain stability (Yellow), eventually leading to inevitable collapse (Red) when the incentive exceeds the system's maximum screening potential.

Mathematically, as $B \to \infty$, the penalty D becomes negligible relative to the reward, and the required screening probability approaches unity:

$$\lim_{B \to \infty} \sigma^*(B) = \lim_{B \to \infty} \frac{\lambda B - C}{D + \lambda B} = 1 \tag{50}$$

This implies that to stop authors from cheating for an infinite reward, the journal must catch 100% of bad papers. The critical submission volume that the journal can support at this maximum screening intensity is fixed by the physical capacity K:

$$V_{crit}^\infty = \lim_{\sigma \to 1} \frac{1}{C_J} \left(\frac{K}{\sigma^*} - \epsilon \right) = \frac{K - \epsilon}{C_J} \tag{51}$$

Under default parameters ($K = 100, \epsilon = 0.01, C_J = 0.5$), this limit is $V_{crit}^\infty = 199.98$. However, even if the entire population is honest ($x = 1$), the minimum "honest" submission volume is $V_{min} = N\alpha = 200$. Since $V_{min} > V_{crit}^\infty$, the system is structurally incapable of providing $\sigma = 1$ screening for the whole population. Consequently, the asymptote x_{inf}^* settles slightly above 1:

$$x_{inf}^* = \frac{1 - V_{crit}^\infty/N}{1 - \alpha} \approx 1.000025 \tag{52}$$

This result signifies that while high rewards drive the system towards corruption, the failure is driven by the **physical saturation** of capacity rather than the economic incentive itself. The journal simply cannot physically review even the honest papers strictly enough to deter the infinite temptation.

7.7 Impact of Rejection Penalty D

The parameter D represents the reputational or administrative penalty imposed on an author when a submitted paper is identified as "Bad" and rejected. This serves as the primary deterrent mechanism in
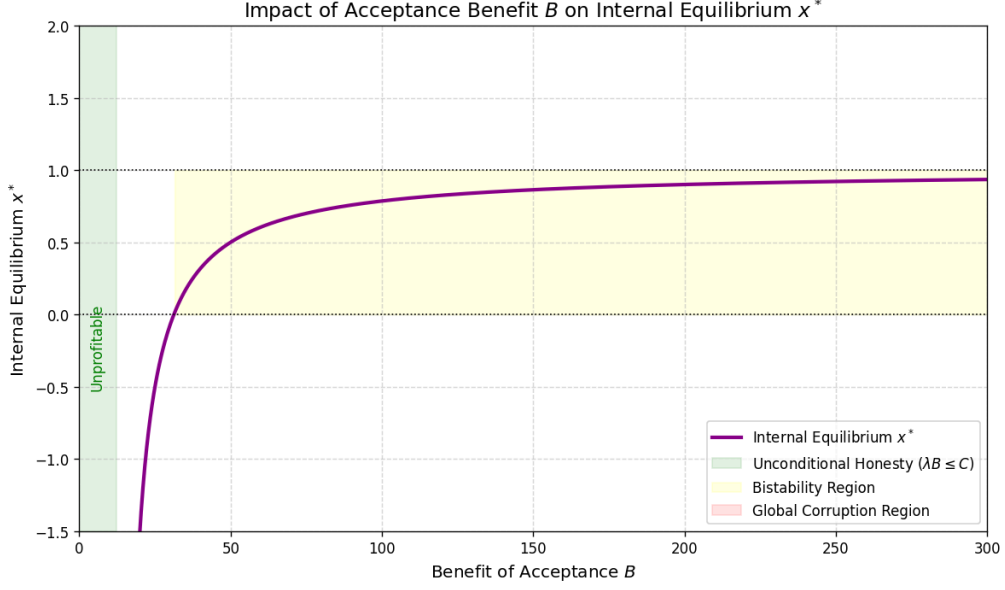
Figure 10: Impact of acceptance benefit B on the internal equilibrium x^*. The Green zone indicates the **Unprofitable Region** where the reward is insufficient to cover submission costs ($\lambda B \leq C$), leading to unconditional honesty. As B increases beyond this threshold, the system enters the Bistable regime (Yellow), requiring strict screening to offset the rising temptation.

the ecosystem. Unlike the reward B, which fuels the incentive to cheat, the penalty D suppresses it. We analyze the sensitivity of the system to this deterrent by examining the critical screening probability σ^* required to maintain equilibrium ($E[U_{Bad}] = 0$).

The required screening intensity is inversely related to the penalty magnitude:

$$\sigma^*(D) = \frac{\lambda B - C}{D + \lambda B} \tag{53}$$

As the penalty D increases, the denominator grows, reducing the necessary screening probability σ^*. Physically, this implies that a severe punishment allows the journal to be less rigorous in detection while still maintaining a sufficient deterrent effect. The relationship between the internal equilibrium x^* and D is governed by the capacity constraint equation:

$$x^*(D) = \frac{1}{1 - \alpha} \left[1 - \frac{1}{NC_J} \left(\frac{K}{\sigma^*(D)} - \epsilon \right) \right] \tag{54}$$

Since x^* is a decreasing function of $1/\sigma^*$, and σ^* decreases with D, the internal equilibrium x^* is strictly decreasing with respect to D. The resulting dynamic landscape is visualized in Figure 11.

The curve in Figure 11 highlights three distinct regimes driven by the magnitude of the penalty. For low values of D, the risk of rejection is trivial compared to the potential reward. Consequently, the journal requires a near-perfect screening rate ($\sigma \to 1$) to neutralize the payoff, which demands an impossibly low submission volume, pushing the system into Global Corruption ($x^* > 1$). As D increases, the "risk premium" of cheating rises. The system enters the Bistable regime, where a moderate level of honesty is sufficient to maintain order. Finally, when D exceeds a critical threshold, the penalty becomes so severe that even a minimal screening probability (which allows for maximum submission volume) is sufficient to render cheating unprofitable. In this regime ($x^* < 0$), the system converges to Global Honesty regardless of the initial population state, as the cost of detection outweighs any potential gain.
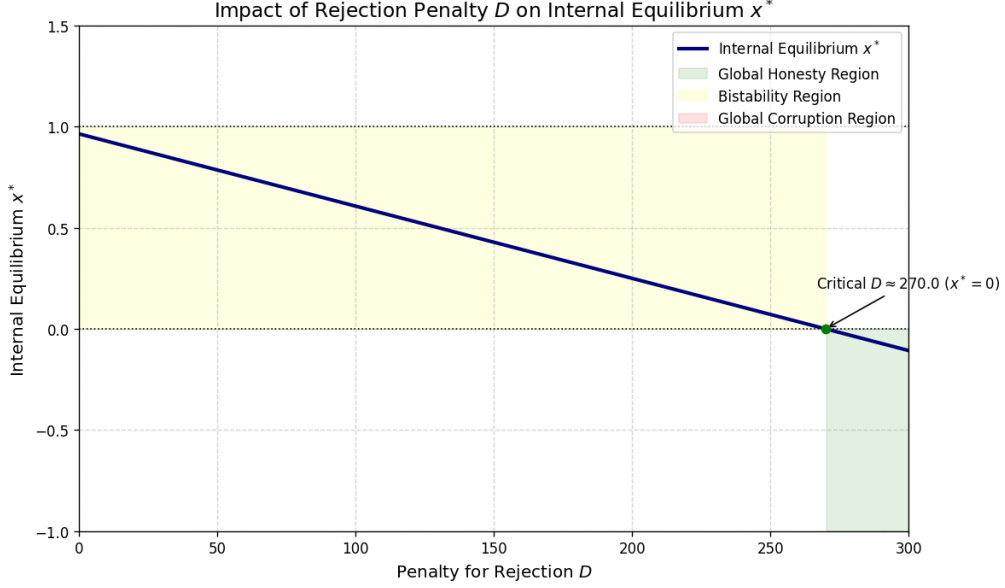
19

Figure 11: Impact of rejection penalty D on the internal equilibrium x^*. The relationship is monotonically decreasing. Low penalties (left) fail to deter cheating, requiring physically impossible screening levels ($x^* > 1$, Global Corruption). High penalties (right) create a strong deterrent even with relaxed screening, stabilizing the system at lower levels of honesty ($x^* < 0$, Global Honesty). The intermediate region (Yellow) represents the zone of Bistability.

7.8 Impact of Submission Cost C

The submission cost C represents the economic or temporal barrier to entry for submitting a manuscript, including submission fees, formatting effort, or the opportunity cost of time. Unlike the capacity parameters which constrain the journal's throughput, C acts as a direct filter on the input side of the system by altering the net profitability of the "Always Submit" strategy, given by $\pi_{bad} = \lambda B - C$.

The relationship between C and the internal equilibrium x^* is strictly monotonically decreasing. As the cost C rises, the net profit of a successful bad paper diminishes, allowing the journal to maintain effective deterrence with a lower screening probability σ^*. Mathematically, this relaxation is expressed as:

$$\sigma^*(C) = \frac{\lambda B - C}{D + \lambda B} \tag{55}$$

A lower required σ^* implies that the system can tolerate a higher volume of submissions while still maintaining equilibrium. With a higher allowable volume, the necessity for a large proportion of honest (silent) authors decreases. Thus, increasing C lowers the stability threshold x^*, making the system more robust against corruption. The bifurcation diagram visualized in Figure 12 illustrates the continuous transition across three distinct stability regimes driven by this cost structure.

At the lower bound where submission is nearly free ($C \to 0$), the profit margin for spamming is maximized. In this scenario, the journal is forced to maintain strict screening ($\sigma^* \to \frac{\lambda B}{D + \lambda B}$), which severely limits the allowable submission volume. If the total population N is large, even a fully honest population ($x = 1$) may generate excessive volume for such strict screening, resulting in $x^* > 1$ and driving the system into a Global Corruption regime. This finding suggests that "free-to-submit" policies in prestige journals are inherently unstable without massive review capacity. As the cost C increases, the burden on the screening system relaxes, and the equilibrium x^* falls into the physical range $(0, 1)$, creating a Bistability regime where the system's fate depends on initial conditions. Ultimately, a critical singularity emerges when the cost fully offsets the expected benefit of a false positive, specifically when $C \geq \lambda B$. In this Unconditional Honesty regime, the expected utility of a bad paper becomes negative
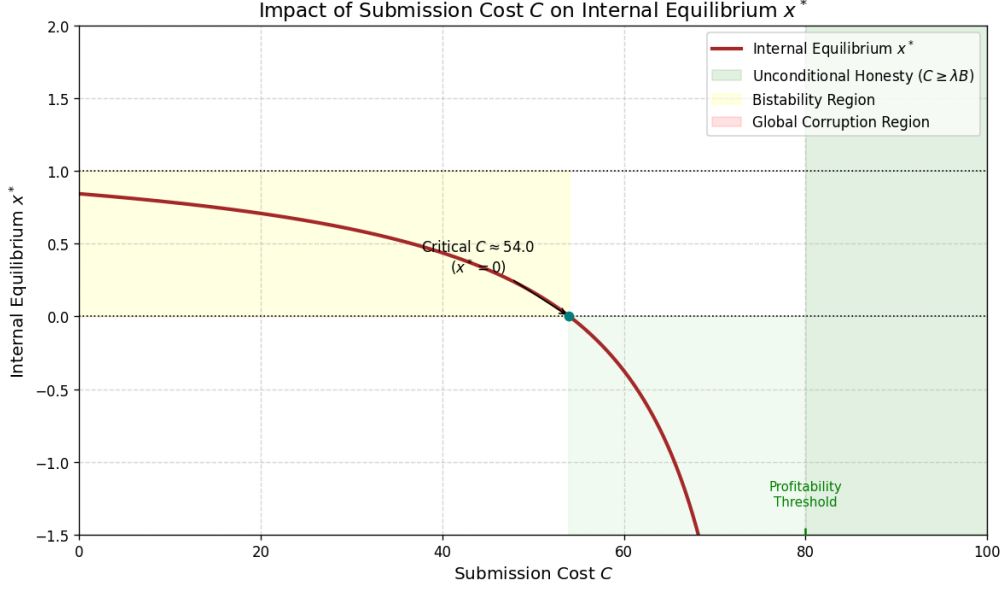
Figure 12: Impact of submission cost C on the internal equilibrium x^*. The relationship is monotonically decreasing. Extremely low costs (left) encourage spamming, pushing the system toward Global Corruption ($x^* > 1$). As costs rise, the system stabilizes. Once C exceeds the profitability threshold (λB), the system achieves Unconditional Honesty (Green) as cheating becomes economically irrational.

even in the absence of screening ($\sigma = 0$). Consequently, rational opportunistic authors voluntarily abstain from submission, leading to $x^* \to -\infty$ and ensuring global stability regardless of journal capacity.

7.9 Impact of Unit Review Cost C_J

The parameter C_J represents the unit cost or resource consumption required to review a single submission. It serves as an inverse proxy for the journal's operational efficiency; a lower C_J implies a highly efficient workflow capable of processing large volumes, while a high C_J indicates a resource-intensive process. We analyze the sensitivity of the internal equilibrium x^* to this cost by rearranging the capacity constraint equation. For the system to maintain the screening probability σ^* required for equilibrium, the submission volume must satisfy $V(x^*) = \frac{1}{C_J}(\frac{K}{\sigma^*} - \epsilon)$. Substituting the volume definition $V(x) = N[1 - (1 - \alpha)x]$, we derive the relationship between stability and review cost:

$$x^*(C_J) = \frac{1}{1 - \alpha}\left[1 - \frac{1}{NC_J}\left(\frac{K}{\sigma^*} - \epsilon\right)\right] \qquad (56)$$

The trajectory of x^* as a function of C_J is illustrated in Figure 13. The curve exhibits a strictly increasing trend, indicating that as the review process becomes more expensive (or less efficient), the system requires a strictly higher proportion of honest authors to maintain stability. In the "High Efficiency" regime (low C_J), the journal can afford to screen a vast number of submissions. Consequently, the critical threshold x^* is low (or negative), meaning the system can remain honest even with a significant number of opportunistic submitters. However, as efficiency degrades (increasing C_J), the allowable volume contracts, forcing the ecosystem to rely increasingly on author self-restraint (honesty) to prevent congestion.

A critical feature of this dynamic is the existence of a horizontal asymptote as the cost becomes arbitrarily large, which we define as the **Saturation Limit**. As $C_J \to \infty$, the term representing the
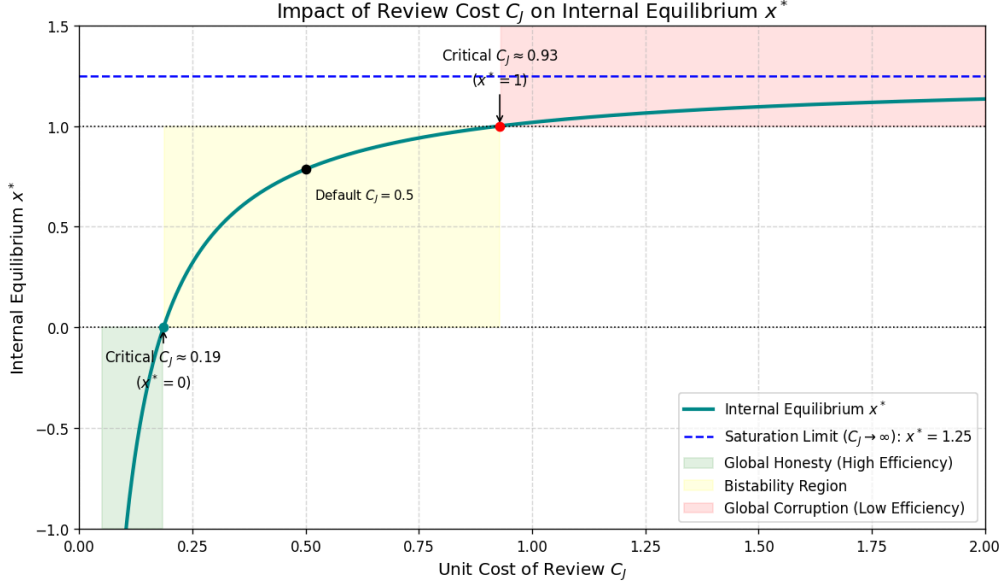
Figure 13: Impact of unit review cost C_J on the internal equilibrium x^*. The relationship is monotonically increasing. Highly efficient journals (Low C_J, Green) can sustain Global Honesty. As costs rise, the system enters Bistability (Yellow). In the limit of extreme inefficiency ($C_J \to \infty$), the curve approaches a **Saturation Limit** (Blue dashed line) determined solely by author productivity α.

journal's processing capability approaches zero ($\frac{1}{C_J} \to 0$). In this limit, the equilibrium converges to:

$$\lim_{C_J \to \infty} x^*(C_J) = \frac{1}{1-\alpha} \tag{57}$$

Under the default parameter $\alpha = 0.2$, this limit is 1.25. The physical interpretation of this saturation limit is profound: if the review process becomes infinitely expensive, the journal's effective capacity drops to zero. At this point, even if the entire population is honest ($x = 1$) and submits only good papers, the resulting volume $V_{min} = N\alpha$ is still infinitely greater than the journal's zero capacity. Therefore, once the cost C_J exceeds the critical threshold where $x^* = 1$, the system is structurally doomed to Global Corruption, not because of spam, but because it lacks the efficiency to handle even the legitimate scientific output of the community.

7.10 Summary of Findings: The Dual Constraints of Capacity and Incentive

Our comprehensive sensitivity analysis reveals that the stability of the academic peer-review ecosystem is not governed by a single factor, but rather by the delicate interplay between two fundamental forces: physical capacity constraints and economic incentive structures.

On the physical side, the journal's ability to maintain order is strictly limited by its processing throughput, determined by capacity K, review cost C_J, and friction ϵ. A counter-intuitive discovery in this domain is the "Productivity Paradox," where an increase in the production of high-quality research (α) ironically destabilizes the system by consuming scarce review resources. Similarly, the identification of a "Saturation Limit" demonstrates that structural inefficiencies (high C_J) cannot be overcome simply by eliminating cheaters; if the review process is too resource-intensive, even a fully honest population can overwhelm the system. These findings suggest that the sustainability of a journal depends as much on its operational efficiency and total submission volume (N) as it does on the ethical conduct of its authors.

On the economic side, the parameters $B, C, D,$ and λ define the "price" of stability. They determine the required screening intensity σ^* needed to deter opportunistic behavior. We observed that while

high penalties (D) and submission costs (C) serve as effective filters that relax the burden on the screening system, high rewards (B) and false positive rates (λ) act as destabilizing amplifiers. Crucially, as the incentive to cheat increases, the system demands a disproportionately higher level of honesty from the community to reduce congestion and allow for stricter screening. When these economic incentives become too strong, they push the required honest fraction beyond physical possibilities ($x^* > 1$), leading to inevitable collapse regardless of capacity.

Ultimately, the prevalence of the Bistability regime across wide parameter ranges underscores the critical role of path dependence. For a journal operating within standard constraints, the long-term outcome is rarely deterministic based on parameters alone but is heavily contingent on the initial state of the community. This implies that policy interventions must be designed not only to adjust parameters (e.g., increasing D or K) but also to actively steer the community's trust level (x) across the tipping point, as a system trapped in the basin of attraction for corruption cannot recover solely through marginal improvements in resources or penalties.