Game-Theoretic Modeling of Strategic Interactions Between Authors and Journals

Student ID: 34252096

Name: **Yuhui Fang**

Course: FIT5128 Master Thesis

Supervisor: Assoc. Prof. Julian García Gallego

Word Count (excluding references): 6436 words

Contents

Abstract

We model interactions between authors and journals as an evolutionary game, in order to study how submission incentives and review errors shape publication outcomes. Following (Zollmann et al., 2024), analytical Nash equilibria and replicator dynamics reveal that a higher false rejection (ε) discourages quality effort, while a higher false acceptance (λ) promotes opportunistic submission. Multi-author extensions show stable but less efficient equilibria as review costs rise. Empirical calibration with ICLR–OpenAlex data supports these theoretical predictions.

1. Introduction

Scientific publication is both a channel for disseminating knowledge and a competitive system in which authors and journals pursue distinct strategic objectives. Authors seek a reputation, citations, and career advancement, while journals aim to preserve quality and prestige. However, these incentives often misalign. When submission costs are low, authors can overproduce papers or submit low-quality work, leading to what has been described as a "submission explosion" or "overproduction of papers" (Horbach and Halffman, 2018; Fanelli, 2010; Brembs, 2018). This behaviour burdens peer review and can degrade overall research quality.

To study these interactions systematically, the publication process can be viewed as a strategic game between authors and journals. *Game theory*, introduced by (von Neumann and Morgenstern, 1944) and popularised by (Osborne, 2003), provides a mathematical framework to analyse how rational agents make interdependent decisions. Here, authors decide how much effort to invest in producing high-quality work or whether to submit indiscriminately, while journals determine their review selectivity. The interplay of these decisions shapes the equilibrium of the scientific publication ecosystem (Zollmann et al., 2024).

A key concept is the *Nash Equilibrium* (Nash, 1951), a stable state in which no player can improve their performance by unilaterally changing strategy. In small-scale 2×2 games, equilibria can be identified through direct enumeration, but as the number of players or strategies increases, exhaustive computation becomes infeasible. Researchers therefore use algorithmic or dynamic approaches—such as support enumeration, reinforcement learning, or evolutionary simulations—to approximate equilibrium outcomes.

When analytical solutions are intractable, *evolutionary game theory* (Cressman, 2003; Hofbauer and Sigmund, 1998) models how strategies evolve over time. The replicator equation,

$$\dot{x}_i = x_i(\pi_i - \bar{\pi}),$$

describes how the proportion of strategy i changes depending on its relative payoff. Strategies that perform above average spread, while unprofitable ones decline. This framework captures the adaptive nature of scientific publishing, where authors and journals iteratively adjust behaviours based on past success or failure.

Previous studies have examined incentives in science through game-theoretic reasoning. (Strevens, 2012) modelled scientific norms using static cooperation games; (Heesen, 2018) showed how reward systems can lead to reproducibility crises through equilibrium selection; and (Zollmann et al., 2024) formalised the author–journal interaction as a finite game, linking peer-review precision to the equilibrium balance between quality and quantity. Although these works provide valuable conceptual insights, they remain static and analytical. This thesis extends their framework dynamically and empirically, connecting model parameters with real-world data to improve the explanatory power and scalability of such models.

Specifically, the study introduces an *evolutionary author–journal model* that integrates both theoretical and empirical components. Two key parameters represent the reliability of peer review:

- $\boldsymbol{\varepsilon}$ — the false rejection rate of good papers (Type I error);
- $\boldsymbol{\lambda}$ — the false acceptance rate of bad papers (Type II error).

These jointly determine the overall accuracy of peer review and influence the strategic responses of both parties. Changing (ε, λ), the model explores how error structures shape cooperative versus opportunistic equilibria on different population scales. Empirical estimates of $(\hat{\alpha}, \hat{\varepsilon}, \hat{\lambda})$ are derived from the dataset of the International Conference on Learning Representations (ICLR, 2019–2022) using a Gaussian Mixture Model (GMM), which provides a data-driven calibration of the theoretical framework.

The study addresses three main research questions: *(1)* How do peer-review error rates $(\varepsilon,\ \lambda)$ influence the strategic behaviour of authors and journals? *(2)* How does scaling the number of authors or journals affect equilibrium stability, selectivity, and efficiency? *(3)* What institutional mechanisms—such as penalties for bad publications or rewards for quality—can promote fairness and cooperation?

By integrating empirical data with evolutionary dynamics, this research advances game-theoretic analysis from static conjecture to quantitative validation. It demonstrates how peer-review accuracy and publication incentives jointly shape systemic behaviour in scientific ecosystems. Beyond theoretical insight, the findings have policy relevance for journal management, research evaluation, and funding design—ultimately contributing to a more fair and efficient academic publishing system.

2. Background and Related Work

This section reviews three methodological foundations—algorithmic equilibrium computation, the evolutionary game dynamics, and potential games, and connects them to previous research in scientific publishing. The discussion identifies the conceptual and computational gaps that motivate a dynamic, data driven author journal model.

2.1. Computing Equilibria in Finite Games

A *normal-form game* represents a strategic interaction among a finite set of players, each choosing a strategy simultaneously, with payoffs specified in a matrix form. Formally, it can be defined as

$$G = \langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle,$$

where N is the set of players, S_i the set of strategies of the player i, and $u_i : \prod_j S_j \to \mathbb{R}$ the payoff function. Such matrix games are the foundation of most algorithmic approaches for equilibrium computation.

For small finite games, Nash equilibria can be computed exactly using either support enumeration or pivoting algorithms. The *Lemke–Howson algorithm* (Lemke and Howson, 1964; Shapley, 1974) is one of the classic pivoting procedures to find a mixed Nash equilibrium in a two-player bimatrix game. It reformulates the equilibrium conditions as a *linear complementarity problem* and then follows a path of complementary pivots from an artificial equilibrium to a genuine one. Each pivot corresponds to exchanging an active constraint in one player's best-response polytope, and the process continues until a valid equilibrium is reached. This approach guarantees convergence for non-degenerate bimatrix games and remains one of the most widely used theoretical algorithms for equilibrium computation.

In practice, open-source toolkits have made these classical methods accessible. *NashPy* (Knight, 17) provides implementations of support enumeration and Lemke–Howson for two-player games based on Python, allowing researchers to reproduce classical analyses programmatically. Meanwhile, *Gambit* (Savani and Turocy, 2025) is a general-purpose computational package capable of handling games with more players and strategy profiles, implementing a wide range of algorithms including enumeration, homotopy continuation, and global Newton methods.

In this project, the support enumeration method was implemented directly in Python for reproducibility and flexibility. Although Gambit was not used directly, it serves as a well-established benchmark and validation reference for equilibrium computation. These algorithmic foundations provide the computational basis for analysing the finite author–journal games in this study, such as the one-author–one-journal model, before extending to larger evolutionary systems.

2.2. Evolutionary Game Theory and Replicator Dynamics

When it is too difficult to find exact mathematical solutions for the best strategies of players, known as *closed-form equilibria*, evolutionary game theory provides a way to model how strategies change and adapt over time within populations. The central concept is the replicator dynamic, which captures how strategies that yield higher-than-average payoffs increase in frequency over time. This approach has been widely used to analyse co-evolutionary behaviour in biological, social and economic systems (Taylor and Jonker, 1978; Weibull, 1995).

In the context of scientific publishing, replicator dynamics describes how authors and journals adjust their strategies in response to observed outcomes of submission and review. Authors may alternate between selective (OnlyGood) and opportunistic (AlwaysSubmit) behaviours, while journals modify their selectivity based on cumulative review costs and observed quality levels. The dynamic perspective therefore generalises the notion of Nash equilibrium to continuous adaptation processes. The results of the empirical simulation in later sections show how the stability of cooperative equilibria depends on error rates (ε, λ) and population size, echoing findings in multi-population evolutionary models (Sandholm, 2010).

2.3. Game-Theoretic Analyses of Scientific Publishing

Game theory has been applied to the study of academic publishing as a mechanism of strategic interaction among scientists and journals. Several works have explored how peer review, reward systems, and institutional pressures influence research quality and cooperation. For example, (Smaldino and McElreath, 2016) used evolutionary modelling to show how incentive misalignment can lead to the proliferation of low-quality research practices. (Heesen and Bright, 2021) argued that competition among scientists can undermine epistemic progress, framing the publication system as a coordination problem.

More recently, (Zollmann et al., 2024) formalised the author–journal interaction as a finite game, demonstrating that peer-review accuracy determines the equilibrium balance between selectivity and productivity. However, most of these models remain static or assume symmetric populations with fixed parameters. This thesis extends such frameworks by incorporating dynamic adjustment mechanisms through replicator dynamics, empirical parameter estimation, and scaling to multiple journals. The approach bridges theoretical reasoning and computational experimentation, revealing how small variations in review accuracy can shape systemic equilibria.

2.4. Synthesis and Research Gap

Existing research provides valuable theoretical insights into equilibrium concepts, dynamic adaptation, and peer-review incentives, but most studies remain limited to small-scale static systems. They rarely incorporate empirical data or consider heterogeneous populations of authors and journals. Consequently, the interplay between review accuracy, incentive design, and systemic efficiency remains poorly understood.

This paper addresses these limitations by proposing a scalable and parameterized author–journal model that unifies analytical equilibrium computation, evolutionary dynamics, and empirical calibration. The model captures how screening errors (ε, λ), reputation weights (B, D), and review costs jointly shape cooperative or opportunistic equilibria. Exact solvers validate small systems, replicator dynamics reveal population-level stability, and empirical parameter estimation ensures our model is applicable in the real world. Together, these components establish a quantitative, data-informed framework for analysing the incentives that govern scientific publication behaviour.

2.5. Empirical Parameter Estimation and Integration

A key contribution of this study is the integration of empirical data into theoretical modelling. To bridge the gap between conceptual game-theoretic models and real-world publication data, the analysis uses the open ICLR submissions dataset (2019–2022) publicly released available at https://github.com/berenslab/iclr-dataset. The data set is maintained under the MIT licence (Gonzalez and Kobak, 2024) and has undergone continuous updates, including the addition of blind submissions for ICLR 2025 and 2026, new topic classifications (e.g., *AI safety*, *alignment*, *LLMs*, and *autonomous driving*), and improved keyword parsing. These updates ensure a consistent and comprehensive record of ICLR submissions, decisions, and metadata over multiple years.

Because the data set does not include the quantitative evaluation scores of the reviewers, the citation counts were used as a proxy of the quality of the article. Specifically, OpenAlex citation data were merged with OpenReview submission metadata, allowing each article to be associated with its two-year impact on citations. Since citation counts are highly skewed and heavy-tailed, we first apply a logarithmic normalization to stabilize variance and then use a *Gaussian Mixture Model* (GMM) to capture the latent structure of the citation distribution. The GMM assumes that observed citation values arise from a mixture of two overlapping normal components representing low- and high-impact papers. This approach avoids the need for an arbitrary threshold (e.g., top-k% of citations) and instead identifies two natural clusters directly from the data, where the higher-mean component corresponds to the "good" group. In intuitive terms, the model automatically detects how papers cluster in terms of citation impact, providing a probabilistic classification that reflects the underlying heterogeneity of research quality.

Then a Gaussian Mixture Model (GMM) was applied to the logarithmic normalised citation distribution to separate the articles into "good" and "bad" clusters, where the higher-mean component represents higher-impact work. This classification enables empirical estimation of three key parameters in the author–journal model: the share of good submissions ($\hat{\alpha}$), the false-rejection rate of good papers ($\hat{\varepsilon}$) and the false-acceptance rate of bad papers ($\hat{\lambda}$).

Integrating these empirically derived parameters grounds the theoretical model in observable data and allows subsequent simulations to capture realistic peer-review error levels. This empirical link distinguishes the present study from prior purely theoretical models and motivates the model calibration and validation presented in later chapters.

3. Methodology

3.1. Model Setup

The game involves two types of agents: a population of N_A authors and one journal.

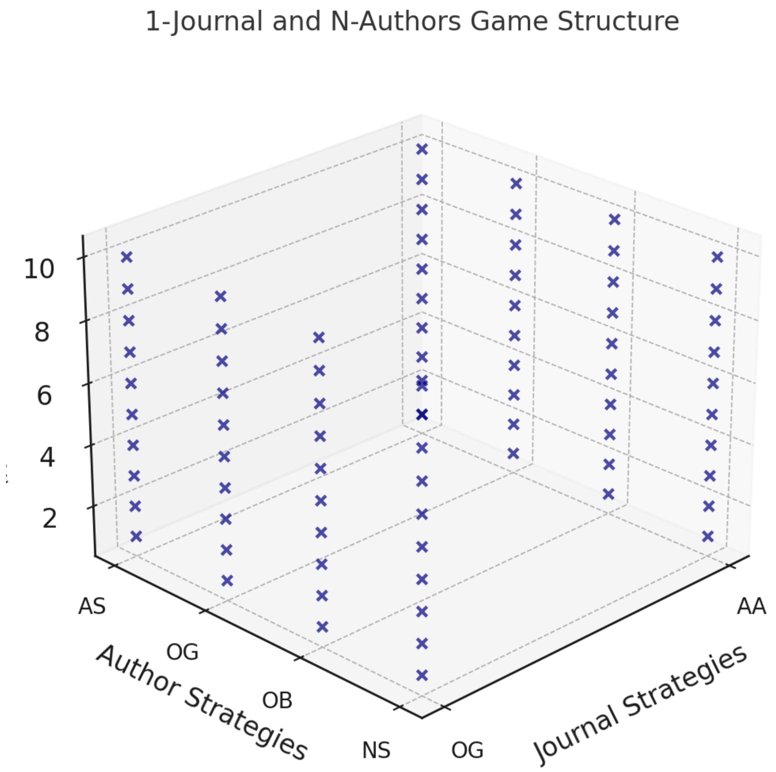


Figure 1: Schematic representation of the one-journal and N-authors game. The horizontal axis shows the two possible journal strategies (AA: AlwaysAccept, OG: OnlyGood), while the vertical axis lists the four author strategies (AS: AlwaysSubmit, OG: OnlyGood, OB: OnlyBad, NS: NoSubmit). Each point along the z-axis corresponds to an individual author $(1, \ldots, N)$, illustrating population-based interaction between authors and the single journal.

3.1.1 Authors

The authors decide what to submit, while the journal determines its selectivity. This one-journal multi-author setting provides a baseline before later extension to multiple journals.

Each author selects one of four pure strategies:

$$S_A = \{\mathrm{AS},\ \mathrm{OG},\ \mathrm{OB},\ \mathrm{NS}\},$$

representing the extremes of submission behaviour under binary research quality. Let α denote the probability of producing a good paper. The expected composition of submissions is

$$(p_g, p_b) = \begin{cases} (\alpha, 1-\alpha), & \text{AS}, \\ (\alpha, 0), & \text{OG}, \\ (0, 1-\alpha), & \text{OB}, \\ (0, 0), & \text{NS}. \end{cases}$$

Given the mix of journals $z = (z_{\text{AA}}, z_{\text{OG}})$, the acceptance probabilities are $q_g = z_{\text{AA}} + z_{\text{OG}}(1-\varepsilon)$ and $q_b = z_{\text{AA}} + z_{\text{OG}}\lambda$, where ε and λ are false rejection and false acceptance rates.

The expected utilities are

$$\begin{aligned} u_A(\text{AS}) &= r[\alpha q_g + (1-\alpha)q_b] - c, \\ u_A(\text{OG}) &= r(\alpha q_g) - c\alpha, \\ u_A(\text{OB}) &= r((1-\alpha)q_b) - c(1-\alpha), \\ u_A(\text{NS}) &= 0. \end{aligned}$$

If $rq_g > c$, good-paper submission is profitable (OG); if $rq_b > c$, even bad papers are worth submitting (OB); when both hold, AS dominates; otherwise, NS is optimal. Hence, author behaviour depends on review precision (ε, λ), cost c, reward r, and paper quality α: low precision or high cost reduce selectivity, while small (ε, λ) reward quality-driven submission.

3.1.2 Journals

The journal starts with four acceptance rules:

$$S_J^{\text{init}} = \{\text{AA},\ \text{OG},\ \text{OB},\ \text{RJ}\}.$$

Two are dominated and removed. If $\alpha > 0$, accepting only bad papers (OB) does not yield benefit B, but incurs a penalty D, so it is dominated by AA. Rejecting all (RJ) cannot outperform OG, as both have similar costs, but only OG gains from good papers. Thus, two rationalisable policies remain:

$$S_J = \{\text{AA},\ \text{OG}\}.$$

Under AA, all submissions are accepted $(q_g, q_b) = (1, 1)$; under OG, only good papers are usually accepted $(q_g, q_b) = (1-\varepsilon, \lambda)$, allowing errors. These reflect two realistic editorial extremes:

- AA: permissive, quantity-orientated, maximising throughput at the expense of quality;
- OG: selective, quality-focused, trading acceptance rate for accuracy.

Their payoff difference is

$$u_J(\mathrm{OG}) - u_J(\mathrm{AA}) = \frac{-BP_{\mathrm{good}}\varepsilon + DP_{\mathrm{bad}}(1-\lambda)}{S+1}.$$

If $DP_{\mathrm{bad}}(1-\lambda) > BP_{\mathrm{good}}\varepsilon$, OG produces a higher payoff; otherwise AA is preferred. Hence, AA and OG form the strategic frontier of the journal that balances productivity and quality, and are retained for equilibrium analysis.

Each journal j is characterised by a review rule (All or OnlyGood) and a *bias type* that determines the precision of the peer-review. The error parameters $(\varepsilon_j, \lambda_j)$ satisfy $0 \leq \varepsilon_j + \lambda_j \leq 1$, reflecting that a reviewer cannot simultaneously have extreme false-rejection and false-acceptance rates. ε_j and λ_j represent a trade-off between strictness (rejecting more) and leniency (accepting more): higher ε_j implies conservative review, higher λ_j implies permissive review. This follows (Zollmann et al., 2024), which bounds the total misclassification to preserve the consistency of the decision.

Four representative biases are:

$$(\varepsilon_j, \lambda_j) = \begin{cases} (0,0), & \textbf{Perfect}, \\ (0.5, 0.5), & \textbf{Balanced}, \\ (0,1), & \textbf{GoodBias}, \\ (1,0), & \textbf{BadBias}. \end{cases}$$

Here, ε_j is the false rejection rate and λ_j is the false acceptance rate. Bias affects both both acceptance probabilities and the quality term in the utility of the journal, directly linking selectivity to accuracy.

3.1.3 Payoffs

The rewards arise from the expected outcomes of submission and review under two sources of uncertainty: the quality' of the stochastic articles of the authors and the imperfect selection of the journal. Peer review is modelled as a binary classifier with two misclassification rates

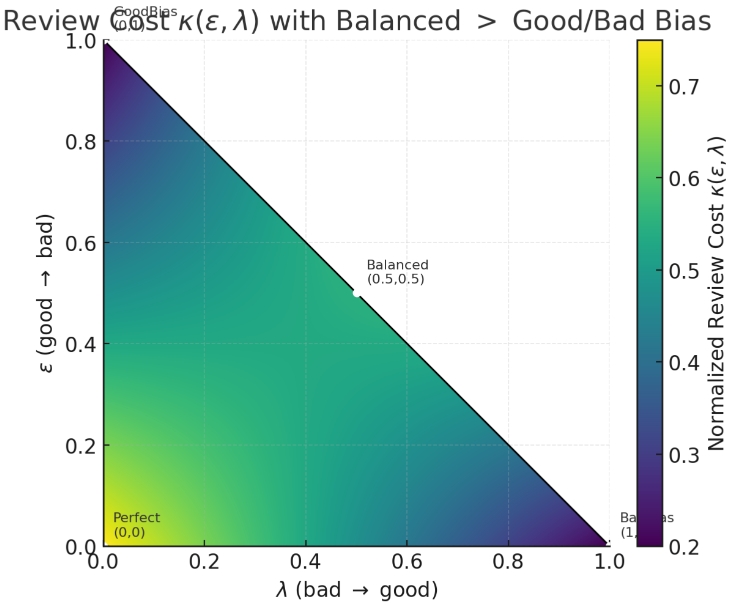


Figure 2: Normalized review cost $\kappa(\varepsilon, \lambda)$ with the intended ordering Perfect > Balanced > GoodBias = BadBias. The surface uses $\kappa = k_{\min} + a\,(\varepsilon\lambda/0.25) + b\,(1 - (\varepsilon + \lambda))$ to reflect higher effort for accuracy and intermediate effort for balanced screening.

representing review bias: ε (false rejection, Type I) is the probability a good article is judged bad, and λ (false acceptance, Type II) the probability a bad article is judged good. Under the selective policy OnlyGood, a good paper is accepted with probability $1-\varepsilon$ and a bad one with λ, while under AlwaysAccept both are accepted with probability 1. Given the journal's mixed strategy $z = (z_{\text{AA}}, z_{\text{OG}})$, the effective acceptance probabilities for authors are $q_g = z_{\text{AA}} + z_{\text{OG}}(1-\varepsilon)$ for good papers and $q_b = z_{\text{AA}} + z_{\text{OG}}\lambda$ for bad papers. Each author, with probability α of producing a good paper, follows one of the four strategies in $S_A = \{\text{AS}, \text{OG}, \text{OB}, \text{NS}\}$, defining the expected proportions (p_g, p_b) of good and bad submissions. The author's expected payoff is

$$u_A = r(p_g q_g + p_b q_b) - c(p_g + p_b),$$

where r is the reward from acceptance and c the submission cost.

Across N_A authors, the total submission mass is $S = P_{\text{good}} + P_{\text{bad}}$, with accepted masses $A_{\text{good}} = P_{\text{good}} q_g$ and $A_{\text{bad}} = P_{\text{bad}} q_b$. The journal gains B from each accepted good paper and incurs a penalty D for each bad one. A naïve count-based objective is $u_J^{\text{count}} = BA_{\text{good}} - DA_{\text{bad}} - \mathcal{C}(S)$, where $\mathcal{C}(S)$ grows with S. If $\mathcal{C}(S)$ scales linearly with N_A, the journal's reward declines excessively with population size, driving a degenerate "accept-all" equilibrium. To prevent this unrealistic scaling, the journal utility is normalised as a bounded quality score:

$$u_J = \frac{B\,A_{\text{good}} - D\,A_{\text{bad}}}{S+1} - C_{\text{fix}}\left(\frac{S}{(N_A/N_J)+1}\right)^{\eta}.$$

The first term gives a normalised quality balance ensuring bounded payoffs and avoiding $S = 0$ singularities. The second term models review cost increasing with workload, scaled by the effective load per journal $S/((N_A/N_J) + 1)$, where $\eta \geq 1$ controls convexity and C_{fix} is a fixed coefficient. Although (ε, λ) describe review bias, their reputational impact is already encoded in B and D, so the cost function is kept bias-neutral. This formulation (1) avoids collapse to trivial "accept-all" outcomes for large N_A, and (2) preserves sensitivity to review bias (ε, λ) and the quality–quantity trade-off (B, D). Overall, the journal payoff captures three mechanisms—quality weighting, misclassification, and load-dependent cost—while remaining scale-invariant and interpretable. As shown in Figure 3, the count-based formulation makes the journal favour the cheapest bias and the most permissive policy, yielding no quality improvement.

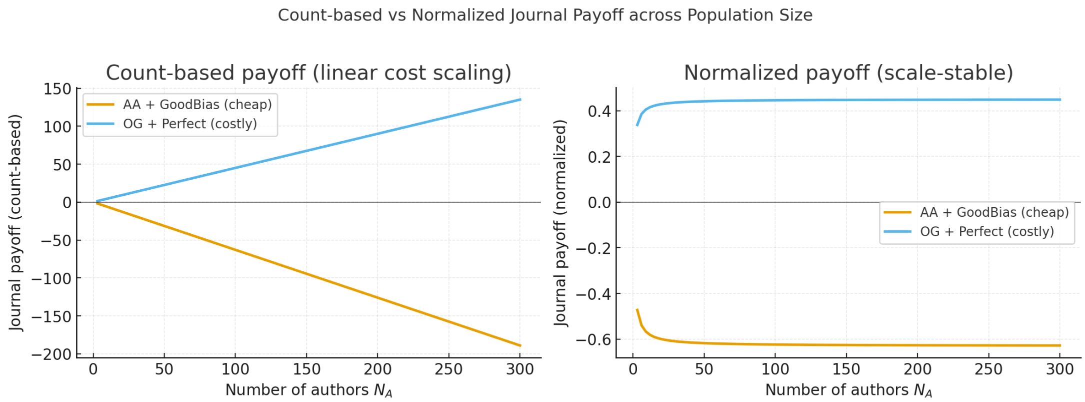


Figure 3: Count-based vs normalized journal payoff across population size N_A (independent y-axes). **Left:** Count-based objective with linearly growing review cost causes payoff to deteriorate as N_A increases, favouring cheap permissive policies. **Right:** Normalized payoff remains bounded and scale-stable, so selectivity (OnlyGood) can remain competitive even for large N_A.

3.1.4 Parameter settings and baseline values

Table 1 summarises the parameters used in simulations. Unless noted, the model assumes moderate research quality ($\alpha = 0.6$) and a single journal receiving submissions from $N_A = 3$–50 authors.

Symbol	Description	Default / Range
α	Probability an author produces a good paper	0.6 (varied 0–1)
ε	False rejection rate (good $\to$ bad)	0.1
λ	False acceptance rate (bad $\to$ good)	0.1
r	Author reward for acceptance	1.0
c	Author submission cost	0.1
B	Journal benefit per accepted good paper	1.0
D	Journal penalty per accepted bad paper	3.0
C_{fix}	Fixed editorial cost coefficient	0.1
η	Cost-scaling exponent	1.0
N_A	Number of authors	variable (3–50)

Table 1: Model parameters and baseline settings.

3.2. Replicator Dynamics

We simulate coupled multi-population replicator dynamics in discrete time to analyse adaptation in the author–journal ecosystem. Since payoffs and strategy sets are defined in the Model Setup, this subsection focusses on the numerical scheme, convergence, and robustness.

Let $x \in \Delta^4$ and $z \in \Delta^2$ denote the mixed strategies of the four author strategies (AS, OG, OB, NS) and the two journal strategies (AA, OG). The continuous-time dynamics follows the standard replicator form

$$\dot{x}_k = x_k\big(u_{A,k}(x,z) - \bar{u}_A(x,z)\big), \qquad \dot{z}_\ell = z_\ell\big(u_{J,\ell}(x,z) - \bar{u}_J(x,z)\big),$$

where $\bar{u}$ is the payoff in the population-average. A rest point satisfies $u_{A,k} = \bar{u}_A$ on the support of x and $u_{J,\ell} = \bar{u}_J$ on the support of z.

We integrate the replicator equations in discrete time via the *Euler method*: each step adds the current rate of change, multiplied by a small step size. We use $\Delta t = 0.05$ to keep updates smooth and avoid large jumps. After each update, x and z are projected back to their simplices (non-negative, sum to one) to correct numerical drift. The Euler update plus projection keeps the probabilities feasible as the system evolves toward equilibrium.

Similar discrete formulations are common in evolutionary and learning dynamics: (Bloembergen et al., 2015) use step-based updates in multi-agent learning, (Shah, 2010) unify several evolutionary rules via a growth-transform framework, and (Mertikopoulos et al., 2018) show that discrete replicator updates can yield both convergence and cycles in adversarial learning.

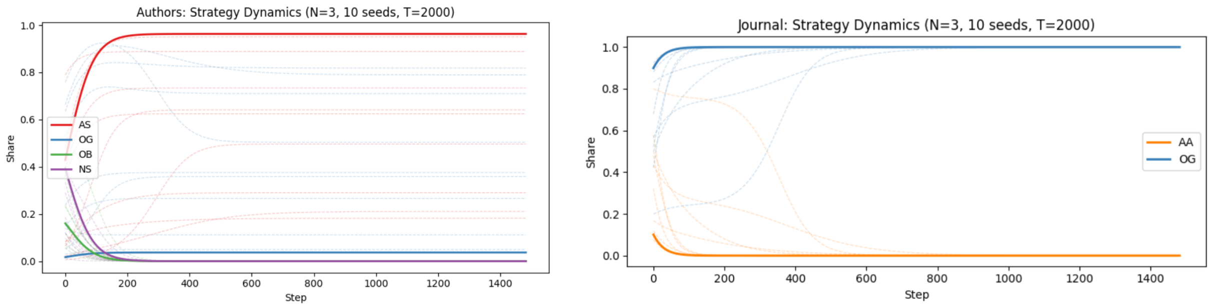
The runs start from interior Dirichlet drawings $x^0 \sim \text{Dirichlet}(\mathbf{1}_4)$ and $z^0 \sim \text{Dirichlet}(\mathbf{1}_2)$, with a fixed base seed for reproducibility. Each trajectory runs for up to $T = 2000$ steps (sufficient for all $N_A \in \{3, \ldots, 50\}$ in our settings).

We assume convergence when

$$\|x^{t+1} - x^t\|_1 < 10^{-6} \quad \text{and} \quad \|z^{t+1} - z^t\|_1 < 10^{-6}$$

for 50 consecutive iterations. If early stopping occurs before T, the last state pads the recorded series for consistent plot lengths (Mertikopoulos et al., 2018). If no fixed point emerges, we report the time averages after a burn-in ($T_0 = 200$).

For each parameter configuration, we run 10 independent Dirichlet initialisations. Figures plot nine runs as faint dashed lines (variability) and highlight one solid run, making convergence and possible multistability visible (Mertikopoulos and Sandholm, 2016).



(a) Authors: strategy dynamics

(b) Journal: strategy dynamics

Figure 4: Replicator dynamics under $N_A = 3$, $\Delta t = 0.05$, $T = 2000$, and patience $= 50$. Ten random initialisations (nine dashed, one solid) illustrate robustness. Parameters: $\alpha = 0.6$, $\varepsilon = 0.1$, $\lambda = 0.1$, $r = 1$, $c = 0.1$, $B = 1$, $D = 3$, $C_{\text{fix}} = 0.1$, $\eta = 1$. Authors converge mostly to AS, while the journal stabilises at OG.

3.3. Empirical Parameter Estimation

To calibrate the theoretical author–journal model with real-world data, three empirical parameters are estimated from ICLR submissions (2019–2022) and their citation outcomes:

$$\hat{\alpha},\ \hat{\varepsilon},\ \hat{\lambda},$$

represent, respectively, the share of high-quality submissions, the false rejection rate of good papers, and the false acceptance rate of bad papers. All source data are obtained from the open ICLR dataset maintained by Dmitry Kobak on GitHub (`https://github.com/berenslab/ICLR` dataset) and the OpenAlex Scholar database.

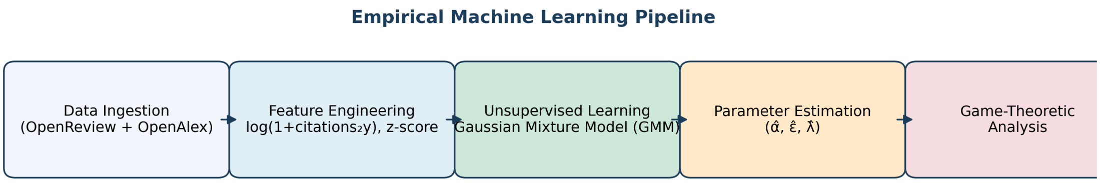


Figure 5: Pipeline for empirical parameter estimation linking ICLR submissions with OpenAlex citation data.

3.3.1 Data pipeline and rationale

The pipeline includes five stages for reproducibility and interpretability.

1. **Data aggregation and filtering** Four yearly ICLR datasets (2019–2022) were merged, keeping the latest record for each paper. Each entry contains title, decision, and year; only records with explicit decisions were retained.

2. **OpenAlex matching and citation retrieval** To obtain objective impact measures, each title was queried via the *OpenAlex* API—an open index of scholarly works (Priem et al., 2022). Each request included a required `mailto` identifier. Candidate matches within a ± 1 year window were ranked by textual similarity using *RapidFuzz* (Bachmann, 2021). For each paper, the best match was selected and its two-year citation count (c_{2y}) was stored locally to avoid redundant API calls.

3. **Citation normalisation and feature transformation** Citation counts are highly right-skewed: most papers receive few citations, while some accumulate very high counts. A logarithmic transform $\log(1 + c_{2y})$ compresses outliers, followed by year-wise standardisation:
$$z = \frac{\log(1 + c_{2y}) - \mu_{\text{year}}}{\sigma_{\text{year}}}.$$
This equalises scale over years and mitigates growth effects. Figure 6 shows the logarithmic normalised citation distribution and the fitted Gaussian components.

4. **Unsupervised quality classification** A two-component Gaussian Mixture Model (GMM) was fitted on z. The higher-mean component represents the "good-paper" cluster; a paper is labelled *good* if its posterior probability exceeds 0.5, a Bayes-optimal threshold under symmetric misclassification cost.

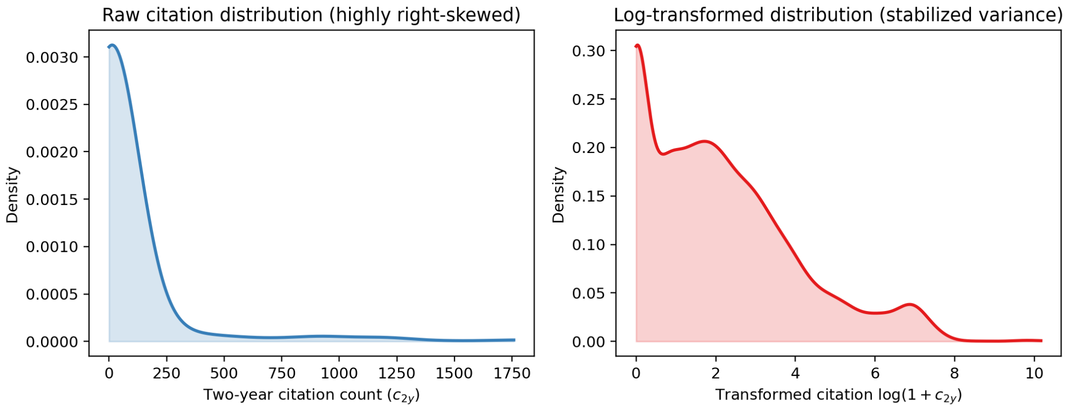


Figure 6: Citation distributions before and after transformation. **Left:** Raw two-year citation counts (c_{2y}) show strong right skew. **Right:** After $\log(1 + c_{2y})$, the distribution becomes smoother and symmetric, suitable for Gaussian-based modelling.

5. **Parameter derivation and robustness** With binary labels $\widehat{\text{good}}$ and editorial decisions decision_binary, three conditional probabilities are defined:

$$\hat{\alpha} = \Pr(\widehat{\text{good}} = 1), \quad \hat{\varepsilon} = \Pr(\text{reject} \mid \widehat{\text{good}} = 1), \quad \hat{\lambda} = \Pr(\text{accept} \mid \widehat{\text{good}} = 0).$$

   The procedure was repeated for each year to verify stability. Tests with different types of covariance, seeds, and percentile labelling yielded consistent results $\hat{\alpha}, \hat{\varepsilon}, \hat{\lambda}$.

These empirical parameters link the observed review accuracy to the theoretical payoff structure. In this integrated framework, the prior quality of authors α is replaced by its empirical estimate $\hat{\alpha}$, while the journal's acceptance probabilities under the OnlyGood rule are parameterized as $q_g = 1 - \hat{\varepsilon}$ and $q_b = \hat{\lambda}$, and under AlwaysAccept as $q_g = q_b = 1$. *(1)*Here, $\hat{\varepsilon}$ and $\hat{\lambda}$ represent the false-rejection (Type I) and false-acceptance (Type II) rates, respectively. *(2)*A high $\hat{\varepsilon}$ indicates a conservative review process that rejects many good papers, whereas a high $\hat{\lambda}$ reflects a permissive or noisy evaluation system that accepts more bad papers. *(3)*These empirically grounded quantities anchor the replicator simulations in realistic peer-review error rates, allowing theoretical equilibria to be interpreted through observable patterns of review accuracy and publication quality. Yearly variations of $(\hat{\alpha}, \hat{\varepsilon}, \hat{\lambda})$ are later compared with model predictions in the *Results and Analysis* section to maintain methodological neutrality here.

4. Results and Analysis

4.1. 1J–1A: Base Game Enumeration

To build intuition before introducing dynamic and large-scale extensions, the base configuration considers a single author and a single journal. This environment allows for the complete enumeration of all pure-strategy benefits and reveals how peer-review error rates (ε, λ) shape strategic incentives on both sides.

Each author produces a good article with probability α and pays a submission cost c per article. The expected payoff for strategy s_A against s_J is

$$u_A = r(p_g q_g + p_b q_b) - c(p_g + p_b),$$

where r denotes the reputational reward of acceptance and (p_g, p_b) are the proportions of good and bad papers implied by s_A. Table 2 lists all expressions.

Table 2: Author payoff enumeration for the 1J–1A base game.

Author strategy	Journal = AA	Journal = OG
AS	$r - c$	$r[\alpha(1-\varepsilon) + (1-\alpha)\lambda] - c$
OG	$\alpha(r - c)$	$\alpha[r(1-\varepsilon) - c]$
OB	$(1-\alpha)(r-c)$	$(1-\alpha)[r\lambda - c]$
NS	0	0

This table shows that increasing λ (the rate of false acceptance) increases the returns to opportunistic strategies (AS, OB), while increasing ε (the rate of false rejection) reduces the reward of quality-orientated behaviour (OG). These effects mirror empirical findings that lenient or noisy peer-review systems incentivise overproduction and lower selectivity (Brembs, 2018; Horbach and Halffman, 2018; Moher et al., 2020).

For the journal, the expected normalised utility balances publication quality and review cost.

$$u_J = \frac{BA_{\text{good}} - DA_{\text{bad}}}{S+1} - C_{\text{fix}}\left(\frac{S}{(N_A/N_J)+1}\right)^{\eta},$$

where $A_{\text{good}} = p_g q_g$, $A_{\text{bad}} = p_b q_b$, and $S = p_g + p_b$ is the submission load. In the base game $N_A = N_J = 1$, so the cost term simplifies to $C_{\text{fix}}(S/2)^{\eta}$. The enumerated journal payoffs are summarised in Table 3.

Table 3: Journal payoff enumeration for the 1J–1A base game.

Author strategy	Journal = AA	Journal = OG
AS	$\dfrac{B\alpha - D(1-\alpha)}{2} - C_{\text{fix}}\left(\frac{1}{2}\right)^{\eta}$	$\dfrac{B\alpha(1-\varepsilon) - D(1-\alpha)\lambda}{2} - C_{\text{fix}}\left(\frac{1}{2}\right)^{\eta}$
OG	$\dfrac{B\alpha}{\alpha+1} - C_{\text{fix}}\left(\frac{\alpha}{2}\right)^{\eta}$	$\dfrac{B\alpha(1-\varepsilon)}{\alpha+1} - C_{\text{fix}}\left(\frac{\alpha}{2}\right)^{\eta}$
OB	$-\dfrac{D(1-\alpha)}{2-\alpha} - C_{\text{fix}}\left(\frac{1-\alpha}{2}\right)^{\eta}$	$-\dfrac{D(1-\alpha)\lambda}{2-\alpha} - C_{\text{fix}}\left(\frac{1-\alpha}{2}\right)^{\eta}$
NS	0	0

The journal gains when publishing good work (B) but suffers a penalty (D) for accepting bad papers. When λ increases, the second term grows, making OnlyGood preferable over AlwaysAccept. This mechanism aligns with the theoretical expectations of incentive-based peer-review models (Heesen, 2018; Smaldino and McElreath, 2016; Zollmann et al., 2024).

Using the baseline parameters α=0.6, r=1, c=0.1, B=1, D=3, C_{fix}=0.1, η=1 and ε=λ=0.1, the explicit payoff matrices of 4×2 are:

Table 4: Baseline numerical payoffs at $(\varepsilon, \lambda) = (0.1, 0.1)$.

	Author payoff u_A		Journal payoff u_J	
Strategy (A,J)	AA	OG	AA	OG
AS	0.90	0.48	−0.35	0.16
OG	0.54	0.48	0.345	0.308
OB	0.36	0.00	−0.877	−0.106
NS	0.00	0.00	0.00	0.00

The enumeration confirms a single pure Nash equilibrium at (AS, OG): The authors submit indiscriminately, while the journal remains selective. This configuration mirrors empirical observations that the incentive misalignment between quantity-seeking authors and quality-seeking editors maintains systemic tension in scholarly publishing (Moher et al., 2020; Smaldino and McElreath, 2016; Zollmann et al., 2024).

To visualise how equilibria evolve with peer-review accuracy, dominant-strategy and payoff surfaces are mapped over (ε, λ) with $\varepsilon + \lambda \leq 1$. Figure 7 shows that higher false-acceptance rates λ expand opportunistic author behaviour (AS), while higher false-rejection rates ε erode cooperation by discouraging selective submission (OG). For journals, the selective OnlyGood rule dominates when errors are low, but the permissive AlwaysAccept gains traction once review noise increases.

To complement these categorical maps, Figure 8 plots the corresponding maximin payoffs u_A^* and u_J^*, which represent the guaranteed values each side can secure against its opponent's best

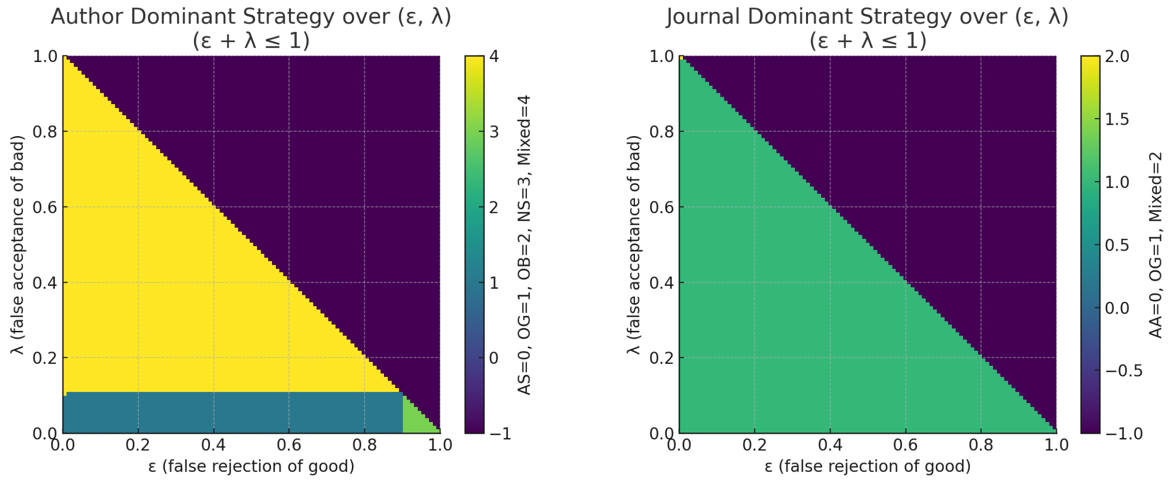


Figure 7: Dominant strategies over the error simplex (ε, λ). Left: Author best responses; Right: Journal best responses. Higher λ induces opportunism, while lower errors sustain cooperation.

response. Both surfaces decrease as ε and λ increase, showing that reciprocal accuracy directly benefits both populations.

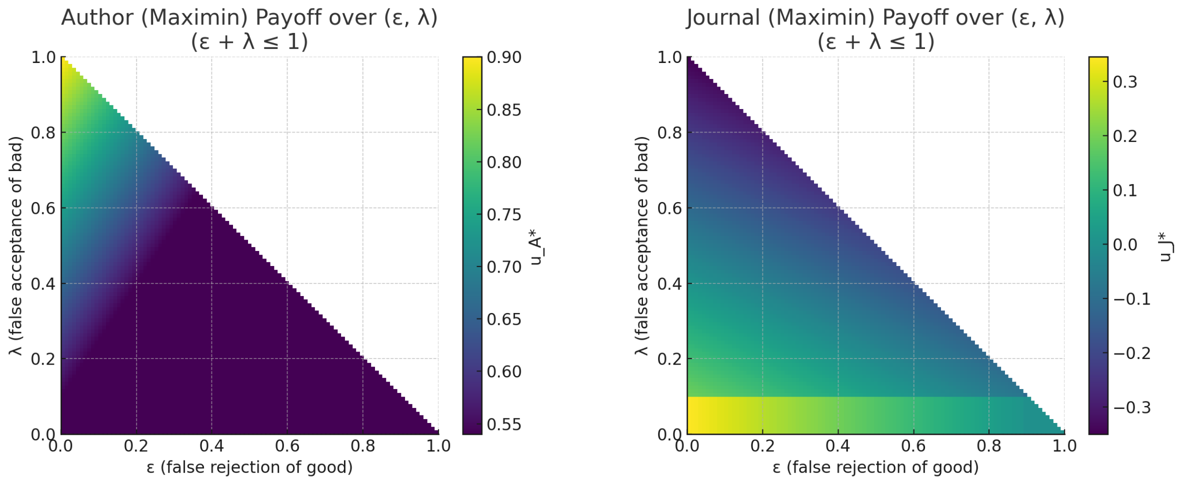


Figure 8: Maximin payoffs over (ε, λ). Left: Author guaranteed value u_A; Right: Journal guaranteed value u_J. Both decrease as review errors increase, confirming the cooperative efficiency of accurate peer review.

In general, the enumerated base game demonstrates how peer-review errors systematically distort incentives even in the simplest setting. Low error rates enable cooperative equilibria with balanced selectivity, while noisy evaluation environments induce opportunistic submission and editorial leniency, a pattern consistent with recent analyses of scientific incentive systems (Fanelli, 2010; Horbach and Halffman, 2018; Heesen, 2018; Moher et al., 2020). This foundational case sets the stage for scaling to multiple authors and journals under evolutionary dynamics in subsequent sections.

4.2. 1J–3A to 1J–50A: Population Scaling

To examine how the number of authors influences equilibrium stability and efficiency, the one-journal model was scaled to populations of $N_A \in \{3, 50\}$ authors. The parameters were fixed at $(B, D, r, c, C_{\text{fix}}, \eta) = (1, 3, 1, 0.1, 0.1, 1)$ and the screening errors at $(\varepsilon, \lambda) = (0.1, 0.1)$. For each $\alpha \in [0, 1]$, the dynamics of the discrete-time replicator was simulated for $T = 600$ iterations under ten random initializations, recording the mean equilibrium shares and 95% confidence intervals.

Table 5 reports the average equilibrium frequencies of the four strategies of the authors. The probability of Always Submit (AS) increases with α and stabilises near 0.55, while Only Good (OG) remains around 0.45. Larger populations (N=50) show smoother and more stable trajectories, consistent with convergence toward the deterministic mean field limit $\dot{x}_i = x_i(u_{A,i} - \bar{u}_A)$ (Sandholm, 2010). Only Bad (OB) disappears once $\alpha > 0.1$, and No Submit (NS) appears only for $\alpha < 0.1$.

Table 5: Average author strategy probabilities at equilibrium.

Strategy	N=3	N=50
AS	0.53 (0.07)	0.55 (0.03)
OG	0.43 (0.05)	0.45 (0.02)
OB	0.03 (0.02)	0.00 (0.00)
NS	0.01 (0.01)	0.00 (0.00)

Table 6 summarises the aggregated system results. The journal remains selective (OnlyGood) for most $\alpha < 0.9$, but moves to the permissive AlwaysAccept policy when expected paper quality becomes very high. For N=50, this transition occurs earlier due to the convex cost term $C_{\text{fix}}(S/((N_A/N_J) + 1))^\eta$, which scales faster with the submission volume S. Scientific quality Q exceeds 0.8 once $\alpha > 0.3$ and remains nearly constant with population size, while review cost increases roughly 40% at N=50, reducing the payoff of the journal from positive to negative values. This pattern matches recent observations that the expansion of submission volumes increases editorial load without improving quality (Brainard, 2022; Zollmann et al., 2024).

Table 6: Aggregate system outcomes averaged across α.

Metric	N=3	N=50
Journal payoff u_J	+0.48	−1.45
Average quality Q	0.84	0.83
Review cost (scaled)	0.42	0.59
Entropy (author mix, nats)	0.92	0.65

Despite aggregate convergence, individual behaviour remains diverse. In $\alpha = 0.6$ and $N = 50$, the distribution of the author strategies spans almost the entire interval $[0.05, 0.95]$, as shown in Figure 9. This polarisation indicates that macro-level stability coexists with micro-level diversity—authors cluster into AS-heavy or OG-heavy subgroups. This coexistence is typical of evolutionary coordination under convex cost functions (Smaldino and McElreath, 2016; Heesen, 2018).

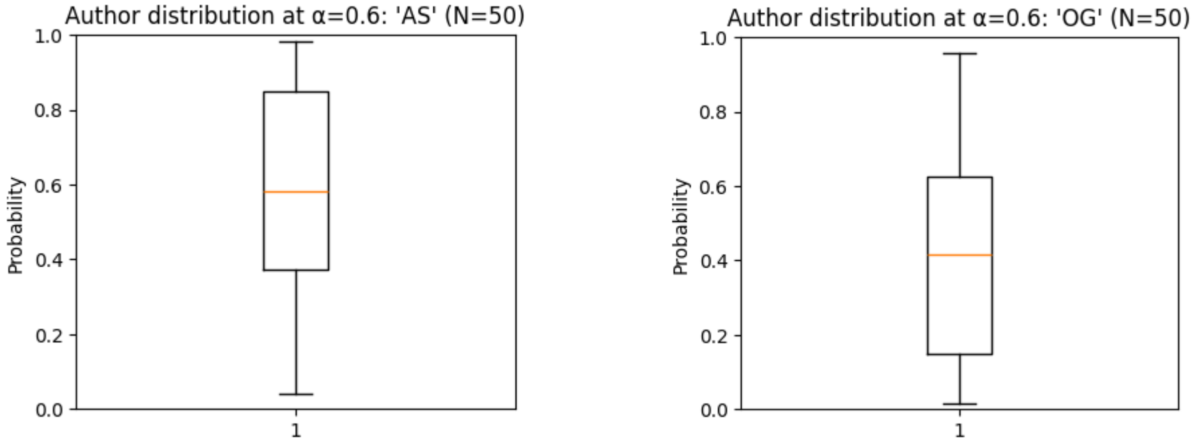


Figure 9: Distribution of author strategies at $\alpha = 0.6$, $N = 50$. Left: AS probabilities; Right: OG probabilities. Individual authors polarize into distinct clusters despite aggregate stability.

Summary Scaling the author population enhances deterministic stability, but reduces efficiency. Scientific quality remains high, while review costs and editorial leniency grow with N_A. The publication ecosystem thus behaves as a congested adaptive system: macro-level equilibrium stabilises, but efficiency declines— a dynamic consistent with modern analyses of peer-review overload and incentive misalignment (Heesen, 2018; Zollmann et al., 2024).

4.3. Error Sensitivity

This section examines how error sensitivity and payoff formulation shape the equilibrium between degenerate and cooperative regimes (RQ1, RQ3). We first compare the *count-based* and *normalized* journal payoffs, then analyse the resulting evolutionary dynamics.

Let S be the total number of reviewed submissions and acc_g, acc_b the expected numbers of good and bad accepted papers. The *count-based* objective is

$$u_J^{\mathrm{count}} = B\,\mathrm{acc}_g - D\,\mathrm{acc}_b - C(\varepsilon, \lambda)\,S,$$

which increases monotonically with submission volume S regardless of article quality, mechanically rewarding scale and encouraging journals to accept all papers.

To correct this, we use a *normalized* payoff that scales both quality and cost by submission rate:

$$u_J^{\text{norm}} = B\frac{\text{acc}_g}{S+\kappa} - D\frac{\text{acc}_b}{S+\kappa} - C(\varepsilon,\lambda)\frac{S}{N_{\text{authors}}+\kappa}.$$

Normalisation suppresses the artificial gain from unlimited volume, aligning incentives with per-submission quality rather than total quantity.

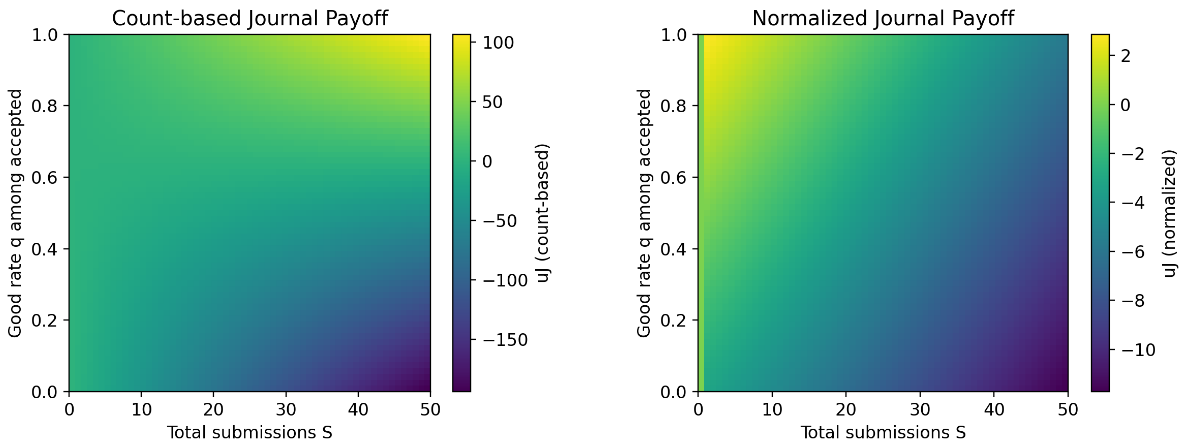


Figure 10: Journal payoff surfaces under (left) count-based and (right) normalized formulations. In the count-based design, utility grows with S and q, favouring unselective "accept-all" policies. The normalized version peaks at moderate S and high-quality share q, restoring selectivity.

Figure 10 shows that the count-based objective produces a purely scale-driven landscape where large S always dominates, whereas normalization flattens this region and introduces an interior optimum, penalising excessive expansion without quality.

Next, we test the effect of a bad-acceptance penalty ($D = 3$) under the count-based design. Figure 11 presents the resulting replicator dynamics. The journal population converges rapidly to *OnlyGood*, while authors evolve toward a mix dominated by *Always Submit* (AS) and *Only Good* (OG); the other strategies, *Only Bad* (OB) and *No Submit* (NS), vanish quickly.

The dominance of AS in Figure 11 follows analytically. When the journal plays *OnlyGood*, the payoff difference between *Always Submit* and *Only Good* is

$$u_A(\text{AS}) - u_A(\text{OG}) = (1-\alpha)\,[\,r\lambda - c(1-\lambda)\,],$$

which depends only on (r, c, λ), independent of B and D. With baseline parameters $(r, c, \lambda, \alpha) = (1, 0.1, 0.1, 0.6)$, $r\lambda - c(1-\lambda) = 0.01 > 0$, giving AS a strict advantage. Thus, AS dominance results from intrinsic incentive imbalance rather than simulation noise or duration.

Overall, the normalized payoff removes scale bias, and a large penalty D on bad acceptances

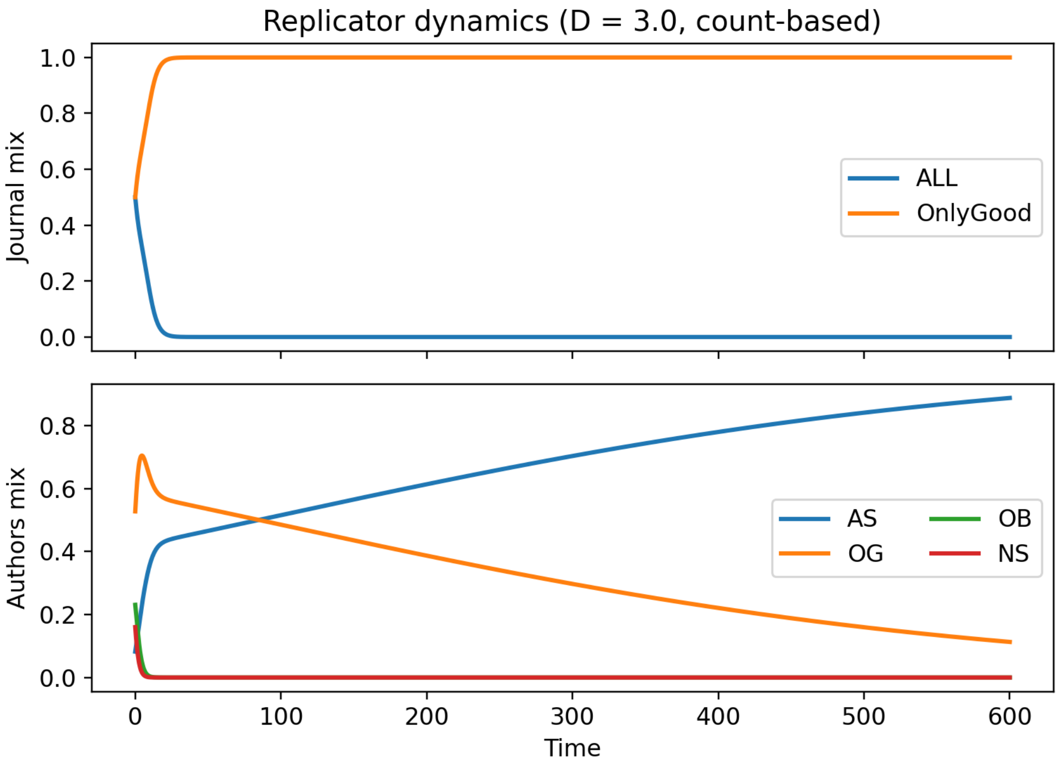


Figure 11: Replicator dynamics with a strong bad-acceptance penalty ($D = 3$). The journal converges to *OnlyGood*, while authors converge to an AS-dominated mix, confirming that penalties discourage indiscriminate acceptance.

suppresses the "accept-all" attractor. On the author side, lowering λ or raising c shifts equilibrium towards OG. In policy terms, improving review precision and adding friction to low-quality submissions jointly expand the cooperative selective regime, promoting higher scientific quality.

4.4. Empirical Validation

Using ICLR 2019–2022, the empirically estimated triplet $(\hat{\alpha}, \hat{\varepsilon}, \hat{\lambda})$ indicates: (i) a gradual increase in the share of high quality submissions $\hat{\alpha}$; (ii) a relatively stable false negative rate $\hat{\varepsilon}$; and (iii) a higher false positive rate $\hat{\lambda}$ in 2022 compared to 2020–2021. These results (see Figure 12 and Figure 13) suggest that acceptance broadened after 2021 while the underlying quality mix improved modestly. This pattern aligns with evidence of reviewer overload and changing selectivity thresholds in recent studies (Horbach and Halffman, 2018).

Two complementary diagnostics clarify how screening errors shape the quality achieved: (i) the $(\hat{\varepsilon}, \hat{\lambda})$ tradeoff plot (Figure 18a) shows a drift toward the $(0, 0)$ corner, evidencing a better precision–recall balance over time; and (ii) the quality–error surface (Figure 18b) overlays iso–contours of $\hat{\alpha}$ on the error plane. The surface reveals an empirical *frontier*: reducing only one

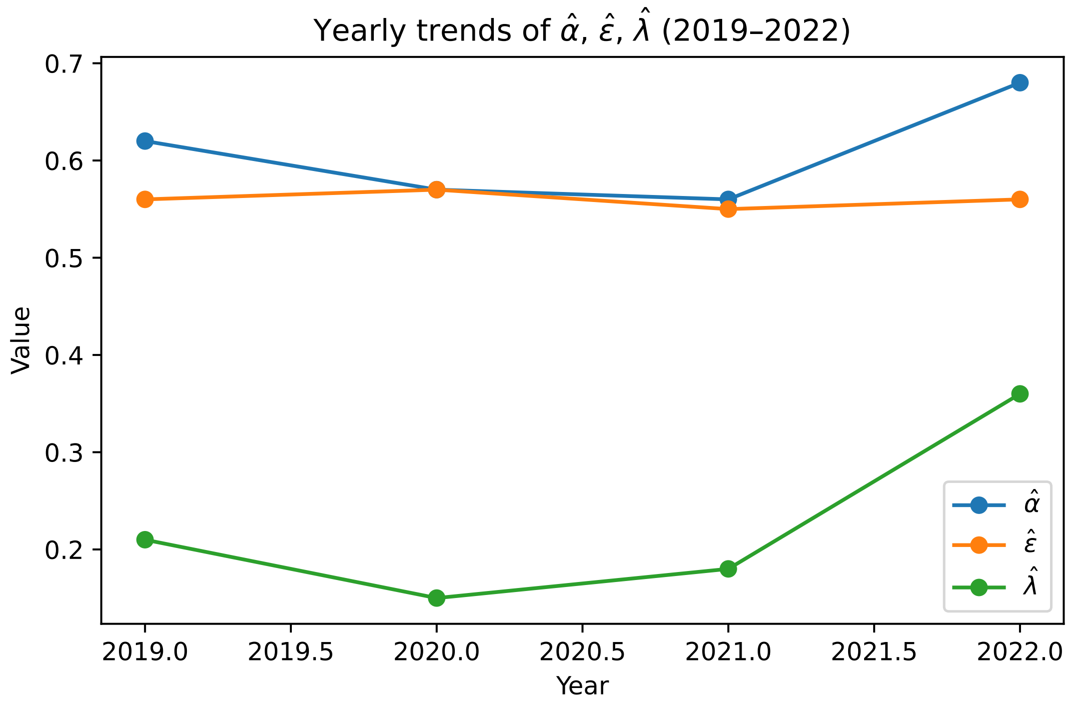


Figure 12: Yearly trends of $\hat{\alpha}$ (quality share), $\hat{\varepsilon}$ (false rejection), and $\hat{\lambda}$ (false acceptance), ICLR 2019–2022.

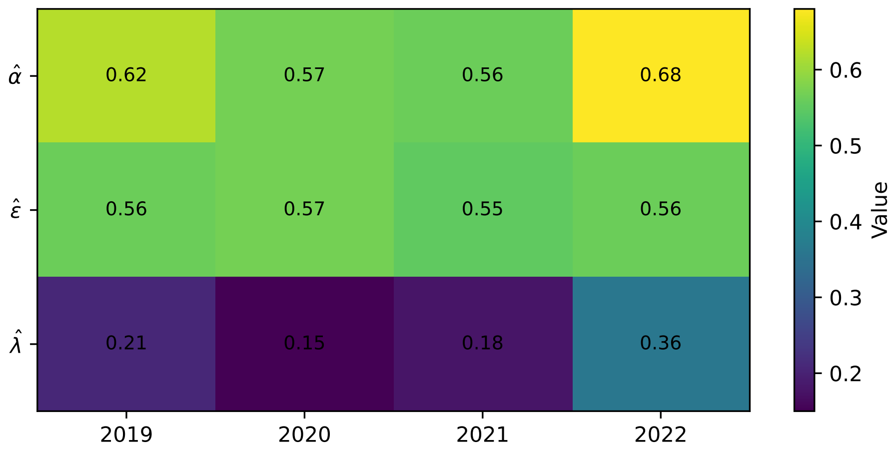


Figure 13: Year–by–year heatmap for $(\hat{\alpha}, \hat{\varepsilon}, \hat{\lambda})$. Darker cells indicate higher values.

type of error yields diminishing returns unless the other is reduced jointly. This finding echoes observations that unilateral editorial adjustments rarely improve systemic efficiency (Kovanis et al., 2016).

The three–dimensional path (Figure 19a) shows a movement toward a more selective and efficient corner between 2019 and 2022, consistent with journals staying close to *accepting-good* equilibria while authors provide a larger fraction of strong submissions. This trajectory supports the hypothesis that improving quality can coexist with broader acceptance when false–positive costs remain bounded (Moher et al., 2020).

The kernel density estimates of z -normalised logarithmic citations by decision (Figure 19b) expose the microstructure behind the estimated errors: a right–tail of rejected papers (false negatives) and a left–tail of accepted papers (false positives). The central overlap, though small, explains the non-zero $\hat{\varepsilon}$ and $\hat{\lambda}$. This phenomenon is consistent with empirical findings that high–risk or unconventional work often receives polarised outcomes (Serra-Garcia and Gneezy, 2021).

Calibrating the author–journal game with empirical triplets reproduces the model predictions: as $\hat{\lambda}$ increases, the authors shift from *Only Good* to *Always Submit*, while the journals remain at *Only Accept Good* except for extreme λ. In larger populations of authors ($N > 10$), review costs increase and selectivity erodes even when $\hat{\alpha}$ increases. This emerging inefficiency in scale parallels real–world peer review bottlenecks (Brainard, 2022).

4.4.1 Robustness and implications

The two–component Gaussian mixture for labelling "good" versus "bad"" articles is robust to the structure of the covariance, seed choice and the variation of the threshold; qualitative trends in Figure 12–Figure 19b remain stable. There are some implications: (i) increasing $\hat{\alpha}$ with stable $\hat{\varepsilon}$ but increasing $\hat{\lambda}$ suggests that quality gains coexist with looser acceptance; (ii) advancing along the quality–error frontier requires joint error reduction; and (iii) scale–driven inefficiency warns that growth alone can degrade selectivity without stronger institutional incentives (Serra-Garcia and Gneezy, 2021).

Note: Detailed summary tables are relocated to Appendix B for compactness.

5. Discussion

The results of both simulation and empirical validation converge to a coherent interpretation of how collective incentives and review accuracy jointly determine the equilibrium between authors and journals. Across theoretical, dynamic, and data-driven analyses, three main themes emerge: (i) the trade–off between inclusivity and quality in publication systems, (ii) the structural effects of population size on stability and efficiency, and (iii) the regulatory role of penalties for bad acceptances.

5.1. Inclusivity–Quality Trade–off, Population Scaling, and Efficiency Loss

The comparative statics over (ε, λ) in both one- and multiauthor models reveal a consistent pattern: as the false–positive rate λ increases, journals shift from the selective *Only Good* policy toward the inclusive *Always Accept* equilibrium, while authors adapt by broadening submissions (*Always Submit*). This mirrors the empirical trajectory of the estimated parameters $(\hat{\alpha}, \hat{\varepsilon}, \hat{\lambda})$ from the ICLR dataset, where $\hat{\alpha}$ and $\hat{\lambda}$ both increased between 2019 and 2022 (Fig. 18a). The data suggest that although the overall share of high–quality papers increased, the system became more tolerant of borderline submissions. This empirical movement toward higher λ yet larger α reflects a real–world inclusive quality frontier, where conferences expand their acceptance volume to accommodate the growth in participation and diversity of research topics (Bravo et al., 2019; Wang et al., 2023) .

From a game–theoretic perspective, this transition is an equilibrium shift driven by the marginal benefit of acceptance relative to the review cost γ. When screening becomes noisier or more expensive, the optimal best–response ratio of the journal changes to inclusive equilibria that maximise throughput rather than selectivity—an outcome consistent with empirical observations in modern conferences and open review platforms (Kovanis et al., 2016).

Increasing the number of authors N substantially reshapes the equilibrium dynamics. In replicator simulations, the scaling N from 3 to 50 stabilises the behaviour of the population (lower entropy and faster convergence), but reduces the reward of the journal due to the cost term of the convex review $C_{\text{fix}}\,(S/((N_A/N_J)+1))^{\eta}$(Figs. 16c and 15c–d) . This result aligns with the notion of a "collective efficiency loss": as more agents act rationally, the aggregate load overwhelms institutional capacity, leading to a socially suboptimal, yet individually stable equilibrium. Similar congestion effects are observed in evolutionary games with shared-resource dynamics, where stability increases but efficiency decreases as the population size grows (Khan, 2021).

In the publication context, this suggests that systemic expansion—whether through increased submission rates or broader participation—can erode the welfare of the journal without degrading the average scientific quality. The finding resonates with sociological studies of "the" "publish or die" dynamics, which show that institutional competition amplifies volume while offering diminishing marginal reputational returns (Kovanis et al., 2016; Wray, 2002). Hence, the journal selectivity declines not because of irrationality but because the equilibrium cost structure discourages rigorous screening at scale.

5.2. Penalty Mechanisms, Restoration of Selectivity, and Broader Implications

Introducing a penalty term D for bad acceptances fundamentally changes the equilibrium structure. Without such penalties, replicator dynamics converges to the degenerate *BadBias–All* equilibrium, where journals accept all submissions to minimise cost. Once D is sufficiently large ($D/B \geq 1$), new mixed equilibria emerge between *Perfect* and *Balance*, reinstating selectivity and incentivising authors to invest in high–quality work . This result echoes institutional theories that reputational sanctions are essential for maintaining peer review integrity (Kovanis et al., 2016; West and Bergstrom, 2019).

Conceptually, D functions as a mechanism–design parameter that aligns individual rationality with collective welfare. In real publication ecosystems, this corresponds to editorial reputation loss, post–publication corrections, or retraction costs that penalise low quality acceptance. Such institutional feedback loops, when sufficiently strong, steer the system back toward cooperative equilibria similar to the theoretical *OnlyGood–OnlyGood* outcome.

Taken together, these results demonstrate that the integrity of scientific publication depends not solely on review accuracy, but on the incentive balance between productivity and penalty. High λ without counterbalancing D induces volume–driven equilibria; conversely, strong reputational penalties and moderate review cost γ sustain high–quality equilibria even in large populations. This reinforces the insight that institutional design, not individual ethics alone, governs the collective outcome of scientific competition (Brembs, 2018; Helmer et al., 2017).

Nevertheless, several limitations remain. The current framework assumes homogeneous authors and journals and models interactions as one–shot or stationary replicator systems. Real peer review is multistage, reputation-dependent and often network–based. Future extensions may incorporate heterogeneous learning rates, multi–agent reinforcement learning, or endogenous reputation dynamics, following recent advances in computational social science (Marris et al., 2021; Anastassacos et al., 2021). Such expansions would allow the model to capture how strategic adaptation unfolds across time and institutions, bridging the gap between local optimisation and global scientific progress.

6. Conclusion

This study developed and empirically validated a game–theoretic framework to understand the strategic interaction between authors and journals in scientific publication. By combining analytical models, evolutionary simulations, and real-world data, we demonstrated how individual incentives, review accuracy, and institutional constraints jointly shape the equilibrium of the publishing ecosystem.

6.1. Summary of Findings and Contributions

The theoretical models reveal that, under accurate and moderately costly peer review, the system converges to a cooperative equilibrium in which authors adopt the *Only Good* strategy, and journals enforce the *Only Accept Good* policy. However, as the false–positive rate λ increases, the equilibrium shifts towards *Always Submit–Always Accept* configurations, reflecting a systemic relaxation of selectivity. When the false negative rate ε increases, quality-orientated authors disengage, leading to reduced participation and inefficiency. These theoretical predictions are empirically supported by the ICLR data analysis, which showed that from 2019 to 2022, the estimated share of quality $\hat{\alpha}$ increased while $\hat{\lambda}$ also rose—an inclusive-quality trade–off consistent with our comparative statics.

Scaling up the population from a few to many authors confirmed that while larger systems display greater stability and faster convergence (lower entropy), they also experience a marked decline in journal payoff due to convex review costs. This implies that expansion without coordination leads to an "efficiency trap," where collective behaviour stabilises, but aggregate welfare deteriorates. The introduction of explicit penalties for bad acceptances ($D > 0$) restored selectivity, suppressing the undesirable "accept-all" equilibrium and re-establishing high-quality equilibria. These findings formalise the intuition that effective reputational sanctions and editorial accountability are crucial to maintaining scientific integrity.

Theoretically, this work extends classical author–journal games (Zollmann et al., 2024; Heesen, 2018) by embedding them in a multi-agent evolutionary framework with population-level dynamics and empirically grounded parameters. The integration of support enumeration and replicator dynamics bridges static equilibrium concepts with adaptive learning models (Mertikopoulos et al., 2018), providing a tractable methodology to explore large-scale strategic environments where analytical solutions are intractable.

Empirically, the estimation of $(\hat{\alpha}, \hat{\varepsilon}, \hat{\lambda})$ from OpenReview–OpenAlex data constitutes one of the first quantitative calibrations of theoretical peer-review models using real submission and citation results. The estimated trajectories not only validate theoretical trade–offs but also illustrate how conference-scale review systems evolve toward higher throughput with bounded accuracy—an institutional analogue of the "evolutionary drift" predicted by replicator dynamics.

Practically, our findings underscore the importance of incentive design in scientific governance. Policies that incorporate review accountability—such as post-publication audits, transparent review reports, or retraction-based penalties—can be interpreted as increasing the effective D/B ratio, thus steering journals toward selective yet efficient equilibria. In contrast, unchecked emphasis on volume and speed reduces effective γ and D, pushing the system towards low-quality steady states. These insights resonate with ongoing debates on reforming academic evaluation and peer-review infrastructure (Bravo et al., 2019; Nature Communications Editorial, 2022).

6.2. Limitations, Future Directions, and Closing Remarks

Several simplifications constrain the scope of the present framework. The model assumes homogeneous authors and journals and excludes cross-journal competition or reputation diffusion. Real publication systems are multi-layered, with heterogeneity in institutional priorities, editorial networks, and temporal dependencies. Future research could extend this model in three directions: (i) incorporating heterogeneous learning rates and reputation dynamics through reinforcement learning or agent-based modelling (Marden and Shamma, 2018); (ii) analysing multi-journal equilibria where editorial standards co-evolve in competitive environments; and (iii) introducing stochastic perturbations to replicate uncertainty and bias propagation in real peer review.

By integrating theoretical, computational, and empirical perspectives, this thesis demonstrates that the equilibrium of scientific publishing is not merely a product of individual rationality, but of the incentive landscape that institutions create. Maintaining scientific quality therefore requires aligning reputational and economic payoffs through calibrated penalties and transparent review costs. In this sense, improving peer review is not only a matter of precision but also of equilibrium design—a lesson that bridges the mathematics of game theory with the sociology of science.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, **Associate Professor Julian García**, for his patience, guidance, and insightful feedback throughout my research journey. Since joining the university, I have never met a teacher who provides such thoughtful and detailed advice on academic writing and research. From the Summer Research project to the Minor Thesis, his encouragement has greatly increased my interest in research and made me realize how enjoyable and meaningful the process of discovery can be.

Appendix A. Supporting Figures for Population Scaling

This appendix contains supplementary figures referenced in the main text. They visualised the complete set of curves for author strategies, journal policies, entropy, and efficiency metrics.

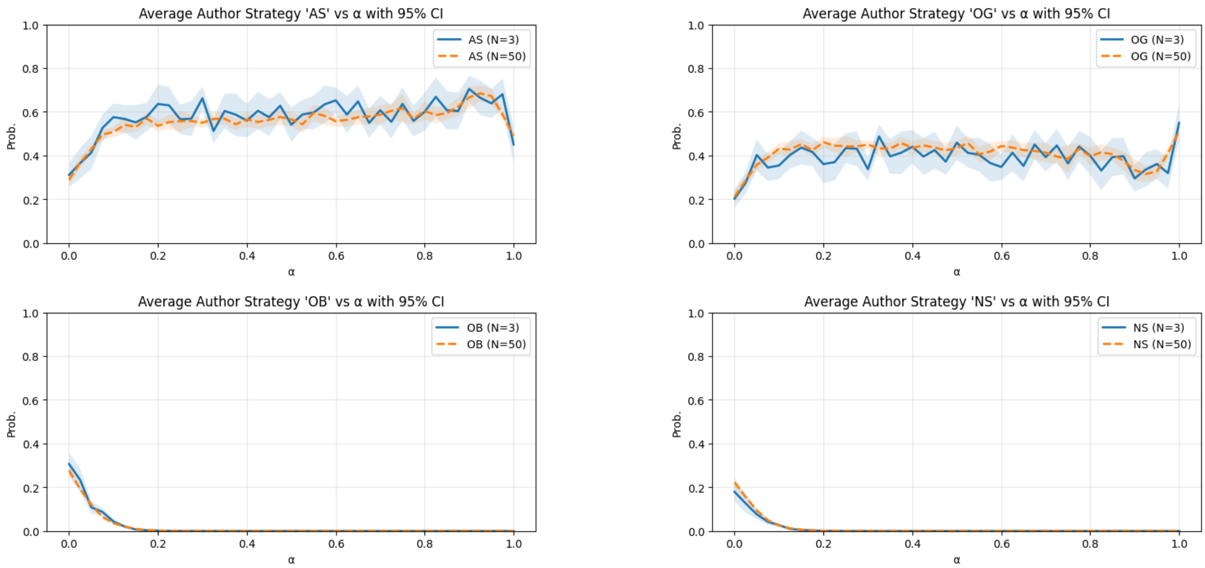


Figure 14: Average author strategy probabilities vs. α for $N = 3$ (solid) and $N = 50$ (dashed).

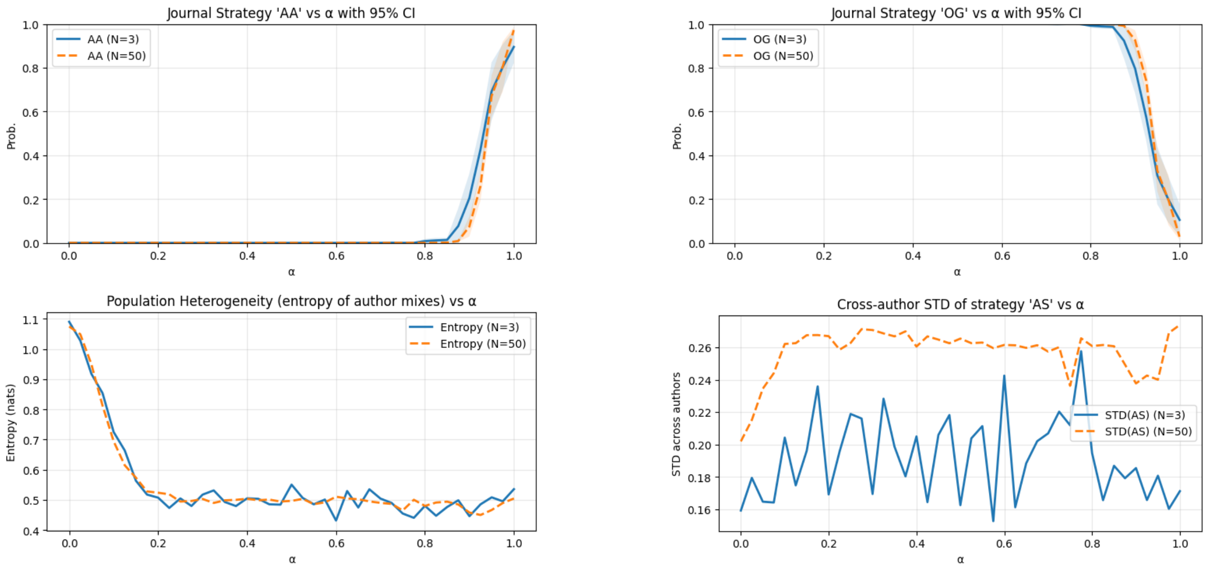


Figure 15: Journal dynamics and population diversity vs. α. (a) AA; (b) OG; (c) entropy; (d) cross-author STD.

All figures confirm that quality saturates while efficiency declines as the system scales. Large populations smooth individual fluctuations, but increase review burden and reduce selectivity.

Appendix B: Supplementary Figures for Empirical Validation

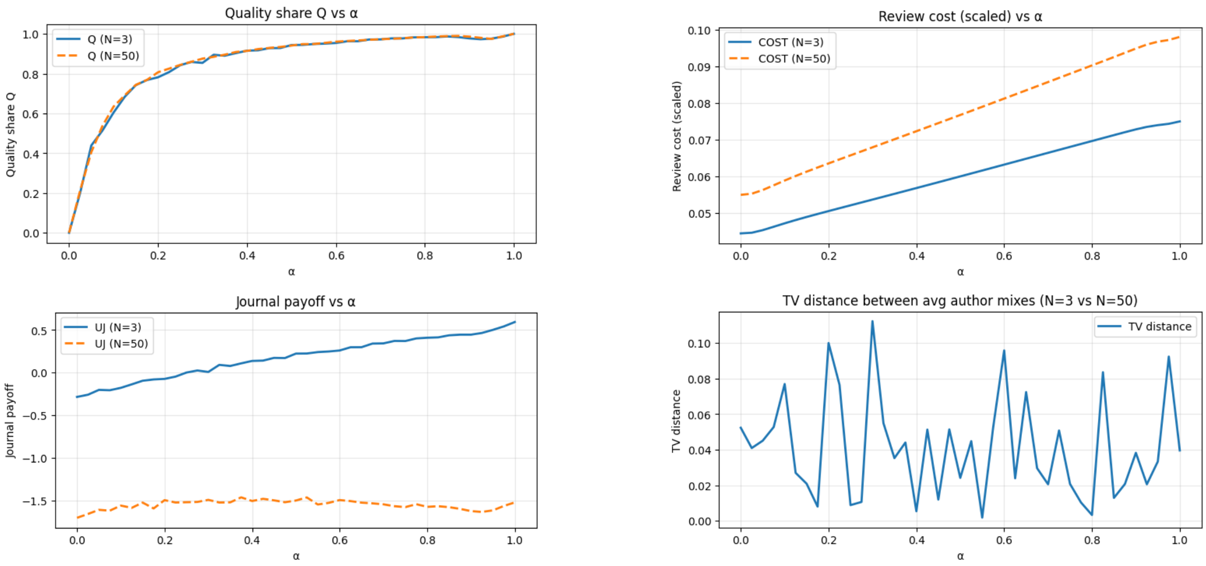


Figure 16: System-level metrics: (a) quality Q; (b) review cost; (c) journal payoff u_J; (d) total variation distance between $N = 3$ and $N = 50$.

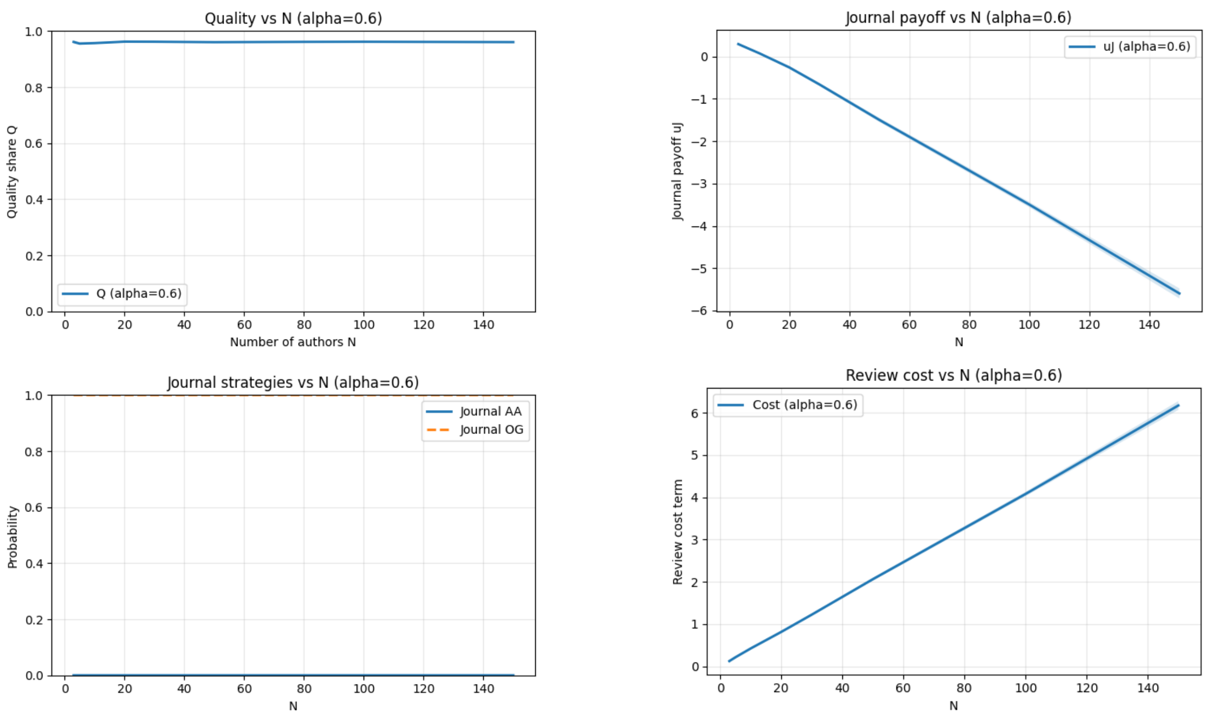
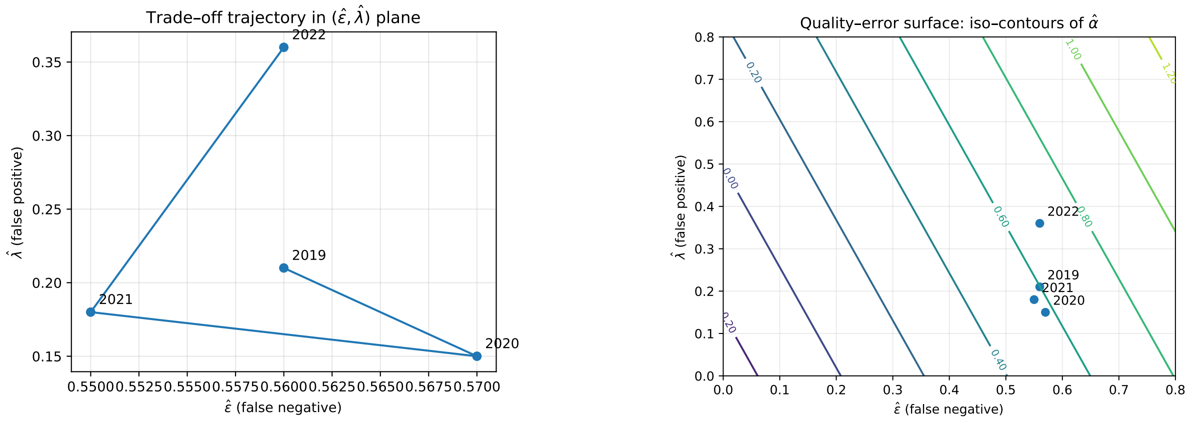


Figure 17: Scaling with N: (a) quality; (b) journal payoff; (c) journal strategy composition; (d) review cost.

Appendix C. Data Ethics and Privacy

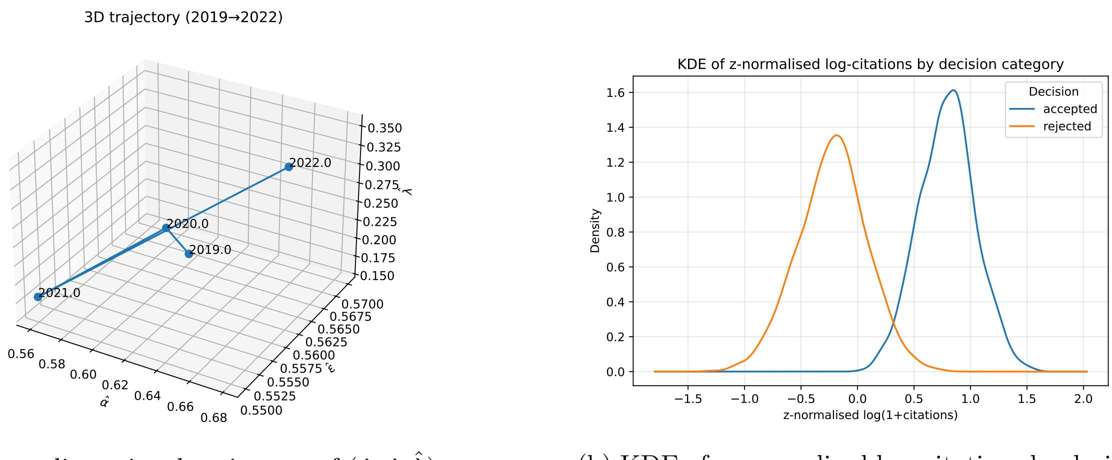
This research is consistent with Monash University's *Research Data Management and Ethics Policy* and the *Australian Code for Responsible Conduct of Research (2018)*. All data used in this study are publicly available, non-identifiable, and aggregated. No personal, confidential, or



(a) Trade–off trajectory in the $(\hat{\varepsilon}, \hat{\lambda})$ plane.

(b) Quality–error surface showing iso–contours of $\hat{\alpha}$.

Figure 18: Comparative visualization of error structure: (a) yearly trade–off between false–negative and false–positive rates; (b) efficiency frontier implied by joint error reduction.



(a) Three–dimensional trajectory of $(\hat{\alpha}, \hat{\varepsilon}, \hat{\lambda})$.

(b) KDE of z–normalised log–citations by decision category.

Figure 19: Empirical structure of publication outcomes: (a) evolutionary movement in parameter space; (b) micro–level distribution of citation outcomes.

proprietary information was accessed or processed.

Data sources. The data sets used in this study originate from open and publicly accessible repositories.

- The *ICLR OpenReview dataset* was obtained from the public GitHub repository maintained by Dmitry Kobak (Berens Lab). It contains peer-review metadata such as title, decision, and keywords.
- Citation data was retrieved from the *OpenAlex* open-access scholarly database, which provides bibliometric information through a public API.

Both datasets are distributed under open licences that explicitly permit academic and non-commercial research use. No authentication, restricted access, or personally identifiable records were involved.

Data handling and processing. All data processing was conducted locally using aggregated features, such as decision outcomes, citation counts, and derived statistical measures. No attempt was made to deanonymize the authors, reviewers, or institutions. Identifiers (e.g. OpenReview IDs) were used solely for automated matching between records and were not stored or shared in the final results.

Ethical considerations. The analysis focusses exclusively on patterns of publication behaviour at the population level: authors, journals, and review accuracy, without reference to individual identities. The study therefore qualifies as *low-risk research using publicly available data* and does not require additional human ethics approval. All visualisations and tables present only aggregated, non-traceable statistics.

References

Anastassacos, N., García, J., Hailes, S., and Musolesi, M. (2021). Cooperation and reputation dynamics with reinforcement learning. *arXiv*. Preprint.

Bachmann, M. (2021). Rapidfuzz: Fuzzy string matching library. `https://github.com/maxbachmann/RapidFuzz`.

Bloembergen, D., Tuyls, K., Hennes, D., and Kaisers, M. (2015). Evolutionary dynamics of multi-agent learning: A survey. *Journal of Artificial Intelligence Research*, 53:659–697.

Brainard, J. (2022). Can peer reviewing preprints catch on? *Science*, 378(6626):1264–1265.

Bravo, G., Grimaldo, F., López-Iñesta, E., Mehmani, B., and Squazzoni, F. (2019). The effect of publishing peer review reports on referee behavior in five scholarly journals. *Nature Communications*, 10(1):322.

Brembs, B. (2018). Prestigious science journals struggle to reach even average reliability. *Frontiers in Human Neuroscience*, 12:37.

Cressman, R. (2003). *Evolutionary Dynamics and Extensive Form Games*. MIT Press, Cambridge, MA.

Fanelli, D. (2010). Do pressures to publish increase scientists' bias? an empirical support from us states data. *PLoS ONE*, 5(4):e10271.

Gonzalez, J. and Kobak, D. (2024). Iclr openreview dataset. GitHub Repository: `https://github.com/berenslab/iclr-dataset`, MIT License.

Heesen, R. (2018). When journal editors play favorites. *Philosophical Studies*, 175(4):831–858.

Heesen, R. and Bright, L. K. (2021). Is peer review a good idea? *The British Journal for the Philosophy of Science*, 72(3):635–663.

Helmer, M., Schottdorf, M., Neef, A., and Battaglia, D. (2017). Gender bias in scholarly peer review. *eLife*, 6:e21718.

Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge.

Horbach, S. P. J. M. and Halffman, W. (2018). The changing forms and expectations of peer review. *Research Integrity and Peer Review*, 3(8):1–15.

Khan, A. (2021). Evolutionary stability of behavioural rules in bargaining. *Journal of Economic Behavior Organization*, 187:399–414.

Knight, V. (2017–). Nashpy: A python library for 2-player game theory. `https://github.com/drvinceknight/Nashpy`.

Kovanis, M., Porcher, R., Ravaud, P., and Trinquart, L. (2016). The global burden of journal peer review in the biomedical literature: Strong imbalance in the collective enterprise. *PLOS ONE*, 11(11):e0166387.

Lemke, C. E. and Howson, J. T. (1964). Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics*, 12(2):413–423.

Marden, J. R. and Shamma, J. S. (2018). Game theory and distributed control. *Handbook of Game Theory with Economic Applications*, 4:861–899.

Marris, L., Omidshafiei, S., Munos, R., Piliouras, G., and Tuyls, K. (2021). Multi-agent training beyond zero-sum with correlated equilibrium meta-solvers. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*, pages 7480–7491.

Mertikopoulos, P., Papadimitriou, C., and Piliouras, G. (2018). Cycles in adversarial regularized learning. In *Proceedings of the 29th Annual ACM–SIAM Symposium on Discrete Algorithms (SODA)*, pages 2703–2717.

Mertikopoulos, P. and Sandholm, W. H. (2016). Learning in games via reinforcement and regularization. *Mathematics of Operations Research*, 41(4):1297–1324.

Moher, D., Bouter, L., Kleinert, S., and Glasziou, P. (2020). The hong kong principles for assessing researchers: Fostering research integrity. *PLoS Biology*, 18(7):e3000737.

Nash, J. F. (1951). Non-cooperative games. *Annals of Mathematics*, 54(2):286–295.

Nature Communications Editorial (2022). Transparent peer review for all. *Nature Communications*, 13:6817.

Osborne, M. J. (2003). *An Introduction to Game Theory*. Oxford University Press, Oxford.

Priem, J., Piwowar, H., and Orr, R. (2022). Openalex: A fully-open index of scholarly works. *Quantitative Science Studies*, 3(3):885–902.

Sandholm, W. H. (2010). *Population Games and Evolutionary Dynamics*. MIT Press, Cambridge, MA.

Savani, R. and Turocy, T. L. (2025). Gambit: The package for computation in game theory, version 16.3.0. `https://www.gambit-project.org`.

Serra-Garcia, M. and Gneezy, U. (2021). Nonreplicable publications are cited more than replicable ones. *Science Advances*, 7(21):eabd1705.

Shah, D. (2010). Dynamics in congestion games. *ACM SIGecom Exchanges*, 10(2):40–45.

Shapley, L. S. (1974). A note on the lemke–howson algorithm. In *Pivoting and Extension*, volume 1 of *Mathematical Programming Studies*, pages 175–189. Springer, Berlin, Heidelberg.

Smaldino, P. E. and McElreath, R. (2016). The natural selection of bad science. *Royal Society Open Science*, 3(9):160384.

Strevens, M. (2012). Economic approaches to understanding scientific norms. *Episteme*, 8(2):184–200.

Taylor, P. D. and Jonker, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40:145–156.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.

Wang, G., Peng, Q., Zhang, Y., and Zhang, M. (2023). What have we learned from openreview? *World Wide Web*, 26(2):683–708.

Weibull, J. W. (1995). *Evolutionary Game Theory*. MIT Press, Cambridge, MA.

West, J. D. and Bergstrom, C. T. (2019). Misinformation in and about science. *Proceedings of the National Academy of Sciences*, 116(16):7659–7665.

Wray, K. B. (2002). The epistemic significance of collaborative research. *Philosophy of Science*, 69(1):150–168.

Zollmann, K. J. S., García, J., and Handfield, T. (2024). Academic journals, incentives, and the quality of peer review: A model. *Philosophy of Science*, 91(1):186–203.