Artificial Intelligence and Quantity Traps

Naveen Gondhi Lin Shen Junyuan Zou *

August 1, 2026

Abstract

We study an equilibrium model of academic publishing in which researchers choose the quality (and quantity) of papers, and journals screen submissions imperfectly. AI expands researchers' productive frontier by increasing the routine capacity, but their deep attention capacity remains unchanged. As AI improves, researchers have stronger incentives to produce more papers, which leads to more congestion at journals. More congestion, in turn, increases the screening noise, which incentivizes authors to write more papers rather than investing in quality. This amplification mechanism leads to (i) more submissions, (ii) lower average quality, and (iii) under identifiable conditions, reduces the published knowledge. Interestingly, the model also highlights how heterogeneity across researchers (in career stage, time constraints, and institutional burdens) reshapes the composition of published work, favoring scalable production over deep contribution for some groups more than others.

Key Words: Artificial Intelligence, peer review, screening congestion, publication quality, submission incentives

JEL Classification: D83, I23, O33

*Naveen Gondhi, INSEAD, naveen.gondhi@insead.edu; Lin Shen, INSEAD, lin.shen@insead.edu; Junyuan Zou, INSEAD, junyuan.zou@insead.edu. We thank Frederico Belo, Guillermo Ordonez, Joel Peress, and conference participants at Finance Theory Group (FTG) and seminar participants at INSEAD for useful feedback. All errors are the authors' alone.

1 Introduction

Generative artificial intelligence (AI) is reshaping the economics of scholarly production. By automating routine tasks, it raises productivity at nearly every stage of research, from ideation and analysis to drafting and revision (Korinek 2023). Yet this productivity gain also creates a tension between quantity and quality. This increase in quantity is visible in the number of working-papers in Finance and Economics: as Figure 1 shows, monthly uploads to SSRN rose from roughly 4,500 papers in late 2023 to more than 7,000 during 2026 and the trend is clearly upwards. Studying the period since the late-2022 release of ChatGPT, Gartenberg et al. (2026) document a 42% rise in journal submissions at Organization Science, together with more AI-assisted writing in both manuscripts and reviews. At the same time, the writing quality and topical diversity of submissions decline, suggesting that the productivity gain raises quantity more than quality.[1]

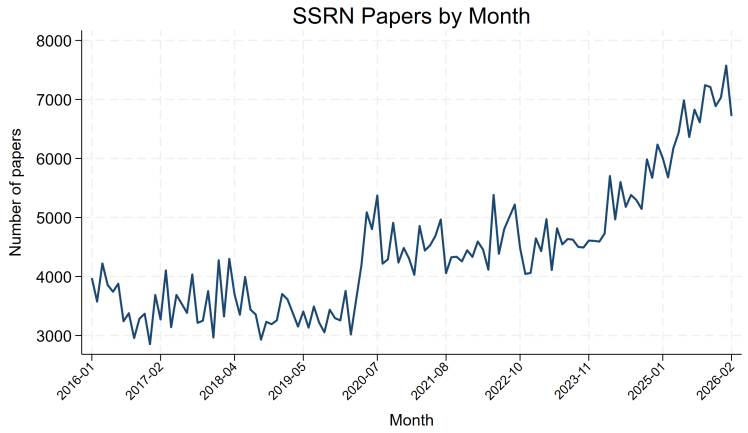


Figure 1: Monthly number of working papers uploaded to SSRN in finance and economics.

This is more concerning once we account for the externalities of rising research volume, and in particular for congestion in peer review. Peer-review labor is scarce, so a growing flow of submissions can erode the quality of the selection that review delivers. Artificial-intelligence conferences provide a clear example. A surge in submissions has raised questions about the credibility and sustainability of their review system (Kim, Lee & Lee 2025). In their most recent cycle, five major AI conferences saw submissions rise by roughly 60% in aggregate relative

[1] Relatedly, Kusumegi et al. (2025) find that AI-assisted research tends to be linguistically more complex but substantively thinner, and Hao et al. (2026) show that AI tools raise scientists' output and citations while narrowing the focus of their work.

to the previous year, and organizers warned that this growth places severe strain on review quality and accountability.[2]

These developments raise questions about knowledge production in the era of AI. First, how does AI reshape researchers' incentives to trade off paper quantity against quality? Second, how do its effects vary across heterogeneous researchers? Third, how should publication policy adjust so that AI-induced productivity gains translate into socially valuable knowledge production?

To answer these questions, we develop a model of research production and publication screening with three key ingredients. First, researchers choose how many papers to produce and face a quantity-quality tradeoff: expanding the research portfolio spreads effort more thinly across papers and therefore reduces the quality of each paper. Thus, although generative AI improves quality for a fixed portfolio, if a researcher writes more papers, her average research quality can decline. Second, review noise increases with aggregate submission volume. Each researcher internalizes quality dilution within her own research portfolio but does not internalize the externality that her submissions make the publication screen less informative for everyone else. Third, researchers have publish-or-perish payoffs: they derive private value only from published papers, while each submission entails the same cost.

We first examine how AI reshapes researchers' research portfolio choices. The publish-or-perish payoff generates a natural quantity-quality tradeoff. By writing one additional paper, a researcher creates another publication opportunity. However, this extra paper dilutes effort and quality, which reduces the publication probability of every paper in the portfolio. The researcher chooses her research portfolio by balancing these two forces, which pins down an optimal target publication probability for each paper.

AI advancement affects this choice through two channels: a direct *productivity channel* and an indirect *review-congestion channel*. The direct productivity channel operates through the researcher's own reoptimization of research portfolio. Holding review noise fixed, AI advancement has no direct influence on researchers' publication incentive. Therefore, researchers maintain their target publication probability by keeping the quality of each paper unchanged and using the AI-induced productivity gain to produce more papers. In other words, the AI-induced productivity gains translate into quantity rather than quality.

The indirect channel operates through equilibrium review congestion. When researchers submit more papers, the review process becomes more congested and publication outcomes

[2]Relative to the previous year, submissions rose by 38% at NeurIPS, 13% at CVPR, 68% at ICLR, 98% at ICML, and more than 80% at AAAI. The submission volume data are from official conference statistics.

become noisier. Higher review noise then reshapes researchers' incentives through two mechanisms. First, publication outcomes become less sensitive to paper quality, so the cost of quality dilution from writing more papers falls. This incentivizes researchers to lower their target publication probability and prioritize quantity over quality. Second, higher review noise alters the mapping from quality to publication probability. A noisier review process raises the publication probability of papers below the publication threshold and lowers that of papers above the threshold. For researchers with low target publication probabilities, this effect reinforces the incentive to expand quantity because noise increases the chance that lower-quality papers are accepted. For researchers with high target publication probabilities, however, higher review noise reduces publication probability, which induces them to cut quantity and concentrate effort on improving quality. As a result, the overall effect of higher review noise on research quantity is ambiguous. We call a researcher *congestion-amplifying* if she writes more papers in response to an increase in review noise, and *congestion-disciplining* if she writes fewer papers.

Motivated by the low acceptance rates at economics and finance journals, we focus on environments in which researchers are congestion-amplifying in the aggregate. In these environments, aggregate submissions increase when the publication screen becomes noisier. When researchers are homogeneous, this implies that AI advancement increases equilibrium research volume while reducing average research quality. Moreover, this also leads to a feedback loop between submission incentives and screening precision. A noisier review process encourages researchers to produce more papers, and the additional submissions further congest the review process and raises review noise.

This feedback loop can give rise to a self-fulfilling *AI-induced quantity trap*. When AI intensity is low, the economy admits a unique equilibrium with low research quantity and high average quality. At intermediate levels of AI intensity, multiple equilibria may emerge: one with low quantity and high average quality, and another with high quantity and low average quality. As AI progresses, the economy may switch from the high-quality equilibrium to the low-quality equilibrium, generating a discrete jump in research volume and a discrete decline in average quality. With further AI advancement, the high-quantity, low-quality equilibrium eventually becomes the unique equilibrium. Importantly, for any fixed research volume, AI advancement raises research quality. The quantity trap is an equilibrium consequence of private research portfolio expansion in a congested review system.

AI advancement increases research volume while lowering average research quality. This raises the question of how AI affects aggregate socially valuable knowledge production, which we proxy by the aggregate quality of published papers, because published papers are most

likely to reach a broad audience. AI affects aggregate published quality through three forces: a scale gain from more submissions, a dilution loss from lower mean quality, and a screening-congestion loss from noisier selection. Since these forces are at play simultaneously, the overall effect depends on the relative strength of the forces and is generally ambiguous.

We then examine how AI advancement affects heterogeneous researchers. Recall that researchers trade off the benefit of additional publication opportunities against the quality dilution that results from writing more papers, which pins down an optimal target publication probability for each paper. Heterogeneity that directly affects publication incentives–such as submission costs, private benefits from publication, or the elasticity of quality dilution–can generate different target publication probabilities across researchers. By contrast, heterogeneity that does not directly affect publication incentives implies that researchers choose the same target publication probability and average paper quality. Such heterogeneity nevertheless matters for research volume. For example, if researchers differ only in their AI skills, AI advancement generates heterogeneous productivity gains. Because AI skill does not directly affect publication incentives, researchers produce papers of the same average quality. However, researchers with stronger AI skills write more papers, and expand disproportionately as AI improves.

In contrast, heterogeneity in the ratio of submission costs to private publication benefits changes publication incentives directly. A pre-tenure researcher with abundant research funding, for example, faces low submission costs and high publication value, and therefore has a low cost-benefit ratio. Such a researcher does not need a high acceptance probability to justify submitting papers, so her target publication probability is low. As a result, she produces more papers of lower average quality than researchers with higher cost-benefit ratios. Moreover, AI-driven increases in aggregate submission volume raise review noise through the congestion externality. This effect has a larger impact on researchers with low cost-benefit ratios, thereby widening quantity and quality gaps across heterogeneous researchers. Interestingly, researchers with high cost-benefit ratios may instead be disciplined by congestion. When the cost-benefit ratio is sufficiently high, the researcher targets a high publication probability. Higher review noise then reduces her publication prospects, which incentives her to cut quantity and concentrate effort on quality. Thus, as AI progresses and aggregate submission volume rises, these researchers may produce fewer papers of higher average quality.

The last source of heterogeneity we examine is the elasticity of quality with respect to effort. For deep researchers with high elasticity, effort translates more effectively into paper quality: a given increase in effort raises quality more than it does for other researchers. This source of heterogeneity combines features of the previous two. Like the cost-benefit ratio, it affects pub-

lication incentives and the implied quantity-quality tradeoff. Because quality is more elastic to effort for deep researchers, scaling up research volume is more costly: writing more papers dilutes effort across projects and reduces quality more sharply. As a result, deep researchers target a higher publication probability and produce papers of higher average quality. This force also pushes them toward lower research volume. At the same time, this heterogeneity also resembles differences in AI skill because it directly affects productivity. Holding research capacity fixed, deep researchers can produce higher-quality papers at any given research volume. This productivity advantage can offset their stronger quality-dilution effect, so the equilibrium comparison of research volume across types is ambiguous. As before, deep researchers who target high publication probabilities may be disciplined by congestion. AI-induced increases in review noise can then induce them to improve paper quality, while congestion-amplifying shallow researchers lower paper quality. AI advancement can therefore widen the quality gap between deep and shallow researchers.

We also derive policy implications for facilitating socially valuable knowledge production. The analysis focuses on policy levers that a journal can use to maximize the aggregate quality of published papers. For a fixed screening technology, constrained efficiency requires combining quantity control with an optimal publication threshold. For any given submission volume, the journal can infer the distribution of paper quality and sets the publication threshold so that the marginal published paper has zero expected social value. Because individual researchers do not internalize the congestion externality, a separate quantity instrument is needed to reduce excessive submissions. The journal can implement this quantity control either by raising submission fees or by directly capping submission volume.

Interestingly, although the optimal submission volume increases monotonically with AI advancement, the submission fee that implements it can be non-monotone. This reflects two opposing forces. On the one hand, AI induces excessive submissions, calling for a higher submission fee to curb congestion. On the other hand, the socially optimal submission volume also rises with AI, calling for a lower fee to accommodate the larger efficient volume. Similarly, the optimal publication threshold can be non-monotone in AI advancement because it balances Type I and Type II errors: accepting low-quality papers and rejecting high-quality papers. Under the optimal policy, however, AI advancement *always* increases aggregate published quality. In this sense, an appropriately adjusted publication policy can translate AI-induced productivity gains into socially valuable knowledge production. This result highlights the importance of updating publication policy in a timely manner as AI advances. Finally, reducing review noise is always beneficial. Thus, if AI can improve review accuracy at low cost, journals should

adopt it as part of the screening technology.

Although we focus on scholarly publication in this paper, the mechanism applies more broadly to innovation systems in which agents allocate scarce creative effort across projects and seek approval from a noisy, capacity-constrained evaluator. Startups and entrepreneurial scientists, for example, choose how many ideas to develop and pitch to investors or grant panels; within firms, product teams choose how many projects to propose to investment committees or senior executives. In each setting, AI can raise the productivity of idea generation, prototyping, and communication for a fixed project portfolio. At the same time, it can induce agents to expand the portfolio and dilute effort across ideas. As the screening process becomes congested and less accurate, private incentives can shift toward generating more proposals rather than developing better ones. The quantity trap is therefore not specific to academia. It captures a general risk in AI-enabled innovation systems: private payoffs depend on passing a screening system, while social value depends on the quality of the ideas ultimately selected.

Literature Review This paper first relates to work on AI and scholarly production. Korinek (2023) describes generative AI as a productivity tool for economic research. Gartenberg et al. (2026) provide the closest empirical motivation by linking ChatGPT to higher submission volume, more AI-heavy writing, and weaker measured writing quality in peer review. Hao et al. (2026) show in a broader scientific setting that AI tools can raise individual output while narrowing the collective focus of science. These papers motivate the productivity shock studied here. Our contribution is to analyze how a technology that improves quality at a fixed portfolio size can reduce equilibrium mean quality when researchers choose submission quantity and the publication screen becomes congested.

A second set of papers studies noisy evaluation and peer-review feedback. Bergstrom & Gross (2026) analyze how author self-screening, journal sorting, and review accuracy can form feedback cycles that imperil peer review. Azar (2015) models informed authors who decide whether to submit when journals observe noisy review signals. Adda & Ottaviani (2024) show that greater evaluation noise can increase applications in nonmarket allocation. We build on these forces by endogenizing the interaction between AI-driven research production, private portfolio choice, and aggregate screening precision. The distinct mechanism is that AI raises productive capacity, researchers expand submission portfolios, and aggregate volume makes the publication screen less informative.

The paper also contributes to the literature on journal screening, thresholds, and publication policy. Zollman, García & Handfield (2024) study journal incentives to choose review

accuracy and acceptance standards, while Gehrig & Stenbacka (2021) and Atal (2010) analyze how journal competition and screening regimes affect published quality. Cotton (2013) shows how submission fees and response times can screen low-probability submissions and manage referee burden. Bayar & Chemmanur (2021) and Card & DellaVigna (2020) provide theoretical and empirical foundations for treating editorial decisions as based on noisy referee information. Frankel & Kasy (2022) study optimal publication rules when findings affect public beliefs and journal space is scarce, and Boudreau et al. (2016) show that evaluator assignment and intellectual distance shape scientific resource allocation. Relative to this work, our policy implication is that quantity instruments, publication thresholds, and screening-informativeness investments operate on different margins of the AI-induced congestion problem.

Finally, the paper speaks to research incentives and the quality of academic publishing. Ellison (2002*b*) documents the slowdown of the economics publication process, and Ellison (2002*a*) models evolving standards for academic publishing. Heckman & Moktan (2020) show that top-journal placement has large career consequences in economics, which motivates high and heterogeneous private values of publication. Hirshleifer (2015) and Welch (2014) highlight distortions and noise in the review process, while Hadavand, Hamermesh & Wilson (2024) documents publication delays and discusses reforms. Kiri, Lacetera & Zirulia (2018) study private incentives and high-quality scientific production, and Kovanis et al. (2016) document the scale and imbalance of peer-review labor. Our contribution is to connect these incentive and capacity frictions through an equilibrium feedback index: in the quantity-trap region, a privately useful research technology can lower equilibrium mean quality because the induced submission expansion weakens the precision of the publication screen.

The rest of the paper proceeds as follows. The model section presents research production, screening congestion, payoffs, and symmetric equilibrium. The equilibrium section characterizes the best response, defines the quantity-trap region, derives the stable-trap comparative static, and decomposes aggregate published quality. The appendix records the outside-trap benchmark. The heterogeneity section extends the mechanism to heterogeneous researchers and derives the sorting and composition predictions. The policy section discusses policy implications based on the separation of quantity, selection, and screening-informativeness instruments. The conclusion summarizes the main lesson.

2 Model Setup

This section specifies the model primitives. The economy is populated with a unit mass of researchers, indexed by $i \in [0, 1]$, and a journal with finite screening capacity. Researchers are ex ante identical in the baseline; we introduce heterogeneity in Section 4. Researchers produce research papers and submit them to the journal, which then screens them for publication. The model is static. To fix ideas, we take the time period to be one year.

Research Production. Research production requires two inputs: routine effort and deep attention. A researcher faces capacity constraints on both inputs. Routine effort encompasses tasks such as drafting, coding, checking references, and producing variants of an idea. These are the activities for which generative AI is most naturally interpreted as a productivity tool; Korinek (2023), for example, emphasizes AI's usefulness for writing, coding, ideation, and research assistance. A researcher's total routine capacity rises with AI according to

$$R(a) = R_0(1 + \gamma a),$$

where $R_0 > 0$ is its baseline level, $a \geq 0$ is the intensity of generative AI, and $\gamma > 0$ measures the exposure of routine research tasks to generative AI. The exact functional form is not important as long as routine capacity $R(a)$ is increasing in AI intensity a. Deep attention, by contrast, captures the scarce researcher time/bandwidth required to choose questions, weigh arguments, interpret results, and judge which ideas deserve sustained effort. The deep-attention capacity is fixed at $Z > 0$, unchanged by generative AI.

A researcher who writes n papers spreads both inputs symmetrically across them.[3] We treat n as a continuous variable: $n = 0.5$, for example, represents one paper every two years. Each paper therefore receives a routine input

$$r(n, a) = \frac{R(a)}{n}$$

and a deep-attention input

$$z(n) = \frac{Z}{n}.$$

[3] Equivalently, the researcher could choose time per paper subject to fixed research time Z: writing $n = Z/\ell$ gives the same problem. If unequal times across ex ante identical papers are allowed, additivity implies that all active papers use the same payoff-per-time-maximizing ℓ, so the continuous-choice optimum is unchanged.

The two inputs determine the average quality of a research paper according to a Cobb–Douglas production function:

$$q(n,a) = \theta + \beta_R \log r(n,a) + \beta_Z \log z(n),$$

with $\beta_R, \beta_Z > 0$. Substituting for $r(n,a)$ and $z(n)$ gives

$$q(n,a) = \Theta(a) - K \log n, \tag{1}$$

where $K \equiv \beta_R + \beta_Z$ is the combined input elasticity of paper quality, which is the *dilution factor* governing how rapidly quality falls as inputs are spread across more papers, and $\Theta(a) \equiv \theta + \beta_R \log R(a) + \beta_Z \log Z$ is the benchmark paper quality for a researcher who chooses $n = 1$, i.e., writes one paper per year.

The Cobb–Douglas specification serves as a tractable benchmark that captures two key features of research production.[4] First, generative AI raises research quality by raising researchers' routine task productivity,

$$\frac{\partial q(n,a)}{\partial a} = \frac{\beta_R \gamma}{1 + \gamma a} > 0. \tag{2}$$

Second, writing more papers (for a fixed AI level) dilute efforts and quality,

$$\frac{\partial q(n,a)}{\partial n} = -\frac{K}{n} < 0. \tag{3}$$

The marginal dilution K/n is decreasing in n, so the quality penalty of an additional submission diminishes as n rises.

For a researcher who writes n papers, the mean quality of each paper is $q(n,a)$. The realized true quality of a paper fluctuates around this mean,

$$\tilde{q} = q(n,a) + \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2), \tag{4}$$

where ε is independent across papers and researchers.

Editorial Screening We model peer review as noisy evaluation.[5] For a paper with true quality $\tilde{q}$, the journal observes a noisy signal

$$s = \tilde{q} + \eta, \qquad \eta \sim \mathcal{N}(0, \sigma_s(S)^2),$$

[4] Neither the Cobb–Douglas form nor the fixed capacities R_0 and Z is essential for the main results. Appendix A generalizes quality production to an arbitrary homogeneous aggregator and lets researchers allocate a common time budget between research and leisure; the equilibrium effects of AI on submissions and mean quality are qualitatively unchanged.

[5] Noisy screening is standard in models of journal submission and evaluation: Azar (2015) model informed authors facing noisy review signals, while Bayar & Chemmanur (2021) and Card & DellaVigna (2020) provide theoretical and empirical foundations for treating editorial decisions as based on imperfect referee information.

where signal noise η is independent across papers and researchers, and independent of ε. We can interpret η as the noise in the referee screening process. Importantly, the variance of the signal noise, $\sigma_s(S)^2$, increases in aggregate submissions $S \geq 0$. In words, a higher volume of submissions congests the journal's screening capacity and produces noisier evaluations. One interpretation is that editors and referees have finite attention, so a paper is more likely to receive careful screening when the submission pool is small and more likely to be evaluated under rushed or thinner review when the pool is large. Ellison (2002*b*), Kovanis et al. (2016), and Hadavand, Hamermesh & Wilson (2024) document publication delays or peer-review labor constraints, and recent journal-level evidence in Gartenberg et al. (2026) links ChatGPT to higher submission volume and more AI-heavy manuscripts and reviews.[6]

We use the bounded screening-variance function

$$\sigma_s^2(S) = \sigma_0^2 + (\bar{\sigma}_s^2 - \sigma_0^2)(1 - e^{-\lambda S}),$$

where $\bar{\sigma}_s > \sigma_0 > 0$ and $\lambda > 0$. For compactness, let $\Delta_s \equiv \bar{\sigma}_s^2 - \sigma_0^2 > 0$. Thus $\sigma_s(0) = \sigma_0$, and screening noise approaches $\bar{\sigma}_s$ as the submission pool becomes large. The parameter λ governs how quickly the screening system moves from its low-congestion variance (σ_0^2) toward its overloaded variance (σ_s^2). Note that the exact functional form is not important for the qualitative results, as long as total signal noise is increasing and bounded.

The journal follows a threshold rule: upon receiving signal s, the journal accepts the paper if and only if $s \geq \bar{q}$, where $\bar{q}$ is a fixed publication threshold. In Section 5, we analyze the optimal choice of $\bar{q}$ for a publisher who cares about aggregate published knowledge.

Researcher's Problem. For an individual researcher, each published paper yields a private value $V > 0$, and each submitted paper costs $c \in (0, V)$.[7] The value V summarizes the career, reputational, and intellectual payoff from publication, a reduced form natural in a setting where journal placement has high private value for researchers; Heckman & Moktan (2020), for example, document the career relevance of top-journal placement in economics. We assume that, the value of an unpublished paper is zero. In the baseline, V is common across researchers; we explore heterogeneity in V in Section 4. The cost c captures the submission fee charged by the journal. In Section 5, we examine the optimal choice of c for the publisher.

[6] Gartenberg et al. (2026) also shows that AI adoption has increased in referee reports, but the reports are characterized by lower writing quality and less topical diversity than human-generated writing.

[7] The lower bound $c > 0$ keeps optimal submissions finite. The upper bound $c < V$ ensures that a certainly accepted paper would be privately worth submitting.

Given aggregate submissions S, an individual researcher solves

$$\max_{n \geq 0} \; U(n, S; a), \qquad U(n, S; a) = n \left[V \Phi(x(n, S; a)) - c \right], \tag{5}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function, and $x(n, S; a)$ is the standardized quality gap,

$$x(n, S; a) \equiv \frac{q(n, a) - \bar{q}}{\sigma_u(S)}. \tag{6}$$

Here the denominator $\sigma_u(S)$ denotes the total noise, combining intrinsic quality noise with screening noise

$$\sigma_u(S) = \sqrt{\sigma_\varepsilon^2 + \sigma_0^2 + (\bar{\sigma}_s^2 - \sigma_0^2)(1 - e^{-\lambda S})}.$$

$\Phi(x(n, S; a))$ is the per-paper acceptance probability when a researcher submits n papers given aggregate submissions S and AI intensity a. (6) implies that, when an individual researcher writes more papers, that is, increases n, her effort and attention are spread more thinly. This lowers the mean quality $q(n, a)$ and therefore reduces each paper's acceptance probability. By contrast, when other researchers write more papers, aggregate submissions S increase, congesting editorial screening and raising the noise $\sigma_u(S)$. This increases the acceptance probability of weak papers with $q(n, a) < \bar{q}$, while compressing the acceptance probabilities of strong papers with $q(n, a) > \bar{q}$.

Equilibrium Definition. We close the model by imposing consistency between individual submissions and aggregate congestion. With a unit mass of researchers, equilibrium requires individual optimality and aggregate consistency.

Definition 1 (Equilibrium) *An equilibrium at AI intensity a is an individual submission profile $\{n_i^*(a)\}_{i \in [0,1]}$ and aggregate submission level $S^*(a)$ such that:*

1. *Each researcher optimizes given the aggregate screening environment:*

$$n_i^*(a) \in \arg\max_{n_i \geq 0} U(n_i, S^*(a); a) \quad \textit{for every } i \in [0, 1].$$

2. *Aggregate submissions are consistent with individual choices:*

$$S^*(a) = \int_0^1 n_i^*(a) \, di.$$

Since researchers are identical in the baseline, as we will prove in the next section, equilibrium is symmetric, i.e. all researchers have the same submission level, $n_i^*(a) = n^*(a) \; \forall i$.

3 Equilibrium Characterization

We now characterize how the individual problem defined above becomes an equilibrium feedback loop. The section first solves the researcher's problem for a fixed screening environment, then imposes the symmetric fixed point, and finally studies how generative AI changes equilibrium quantity and quality.

3.1 Researcher's Problem

Fix aggregate submissions S and AI intensity a. A researcher takes the screening environment $\sigma_u(S)$ as given and chooses portfolio size n to maximize the payoff in (5). The first-order condition can be expressed in terms of the standardized quality gap alone. Differentiating the objective in (5) with respect to n gives

$$\frac{\partial U(n,S;a)}{\partial n} = V\Phi(x(n,S;a)) - c - \frac{VK}{\sigma_u(S)}\phi(x(n,S;a)).$$

Any interior optimum therefore has a standardized quality gap $x^{\text{br}}(S)$ satisfying

$$\Phi(x^{\text{br}}(S)) - \frac{K}{\sigma_u(S)}\phi(x^{\text{br}}(S)) = \frac{c}{V}. \tag{7}$$

The left-hand side of (7) decomposes the marginal effect of one more submission on the expected number of publications. The first term, $\Phi(x^{\text{br}}(S))$, is the direct gain: holding the per-paper acceptance probability fixed, an additional paper contributes one more chance of acceptance. The second term is the attention-dilution loss: the marginal submission spreads deep attention more thinly across the portfolio, lowering each existing paper's quality and hence its acceptance probability. At an interior optimum, the net marginal benefit equals the cost-benefit ratio c/V.

The best response has a useful two-step interpretation. The researcher first chooses a target standardized gap $x^{\text{br}}(S)$, or equivalently a target per-paper acceptance probability $\Phi(x^{\text{br}}(S))$. This target pins down the mean submitted quality

$$q^{\text{br}}(S;a) = \bar{q} + \sigma_u(S)x^{\text{br}}(S), \tag{8}$$

because reaching the target gap requires submissions whose mean quality is exactly $\sigma_u(S)x^{\text{br}}(S)$ above the screening bar $\bar{q}$. The Cobb–Douglas production technology $q = \Theta(a) - K\log n$ then determines the submission count that delivers this mean quality:

$$n^{\text{br}}(S;a) = \exp\left(\frac{\Theta(a) - q^{\text{br}}(S;a)}{K}\right) = \exp\left(\frac{\Theta(a) - \bar{q} - \sigma_u(S)x^{\text{br}}(S)}{K}\right). \tag{9}$$

A more ambitious target gap thus translates directly into higher mean quality and, through attention dilution, into a smaller portfolio.

Lemma 1 (Private best response) *Given any S and a, the researcher has a unique interior best response $n^{\mathrm{br}}(S;a)$, characterized by Equations (7) and (9). The following comparative statics hold:*

- *Holding S fixed, AI leaves best-response mean quality unchanged but raises the individual submission count:*
$$\frac{\partial q^{\mathrm{br}}(S;a)}{\partial a} = 0, \qquad \frac{\partial n^{\mathrm{br}}(S;a)}{\partial a} > 0.$$
- *Holding AI intensity a fixed, aggregate submissions move mean quality and the individual submission count in opposite directions, according to the cutoff*
$$\hat{x}(S) \equiv \frac{-\sigma_u(S) + \sqrt{\sigma_u(S)^2 + 4K^2}}{2K} > 0. \tag{10}$$
Specifically,
$$\frac{\partial q^{\mathrm{br}}(S;a)}{\partial S} \lesseqgtr 0 \quad \textit{and} \quad \frac{\partial n^{\mathrm{br}}(S;a)}{\partial S} \gtreqless 0 \quad \textit{if and only if} \quad x^{\mathrm{br}}(S) \lesseqgtr \hat{x}(S).$$

The proof is in Appendix G. The uniqueness statement ensures that the researcher's problem has a single interior solution at any (S,a). For n close to zero the marginal submission is valuable, since the cost-benefit assumption keeps the per-paper acceptance probability sufficiently large; for n large, attention dilution drives the per-paper acceptance probability toward zero and the marginal payoff toward $-c$. Between these two limits the first-order condition admits a unique solution, with a standardized quality gap $x^{\mathrm{br}}(S) > -\sigma_u(S)/K$ that places the optimum in the locally concave region of the payoff.

The response to AI intensity is direct. Generative AI raises routine research productivity and makes a larger individual submission count privately attractive. Notably, a does not enter the first-order condition (7), so the target standardized gap $x^{\mathrm{br}}(S)$ and the per-paper acceptance probability $\Phi(x^{\mathrm{br}}(S))$ are unaffected. AI therefore raises the researcher's individual submission count without lowering per-paper quality: the productivity gain in $\Theta(a)$ is fully absorbed by the additional attention dilution from a larger n, leaving best-response mean quality at $q^{\mathrm{br}}(S;a) = \bar{q} + \sigma_u(S)x^{\mathrm{br}}(S)$.

The response to aggregate submissions is more subtle. When S rises, screening becomes noisier. Holding the standardized gap fixed, a larger $\sigma_u(S)$ weakens the coefficient $K/\sigma_u(S)$ on the attention-dilution term in (7): a given reduction in mean quality has a smaller effect

on the acceptance probability when the signal is noisier. The first-order condition therefore rebalances at a smaller standardized gap $x^{\mathrm{br}}(S)$. Economically, the journal threshold disciplines the marginal paper less tightly, so the researcher is willing to operate with a lower per-paper acceptance probability.

This fall in the target gap pushes quantity upward, but it is not the only force. The absolute mean-quality target is $\bar{q} + \sigma_u(S)x^{\mathrm{br}}(S)$. A lower $x^{\mathrm{br}}(S)$ lowers this target and, through $q = \Theta(a) - K \log n$, requires a larger portfolio. At the same time, the increase in $\sigma_u(S)$ mechanically moves the absolute target away from $\bar{q}$ for any fixed standardized gap: it raises the target when $x^{\mathrm{br}}(S) > 0$ and lowers it when $x^{\mathrm{br}}(S) < 0$. Thus noisier screening directly reinforces quantity expansion when desired mean submitted quality is below the publication threshold, but works against the gap adjustment when desired mean submitted quality is above the threshold.

The cutoff $\hat{x}(S)$ is exactly the point at which these forces balance. If $x^{\mathrm{br}}(S) < \hat{x}(S)$, the gap adjustment dominates and the best response is *congestion-amplifying*: higher aggregate submissions lower best-response mean quality and raise the individual submission count. If $x^{\mathrm{br}}(S) > \hat{x}(S)$, the level effect dominates and the best response is *congestion-disciplining*: higher aggregate submissions raise best-response mean quality and reduce the individual submission count. The same screening congestion can therefore either amplify or discipline individual quantity choices, depending on where the privately optimal standardized gap lies relative to the cutoff.

Although the model admits both regimes, the congestion-amplifying case is the empirically relevant one. Public journal statistics report low acceptance rates in finance: the *Journal of Finance* reports a 5.6% acceptance rate for editorial decisions since July 2016; the *Review of Financial Studies* reports an 8.1% acceptance rate for 2025, following 6.0% rates in 2023 and 2024; and evidence from the *Journal of Financial Economics* places annual rejection rates in the 85%–91% range over active referee years.[8] These facts are only suggestive, but they point to an environment in which the average submitted paper is plausibly well below the publication threshold. In the model, that is the environment in which the desired quality gap $x^{\mathrm{br}}(S)$ is low and can lie below the positive cutoff $\hat{x}(S)$. Motivated by this case, the main text imposes a global version of the congestion-amplification condition.

[8] For the *Journal of Finance* statistics, see `https://afajof.org/journal-statistics/`. For the *Review of Financial Studies* statistics, see `https://sfs.org/rfs-statistics/`. The *Journal of Financial Economics* rejection-rate evidence is from Hrazdil et al. (2026).

Assumption 1 (Global congestion amplification) *The cost-benefit ratio satisfies*

$$\tau \equiv \frac{c}{V} < \Phi(\hat{x}(0)) - \frac{K}{\sigma_u(0)}\phi(\hat{x}(0)), \tag{11}$$

where $\sigma_u(\cdot)$ *and* $\hat{x}(\cdot)$ *are defined in* (2) *and* (10)*, respectively.*

Assumption 1 ensures global congestion amplification. To see this, fix S. Because the first-order condition is increasing for $x > -\sigma_u(S)/K$, Lemma 1 implies that $\partial n^{\text{br}}(S;a)/\partial S > 0$ if and only if

$$\tau < \Phi(\hat{x}(S)) - \frac{K}{\sigma_u(S)}\phi(\hat{x}(S)).$$

The right-hand side is minimized when total noise is lowest, at $S = 0$. Thus the weak cutoff obtained by replacing the strict inequality in Assumption 1 with a weak inequality is necessary and sufficient for

$$\frac{\partial n^{\text{br}}(S;a)}{\partial S} > 0$$

for every $S > 0$ and every $a \geq 0$. This sign does not depend on a: AI and other quality-level shifters move the level of the best response, but they do not change the effect of congestion that determines its slope with respect to S. We maintain Assumption 1 throughout the main text. The congestion-disciplining case, in which $\partial n^{\text{br}}(S;a)/\partial S < 0$, is analytically useful as a benchmark and is developed separately in Appendix B.

The next step is to characterize the fixed point. Each researcher treats S as fixed, but in a symmetric equilibrium all researchers' portfolio choices determine S.

3.2 Feedback Loop

We now turn the best response of individual researchers in Lemma 1 into an equilibrium feedback loop. By Definition 1, a symmetric equilibrium at AI intensity a satisfies the fixed-point condition:

$$n^*(a) = n^{\text{br}}(n^*(a);a). \tag{12}$$

It is useful to rewrite the best response function in Equation (9) as

$$n^{\text{br}}(S;a) = C(a)B(S),$$

where $C(a) = \exp((\Theta(a) - \bar{q})/K)$ and $B(S) = \exp\big(-\sigma_u(S)x^{\text{br}}(S)/K\big)$. The term $C(a)$ is the direct productivity shifter, while $B(S)$ is the congestion component of the best response. This fixed point is where the screening externality enters. When a rises, Lemma 1 implies that each

researcher wants to submit more papers at the old screening environment. In a symmetric equilibrium, this direct increase in the desired individual submission count raises aggregate submissions. The larger submission pool increases total screening noise $\sigma_u(S)$, and under Assumption 1, the noisier screening environment raises the best response again. This is because congestion makes the publication signal noisier, tilting incentives toward quantity and against quality.

To measure the local strength of this loop, define the feedback index at a fixed AI intensity a by

$$J(S) \equiv \frac{\partial \log n^{\mathrm{br}}(S;a)}{\partial \log S} = \frac{S}{n^{\mathrm{br}}(S;a)} \frac{\partial n^{\mathrm{br}}(S;a)}{\partial S} = S \frac{B'(S)}{B(S)}. \tag{13}$$

The index is the elasticity of desired individual submissions with respect to aggregate submissions, holding the direct AI shifter fixed. Economically, it measures how strongly the screening environment created by aggregate submissions feeds back into an individual researcher's desired submissions. At a symmetric equilibrium, where $n^{\mathrm{br}}(n^*(a);a) = n^*(a)$, the elasticity coincides with the marginal response of n^{br} to aggregate submissions:

$$\left. \frac{\partial n^{\mathrm{br}}(S;a)}{\partial S} \right|_{S=n^*(a)} = J(n^*(a)).$$

Thus the sign of $J(n^*(a))$ records whether screening congestion reinforces or offsets the direct AI shift. Because the main text maintains Assumption 1, Lemma 1 implies $J(n^*(a)) > 0$ at the equilibria studied here. We focus on simple locally stable symmetric equilibria. Under the standard one-dimensional adjustment in which aggregate submissions move in the direction of the excess-submission gap $n^{\mathrm{br}}(S;a) - S$, local stability requires the best-response curve to be flatter than the 45-degree line at the fixed point; equivalently, $J(n^*(a)) < 1$. Appendix F records the corresponding adjustment calculation and the unstable case $J(n^*(a)) > 1$. Thus the stable congestion-amplifying equilibria in the main text satisfy $0 < J(n^*(a)) < 1$.

The equilibrium effect of generative AI follows by differentiating the fixed point in (12). The next proposition states the local multiplier form.

Proposition 1 *In a stable equilibrium with $0 < J(n^*(a)) < 1$, we have the equilibrium effect of AI intensity:*

$$\frac{dn^*(a)}{da} = \left. \frac{\partial n^{\mathrm{br}}(S;a)}{\partial a} \right|_{S=n^*(a)} \frac{1}{1 - J(n^*(a))} > \left. \frac{\partial n^{\mathrm{br}}(S;a)}{\partial a} \right|_{S=n^*(a)} > 0. \tag{14}$$

The first term in (14) is the direct private response from Lemma 1: AI raises routine capacity and shifts the desired individual submission count upward. The factor $1/(1-J)$ is the

equilibrium feedback multiplier. Each direct increase in the individual submission count raises aggregate submissions and worsens screening precision; in the congestion-amplifying region this induces another round of submissions. This positive feedback makes the equilibrium response larger than the partial-equilibrium response. If $J < 0$, congestion disciplines quantity and the same formula attenuates the direct AI effect; Appendix B develops that outside-trap benchmark.

3.3 Quantity Trap

Proposition 1 explains how the feedback effect amplifies researchers' equilibrium response to changes in AI intensity. We now show that when the feedback effect is strong enough, it can lead to an AI-induced quantity trap.

Proposition 2 *There exists a cutoff $\tau^\star$ such that if $\tau \equiv c/V < \tau^\star$, then there exist $a_L(\tau) < a_H(\tau)$ such that the model has three equilibria if*

$$a \in (a_L(\tau), a_H(\tau)) \cap [0, \infty),$$

and a unique positive symmetric equilibrium otherwise.

The proposition shows that multiplicity arises when the private cost of submission is sufficiently low relative to the value of publication.[9] In such cases, researchers's incentive to scale up submission is high enough that screening noise substantially weakens publication discipline. As illustrated in Figure 2, an increase in a can shift the best-response curve into the folded part of the geometry.[10] At low AI intensity, the shifted curve crosses the 45-degree line only once on the low-submission side. At intermediate AI intensity, it crosses three times: the low and high crossings are locally stable, and the middle crossing is unstable because the best-response curve is locally steeper than the 45-degree line. At high AI intensity, the low-submission crossings disappear and only the high-submission equilibrium remains.

This multiplicity reflects a self-reinforcing dynamic: a large submission pool makes screening noisier, and noisier screening makes a large individual count privately attractive, sustaining the very congestion that motivated it. We call this locally stable, high-submission equilibrium

[9] At admissible endpoints $a = a_L(\tau)$ or $a = a_H(\tau)$, two positive equilibria remain, one with $J = 1$.

[10] The figure plots best-response curves against the 45-degree line for three AI intensities. The figure uses the following parameters: $\theta = -0.200$, $\beta_R = \beta_Z = 0.04$, $R_0 = Z = V = 1$, $\sigma_\varepsilon = 0.20$, $\sigma_0 = 0.05$, $\bar{\sigma}_s = 0.30$, and $\lambda = 0.20$. The fixed policy, $\bar{q} = 0.091805$, and $c = 0.049846$, is chosen such that the equilibrium submission $S = 1$ when $a = 0$.

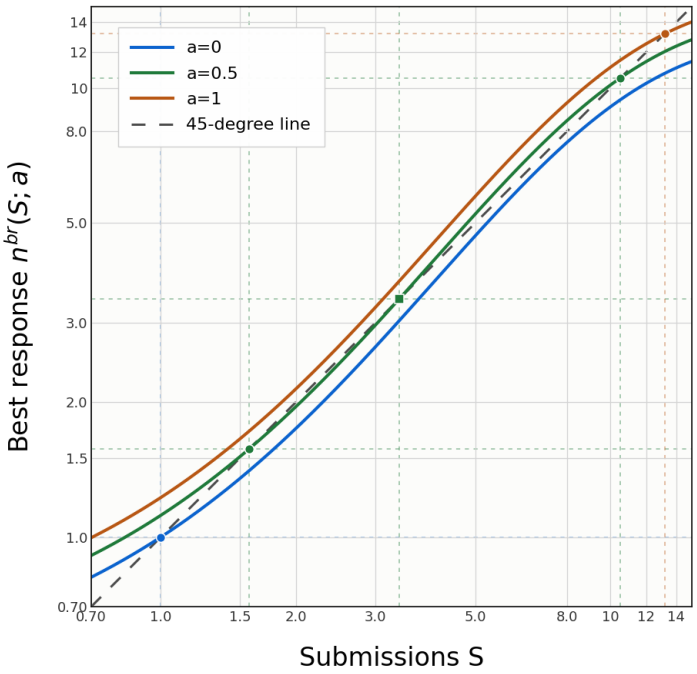


Figure 2: Best-response functions and fixed-point crossings

the *quantity trap*. The trap is an equilibrium property—not an assumption that researchers make mistakes or that AI directly lowers paper quality.

The next result records the quality implication. It combines the cross-equilibrium quality ranking from the production function in Equation (1) with the feedback logic under $0 < J(n^*(a)) < 1$ from Proposition 1.

Proposition 3 (Quantity-trap quality) *Suppose three symmetric equilibria co-exist at AI intensity a with equilibrium submission levels $n_L < n_M < n_H$. Then their mean qualities satisfy*

$$q(n_H, a) < q(n_M, a) < q(n_L, a).$$

Moreover, in a stable equilibrium with $0 < J(n^*(a)) < 1$, as AI intensity rises, equilibrium submissions increase and mean quality decreases: $dn^*(a)/da > 0$ and $dq^*(a)/da < 0$.

The quality ranking across simultaneous equilibria is immediate from Equation (1): holding a fixed, higher submissions dilute routine effort and deep attention, so average quality is decreasing in n. The local comparative static is different. At fixed n, AI raises paper quality by Equation (2). In equilibrium, however, the increase in a raises submissions. In the stable

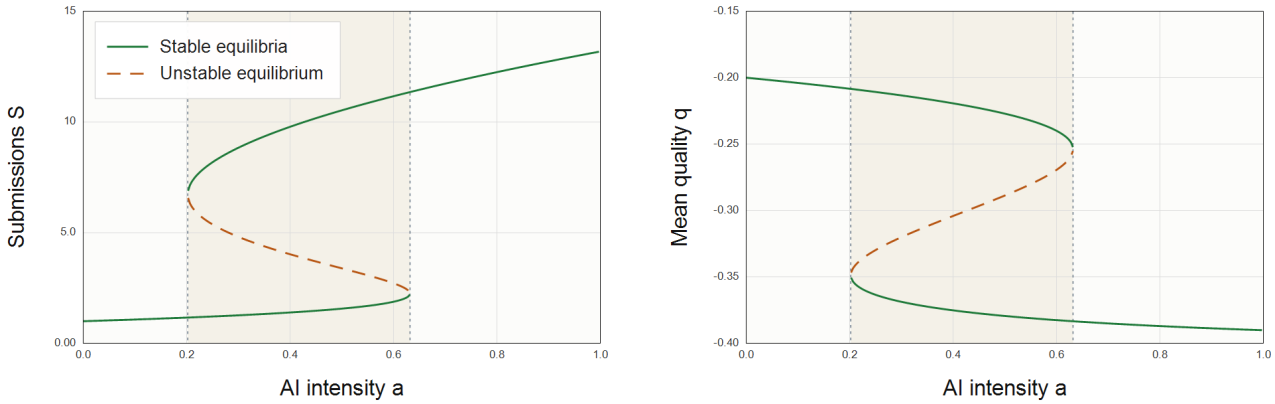


Figure 3: Equilibrium submissions and mean quality under the quantity trap

congestion-amplifying region, the induced increase in submissions is amplified by the screening feedback. The attention-dilution then more than offsets the direct productivity gain, giving the negative derivative of equilibrium quality with respect to a.

Figure 3 traces all three equilibrium branches over $0 \leq a \leq 1$, using the calibration from Figure 2, with the multiplicity window shaded. The left panel shows equilibrium submission levels and the right panel the associated mean quality. The middle branch is unstable; the low- and high-submission branches are stable. Along both stable branches, higher AI intensity expands submissions and compresses mean quality. The pattern captures the central trade-off: AI raises each researcher's productivity at a fixed submission count, but the equilibrium response is to submit more—and in the quantity trap, the congestion from that additional volume more than offsets the direct quality gain. Quantity rises, but average quality falls.

3.4 Aggregate Published Knowledge

Proposition 3 establishes that the quantity trap lowers mean paper quality, but it also brings more submissions—and more submissions open more chances for publication. Whether this volume effect compensates for the quality decline is a separate and important question. We address it by turning from mean quality to *aggregate published quality*: the total knowledge that clears the journal's acceptance threshold and enters the published record.

Aggregate published quality is the total true quality of papers that clear the journal's acceptance threshold:

$$A^{\text{pub}}(a) \equiv n^*(a)\,\mathbb{E}[\tilde{q}\,\mathbf{1}\{s \geq \bar{q}\}] = n^*(a)\,p(q^*(a), \sigma_u(n^*(a)))\,, \tag{15}$$

where $p(q, \sigma_u)$ is the per-submission true published quality:

$$p(q, \sigma_u) \equiv q\Phi\left(\frac{q-\bar{q}}{\sigma_u}\right) + \frac{\sigma_\varepsilon^2}{\sigma_u}\phi\left(\frac{q-\bar{q}}{\sigma_u}\right). \tag{16}$$

The first term is the expected true quality of a published paper: mean quality q weighted by the acceptance probability $\Phi((q-\bar{q})/\sigma_u)$. The second term is the selection premium—the additional true quality captured by conditioning on a favorable signal. The premium shrinks as screening noise σ_ε^2 rises relative to σ_u.

$A^{\text{pub}}(a)$ is a natural output metric for a journal or research field, though not a complete welfare measure: it omits review costs, editorial costs, and author opportunity costs.

For the comparative static, let p_q and p_{σ_u} denote the partial derivatives of p evaluated at $(q^*(a), \sigma_u(n^*(a)))$, and write

$$\sigma_S \equiv \left.\frac{d\sigma_u(S)}{dS}\right|_{S=n^*(a)} = \frac{\lambda\Delta_s e^{-\lambda n^*(a)}}{2\sigma_u(n^*(a))} > 0 \tag{17}$$

for the marginal increase in screening noise from one additional submission. The next result separates the three margins.

Proposition 4 (Aggregate published-quality sign condition) *If $n^*(a)$ is differentiable, is a simple symmetric equilibrium for each nearby AI intensity, and satisfies $J(n^*) < 1$,*

$$\frac{dA^{\text{pub}}(a)}{da} = \frac{\beta_R\gamma}{1+\gamma a}\frac{n^*}{1-J(n^*)}\left[\frac{p}{K} - J(n^*)p_q + \frac{n^*\sigma_S}{K}p_{\sigma_u}\right]. \tag{18}$$

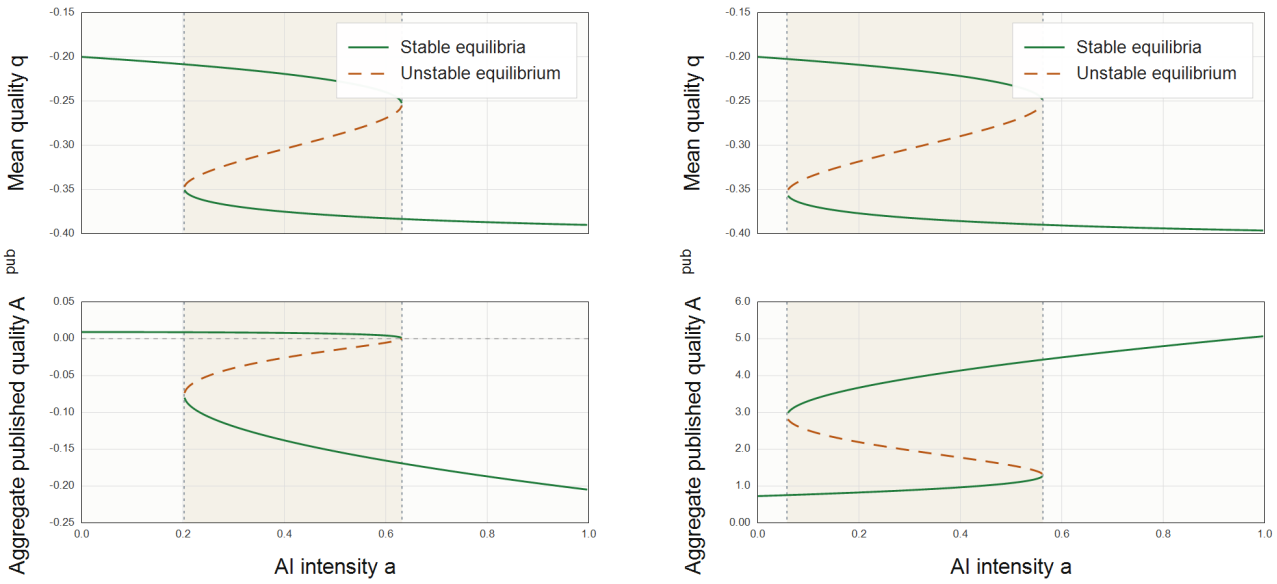
Therefore,

$$\operatorname{sign}\left(\frac{dA^{\text{pub}}(a)}{da}\right) = \operatorname{sign}\left(\frac{p}{K} - J(n^*)p_q + \frac{n^*\sigma_S}{K}p_{\sigma_u}\right). \tag{19}$$

In particular, aggregate published quality falls with AI if and only if

$$J(n^*)p_q - \frac{n^*\sigma_S}{K}p_{\sigma_u} > \frac{p}{K}. \tag{20}$$

Proposition 4 is intentionally conditional. The term p/K is the extensive-margin scale term from the additional submissions induced by AI. The term $-J(n^*)p_q$ is the mean-quality margin: in the stable quantity-trap region, $J(n^*) > 0$, so the decline in q^* lowers aggregate published quality whenever higher submitted-paper quality raises per-submission published quality. The term $(n^*\sigma_S/K)p_{\sigma_u}$ is the screening-congestion margin. When $p_{\sigma_u} < 0$, noisier screening weakens the true-quality content of publication selection and pushes aggregate published quality



(a) Baseline calibration

(b) Scale-dominant variant

Figure 4: Aggregate published quality under alternative sign-condition outcomes

down. The quantity trap therefore guarantees the direction of mean quality, not the direction of aggregate published quality. Total true published quality falls only when the quality-dilution and screening-congestion losses dominate the scale gain, as in (20).

Figure 4 illustrates the conditional nature of the result. Both subfigures use the same stable and unstable equilibrium curves as Figure 3. The left subfigure uses the baseline variance-saturation calibration from the earlier figures. There, aggregate published quality declines along the stable equilibrium curves because lower mean quality and noisier selection dominate the additional scale. The right subfigure shows a calibration in which aggregate published quality rises instead. The high-submission equilibrium still has lower mean quality, but the extra scale is strong enough that $A^{\text{pub}}(a)$ increases.

Thus the paper's sharp unconditional claim is the mean-quality result in Proposition 3. The stronger claim that the publication system receives less true quality despite processing more papers requires the sign condition in (20). If that condition holds, adding review and editorial costs would only strengthen the welfare concern, since those costs rise with the submission pool. The main-text result, however, remains a statement about aggregate true published quality rather than full social welfare. The next section turns from the representative-agent result to the types of researchers and projects most likely to drive this aggregate-quality tradeoff.

4 Researcher Heterogeneity

Section 3 studied a representative researcher. This section shows how the same congestion feedback becomes a composition result when researchers are heterogeneous. The goal is not to re-solve the individual problem from scratch. Instead, we take the best-response objects from Section 3, index them by type, and ask which primitives change the sign of the congestion response. The main message is that heterogeneity matters for quality responses only when it changes type-specific congestion slopes; the clean primitive sources below are private value-cost ratios and own-quality dilution. Pure quality-level heterogeneity changes how much different researchers submit, but it does not by itself determine whose response is congestion amplifying.

4.1 Heterogeneous Equilibrium and Type-Specific Feedback

The heterogeneous setup follows the baseline model naturally. Instead of a unit mass of identical researchers, consider a finite set of active researcher types $i = 1, \dots, N$. Type i has mass $\mu_i > 0$ and chooses per-researcher submissions n_i. Aggregate submissions are

$$S = \sum_{i=1}^{N} \mu_i n_i.$$

The publication screen is common across types, with total signal noise $\sigma_u(S)$ as in the baseline model. Type i has mean quality

$$q_i(n_i, a) = \Theta_i(a) - K_i \log n_i, \qquad K_i \equiv \beta_{Ri} + \beta_{Zi},$$

where $\Theta_i(a)$ collects type-specific quality-level shifters such as baseline quality, routine capacity, AI exposure, and deep-attention capacity. Type i receives private value V_i from an accepted paper and pays submission cost c_i. Let

$$\tau_i \equiv \frac{c_i}{V_i}, \qquad 0 < \tau_i < 1.$$

Definition 2 (Heterogeneous equilibrium) *A heterogeneous equilibrium at AI intensity a is a vector of type-level submissions $\{n_i^*(a)\}_{i=1}^N$ and an aggregate submission level $S^*(a)$ such that:*

1. *Each type optimizes given the aggregate screening environment:*

$$n_i^*(a) \in \arg\max_{n_i \geq 0} \; n_i \left[V_i \Phi\left(\frac{q_i(n_i, a) - \bar{q}}{\sigma_u(S^*(a))} \right) - c_i \right] \quad \text{for every } i.$$

2. *Aggregate submissions are consistent with type-level choices:*

$$S^*(a) = \sum_{i=1}^{N} \mu_i n_i^*(a).$$

This is the finite-type analogue of Definition 1: the individual optimality condition is type-specific, and the aggregate consistency condition replaces the baseline integral with a mass-weighted sum over types. Taking aggregate submissions as given, each type's first-order condition is the type-indexed version of Equation (7). Let $x_i^{\mathrm{br}}(S)$ denote the relevant interior solution to

$$\Phi(x_i^{\mathrm{br}}(S)) - \frac{K_i}{\sigma_u(S)}\phi(x_i^{\mathrm{br}}(S)) = \tau_i. \tag{21}$$

Then type i's best response factorizes as

$$n_i^{\mathrm{br}}(S;a) = C_i(a)B_i(S), \qquad C_i(a) \equiv \exp\left(\frac{\Theta_i(a) - \bar{q}}{K_i}\right), \qquad B_i(S) \equiv \exp\left(-\frac{\sigma_u(S)x_i^{\mathrm{br}}(S)}{K_i}\right). \tag{22}$$

As in the baseline factorization $n^{\mathrm{br}}(S;a) = C(a)B(S)$, $C_i(a)$ is the direct productivity shifter and $B_i(S)$ is the congestion component. Parallel to the baseline feedback index $J(S) = SB'(S)/B(S)$, define the type-specific congestion elasticity

$$J_i(S) \equiv \frac{\partial \log n_i^{br}(S;a)}{\partial \log S} = S\frac{B_i'(S)}{B_i(S)}.$$

The underlying partial derivative of type i's best response with respect to aggregate submissions is

$$\frac{\partial n_i^{\mathrm{br}}(S;a)}{\partial S} = \frac{n_i^{\mathrm{br}}(S;a)}{S}J_i(S).$$

Thus $J_i(S)$ has the same sign as type i's best-response slope with respect to aggregate congestion. Appendix C records the closed-form expression for $J_i(S)$ under the variance-saturation screening technology. When $J_i(S) > 0$, higher aggregate submissions make screening noisier in a way that raises type i's best response; type i is congestion amplifying at S. When $J_i(S) < 0$, noisier screening lowers type i's best response; type i is congestion disciplining at S. This terminology mirrors Section 3: the quantity trap is an equilibrium outcome, whereas $J_i(S)$ is a local type-level elasticity.

For an interior heterogeneous equilibrium, Definition 2 and the best-response factorization imply

$$n_i^* = C_i(a)B_i(S^*) \quad \text{for every } i, \qquad S^* = \sum_{i=1}^{N} \mu_i C_i(a)B_i(S^*). \tag{23}$$

Define type i's equilibrium submission share and the aggregate stability index by

$$\omega_i \equiv \frac{\mu_i n_i^*}{S^*}, \qquad \Omega(S^*) \equiv \sum_{i=1}^{N} \omega_i J_i(S^*).$$

As in the representative-agent model, the aggregate fixed point is locally stable when the aggregate best-response curve is flatter than the 45-degree line. Here the derivative of the aggregate best-response map is the share-weighted average Ω, so a simple locally stable heterogeneous equilibrium satisfies

$$\Omega(S^*) < 1.$$

We focus on stable heterogeneous equilibria satisfying this condition.

The comparative statics also preserve the Section 3 structure. Changes in a move type i's best response directly through $C_i(a)$ and indirectly through aggregate congestion. To avoid confusion with the primitive routine-AI exposure parameter γ in $R(a)$, define the type-specific direct shifter response by

$$\chi_i(a) \equiv \frac{\partial \log C_i(a)}{\partial a} = \frac{\Theta_i'(a)}{K_i}, \qquad \bar{\chi}(a) \equiv \sum_{i=1}^{N} \omega_i \chi_i(a).$$

At any differentiable heterogeneous equilibrium satisfying $\Omega(S^*) < 1$,

$$\frac{dS^*}{da} = S^* \frac{\bar{\chi}(a)}{1 - \Omega(S^*)},$$

and type-level quantities satisfy

$$\frac{dn_i^*}{da} = n_i^* \left[\chi_i(a) + J_i(S^*) \frac{\bar{\chi}(a)}{1 - \Omega(S^*)} \right]. \tag{24}$$

Mean quality for type i, $q_i^*(a) \equiv q_i(n_i^*(a), a)$, satisfies

$$\frac{dq_i^*(a)}{da} = -K_i J_i(S^*) \frac{\bar{\chi}(a)}{1 - \Omega(S^*)}. \tag{25}$$

Appendix C derives these comparative-static formulas. Thus the sign of the type-level quality response is governed by the type-specific congestion elasticity. If AI raises aggregate submissions, the formula has a simple sign implication. Types with $J_i(S^*) > 0$ have falling mean quality, while types with $J_i(S^*) < 0$ have rising mean quality.

The remaining subsections use this machinery one heterogeneity dimension at a time. This separation keeps the economic source of slope heterogeneity transparent. It also marks where

the heterogeneous analysis departs from the maintained representative-researcher restriction in Assumption 1. That assumption imposes global congestion amplification in Section 3.1; here we do not impose it on every type. The heterogeneous economy may therefore contain both congestion-amplifying and congestion-disciplining type responses at the same aggregate state.

4.2 Pure Quality-Level Heterogeneity

We begin with a benchmark: researchers who differ only in quality levels. Examples include type-specific baseline quality θ_i, routine capacity R_{0i}, deep-attention stock Z_i, or AI exposure that enters only through $\Theta_i(a)$. These quality-level parameters do not enter the screening first-order condition (21). Hence, when K_i and τ_i are common, the optimal standardized quality gap $x_i^{\mathrm{br}}(S)$ is the same for all groups. The parameters enter the best response in (22) only through the level shifter $C_i(a)$, which is pinned down by $\Theta_i(a)$. They can make some researchers submit more than others, but they do not, by themselves, make one type congestion amplifying and another type congestion disciplining.

Lemma 2 (Pure quality-level heterogeneity) *Suppose $K_i = K$ and $\tau_i = \tau$ for all active types. If heterogeneity enters only through the quality-level shifter $\Theta_i(a)$, then for every aggregate submission level S,*

$$x_i^{\mathrm{br}}(S) = x^{\mathrm{br}}(S), \qquad B_i(S) = B(S) \qquad \text{and} \qquad J_i(S) = J(S) \quad \text{for all } i.$$

At any interior heterogeneous equilibrium, all active types have the same mean quality $q_i^*(a) = q^*(a)$, and $n_i^*(a) > n_j^*(a)$ if and only if $\Theta_i(a) > \Theta_j(a)$. Hence all active types have the same local congestion classification at S. If the equilibrium is stable, $\chi_i(a) \geq 0$ for all active types, $\bar{\chi}(a) > 0$, and $J(S^*) > 0$, then $\frac{dq_i^*(a)}{da} < 0$ and $\frac{dn_i^*(a)}{da} > 0$ for every active type.

The reason is visible in the first-order condition for the standardized quality gap. The level shifter $\Theta_i(a)$ affects the mapping from the optimal standardized quality gap back to the number of submissions, but it does not affect the quality gap $x_i^{\mathrm{br}}(S)$ itself when K and τ are common. It therefore shifts $C_i(a)$ without changing the congestion shape $B_i(S)$. High-quality types submit more because their best-response curves are higher, not because they choose a higher submitted-paper quality in equilibrium. The common screening condition pins all groups to the same standardized gap, so the group with the higher quality level adjusts on the quantity margin until its submitted mean quality equals the common equilibrium quality.

The comparative-static implication is also common across groups. When $J(S^*) > 0$, the shared congestion response is amplifying: a higher aggregate submission level makes the noisy

screen sufficiently less selective that each type wants to submit more. If AI raises the direct level shifter on average and does not lower any type's direct shifter, the stable-equilibrium formulas from Section 4.1 imply that submissions rise for every group while the common submitted mean quality falls. The quality decline is therefore not a sorting result across types. It is the same feedback force as in the homogeneous benchmark, with quality-level heterogeneity changing who submits more, but not which types are exposed to congestion amplification.

This benchmark disciplines the rest of the section. Cross-sectional differences in submission levels are not enough to identify the congestion-amplification mechanism. What matters is whether a primitive changes the slope of the best response with respect to screening congestion. The two clean primitives below are the private value-cost ratio and the strength of own-quality dilution.

4.3 Heterogeneity in Publication Value Relative to Cost

The most direct source of slope heterogeneity is the private value of publication relative to submission cost. To isolate this force, suppose production primitives are common across types, so $K_i = K$ and $\Theta_i(a) = \Theta(a)$, while $\tau_i = c_i/V_i$ may differ. Using the cutoff $\hat{x}(S)$ from (10), define the value-cost cutoff

$$H(S) \equiv \Phi(\hat{x}(S)) - \frac{K}{\sigma_u(S)}\phi(\hat{x}(S)). \tag{26}$$

The cutoff $H(S)$ translates the best-response slope criterion in Section 3.1 into a condition on the type-specific cost-benefit ratio. For the common-production value-cost environment,

$$J_i(S) > 0 \iff \tau_i < H(S), \qquad J_i(S) < 0 \iff \tau_i > H(S).$$

Thus lower-τ_i types are more likely to be congestion amplifying. When submission costs are common, higher publication values lower $\tau_i = c/V_i$ and therefore move a type toward the congestion-amplifying side of the cutoff.

Lemma 3 (Cross-sectional sorting by normalized cost) *Suppose active types share common production primitives. They differ only in normalized costs.*

If $\tau_i > \tau_j$, then at any interior heterogeneous equilibrium,

$$n_i^* < n_j^* \qquad \text{and} \qquad q_i^* > q_j^*.$$

More precisely,

$$q_i^* - q_j^* = K \log\left(\frac{n_j^*}{n_i^*}\right) > 0.$$

The proof is in Appendix G. The lemma gives the static composition effect of value-cost heterogeneity. A lower normalized cost means that an additional acceptance is privately more valuable relative to the cost of submitting. At the same aggregate congestion level, the lower-τ type optimally accepts a lower standardized quality gap and uses the production technology to generate a larger portfolio. Because production primitives are common, that larger portfolio comes with lower mean paper quality. The masses μ_i determine how much each type contributes to aggregate submissions, but they do not change the per-researcher ordering in the lemma.

Proposition 5 (Value-cost composition under AI) *Suppose active types share common production primitives and differ only in normalized costs. Consider a differentiable interior heterogeneous equilibrium satisfying $\Omega(S^*) < 1$. Then:*

1. *If $J_i(S^*) > 0$, so type i is congestion amplifying, then*

$$\frac{dq_i^*}{da} < 0 \qquad \text{and} \qquad \frac{dn_i^*}{da} > 0.$$

2. *If $J_i(S^*) < 0$, so type i is congestion disciplining, then*

$$\frac{dq_i^*}{da} > 0.$$

   Moreover, $\frac{dn_i^*}{da} < 0$ if and only if $J_i(S^*)$ is sufficiently negative, with the exact N-type threshold in equation (G3) of Appendix G.

3. *Cross-sectional gaps widen with AI: if $\tau_i > \tau_j$, then*

$$\frac{dn_i^*}{da} < \frac{dn_j^*}{da} \qquad \text{and} \qquad \frac{dq_i^*}{da} > \frac{dq_j^*}{da}.$$

The proof is in Appendix G. The sign cases clarify what the congestion elasticity means economically. For a type with $J_i(S^*) > 0$, additional screening noise makes marginal submissions more attractive. The indirect congestion response therefore reinforces the direct AI effect. Its portfolio expands and its mean paper quality falls. For a type with $J_i(S^*) < 0$, noisier screening discourages quantity and raises mean quality by pushing the type toward a smaller, higher-quality portfolio. Whether its submissions actually fall depends on whether this negative congestion effect is strong enough to dominate the common direct AI shifter. In the N-type economy, that threshold depends on the whole equilibrium composition through $\Omega(S^*)$ and the submission-share weights, not on a single comparison with one other type.

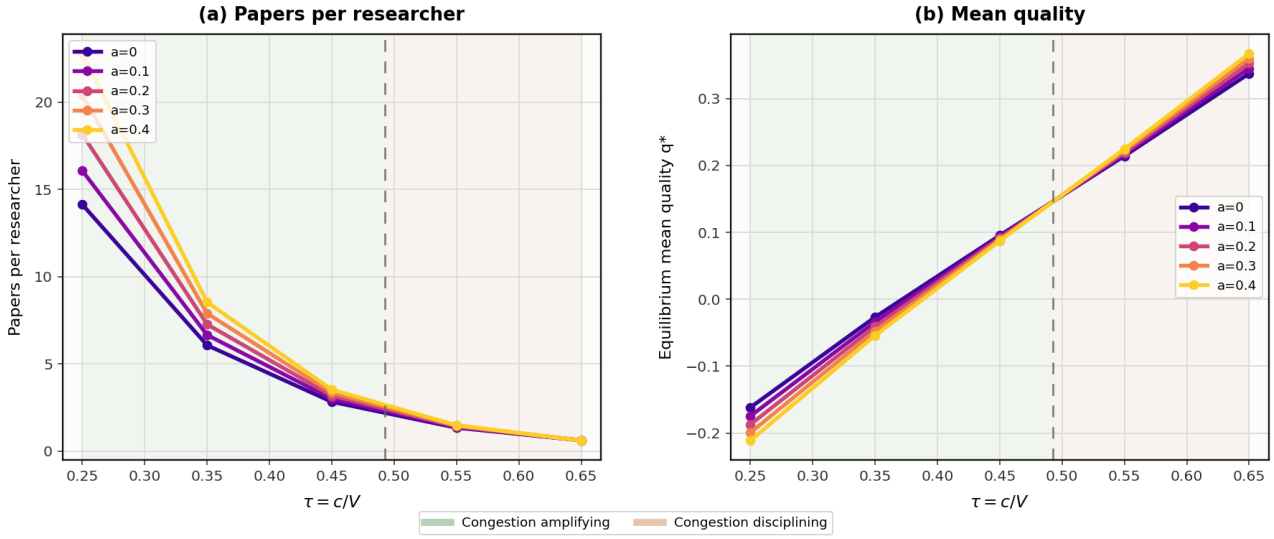


Figure 5: Submissions and mean quality across normalized-cost groups

The proposition also turns the static sorting in the lemma into a comparative-static statement. Since all types have the same direct AI shifter in this subsection, differences in their AI responses come from the induced change in aggregate congestion. Lower-τ types have larger congestion slopes: they are more willing to respond to a noisier screen by expanding quantity. As AI raises the common productivity shifter, the quantity gap between low- and high-τ types therefore widens, and the quality gap widens in the opposite direction.

Figure 5 illustrates these two implications in a five-type calibration along the same stable equilibrium branch. The horizontal axis is $\tau = c/V$. The vertical cutoff separates types whose local response to congestion is amplifying from types whose response is disciplining. As in Lemma 3, lower-τ groups submit more papers per researcher and have lower mean submitted-paper quality, while higher-τ groups submit fewer papers and have higher mean quality. As in Proposition 5, an increase in AI widens the cross-sectional quantity-quality tradeoff: the low-τ groups expand more strongly and experience lower mean quality, whereas the high-τ groups are disciplined by the noisier screen and move toward higher mean quality.

The normalized cost τ has several empirical interpretations. A high-τ group can have high submission costs, low private value of publication, or both. With common costs, pre-tenure researchers are naturally represented as lower-τ types because an additional publication has high career value; senior tenured researchers may be closer to high-τ types because the private career value of another publication is lower. The model then predicts that pre-tenure researchers

write more papers, but that their larger portfolios dilute mean paper quality when production primitives are otherwise similar. The same logic applies to the cost side. Lower c can stand for abundant research funding, research assistance, data access, or computing support that makes additional projects cheaper to carry forward. Holding the quality technology fixed, such support moves researchers toward the low-τ region: it raises submissions and can lower mean quality per paper through the same portfolio-dilution force.

These results concern submitted quantity and submitted mean quality. The corresponding published quantity and published quality are in general not monotone in τ. A lower-τ type submits more, but its papers also have lower mean quality and lower acceptance probabilities, so the number of accepted papers per researcher and the true-quality content of published work need not be ordered monotonically in τ. Their AI responses are also ambiguous because AI changes three margins at once: direct submission scale, congestion-driven portfolio adjustment, and screening-noise erosion of acceptance probabilities and published-paper quality. When screening noise rises rapidly with AI through the induced increase in aggregate submissions, the erosion margin becomes stronger, especially for low-τ types whose portfolios expand most. In that case AI can still widen submitted-paper gaps while narrowing, flattening, or even reversing cross-sectional gaps in published-paper counts or true published quality.

4.4 Heterogeneity in Deep-Attention Sensitivity

A second source of slope heterogeneity comes from the production side. Routine tasks are relatively standardized: drafting, coding, checking references, and producing variants of an idea are likely to improve paper quality in similar ways across researchers. It is therefore natural to keep the routine elasticity β_R common. Deep attention is different. Researchers can differ substantially in how much scarce thought improves a paper's question, argument, identification, or interpretation. We capture this difference by allowing deep-attention sensitivity β_{Zi} to vary while payoff primitives are common, so $\tau_i = \tau = c/V$. Then

$$K_i = \beta_R + \beta_{Zi}.$$

Following the production function in (1), this heterogeneity also enters the quality-level shifter through

$$\Theta_i(a) = \theta + \beta_R \log R(a) + \beta_{Zi} \log Z.$$

Thus β_{Zi} is not only a level parameter. A higher β_{Zi} raises the return to deep attention at a fixed portfolio size, but it also raises K_i, the rate at which quality falls when attention is spread across more papers.

The type-specific analogue of the cutoff $\hat{x}(S)$ in (10) is

$$\hat{x}_i(S) \equiv \frac{-\sigma_u(S) + \sqrt{\sigma_u(S)^2 + 4K_i^2}}{2K_i}. \tag{27}$$

Using the same translation as the value-cost cutoff $H(S)$ in (26), define

$$H_i(S) \equiv \Phi(\hat{x}_i(S)) - \frac{K_i}{\sigma_u(S)}\phi(\hat{x}_i(S)). \tag{28}$$

The cutoff is now type-specific because K_i changes the strength of own-quantity discipline.

Lemma 4 (Cross-sectional sorting by deep-attention sensitivity) *Suppose active types share all primitives except β_{Zi}. Types differ only in β_{Zi}. If $\beta_{Zi} > \beta_{Zj}$, then at any interior heterogeneous equilibrium,*

$$q_i^* > q_j^*.$$

In general, no ordering of n_i^* and n_j^* follows from $\beta_{Zi} > \beta_{Zj}$ alone.

The proof is in Appendix G. The lemma gives the static cross-sectional implication of deep-attention heterogeneity. At a common aggregate screening environment, a higher β_Z type chooses a higher standardized quality gap and therefore a higher mean submitted quality. The quantity comparison is different. A higher β_Z raises baseline quality through the return to deep attention, which tends to support more submissions. At the same time, it raises the dilution factor K_i, which makes additional submissions more costly in quality terms and tends to discipline quantity. This is why deep-attention heterogeneity combines the level force from Section 4.2 with the incentive force from Section 4.3: it changes both where a type's best response sits and how that best response reacts to congestion.

Using the cutoff $H_i(S)$, the type-specific slope criterion is

$$J_i(S) > 0 \iff \tau < H_i(S), \qquad J_i(S) < 0 \iff \tau > H_i(S).$$

This is the deep-attention analogue of the value-cost cutoff in Section 4.3, but the source of heterogeneity is different. In Section 4.3, the cutoff $H(S)$ is common and types differ in their normalized cost τ_i. Here the normalized cost τ is common, while the cutoff itself is type-specific: $H_i(S)$ varies with K_i. For fixed S, $H_i(S)$ is decreasing in K_i. Hence lower-β_Z types are more likely to be congestion amplifying. The reason is that a lower β_Z weakens own-quantity discipline: additional submissions do less damage to the type's own quality margin, making the type more willing to expand when screening becomes noisier.

Proposition 6 (Deep-attention sorting) *Suppose active types share all primitives except β_{Zi}. Types differ only in β_{Zi}, with $K_i = \beta_R + \beta_{Zi}$. Consider a differentiable interior heterogeneous equilibrium satisfying $\Omega(S^*) < 1$. Then:*

1. *If $J_i(S^*) > 0$, so type i is congestion amplifying, then*

$$\frac{dq_i^*}{da} < 0 \qquad \text{and} \qquad \frac{dn_i^*}{da} > 0.$$

2. *If $J_i(S^*) < 0$, so type i is congestion disciplining, then*

$$\frac{dq_i^*}{da} > 0.$$

Moreover, $\frac{dn_i^*}{da} < 0$ if and only if $J_i(S^*)$ is sufficiently negative, with the exact N-type threshold in equation (G4) of Appendix G.

The proof is in Appendix G. The proposition separates cross-sectional sorting from AI responses. High-β_Z types have higher equilibrium mean quality by Lemma 4, but they are not necessarily the types whose submissions expand most under AI. The reason is that β_Z changes both quality levels and incentives. Through $H_i(S)$, it changes whether congestion amplifies or disciplines the type's best response. Through K_i, it also changes direct AI exposure, since

$$\chi_i(a) = \frac{\Theta_i'(a)}{K_i}.$$

When β_R is common and AI raises routine capacity, lower-K_i types have larger direct exposure because the same routine-productivity gain is divided by a smaller own-dilution parameter. For this reason the congestion-disciplining type's crowd-out condition does not simplify to the common-exposure threshold in equation (G3).

The economic interpretation is nevertheless parallel. Low-β_Z researchers are more willing to expand quantity because additional papers do less damage to their own quality margin. This makes positive congestion feedback more likely for them. High-β_Z researchers are disciplined more strongly by the quality loss from spreading deep attention, so congestion is more likely to reduce their desired submissions. When a high-β_Z type is congestion disciplining, AI still raises routine productivity directly, but the induced increase in aggregate congestion pushes the type toward a smaller, higher-quality portfolio. Its submissions fall only when that disciplining force is strong enough to dominate the direct productivity effect. AI can therefore shift output toward researchers, fields, or paper types for which additional manuscripts require relatively little scarce deep attention.

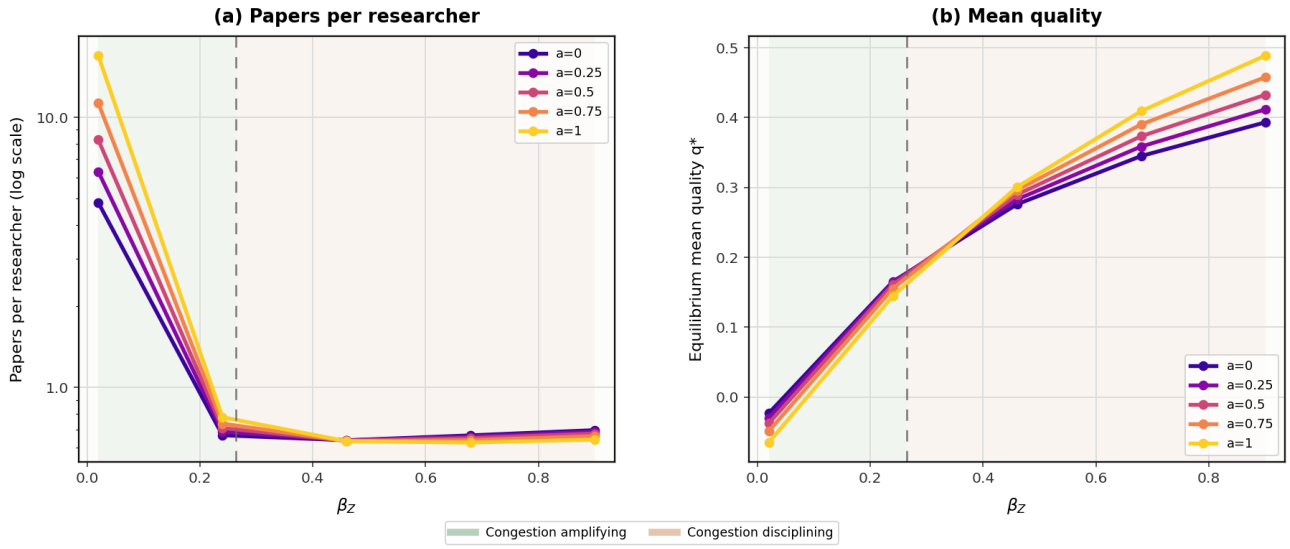


Figure 6: Submissions and mean quality across deep-attention sensitivities

Figure 6 illustrates the distinction. The horizontal axis is β_Z, so moving to the right means that deep attention is more productive for paper quality but also more costly to spread across a larger portfolio. The right panel confirms Lemma 4: higher-β_Z groups have higher mean submitted-paper quality. The left panel shows why the lemma does not impose a monotone quantity ranking. Low-β_Z groups can expand sharply because additional manuscripts impose relatively little own-quality dilution, whereas high-β_Z groups are disciplined by the loss from reallocating scarce deep attention across more papers. In empirical terms, high-β_Z researchers are those whose contribution depends heavily on sustained conceptual work, interpretation, or problem selection. AI may help them with routine tasks, but if it also raises aggregate congestion, their equilibrium response can be to protect quality, even when that means writing fewer papers.

The published record adds another layer. Published quantity and published quality are in general not monotone in β_Z. High-β_Z types have higher mean quality, but their submission counts need not be higher; low-β_Z types may submit many more papers, but those papers face lower acceptance probabilities. Their AI responses are also ambiguous because AI changes type-specific direct exposure, congestion-driven portfolio choices, and the common screening-noise environment at the same time. When screening noise rises rapidly with the AI-induced increase in aggregate submissions, the acceptance-rate erosion and published-quality erosion terms become more important. The result is that AI can raise submitted quantity for low-β_Z

groups and raise submitted mean quality for high-β_Z groups while published-paper counts or true published quality move in a weaker, nonmonotone, or even opposite cross-sectional direction.

5 Policy Implications

Sections 3 and 4 identify the private congestion feedback created by generative AI. This section translates that feedback into policy implications. The purpose is not to solve a full welfare model. Instead, we ask how an institution that cares about the true quality entering the published record should think about quantity control, publication selection, and screening informativeness. The main lesson is a constrained-efficiency implementation result. An institution can attain the best aggregate published-quality outcome available within the model's policy instruments by pairing an optimal publication threshold with a quantity instrument that implements the optimal submission level. The quantity instrument can be an effective submission cost or a binding submission cap.

5.1 The Publisher's Objective and Externality

We use "publisher" as shorthand for the institution that chooses rules for the publication system. The model's direct policy object is aggregate true published quality, not total social surplus. For a generic submission level S and threshold $\bar{q}$, define

$$q(S,a) \equiv \Theta(a) - K \log S, \qquad \sigma(S) \equiv \sigma_u(S), \qquad x(S,\bar{q};a) \equiv \frac{q(S,a)-\bar{q}}{\sigma(S)}.$$

The publisher's aggregate true published-quality objective is

$$\mathcal{A}^{\text{pub}}(S,\bar{q};a) = S\left[q(S,a)\Phi(x(S,\bar{q};a)) + \frac{\sigma_\varepsilon^2}{\sigma(S)}\phi(x(S,\bar{q};a))\right]. \tag{29}$$

This is the Section 3.4 object evaluated away from equilibrium and written with a calligraphic symbol to distinguish it from the private acceptance probability $A(n,S;a)$.

Equation (29) makes the externality explicit. An individual researcher takes S and therefore $\sigma(S)$ as fixed when choosing her portfolio. The publisher internalizes that a larger submission pool changes both the mean quality of submitted papers, through $q(S,a)$, and the informativeness of the publication signal, through $d\sigma(S)/dS > 0$. The decentralized equilibrium can

therefore have too many submissions relative to the publisher's aggregate-quality target, especially in the quantity-trap region where noisier screening raises private desired submissions. This statement is narrower than a welfare claim. It says that private portfolio choices need not maximize the true-quality content of the published record because researchers do not internalize the screening-noise externality imposed on all other submissions.

The heterogeneity results sharpen this policy problem. Section 4 shows that generative AI may shift output toward types whose submission incentives are amplified by congestion and away from types for whom congestion remains disciplining. Thus the externality is not only a representative-agent overproduction problem. It can also be a composition problem, because the types that expand most after an AI improvement need not be the types whose mean quality improves. This is why a useful policy discussion must keep quantity, selection, and screening informativeness separate.

5.2 Constrained Efficiency

Consider a fixed screening technology and AI intensity a. The publisher's constrained problem is to maximize aggregate true published quality using two policy margins: the publication threshold and an instrument that controls aggregate submissions. For any target submission level S, the threshold should be chosen to select papers with positive posterior expected true quality. The marginal effect of the threshold can be written as

$$\frac{\partial \mathcal{A}^{\text{pub}}(S, \bar{q}; a)}{\partial \bar{q}} = -S f_s(\bar{q})\, \mathbb{E}[\tilde{q} \mid s = \bar{q}],$$

where $f_s(\bar{q})$ is the density of the publication signal at the cutoff. Hence an interior selection optimum sets the posterior expected true quality of the marginal published paper equal to zero:

$$\mathbb{E}[\tilde{q} \mid s = \bar{q}^{\circ}(S; a)] = 0. \tag{30}$$

Under the normal signal structure, this rule is equivalent to

$$\bar{q}^{\circ}(S; a) = -\frac{\sigma_s(S)^2}{\sigma_\varepsilon^2}\, q(S, a). \tag{31}$$

The threshold rule is a selection rule. It says which signal cutoff maximizes true published quality conditional on the chosen quantity and screening technology.

Substituting the optimal threshold into the publisher's objective leaves a one-dimensional constrained problem:

$$S^{\circ}(a) \in \arg\max_{S \geq 0}\ \mathcal{A}^{\text{pub}}(S, \bar{q}^{\circ}(S; a); a). \tag{32}$$

Let

$$\bar{q}^{\circ}(a) \equiv \bar{q}^{\circ}(S^{\circ}(a); a), \qquad x^{\circ}(a) \equiv x(S^{\circ}(a), \bar{q}^{\circ}(a); a).$$

The pair $(S^{\circ}(a), \bar{q}^{\circ}(a))$ is the second-best, or constrained-efficient, allocation for the publisher's objective under the maintained screening technology. No alternative feasible pair of aggregate submissions and publication threshold delivers higher true published quality within this restricted policy problem.

Implementation then requires a quantity instrument. If the publisher can choose the effective submission cost, the researcher's private first-order condition at $n = S^{\circ}(a)$ is

$$c^{\circ}(a) = V\left[\Phi(x^{\circ}(a)) - \frac{K}{\sigma(S^{\circ}(a))}\phi(x^{\circ}(a))\right],$$

subject to implementability, local stability, and the intended equilibrium selection when multiple equilibria exist. If the primitive cost is c and the policy uses a fee, the implementing fee is $\kappa^{\circ}(a) = c^{\circ}(a) - c$. Equivalently, a binding submission cap can implement the quantity target directly by setting the per-researcher cap equal to $S^{\circ}(a)$ in the representative-agent economy. In either implementation, the threshold remains $\bar{q}^{\circ}(a)$. The cost or cap delivers the constrained-efficient scale, and the threshold delivers the constrained-efficient selection rule from that submission pool.

Finally, policy should adapt when generative AI changes the production environment. Appendix E records the corresponding reoptimization result. Along a smooth strict interior optimum, generative AI raises optimized aggregate true published quality by the envelope logic and raises the optimal submission target. However, the optimal threshold and implementing quantity instrument need not move monotonically, because they respond to the chosen submission level, the induced mean quality, and the informativeness of screening. The robust policy message is therefore not to raise the threshold, fee, or cap mechanically. It is to recompute the constrained-efficient pair $(S^{\circ}(a), \bar{q}^{\circ}(a))$ as AI changes private research productivity. Appendix D derives the threshold rule and implementing-cost formula used in this subsection.

6 Conclusion

This paper studies how generative AI changes the equilibrium quality of a publication system with finite screening capacity. The central lesson is that a privately productivity-enhancing technology can reduce equilibrium mean research quality when it expands submissions into a congested screen. At fixed quantity, AI raises routine research capacity and improves mean

paper quality. In equilibrium, however, researchers respond by submitting more papers. When the resulting screening congestion weakens the discipline imposed by the publication threshold, the submission response is amplified. In a locally stable quantity-trap equilibrium, this feedback makes the induced attention dilution strong enough to lower mean paper quality. The mechanism is therefore not that AI directly makes research worse, nor that researchers behave irrationally. It is that individually optimal submission decisions do not internalize how aggregate volume reduces the informativeness of journal screening.

The same analysis also clarifies the boundary of the result. Outside the quantity trap, screening congestion disciplines submissions rather than encouraging them, and generative AI raises submissions, mean quality, and a benchmark measure of useful knowledge. Aggregate true published quality is a further object with its own sign condition: more submissions create a scale gain, but they can also lower mean submitted quality and weaken selection through noisier screening. The model therefore predicts that the relevant empirical objects are the strength of screening congestion and the distribution of researcher types whose submission incentives are amplified by noise. Heterogeneity turns the representative-agent trap into a composition prediction, because AI can shift output toward high private-value or low attention-sensitivity types whose mean quality falls. For policy, the implication is to separate instruments by margin. Submission fees or caps control quantity, publication thresholds control selection, and screening improvements control informativeness. Future work can add richer microfoundations, robustness checks, and quantitative policy calibration, but the main message is already visible in the baseline model: the effect of generative AI on published knowledge depends on the equilibrium interaction between private submission incentives and scarce screening capacity.

Appendix A. Robustness: General Homogeneous Production and Endogenous Research Time

The main text derives the quantity trap under two simplifying assumptions: paper quality is produced by a Cobb–Douglas technology, and input capacities are fixed, so generative AI cannot change how much total time a researcher devotes to research. This appendix relaxes both assumptions. Per-paper quality is generated by a general homogeneous aggregator over routine effort and deep attention, and each researcher allocates a common time budget across leisure, routine effort, and deep attention. The main result is qualitatively unchanged. The best response retains the standardized-gap structure of Lemma 1, generative AI raises equilibrium submissions along every stable equilibrium branch, and equilibrium mean quality falls with AI exactly when the screening-congestion feedback of the main text is positive. The leisure margin adds a single new feedback term, which scales the size of the equilibrium response but leaves the sign of the quality effect unchanged.

Setup. A researcher divides a time budget $T > 0$ between leisure L, routine effort e_R, and deep attention e_Z, subject to $L + e_R + e_Z = T$. As in the main text, generative AI augments routine tasks only: e_R units of routine time yield $(1 + \gamma a)e_R$ units of effective routine input. A researcher who writes n papers spreads both inputs evenly across them, so each paper receives routine input $r = (1 + \gamma a)e_R/n$ and deep attention $z = e_Z/n$. Mean per-paper quality is

$$q = \theta + \log Y(r, z),$$

where the aggregator Y is positive, continuously differentiable, strictly increasing in both inputs, homogeneous of degree $K > 0$, and strictly log-concave, so that $\log Y$ is strictly concave on the relevant input range.[11] The Cobb–Douglas technology in (1) is the special case $Y(r, z) = r^{\beta_R} z^{\beta_Z}$ with fixed inputs $e_R = R_0$ and $e_Z = Z$; its homogeneity degree is $K = \beta_R + \beta_Z$.

Leisure yields utility $u(L)$, where u is strictly increasing, strictly concave, and twice differentiable, with boundary conditions $u'(0) = +\infty$ and $u'(T) = 0$. The first condition makes zero leisure infinitely costly at the margin; the second makes the first unit of research time costless at the margin. Together they keep the time allocation interior, so no boundary cases arise below. Everything else is as in Section 2: given aggregate submissions S, a submitted paper is

[11] Strict log-concavity is compatible with any homogeneity degree, including $K > 1$; the Cobb–Douglas and CES families satisfy it for every $K > 0$.

accepted when its noisy signal clears the threshold $\bar{q}$, total signal noise $\sigma_u(S)$ is increasing in S, each accepted paper is worth V, and each submission costs $c \in (0, V)$.

Homogeneity and the Dilution Form. The Cobb–Douglas form matters for the main text only through the log-linear dilution expression in (1), and homogeneity alone delivers that expression. Let $E \equiv e_R + e_Z = T - L$ denote total research effort and $\omega \equiv e_R/E$ the routine share of that effort. Homogeneity of degree K implies

$$Y\left(\frac{(1+\gamma a)\omega E}{n}, \frac{(1-\omega)E}{n}\right) = \left(\frac{E}{n}\right)^K Y\big((1+\gamma a)\omega, 1-\omega\big).$$

The routine share enters only the last factor, so the optimal split maximizes $Y\big((1+\gamma a)\omega, 1-\omega\big)$ regardless of n, E, and the screening environment. Let $\omega_Y(a)$ denote the maximizing share, which we take to be interior; it is unique because maximizing Y is equivalent to maximizing $\log Y$, and both inputs are linear in ω, so $\omega \mapsto \log Y\big((1+\gamma a)\omega, 1-\omega\big)$ inherits the strict concavity of $\log Y$. Substituting the optimal split and defining the productivity shifter

$$\Theta_Y(a) \equiv \theta + \log Y\big((1+\gamma a)\omega_Y(a), 1-\omega_Y(a)\big)$$

gives

$$q(n, E, a) = \Theta_Y(a) + K \log E - K \log n. \tag{A1}$$

By the envelope theorem, the shifter rises with AI:

$$\Theta_Y'(a) = \frac{\gamma\, \omega_Y(a)\, Y_r\big((1+\gamma a)\omega_Y(a), 1-\omega_Y(a)\big)}{Y\big((1+\gamma a)\omega_Y(a), 1-\omega_Y(a)\big)} > 0,$$

where Y_r denotes the partial derivative of Y with respect to its routine input. AI makes each unit of routine time more productive, and the re-optimized input mix converts that gain into higher benchmark quality.

Equation (A1) is the exact analog of the quality expression in (1), with total research effort now endogenous. The comparison identifies the property that the main analysis actually uses: a constant elasticity of per-paper quality with respect to spreading inputs across more papers. Homogeneity is precisely that property. Doubling research effort and submissions together leaves per-paper quality unchanged, while raising submissions alone dilutes quality at the constant rate K. In the general technology, the homogeneity degree therefore plays the role of the dilution factor $K = \beta_R + \beta_Z$ in the main text.

The Researcher's Problem. Given aggregate submissions S, a researcher chooses submissions and total effort to maximize

$$U(n, E, S; a) = n\left[V\Phi\big(x(n, E, S; a)\big) - c\right] + u(T - E), \qquad x(n, E, S; a) \equiv \frac{q(n, E, a) - \bar{q}}{\sigma_u(S)}.$$

Relative to the objective in (5), the problem adds one choice variable, total research effort, and one payoff term, the utility of the leisure time left over.

The submission margin is unchanged. Holding E fixed, the derivative of the objective with respect to n has the same form as in Section 3.1, so any interior optimum satisfies the standardized-gap condition (7),

$$\Phi(x) - \frac{K}{\sigma_u(S)}\phi(x) = \frac{c}{V},$$

with K reinterpreted as the homogeneity degree. The fixed-congestion apparatus of the main text therefore applies verbatim: the target gap $x^{\text{br}}(S)$ is the unique economically relevant solution, best-response mean quality is $q^{\text{br}}(S; a) = \bar{q} + \sigma_u(S)x^{\text{br}}(S)$ as in (8), the congestion cutoff $\hat{x}(S)$ in (10) is unchanged, and Assumption 1 again delivers congestion amplification at every submission level.[12]

The new margin is total effort. Differentiating the objective with respect to E gives the effort–leisure first-order condition

$$u'(T - E) = \frac{nVK}{E\,\sigma_u(S)}\,\phi\big(x(n, E, S; a)\big). \tag{A2}$$

The left-hand side is the marginal utility cost of research time. The right-hand side is its marginal value: an additional unit of effort raises every paper's quality at rate K/E, and each quality gain converts into acceptance probability at rate $\phi(x)/\sigma_u(S)$ across the researcher's n papers, each worth V if accepted.

The two conditions jointly determine the fixed-S best response. Substituting (A1) into the mean-quality target $q^{\text{br}}(S; a)$ gives the submission rule

$$n = E\,C_Y(a)B(S), \qquad C_Y(a) \equiv \exp\left(\frac{\Theta_Y(a) - \bar{q}}{K}\right), \tag{A3}$$

where $B(S) = \exp\big(-\sigma_u(S)x^{\text{br}}(S)/K\big)$ is the congestion component defined in Section 3.2. Substituting the submission rule into (A2) reduces the effort choice to a single condition in E alone:

$$u'(T - E) = VK\,C_Y(a)B(S)\,\frac{\phi\big(x^{\text{br}}(S)\big)}{\sigma_u(S)}. \tag{A4}$$

[12] The own-margin second-order conditions also carry over: on the relevant interval $x > -\sigma_u(S)/K$, the objective is locally strictly concave in n and in E, as in the proof of Lemma 1.

As E rises from 0 to T, the left-hand side increases strictly from 0 to $+\infty$ by the concavity and boundary conditions on u, while the right-hand side does not depend on E. The first-order condition in (A4) therefore has a unique solution $E^{\text{br}}(S,a) \in (0,T)$, and the fixed-$S$ best response is

$$n^{\text{br}}(S;a) = E^{\text{br}}(S,a)\, C_Y(a) B(S).$$

The best response inherits the productivity–congestion factorization of Section 3.2, with endogenous research effort as a third multiplicative factor.

The Response to AI at Fixed Congestion. How strongly effort responds to AI depends on the curvature of leisure utility. Define

$$\alpha(E) \equiv -\frac{E\, u''(T-E)}{u'(T-E)} > 0,$$

the elasticity of the marginal utility of leisure with respect to research effort. Its reciprocal $1/\alpha(E)$ measures how elastically research time responds to its marginal return: a nearly linear u makes effort supply highly elastic, while $\alpha(E) \to \infty$ approximates the fixed research time of the main text. Log-differentiating (A4) with respect to a at fixed S, the left-hand side contributes $\alpha(E)$ times the log change in effort and the right-hand side contributes $C_Y'(a)/C_Y(a) = \Theta_Y'(a)/K$, so $\partial \log E^{\text{br}}(S,a)/\partial a = (1/\alpha(E))\,\Theta_Y'(a)/K$. By the submission rule in (A3),

$$\frac{\partial \log n^{\text{br}}(S;a)}{\partial a} = \left(1 + \frac{1}{\alpha(E)}\right)\frac{\Theta_Y'(a)}{K} > 0, \qquad \frac{\partial q^{\text{br}}(S;a)}{\partial a} = 0. \tag{A5}$$

Both parts of the fixed-congestion response in Lemma 1 survive. AI leaves best-response mean quality unchanged, because a does not enter the standardized-gap condition. Moreover, AI raises the desired submission count, now through two channels: the direct productivity channel $\Theta_Y'(a)/K$ from the main text, and a new reallocation channel $(1/\alpha(E))\,\Theta_Y'(a)/K$, as the higher return to research pulls time out of leisure. The leisure margin amplifies the private quantity response to AI without touching the quality target.

Equilibrium Comparative Statics. In a symmetric equilibrium, the common submission level satisfies the fixed point $n^*(a) = n^{\text{br}}(n^*(a);a)$, with equilibrium effort $E^*(a) = E^{\text{br}}(n^*(a),a)$. The congestion feedback now has two parts. Let

$$J(S) \equiv S\frac{B'(S)}{B(S)},$$

which is the feedback index in (13). As in the main text, J records how screening congestion moves the quality target and hence the desired submission count. The new part is the effort response to congestion,

$$\zeta_E(S,E) \equiv \frac{\partial \log E^{\mathrm{br}}(S,a)}{\partial \log S} = -\frac{S\,\sigma_u'(S)\,x^{\mathrm{br}}(S)}{K\,\alpha(E)}, \tag{A6}$$

computed in the proof of Proposition A1 below: noisier screening also changes the marginal return to research time and pulls effort out of, or back into, leisure. The elasticity of desired submissions with respect to aggregate submissions is the sum of the two parts,

$$\mathcal{J}(S,E) \equiv \frac{\partial \log n^{\mathrm{br}}(S;a)}{\partial \log S} = J(S) + \zeta_E(S,E), \tag{A7}$$

At a symmetric equilibrium, this elasticity coincides with the marginal response of n^{br} to aggregate submissions, so the one-dimensional adjustment argument of Appendix F applies with $\mathcal{J}$ in place of J: a simple symmetric equilibrium is locally stable when $\mathcal{J}(n^*(a), E^*(a)) < 1$. The next proposition records the equilibrium effects of AI.

Proposition A1 (Robustness of the equilibrium effects of AI) *Let $n^*(a)$ be a differentiable symmetric equilibrium branch with an interior time allocation and $\mathcal{J}(n^*(a), E^*(a)) \neq 1$. Then equilibrium submissions and mean quality respond to AI according to*

$$\frac{dn^*(a)}{da} = \left(1 + \frac{1}{\alpha(E^*(a))}\right) \frac{\Theta_Y'(a)}{K} \frac{n^*(a)}{1 - \mathcal{J}(n^*(a), E^*(a))} \tag{A8}$$

and

$$\frac{dq^*(a)}{da} = -\left(1 + \frac{1}{\alpha(E^*(a))}\right) \Theta_Y'(a) \frac{J(n^*(a))}{1 - \mathcal{J}(n^*(a), E^*(a))}. \tag{A9}$$

On a locally stable branch, $\mathcal{J}(n^*(a), E^*(a)) < 1$, so $dn^*(a)/da > 0$ and

$$\operatorname{sign}\left(\frac{dq^*(a)}{da}\right) = -\operatorname{sign}\big(J(n^*(a))\big).$$

In particular, under Assumption 1, equilibrium submissions rise and equilibrium mean quality falls with AI, as in Proposition 3.

Proof. We first compute ζ_E. Applying the implicit function theorem to the standardized-gap condition (7), as in the proof of Lemma 1,

$$\frac{dx^{\mathrm{br}}(S)}{dS} = -\frac{K\sigma_u'(S)}{\sigma_u(S)\big(\sigma_u(S) + Kx^{\mathrm{br}}(S)\big)}.$$

Write $x = x^{\mathrm{br}}(S)$ and $\sigma = \sigma_u(S)$ for brevity. Then $\log B(S) = -\sigma x/K$ gives

$$J(S) = -\frac{S}{K}\left(\sigma_u'(S)x + \sigma\frac{dx^{\mathrm{br}}(S)}{dS}\right) = \frac{S\sigma_u'(S)}{K}\,\frac{K - \sigma x - Kx^2}{\sigma + Kx},$$

and

$$\frac{d}{dS}\log\frac{\phi(x^{\mathrm{br}}(S))}{\sigma_u(S)} = -x\frac{dx^{\mathrm{br}}(S)}{dS} - \frac{\sigma_u'(S)}{\sigma} = -\frac{\sigma_u'(S)}{\sigma + Kx}.$$

Log-differentiating (A4) with respect to $\log S$ at fixed a, the left-hand side contributes $\alpha(E)\zeta_E(S,E)$ and the right-hand side contributes the two elasticities above:

$$\alpha(E)\zeta_E(S,E) = J(S) - \frac{S\sigma_u'(S)}{\sigma + Kx} = \frac{S\sigma_u'(S)}{\sigma + Kx}\left(\frac{K - \sigma x - Kx^2}{K} - 1\right) = -\frac{S\sigma_u'(S)\,x}{K},$$

which is (A6). By (A3), $\log n^{\mathrm{br}}(S;a) = \log E^{\mathrm{br}}(S,a) + \log C_Y(a) + \log B(S)$, so the elasticity of the best response with respect to S is $\zeta_E(S,E) + J(S) = \mathcal{J}(S,E)$, as in (A7).

Differentiating the fixed point $n^*(a) = n^{\mathrm{br}}(n^*(a);a)$ with respect to a and solving,

$$\frac{d\log n^*(a)}{da} = \left.\frac{\partial \log n^{\mathrm{br}}(S;a)/\partial a}{1 - \mathcal{J}(S,E)}\right|_{S=n^*(a),\,E=E^*(a)},$$

and substituting the fixed-congestion response (A5) gives (A8).

For quality, equilibrium mean quality equals the best-response quality target evaluated at equilibrium congestion, $q^*(a) = Q(n^*(a))$ with $Q(S) \equiv \bar{q} + \sigma_u(S)x^{\mathrm{br}}(S)$. Since $\log B(S) = -(Q(S) - \bar{q})/K$, the slope of the target is

$$Q'(S) = -\frac{K}{S}J(S).$$

The chain rule $dq^*(a)/da = Q'(n^*(a))\,dn^*(a)/da$ then yields (A9). On a locally stable branch, $1 - \mathcal{J} > 0$, and $\Theta_Y'(a) > 0$ and $\alpha(E^*(a)) > 0$ imply $dn^*(a)/da > 0$; the sign of $dq^*(a)/da$ is then opposite to that of $J(n^*(a))$. Finally, the closed form for $J(S)$ above is positive exactly when $x^{\mathrm{br}}(S) < \hat{x}(S)$, which Assumption 1 guarantees at every submission level, so $J(n^*(a)) > 0$ and $dq^*(a)/da < 0$. □

The reason the sign criterion survives is instructive. Equilibrium mean quality is read off the quality target $Q(S) = \bar{q} + \sigma_u(S)x^{\mathrm{br}}(S)$, and the slope of that target, $Q'(S) = -(K/S)J(S)$, depends only on the screening-congestion index J. The two generalizations enter elsewhere. The general aggregator changes the direct productivity shifter, with $\Theta_Y'(a)$ replacing $\beta_R\gamma/(1+\gamma a)$,

and the leisure margin changes how far equilibrium submissions move, through the amplification factor $1 + 1/\alpha$ and through ζ_E in the feedback denominator. Neither touches the map from congestion to quality. The direction of the equilibrium quality effect is therefore decided by the same question as in the main text: does screening congestion weaken publication discipline, $J > 0$, or strengthen it, $J < 0$?

The leisure margin also reinforces, rather than offsets, the congestion loop in the empirically relevant region. When mean submitted quality lies below the publication threshold, $x^{\mathrm{br}}(S) < 0$—the case suggested by the acceptance-rate evidence discussed in Section 3.1—the elasticity in (A6) is positive: noisier screening raises the acceptance prospects of marginal papers, raises the marginal return to research time, and pulls effort out of leisure.

Relation to the Main Text. Two special cases connect the formulas to the main text. First, as effort supply becomes inelastic, $\alpha(E) \to \infty$, the reallocation terms vanish: $1 + 1/\alpha \to 1$, $\zeta_E \to 0$, and $\mathcal{J} \to J$. The comparative statics in (A8) and (A9) then collapse to those of Propositions 1 and 3, with the constant $K \log E$ absorbed into the productivity shifter. Second, the Cobb–Douglas technology $Y(r, z) = r^{\beta_R} z^{\beta_Z}$ has homogeneity degree $K = \beta_R + \beta_Z$, optimal routine share $\omega_Y(a) = \beta_R / K$, and $\Theta'_Y(a) = \beta_R \gamma / (1 + \gamma a)$, so the direct AI shifter matches the main text exactly.

The quantity trap therefore rests neither on the Cobb–Douglas functional form nor on the assumption that total research time is fixed. Any quality technology with a constant dilution elasticity, combined with any smooth interior leisure choice, delivers the same conclusion: generative AI raises equilibrium submissions, and it lowers equilibrium mean quality precisely when screening congestion weakens publication discipline.

Appendix B. Outside-Trap Benchmark

The same model also gives a constructive benchmark in which generative AI is beneficial for the quality measures studied here. Outside the quantity trap, additional congestion lowers the individual best response. The feedback from screening congestion then dampens, rather than amplifies, the direct increase in submissions caused by AI.

Let

$$\underline{\sigma} \equiv \sqrt{\sigma_\varepsilon^2 + \sigma_0^2}, \qquad \overline{\sigma} \equiv \sqrt{\sigma_\varepsilon^2 + \bar{\sigma}_s^2}.$$

For any $\sigma \in [\underline{\sigma}, \overline{\sigma}]$, define

$$\hat{x}(\sigma) \equiv \frac{-\sigma + \sqrt{\sigma^2 + 4K^2}}{2K}.$$

The global outside-trap condition is

$$\frac{c}{V} > \max_{\sigma \in [\underline{\sigma}, \overline{\sigma}]} \left[\Phi(\hat{x}(\sigma)) - \frac{K}{\sigma} \phi(\hat{x}(\sigma)) \right]. \tag{B1}$$

Because the best-response equation is increasing for $x > -\sigma_u(S)/K$, this condition implies $x^{\text{br}}(S) > \hat{x}(\sigma_u(S))$ for every S. By Lemma 1, the best response is decreasing in aggregate submissions everywhere.

Proposition B1 (Outside-trap benchmark) *Suppose $0 < c < V$ and the outside-trap condition above holds. Then for every AI intensity a, there is a unique symmetric equilibrium. Along this equilibrium,*

$$\frac{dn^*(a)}{da} > 0, \qquad \frac{dq^*(a)}{da} > 0.$$

Moreover, threshold-excess useful knowledge,

$$X(a) \equiv n^*(a)\, \mathbb{E}[(\tilde{q} - \bar{q})_+],$$

is increasing in a.

Thus the model does not imply that generative AI must reduce research quality. It identifies the boundary condition: AI raises these quality measures when screening congestion disciplines quantity rather than amplifying it.

Appendix C. Heterogeneous Comparative-Static Formulas

For each active type, the private problem has the same one-dimensional structure as in the representative-agent case. The only substitutions are K_i for K, τ_i for τ, and Θ_i for Θ. Thus

$$n_i^* = C_i(a) B_i(S^*)$$

and the aggregate fixed point is

$$S^* = \sum_{i=1}^{N} \mu_i C_i(a) B_i(S^*).$$

The type-specific congestion elasticity is

$$J_i(S) \equiv S\frac{B_i'(S)}{B_i(S)}.$$

Equivalently,

$$\frac{\partial n_i^{\mathrm{br}}(S;a)}{\partial S} = \frac{n_i^{\mathrm{br}}(S;a)}{S} J_i(S).$$

Differentiating the type-specific first-order condition as in the representative-agent case gives

$$J_i(S) = \frac{S\sigma_u'(S)}{K_i}\frac{K_i - \sigma_u(S)x_i^{\mathrm{br}}(S) - K_i x_i^{\mathrm{br}}(S)^2}{\sigma_u(S) + K_i x_i^{\mathrm{br}}(S)}.$$

Under the variance-saturation screening technology, this is

$$J_i(S) = \frac{S\lambda\Delta_s e^{-\lambda S}}{2K_i\sigma_u(S)}\frac{K_i - \sigma_u(S)x_i^{\mathrm{br}}(S) - K_i x_i^{\mathrm{br}}(S)^2}{\sigma_u(S) + K_i x_i^{\mathrm{br}}(S)}.$$

The derivative of the aggregate fixed-point gap with respect to S is

$$\sum_{i=1}^{N} \mu_i C_i(a) B_i'(S^*) - 1 = \sum_{i=1}^{N} \omega_i J_i(S^*) - 1 = \Omega(S^*) - 1.$$

Hence a simple local equilibrium is stable when $\Omega(S^*) < 1$.

Let

$$\chi_i(a) \equiv \frac{\partial \log C_i(a)}{\partial a} = \frac{\Theta_i'(a)}{K_i}, \qquad \bar{\chi}(a) \equiv \sum_{i=1}^{N} \omega_i \chi_i(a).$$

Differentiate the aggregate fixed point with respect to a:

$$\frac{dS^*}{da} = \sum_{i=1}^{N} \mu_i n_i^* \chi_i(a) + \sum_{i=1}^{N} \mu_i n_i^* \frac{B_i'(S^*)}{B_i(S^*)}\frac{dS^*}{da}.$$

Dividing by S^* and using $J_i(S) = SB_i'(S)/B_i(S)$ gives

$$\frac{1}{S^*}\frac{dS^*}{da} = \bar{\chi}(a) + \Omega(S^*)\frac{1}{S^*}\frac{dS^*}{da}.$$

Solving yields

$$\frac{dS^*}{da} = S^*\frac{\bar{\chi}(a)}{1 - \Omega(S^*)}.$$

For each type,

$$\frac{dn_i^*}{da} = n_i^*\left[\chi_i(a) + J_i(S^*)\frac{1}{S^*}\frac{dS^*}{da}\right] = n_i^*\left[\chi_i(a) + J_i(S^*)\frac{\bar{\chi}(a)}{1 - \Omega(S^*)}\right].$$

Finally, $q_i^*(a) = \Theta_i(a) - K_i \log n_i^*(a)$ and $\Theta_i'(a) = K_i\chi_i(a)$, so

$$\frac{dq_i^*(a)}{da} = -K_i J_i(S^*)\frac{\bar{\chi}(a)}{1 - \Omega(S^*)}.$$

Appendix D. Policy Instrument Derivations

For fixed (S, a), write the publisher's objective as

$$\mathcal{A}^{\text{pub}}(S, \bar{q}; a) = S \int_{\bar{q}}^{\infty} \mathbb{E}[\tilde{q} \mid s = y] f_s(y)\, dy.$$

Differentiating with respect to the lower limit gives

$$\frac{\partial \mathcal{A}^{\text{pub}}(S, \bar{q}; a)}{\partial \bar{q}} = -S f_s(\bar{q}) \mathbb{E}[\tilde{q} \mid s = \bar{q}].$$

An interior threshold optimum therefore sets the posterior mean of the marginal published paper equal to zero.

Under the normal signal structure, $\tilde{q} = q(S, a) + \varepsilon$ and $s = \tilde{q} + \eta$. Thus

$$\mathbb{E}[\tilde{q} \mid s] = q(S, a) + \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_s(S)^2}\big(s - q(S, a)\big),$$

or equivalently

$$\mathbb{E}[\tilde{q} \mid s] = \frac{\sigma_s(S)^2}{\sigma_\varepsilon^2 + \sigma_s(S)^2} q(S, a) + \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_s(S)^2} s.$$

Setting this expression equal to zero at $s = \bar{q}^\circ(S; a)$ yields

$$\bar{q}^\circ(S; a) = -\frac{\sigma_s(S)^2}{\sigma_\varepsilon^2} q(S, a).$$

If the publisher wants to implement target submissions S at threshold $\bar{q}$, the representative researcher's first-order condition at $n = S$ with a per-submission charge κ is

$$V\Phi(x(S, \bar{q}; a)) - c - \kappa - \frac{VK}{\sigma(S)}\phi(x(S, \bar{q}; a)) = 0.$$

Solving for κ gives the implementation equation stated in the main text:

$$c + \kappa = V\left[\Phi(x(S, \bar{q}; a)) - \frac{K}{\sigma(S)}\phi(x(S, \bar{q}; a))\right].$$

The formula is local and must be paired with implementability, local stability, and the intended equilibrium selection when multiple equilibria exist. A binding cap implements the same quantity margin directly by restricting feasible submissions to the target level, but it still leaves the threshold to determine which submitted papers are published.

The constrained-efficiency implementation follows from the posterior-zero rule and the implementing-fee derivation. The posterior-zero threshold rule is a selection condition conditional on (S, σ_s). The implementing fee or a binding cap chooses the constrained-efficient submission quantity. A screening instrument m changes the posterior mapping by replacing $\sigma_s^2(S)$ with $\sigma_s^2(S, m)$, where

$$\frac{\partial \sigma_s^2(S, m)}{\partial m} < 0.$$

Thus the fee or cap implements the submission target, the threshold implements the selection rule at that target, and screening improvements change the informativeness of the signal.

Appendix E. Policy Reoptimization Under Generative AI

Consider the baseline threshold and fee portfolio with a given screening-congestion technology. After substituting the posterior-zero threshold, define

$$x^\circ(S; a) = \frac{q(S, a) - \bar{q}^\circ(S; a)}{\sigma(S)}, \qquad p^\circ(S; a) = q(S, a)\Phi(x^\circ(S; a)) + \frac{\sigma_\varepsilon^2}{\sigma(S)}\phi(x^\circ(S; a)),$$

and

$$W(S; a) = S p^\circ(S; a).$$

Let $S^\circ(a)$ be a smooth strict interior maximizer, so $W_S(S^\circ(a); a) = 0$ and $W_{SS}(S^\circ(a); a) < 0$. Then

$$\frac{dS^\circ(a)}{da} = -\frac{W_{Sa}(S^\circ(a); a)}{W_{SS}(S^\circ(a); a)}.$$

At the posterior-zero threshold, the envelope derivatives of optimized per-submission value are

$$p_q^\circ = \Phi(x^\circ), \qquad p_\sigma^\circ = -\frac{\sigma_\varepsilon^2}{\sigma^2}\phi(x^\circ) < 0.$$

The target first-order condition is

$$W_S = p^\circ - K\Phi(x^\circ) + S p_\sigma^\circ \sigma_S = 0.$$

Let

$$e = \frac{S\sigma_S}{\sigma} \geq 0, \qquad \ell(x) = \frac{\phi(x)}{\Phi(x)}, \qquad x = x^\circ(S; a).$$

Using

$$p^\circ = \frac{\sigma_\varepsilon^2}{\sigma}\left[x\Phi(x) + \phi(x)\right],$$

the target first-order condition implies

$$\frac{K\sigma}{\sigma_\varepsilon^2} = x + (1-e)\ell(x).$$

At fixed S, research AI shifts submitted-paper quality by

$$q_a(S,a) = \frac{\partial q(S,a)}{\partial a} = \frac{\beta_R \gamma}{1+\gamma a} > 0.$$

Differentiating the target condition with respect to a and substituting the preceding identity gives

$$\frac{W_{Sa}}{q_a} = \Phi(x)\left[1 - (1-e)\ell(x)\big(x+\ell(x)\big)\right].$$

The lower-tail inverse-Mills inequality gives

$$0 < \ell(x)\big(x+\ell(x)\big) < 1.$$

Since $e \geq 0$, the bracket is strictly positive. Thus $W_{Sa} > 0$, and the strict second-order condition $W_{SS} < 0$ implies

$$\frac{dS^\circ(a)}{da} > 0.$$

The optimized value is

$$\mathcal{A}^{\text{pub},\circ}(a) = W(S^\circ(a);a).$$

By the envelope theorem,

$$\frac{d\mathcal{A}^{\text{pub},\circ}(a)}{da} = W_a(S^\circ(a);a) = S^\circ(a)\Phi(x^\circ(S^\circ(a);a))q_a(S^\circ(a),a) > 0.$$

The optimized threshold satisfies

$$\bar{q}^\circ(S;a) = -r(S)q(S,a), \qquad r(S) = \frac{\sigma_s(S)^2}{\sigma_\varepsilon^2},$$

so along the optimized target

$$\frac{d\bar{q}^\circ(a)}{da} = -rq_a - (r_S q + r q_S)\frac{dS^\circ(a)}{da}.$$

The two terms can have opposing signs. Similarly, the implementing fee combines a direct fixed-target term with the endogenous target response:

$$\frac{d\kappa^\circ(a)}{da} = \kappa_a^\circ + \kappa_S^\circ \frac{dS^\circ(a)}{da}.$$

Holding S fixed, the direct term is positive at the optimizing solution satisfying $1 + Kx/\sigma(S) > 0$ because research AI makes submissions privately more attractive, but the target-adjustment term may offset or reinforce it. Hence the optimized target and optimized aggregate published quality rise locally, while the threshold and implementing fee do not have unconditional signs.

Appendix F. Equilibrium Stability Details

This appendix records the fixed-point stability calculation used in Section 3.2. Fix AI intensity a and write the symmetric best response in factorized form,

$$n^{\mathrm{br}}(S;a) = C(a)B(S).$$

A symmetric equilibrium is a fixed point S^* satisfying $S^* = C(a)B(S^*)$. Define the excess-submission gap

$$F_a(S) \equiv n^{\mathrm{br}}(S;a) - S = C(a)B(S) - S.$$

The local adjustment considered in the main text is the standard one-dimensional rule under which aggregate submissions move in the direction of excess desired submissions:

$$\dot{S}_t = \rho F_a(S_t), \qquad \rho > 0.$$

The same local condition is obtained from the corresponding discrete best-response geometry.

At a fixed point,

$$F_a'(S^*) = C(a)B'(S^*) - 1.$$

Using the equilibrium identity $C(a)B(S^*) = S^*$,

$$C(a)B'(S^*) = S^*\frac{B'(S^*)}{B(S^*)} = J(S^*).$$

Therefore

$$F_a'(S^*) = J(S^*) - 1.$$

Linearizing the adjustment equation around S^* gives

$$\dot{\Delta}_t = \rho\,[J(S^*) - 1]\,\Delta_t + o(\Delta_t), \qquad \Delta_t \equiv S_t - S^*.$$

Thus a simple fixed point is locally stable when $J(S^*) < 1$ and locally unstable when $J(S^*) > 1$. The boundary case $J(S^*) = 1$ is non-simple and corresponds to the fold case in the fixed-point geometry. In the congestion-amplifying region maintained in the main text, $J(S^*) > 0$. Combining amplification with local stability gives the stable feedback range $0 < J(S^*) < 1$: congestion reinforces the direct AI-induced increase in desired submissions, but the feedback is not strong enough to make the fixed point locally unstable.

Appendix G. Proofs

Proof of Lemma 1. Fix $S \geq 0$ and $a \geq 0$. Write $\sigma = \sigma_u(S) > 0$ and define the cost-benefit ratio $\tau \equiv c/V \in (0,1)$. The proof has three steps. We first solve the researcher's optimization problem by characterizing the unique optimizing standardized quality gap and the associated global best response. We then differentiate the best response and associated best-response quality with respect to AI intensity, and finally with respect to aggregate submissions.

Step 1: unique global best response. For a candidate standardized quality gap x, define

$$G_\sigma(x) \equiv \Phi(x) - \frac{K}{\sigma}\phi(x).$$

The derivative of the objective can be written as

$$\frac{\partial U(n,S;a)}{\partial n} = V\Phi(x(n,S;a)) - c - \frac{VK}{\sigma_u(S)}\phi(x(n,S;a)) = V\left[G_\sigma(x(n,S;a)) - \tau\right].$$

Because x is strictly monotone in n, critical points of U in n correspond exactly to solutions of $G_\sigma(x) = \tau$.

The derivative of G_σ is

$$G'_\sigma(x) = \phi(x)\left(1 + \frac{Kx}{\sigma}\right),$$

so G_σ decreases on $(-\infty, -\sigma/K)$ and increases on $(-\sigma/K, \infty)$. At the minimum,

$$G_\sigma\left(-\frac{\sigma}{K}\right) = \Phi\left(-\frac{\sigma}{K}\right) - \frac{K}{\sigma}\phi\left(\frac{\sigma}{K}\right) < 0,$$

where the inequality follows from the standard Mills-ratio bound $\Phi(-z) < \phi(z)/z$ for $z > 0$. Moreover, $\lim_{x\to-\infty} G_\sigma(x) = 0$ and $\lim_{x\to\infty} G_\sigma(x) = 1$. Because $\tau > 0$, the equation $G_\sigma(x) = \tau$ has no solution on $(-\infty, -\sigma/K]$, where $G_\sigma \leq 0$. Because $\tau < 1$, it has exactly one solution on $(-\sigma/K, \infty)$, denoted $x^{\text{br}}(S)$, which therefore satisfies

$$x^{\text{br}}(S) > -\frac{\sigma_u(S)}{K}.$$

The map

$$x(n,S;a) = \frac{\Theta(a) - \bar{q} - K\log n}{\sigma_u(S)}$$

is a strictly decreasing bijection from $n \in (0,\infty)$ onto $x \in \mathbb{R}$, with $x(n,S;a) \to \infty$ as $n \to 0^+$ and $x(n,S;a) \to -\infty$ as $n \to \infty$. The bijection then yields the unique critical point

$$n^{\text{br}}(S;a) = \exp\left(\frac{\Theta(a) - \bar{q} - \sigma_u(S)x^{\text{br}}(S)}{K}\right),$$

which is the best-response expression in (9). To confirm that this critical point is the global maximum, we trace the sign of $\partial U/\partial n$ along the path $n : 0^+ \to \infty$, equivalently $x : \infty \to -\infty$. As x decreases from $+\infty$ to $-\sigma/K$, $G_\sigma(x)$ decreases from 1 to a value strictly below zero, crossing τ exactly once at $x^{\text{br}}(S)$. As x continues from $-\sigma/K$ to $-\infty$, $G_\sigma(x)$ rises from its minimum back to $0 < \tau$ and never reaches τ. Thus $\partial U/\partial n$ is strictly positive for $n < n^{\text{br}}(S;a)$ and strictly negative for $n > n^{\text{br}}(S;a)$. Taking the natural continuous extension $U(0,S;a) = 0$, this critical point is the unique global maximizer over $n \geq 0$.

Step 2: Response to AI Intensity. At fixed S, the first-order condition $G_{\sigma_u(S)}(x^{\text{br}}(S)) = \tau$ does not involve a, so $x^{\text{br}}(S)$ is independent of a. The mean-quality identity $q^{\text{br}}(S;a) = \bar{q} + \sigma_u(S)x^{\text{br}}(S)$ then gives

$$\frac{\partial q^{\text{br}}(S;a)}{\partial a} = 0.$$

In the production identity $n^{\text{br}}(S;a) = \exp\big((\Theta(a) - q^{\text{br}}(S;a))/K\big)$, the right-hand side depends on a only through $\Theta(a)$, which is strictly increasing. Hence $n^{\text{br}}(S;a)$ is strictly increasing in a, so $\partial n^{\text{br}}(S;a)/\partial a > 0$.

Step 3: Response to Aggregate Submissions. Apply the implicit function theorem to $G_{\sigma_u(S)}(x^{\text{br}}(S)) = \tau$ to obtain

$$\frac{dx^{\text{br}}(S)}{dS} = -\frac{K\sigma_u'(S)}{\sigma_u(S)\big(\sigma_u(S) + Kx^{\text{br}}(S)\big)}.$$

Differentiating the mean-quality identity $q^{\text{br}}(S;a) = \bar{q} + \sigma_u(S)x^{\text{br}}(S)$ and substituting this expression yields

$$\frac{\partial q^{\text{br}}(S;a)}{\partial S} = \sigma_u'(S)x^{\text{br}}(S) + \sigma_u(S)\frac{dx^{\text{br}}(S)}{dS} = \frac{\sigma_u'(S)}{\sigma_u(S) + Kx^{\text{br}}(S)}\Big(K\big(x^{\text{br}}(S)\big)^2 + \sigma_u(S)x^{\text{br}}(S) - K\Big), \tag{G1}$$

The variance-saturation technology gives $\sigma_u'(S) > 0$, and Step 1 gives $\sigma_u(S) + Kx^{\text{br}}(S) > 0$, so $\partial q^{\text{br}}(S;a)/\partial S$ has the same sign as the quadratic in $x^{\text{br}}(S)$ on the right-hand side of (G1). This quadratic opens upward in $x^{\text{br}}(S)$, and its positive root is the threshold $\hat{x}(S)$ defined in (10); the negative root lies below $-\sigma_u(S)/K$ and is outside the feasible interval $x^{\text{br}}(S) > -\sigma_u(S)/K$. On the feasible interval, it is negative below $\hat{x}(S)$, zero at $\hat{x}(S)$, and positive above $\hat{x}(S)$. Therefore $\partial q^{\text{br}}(S;a)/\partial S$ is negative, zero, or positive according as $x^{\text{br}}(S)$ is below, equal to, or above $\hat{x}(S)$.

The submission-count response follows immediately from the production identity $n^{\text{br}}(S;a) = \exp\big((\Theta(a) - q^{\text{br}}(S;a))/K\big)$:

$$\frac{\partial n^{\text{br}}(S;a)}{\partial S} = -\frac{n^{\text{br}}(S;a)}{K}\frac{\partial q^{\text{br}}(S;a)}{\partial S}. \tag{G2}$$

Since $n^{\mathrm{br}}(S;a) > 0$, $\partial n^{\mathrm{br}}(S;a)/\partial S$ has the opposite sign of $\partial q^{\mathrm{br}}(S;a)/\partial S$, so it is positive, zero, or negative according as $x^{\mathrm{br}}(S)$ is below, equal to, or above $\hat{x}(S)$. □

Proof of Proposition 1. Differentiate the fixed-point identity $n^*(a) = n^{\mathrm{br}}(n^*(a);a)$ with respect to a:

$$\frac{dn^*(a)}{da} = \left.\frac{\partial n^{\mathrm{br}}(S;a)}{\partial a}\right|_{S=n^*(a)} + \left.\frac{\partial n^{\mathrm{br}}(S;a)}{\partial S}\right|_{S=n^*(a)} \frac{dn^*(a)}{da}.$$

At the symmetric equilibrium, the partial in S equals $J(n^*(a))$ by (13), and simplicity rules out $J(n^*(a)) = 1$. Solving for $dn^*(a)/da$ then gives (14). In the congestion-amplifying region $0 < J(n^*(a)) < 1$, so the feedback multiplier $1/(1 - J(n^*(a)))$ exceeds one. The direct AI response is positive by Lemma 1, and the inequality in (14) follows. □

Proof of Proposition 2. From Equation (9),

$$n^{\mathrm{br}}(S;a) = \exp\left(\frac{\Theta(a) - \bar{q}}{K}\right) \exp\left(-\frac{\sigma_u(S)x^{\mathrm{br}}(S)}{K}\right) = C(a)B(S).$$

Thus positive symmetric equilibria solve

$$C(a) = M(S), \qquad M(S) \equiv \frac{S}{B(S)}.$$

The shifter $C(a)$ moves with AI, while the congestion shape M determines the number of fixed points.

The turning points of M are exactly the points where the local feedback index equals one:

$$\frac{M'(S)}{M(S)} = \frac{1}{S} - \frac{B'(S)}{B(S)} = \frac{1 - J(S)}{S}.$$

Because $B(S)$ is positive and bounded away from zero and infinity over the bounded noise range,

$$\lim_{S\downarrow 0} M(S) = 0, \qquad \lim_{S\to\infty} M(S) = \infty.$$

We now record the fold construction under variance saturation. Set $t = \lambda S$ and write

$$\sigma(t) = \sqrt{\sigma_\varepsilon^2 + \sigma_0^2 + \Delta_s(1 - e^{-t})}, \qquad k(t) \equiv \frac{K}{\sigma(t)}, \qquad r(t) \equiv t\frac{\sigma_t(t)}{\sigma(t)}.$$

Under the variance-saturation law,

$$0 < r(t) \le \frac{1}{2}, \qquad \lim_{t\downarrow 0} r(t) = \lim_{t\to\infty} r(t) = 0.$$

Fix the cost-benefit ratio $\tau \equiv c/V \in (0,1)$. Let $x(t;\tau)$ be the relevant solution of

$$\Phi(x) - k(t)\phi(x) = \tau, \qquad 1 + k(t)x > 0.$$

Using the slope formula in Equation (G1), define

$$Q(k,x) \equiv \frac{k - x - kx^2}{k(1+kx)}.$$

Then the feedback index can be written as

$$J(t;\tau) = r(t)Q(k(t), x(t;\tau)).$$

For fixed t, $Q(k(t),x)$ is strictly decreasing on the relevant branch $x > -1/k(t)$:

$$\frac{\partial Q(k,x)}{\partial x} = -\frac{(1+kx)^2 + k^2}{k(1+kx)^2} < 0.$$

Since $Q(k,0) = 1$ and $Q(k,x) \to \infty$ as $x \downarrow -1/k$, the equation $r(t)Q(k(t),x) = 1$ has a unique fold-implied solution $x_f(t) \in (-1/k(t), 0)$. Define the fold-implied cost-benefit ratio

$$\tau_f(t) \equiv \Phi(x_f(t)) - k(t)\phi(x_f(t)).$$

For any fixed τ, a fold occurs at t if and only if $\tau_f(t) = \tau$.

The variance-saturation fold locus is single-peaked on its positive region. To see the source of the single crossing, write $y(t) = 1 + k(t)x_f(t)$. Since $x_f(t) \in (-1/k(t), 0)$, $0 < y(t) < 1$. The fold identity can be written as

$$k(t)^2 = \frac{r(t)y(t)(1-y(t))}{y(t) - r(t)},$$

so $y(t) > r(t)$. Differentiating the fold-implied ratio along the fold locus gives

$$\operatorname{sign}\left(\tau_f'(t)\right) = \operatorname{sign}(\Lambda(t)),$$

where

$$\Lambda(t) \equiv \frac{r(t)^2(1-y(t))(2r(t)-y(t))}{y(t)^2(y(t)-r(t))} - \big(t - 1 + r(t)\big).$$

For the variance-saturation path, $r'(t)/r(t) = 1/t - 1 - 2r(t)/t$. Differentiating the fold identity gives

$$\frac{dy}{dt} = \frac{y(1-y)}{y^2 - 2ry + r}\left[y\left(\frac{1-2r}{t} - 1\right) + \frac{2r(y-r)}{t}\right],$$

where r and y are evaluated at t. Substituting this expression into $\Lambda'(t)$ gives a rational expression whose denominator is positive because $0 < r(t) < y(t) < 1$ and $r(t) \leq 1/2$. On the region where $\tau_f(t) > 0$, the numerator is negative, so $\Lambda(t)$ is strictly decreasing there. At the left edge of the positive region $\Lambda(t) > 0$, while in the tail $\Lambda(t) < 0$. Hence τ_f rises once and then falls; it is strictly single-peaked on its positive region. Let

$$\tau^\star \equiv \sup_{t>0} \tau_f(t).$$

If $0 < \tau < \tau^\star$, the single-peakedness of τ_f gives exactly two fold points $t_1(\tau) < t_2(\tau)$. The sign equivalence

$$J(t;\tau) > 1 \quad \Longleftrightarrow \quad \tau < \tau_f(t)$$

then implies that M is increasing on $(0, t_1/\lambda)$, decreasing on $(t_1/\lambda, t_2/\lambda)$, and increasing on $(t_2/\lambda, \infty)$.

Define the scaled fold levels

$$\overline{m}(\tau) \equiv t_1(\tau) \exp\left(\frac{x_f(t_1(\tau))}{k(t_1(\tau))}\right), \qquad \underline{m}(\tau) \equiv t_2(\tau) \exp\left(\frac{x_f(t_2(\tau))}{k(t_2(\tau))}\right).$$

The first fold is a local maximum of the scaled equilibrium shape and the second fold is a local minimum, so $\overline{m}(\tau) > \underline{m}(\tau)$. Therefore $C(a) = M(S)$ has three positive solutions exactly when

$$\underline{m}(\tau) < \lambda C(a) < \overline{m}(\tau).$$

Since

$$C(a) = C_0(1 + \gamma a)^{\beta_R/K}$$

with $C_0 > 0$, this inequality is equivalent to

$$a_L(\tau) < a < a_H(\tau),$$

where

$$a_L(\tau) \equiv \frac{1}{\gamma}\left[\left(\frac{\underline{m}(\tau)}{\lambda C_0}\right)^{K/\beta_R} - 1\right], \qquad a_H(\tau) \equiv \frac{1}{\gamma}\left[\left(\frac{\overline{m}(\tau)}{\lambda C_0}\right)^{K/\beta_R} - 1\right].$$

Intersecting with the admissible domain $a \geq 0$ gives the stated multiplicity window. At either admissible endpoint, a horizontal line is tangent to M, so two crossings merge into a non-simple fold with $J = 1$. Outside the closed admissible interval, the horizontal line intersects only one of the monotone intervals of M. □

Proof of Proposition 3. Suppose three symmetric equilibria exist at the same a and write them as $n_L < n_M < n_H$. At a symmetric equilibrium, aggregate submissions equal the common individual submission level, and mean quality is

$$q(n, a) = \Theta(a) - K \log n.$$

Holding a fixed, this expression is strictly decreasing in n. Therefore

$$q(n_H, a) < q(n_M, a) < q(n_L, a).$$

For the local comparative static, combine the fixed-point derivative from Proposition 1 with

$$\frac{C'(a)}{C(a)} = \frac{\beta_R \gamma}{K(1 + \gamma a)}$$

yields

$$\frac{dn^*(a)}{da} = \frac{\beta_R \gamma}{K(1 + \gamma a)} \frac{n^*(a)}{1 - J(n^*(a))}.$$

When $J(n^*(a)) < 1$, this derivative is positive.

Mean equilibrium quality is $q^*(a) = \Theta(a) - K \log n^*(a)$. Therefore

$$\frac{dq^*(a)}{da} = \frac{\beta_R \gamma}{1 + \gamma a} - \frac{K}{n^*(a)} \frac{dn^*(a)}{da}.$$

Substituting the expression for dn^*/da gives

$$\frac{dq^*(a)}{da} = \frac{\beta_R \gamma}{1 + \gamma a} \left[1 - \frac{1}{1 - J(n^*(a))} \right] = -\frac{\beta_R \gamma}{1 + \gamma a} \frac{J(n^*(a))}{1 - J(n^*(a))}.$$

If $0 < J(n^*(a)) < 1$, this derivative is strictly negative. □

Proof of Proposition B1. For each feasible σ, $\hat{x}(\sigma) > -\sigma/K$. For $x > -\sigma/K$, $G_\sigma(x) = \Phi(x) - (K/\sigma)\phi(x)$ is strictly increasing. The condition in Equation (B1) therefore implies

$$x^{\text{br}}(S) > \hat{x}(\sigma_u(S)) \qquad \text{for every } S.$$

By Lemma 1, $B'(S) < 0$ and hence $J(S) < 0$ for every S. For fixed a, define the fixed-point gap

$$F_a(S) \equiv C(a)B(S) - S.$$

Then $F_a'(S) = C(a)B'(S) - 1 < 0$. Also $F_a(0) = C(a)B(0) > 0$, while $F_a(S) \to -\infty$ because $B(S)$ is bounded over the bounded noise range. Thus $F_a(S) = 0$ has exactly one solution for every a. The solution is simple because $F_a'(S) < 0$.

The same log-differentiation used above applies because $J(n^*(a)) < 0 < 1$. The equilibrium branch satisfies

$$n^*(a) = C(a)B(n^*(a)).$$

Log-differentiating gives

$$\frac{1}{n^*(a)}\frac{dn^*(a)}{da} = \frac{C'(a)}{C(a)} + \frac{J(n^*(a))}{n^*(a)}\frac{dn^*(a)}{da}.$$

Since $C'(a)/C(a) = \beta_R\gamma/[K(1+\gamma a)]$,

$$\frac{dn^*(a)}{da} = \frac{\beta_R\gamma}{K(1+\gamma a)}\frac{n^*(a)}{1-J(n^*(a))} > 0.$$

Using $q^*(a) = \Theta(a) - K\log n^*(a)$ gives

$$\frac{dq^*(a)}{da} = -\frac{\beta_R\gamma}{1+\gamma a}\frac{J(n^*(a))}{1-J(n^*(a))} > 0.$$

For threshold-excess useful knowledge, let

$$d(a) = \frac{q^*(a) - \bar{q}}{\sigma_\varepsilon}.$$

Then

$$X(a) = n^*(a)\left[(q^*(a) - \bar{q})\Phi(d(a)) + \sigma_\varepsilon\phi(d(a))\right].$$

The bracket is the expected positive part of a nondegenerate normal variable, so it is strictly positive. Its derivative with respect to q^* is $\Phi(d(a)) > 0$. Since both dn^*/da and dq^*/da are positive outside the trap, $dX/da > 0$. □

Proof of Proposition 4. For a differentiable symmetric equilibrium $n^*(a)$, write $q^* = q(n^*(a), a)$, $\sigma = \sigma_u(n^*(a))$, and $x = (q^* - \bar{q})/\sigma$. Let $u = \varepsilon + \eta$. Then $u \sim \mathcal{N}(0, \sigma^2)$, $s = q^* + u$, and

$$\mathbb{E}[\varepsilon \mid u] = \frac{\sigma_\varepsilon^2}{\sigma^2}u.$$

Therefore

$$\begin{aligned}
\mathbb{E}[\tilde{q}\,\mathbf{1}\{s \geq \bar{q}\}] &= q^*\Pr(u \geq \bar{q} - q^*) + \mathbb{E}[\varepsilon\,\mathbf{1}\{u \geq \bar{q} - q^*\}] \\
&= q^*\Phi(x) + \frac{\sigma_\varepsilon^2}{\sigma^2}\mathbb{E}[u\,\mathbf{1}\{u \geq -\sigma x\}] \\
&= q^*\Phi(x) + \frac{\sigma_\varepsilon^2}{\sigma}\phi(x).
\end{aligned}$$

This gives the closed form $A^{\text{pub}}(a) = n^*(a)p(q^*, \sigma)$.

With $\rho_\varepsilon = \sigma_\varepsilon^2/\sigma^2$ and $\rho_s = 1 - \rho_\varepsilon$,

$$p_q(q, \sigma) = \Phi(x) + \phi(x)\left(\frac{\bar{q}}{\sigma} + \rho_s x\right)$$

and

$$p_\sigma(q, \sigma) = -\phi(x)\left(\rho_\varepsilon + \rho_s x^2 + \frac{\bar{q}}{\sigma} x\right).$$

These formulas identify the two partial derivatives used below; the decomposition itself is just the chain rule. Since

$$A^{\text{pub}}(a) = n^*(a)p(q^*(a), \sigma_u(n^*(a))),$$

differentiating with respect to a gives

$$\frac{dA^{\text{pub}}(a)}{da} = p\frac{dn^*(a)}{da} + n^* p_q \frac{dq^*(a)}{da} + n^* p_\sigma \sigma_S \frac{dn^*(a)}{da}.$$

Under variance saturation,

$$\sigma_S = \frac{\lambda \Delta_s e^{-\lambda n^*(a)}}{2\sigma_u(n^*(a))}.$$

The needed branch derivatives follow from the fixed point itself. Log-differentiating $n^*(a) = C(a)B(n^*(a))$ gives

$$\frac{1}{n^*(a)}\frac{dn^*(a)}{da} = \frac{C'(a)}{C(a)} + \frac{J(n^*)}{n^*}\frac{dn^*(a)}{da}.$$

Using $C'(a)/C(a) = \beta_R\gamma/[K(1+\gamma a)]$ and $J(n^*) < 1$,

$$\frac{dn^*(a)}{da} = \frac{\beta_R\gamma}{K(1+\gamma a)}\frac{n^*}{1 - J(n^*)}.$$

Since $q^*(a) = \Theta(a) - K\log n^*(a)$,

$$\frac{dq^*(a)}{da} = -\frac{\beta_R\gamma}{1+\gamma a}\frac{J(n^*)}{1 - J(n^*)}.$$

Substituting these derivatives and collecting terms yields

$$\frac{dA^{\text{pub}}(a)}{da} = \frac{\beta_R\gamma}{1+\gamma a}\frac{n^*}{1 - J(n^*)}\left[\frac{p}{K} - J(n^*)p_q + \frac{n^*\sigma_S}{K}p_\sigma\right].$$

When $J(n^*) < 1$, $1 - J(n^*) > 0$. The sign is therefore the sign of the bracket, which proves Equation (19). The bracket is negative exactly when

$$J(n^*)p_q - \frac{n^*\sigma_S}{K}p_\sigma > \frac{p}{K},$$

which gives Equation (20). □

Proof of Lemma 2. With common K and common τ, every type solves the same equation for the standardized quality gap:

$$\Phi(x_i^{\mathrm{br}}(S)) - \frac{K}{\sigma_u(S)}\phi(x_i^{\mathrm{br}}(S)) = \tau.$$

The unique relevant solution is therefore common across types. Hence $x_i^{\mathrm{br}}(S) = x^{\mathrm{br}}(S)$ for all i,

$$B_i(S) = \exp\left(-\frac{\sigma_u(S)x^{\mathrm{br}}(S)}{K}\right) = B(S),$$

and $J_i(S) = SB_i'(S)/B_i(S) = J(S)$ for all active types. The level shifter $\Theta_i(a)$ changes only $C_i(a)$, not the congestion shape. At an interior equilibrium, the definition of the standardized gap gives

$$q_i^*(a) = \bar{q} + \sigma_u(S^*)x^{\mathrm{br}}(S^*) \equiv q^*(a)$$

for every active type. Moreover, because $B_i(S^*) = B(S^*)$,

$$\frac{n_i^*(a)}{n_j^*(a)} = \frac{C_i(a)}{C_j(a)} = \exp\left(\frac{\Theta_i(a) - \Theta_j(a)}{K}\right),$$

so $n_i^*(a) > n_j^*(a)$ if and only if $\Theta_i(a) > \Theta_j(a)$. Finally, common $J_i(S^*) = J(S^*)$ implies $\Omega(S^*) = J(S^*)$. Under stability, $1 - J(S^*) > 0$. If $\bar{\chi}(a) > 0$ and $J(S^*) > 0$, the heterogeneous quality derivative becomes

$$\frac{dq_i^*(a)}{da} = -KJ(S^*)\frac{\bar{\chi}(a)}{1 - J(S^*)} < 0.$$

The type-level submission derivative becomes

$$\frac{dn_i^*(a)}{da} = n_i^*(a)\left[\chi_i(a) + J(S^*)\frac{\bar{\chi}(a)}{1 - J(S^*)}\right],$$

which is strictly positive for every active type when $\chi_i(a) \geq 0$ for all active types. □

Proof of Lemma 3. With common production primitives, the only type-specific object in the first-order condition is τ_i. Let

$$G_\sigma(x) = \Phi(x) - \frac{K}{\sigma}\phi(x).$$

For $x > -\sigma/K$, G_σ is strictly increasing. Thus $\tau_i > \tau_j$ implies $x_i^*(S^*) > x_j^*(S^*)$ at any common equilibrium aggregate S^*. Using the common-production best-response formula,

$$\frac{n_i^*}{n_j^*} = \exp\left(-\frac{\sigma_u(S^*)\left[x_i^*(S^*) - x_j^*(S^*)\right]}{K}\right) < 1.$$

Equilibrium mean quality satisfies

$$q_i^* = \bar{q} + \sigma_u(S^*)x_i^*(S^*),$$

so $q_i^* > q_j^*$. Combining these two expressions gives

$$q_i^* - q_j^* = K \log\left(\frac{n_j^*}{n_i^*}\right) > 0.$$

□

Proof of Proposition 5. With common production primitives, the only type-specific object in the first-order condition is τ_i. Let $G_\sigma(x) = \Phi(x) - K\phi(x)/\sigma$. For $x > -\sigma/K$, G_σ is strictly increasing. By the type-specific best-response slope formula in Appendix C, type i is congestion amplifying exactly when $x_i^{\text{br}}(S) < \hat{x}(S)$. Using the definition of $H(S)$ in (26), this condition is equivalent to

$$J_i(S) > 0 \quad \Longleftrightarrow \quad \tau_i = G_{\sigma_u(S)}(x_i^{\text{br}}(S)) < G_{\sigma_u(S)}(\hat{x}(S)) = H(S).$$

At equality, $J_i(S) = 0$, so the type is on the knife-edge boundary between congestion amplification and congestion discipline. The reverse inequality gives $J_i(S) < 0$.

Common production primitives imply $\chi_i(a) = \bar{\chi}(a) \equiv \chi(a) = \beta_R\gamma/[K(1+\gamma a)] > 0$. Let $D \equiv 1 - \Omega(S^*)$. Stability gives $D > 0$. The heterogeneous comparative-static formulas reduce to

$$\frac{1}{n_i^*}\frac{dn_i^*}{da} = \chi(a)\left(1 + \frac{J_i(S^*)}{D}\right), \qquad \frac{dq_i^*(a)}{da} = -KJ_i(S^*)\frac{\chi(a)}{D}.$$

If $J_i(S^*) > 0$, then the quality derivative is negative. The quantity derivative is positive because

$$\frac{1}{n_i^*}\frac{dn_i^*}{da} = \chi(a)\left(1 + \frac{J_i(S^*)}{D}\right) > 0.$$

If $J_i(S^*) < 0$, then the quality derivative is positive. The quantity derivative is negative if and only if

$$1 + \frac{J_i(S^*)}{D} < 0,$$

or equivalently

$$J_i(S^*) < -\left[1 - \Omega(S^*)\right].$$

Equivalently, using $\Omega(S^*) = \sum_k \omega_k J_k(S^*)$, type i's quantity falls if and only if

$$\sum_{k \neq i} \omega_k \left[J_k(S^*) - J_i(S^*)\right] > 1. \tag{G3}$$

If $\tau_i > \tau_j$, then $x_i^*(S^*) > x_j^*(S^*)$ and $J_i(S^*) < J_j(S^*)$. Therefore

$$\frac{dq_i^*}{da} - \frac{dq_j^*}{da} = K\frac{\chi(a)}{D}\left[J_j(S^*) - J_i(S^*)\right] > 0.$$

The same ordering of the common-production best-response levels and slopes gives

$$\frac{dn_j^*}{da} - \frac{dn_i^*}{da} = \chi(a)\left[n_j^* - n_i^* + \frac{n_j^* J_j(S^*) - n_i^* J_i(S^*)}{D}\right] > 0,$$

so the cross-sectional quantity gap widens as well. □

Proof of Lemma 4. Fix an interior heterogeneous equilibrium and write $\sigma^* = \sigma_u(S^*)$. For type i, the standardized quality gap solves

$$\Phi(x_i^*) - \frac{K_i}{\sigma^*}\phi(x_i^*) = \tau.$$

On the relevant interior branch, the left-hand side is strictly increasing in x. If $\beta_{Zi} > \beta_{Zj}$, then $K_i > K_j$. Evaluating type i's left-hand side at type j's solution gives

$$\Phi(x_j^*) - \frac{K_i}{\sigma^*}\phi(x_j^*) < \Phi(x_j^*) - \frac{K_j}{\sigma^*}\phi(x_j^*) = \tau.$$

Hence $x_i^* > x_j^*$. Equilibrium mean quality satisfies $q_i^* = \bar{q} + \sigma^* x_i^*$, so $q_i^* > q_j^*$.

The production identity gives

$$\log n_i^* = \frac{\Theta_i(a) - \bar{q} - \sigma^* x_i^*}{K_i}.$$

This expression is not ordered by β_{Zi} alone. A higher β_{Zi} raises the level shifter $\Theta_i(a)$ and the target quality gap, and it also changes the denominator K_i. Without additional restrictions on these forces, the sign of $\log n_i^* - \log n_j^*$ is not pinned down. □

Proof of the cutoff characterization and Proposition 6. The cutoff characterization stated before Proposition 6 follows from the type-specific best-response slope formula in Appendix C applied with K_i. Equivalently, type i is congestion amplifying exactly when

$$\tau < H_i(S).$$

At equality, $J_i(S) = 0$, so the type is on the knife-edge boundary. It remains to show that $H_i(S)$ decreases with K_i for fixed S. Let $\sigma = \sigma_u(S)$ and $\kappa = K_i/\sigma$. The positive root $\hat{x}_i(S)$ solves

$$\kappa = \frac{\hat{x}_i(S)}{1 - \hat{x}_i(S)^2}, \qquad \hat{x}_i(S) \in (0,1).$$

The right-hand side is strictly increasing in $\hat{x}_i(S)$, so $\hat{x}_i(S)$ is increasing in κ and hence in K_i. Writing $H_i(S)$ as a function of $x \in (0,1)$ gives

$$H(x) = \Phi(x) - \frac{x}{1-x^2}\phi(x).$$

Differentiating,

$$H'(x) = -\frac{2x^2}{(1-x^2)^2}\phi(x) < 0.$$

Therefore $H_i(S)$ is decreasing in K_i. Since $K_i = \beta_R + \beta_{Zi}$, lower-$\beta_Z$ types have higher $H_i(S)$ and are more likely to satisfy the congestion-amplification inequality.

We now prove the comparative-static statements in the proposition. Common β_R and $\gamma > 0$ imply $\Theta_i'(a) = \beta_R\gamma/(1+\gamma a) > 0$, so $\chi_i(a) > 0$ for every active type and $\bar{\chi}(a) > 0$. Let $D \equiv 1 - \Omega(S^*)$. Stability gives $D > 0$. The heterogeneous comparative-static formulas imply

$$\frac{1}{n_i^*}\frac{dn_i^*}{da} = \chi_i(a) + J_i(S^*)\frac{\bar{\chi}(a)}{D}, \qquad \frac{dq_i^*(a)}{da} = -K_i J_i(S^*)\frac{\bar{\chi}(a)}{D}.$$

If $J_i(S^*) > 0$, then the quality derivative is negative and the quantity derivative is positive. If $J_i(S^*) < 0$, then the quality derivative is positive. The quantity derivative is negative if and only if

$$-J_i(S^*)\bar{\chi}(a) > \chi_i(a)\big(1 - \Omega(S^*)\big). \tag{G4}$$

□

References

Adda, Jérôme, and Marco Ottaviani. 2024. "Grantmaking, Grading on a Curve, and the Paradox of Relative Evaluation in Nonmarkets." *The Quarterly Journal of Economics*, 139(2): 1255–1319. 1

Atal, Vidya. 2010. "Do Journals Accept Too Many Papers?" *Economics Letters*, 107(2): 229–232. 1

Azar, Ofer H. 2015. "A Model of the Academic Review Process with Informed Authors." *The B.E. Journal of Economic Analysis & Policy*, 15(2): 865–889. 1, 5

Bayar, Onur, and Thomas J. Chemmanur. 2021. "A Model of the Editorial Process in Academic Journals." *Research Policy*, 50(9): 104339. 1, 5

Bergstrom, Carl T., and Kevin Gross. 2026. "Screening, Sorting, and the Feedback Cycles That Imperil Peer Review." *PLOS Biology*, 24(2): e3003650. 1

Boudreau, Kevin J., Eva C. Guinan, Karim R. Lakhani, and Christoph Riedl. 2016. "Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science." *Management Science*, 62(10): 2765–2783. 1

Card, David, and Stefano DellaVigna. 2020. "What Do Editors Maximize? Evidence from Four Economics Journals." *The Review of Economics and Statistics*, 102(1): 195–217. 1, 5

Cotton, Christopher. 2013. "Submission Fees and Response Times in Academic Publishing." *American Economic Review*, 103(1): 501–509. 1

Ellison, Glenn. 2002*a*. "Evolving Standards for Academic Publishing: A *q-r* Theory." *Journal of Political Economy*, 110(5): 994–1034. 1

Ellison, Glenn. 2002*b*. "The Slowdown of the Economics Publishing Process." *Journal of Political Economy*, 110(5): 947–993. 1, 2

Frankel, Alexander, and Maximilian Kasy. 2022. "Which Findings Should Be Published?" *American Economic Journal: Microeconomics*, 14(1): 1–38. 1

Gartenberg, Claudine, Sharique Hasan, Alex Murray, and Lamar Pierce. 2026. "More Versus Better: Artificial Intelligence, Incentives, and the Emerging Crisis in Peer Review." *Organization Science*, 0(0): orsc.2026.ed.v37.n3. Articles in Advance. 1, 1, 2, 6

Gehrig, Thomas, and Rune Stenbacka. 2021. "Journal Competition and the Quality of Published Research: Simultaneous versus Sequential Screening." *International Journal of Industrial Organization*, 76: 102718. 1

Hadavand, Aboozar, Daniel S. Hamermesh, and Wesley W. Wilson. 2024. "Publishing Economics: How Slow? Why Slow? Is Slow Productive? How to Fix Slow?" *Journal of Economic Literature*, 62(1): 269–293. 1, 2

Hao, Qianyue, Fengli Xu, Yong Li, and James Evans. 2026. "Artificial Intelligence Tools Expand Scientists' Impact but Contract Science's Focus." *Nature*, 649(8099): 1237–1243. 1, 1

Heckman, James J., and Sidharth Moktan. 2020. "Publishing and Promotion in Economics: The Tyranny of the Top Five." *Journal of Economic Literature*, 58(2): 419–470. 1, 2

Hirshleifer, David. 2015. "Editorial: Cosmetic Surgery in the Academic Review Process." *Review of Financial Studies*, 28(3): 637–649. 1

Hrazdil, Karel, Pavel Král, Jiri Novak, and Nattavut Suwanyangyuan. 2026. "On the Determinants of Journal Rejection Rates: Evidence from the Journal of Financial Economics." *Financial Innovation*, 12: 52. 8

Kim, Jaeho, Yunseok Lee, and Seulki Lee. 2025. "Position: The AI Conference Peer Review Crisis Demands Author Feedback and Reviewer Rewards." *arXiv preprint arXiv:2505.04966*. ICML 2025, Position Paper Track. 1

Kiri, Bralind, Nicola Lacetera, and Lorenzo Zirulia. 2018. "Above a Swamp: A Theory of High-Quality Scientific Production." *Research Policy*, 47(5): 827–839. 1

Korinek, Anton. 2023. "Generative AI for Economic Research: Use Cases and Implications for Economists." *Journal of Economic Literature*, 61(4): 1281–1317. 1, 1, 2

Kovanis, Michail, Raphaël Porcher, Philippe Ravaud, and Ludovic Trinquart. 2016. "The Global Burden of Journal Peer Review in the Biomedical Literature: Strong Imbalance in the Collective Enterprise." *PLOS ONE*, 11(11): e0166387. 1, 2

Kusumegi, Keigo, Xinyu Yang, Paul Ginsparg, Mathijs de Vaan, Toby Stuart, and Yian Yin. 2025. "Scientific Production in the Era of Large Language Models." *Science*, 390(6779): 1240–1243. 1

Welch, Ivo. 2014. "Referee Recommendations." *Review of Financial Studies*, 27(9): 2773–2804. 1

Zollman, Kevin J. S., Julian García, and Toby Handfield. 2024. "Academic Journals, Incentives, and the Quality of Peer Review: A Model." *Philosophy of Science*, 91(1): 186–203. 1