Paying Creators for Training Data Can Increase AI Displacement

Lijuan Luo* Cornell University
Emaad Manzoor Cornell University
Nathan Yang UIUC
Fan Yao UNC Chapel Hill

November 15, 2026

Abstract

Generative AI increasingly competes with creators while using their content as training data. In response, policymakers have called for platforms to compensate creators for training data use and allow opting out. We study how such policies affect creator displacement in a model of a platform where creators choose whether to contribute training data and consumers choose between purchasing from creators and prompting generative AI trained on their content. When compensation is required, the low-quality creator chooses to remain solely as a training data supplier while AI fully displaces her demand, despite AI quality and the price the platform charges for AI-generated content being identical with and without compensation. Without compensation, the risk of losing creators and their training data constrains the platform to set the price of creators' content to preserve their demand. This constraint also prevents further AI displacement as AI improves, with the platform raising AI prices while keeping creators' prices and demand unchanged to prevent exit. Compensation increases creators' payoffs and total surplus but reduces consumer surplus. Our findings highlight how compensation policies for training data can shape creators' economic roles in a future with more capable generative AI.

Keywords: Generative AI; Platform Design; Creator Economy; Future of Work; Ethical AI

*Lijuan Luo (ll954@cornell.edu) is the corresponding author. She is a doctoral student at Cornell University. Emaad Manzoor is an Assistant Professor of Marketing at Cornell University. Nathan Yang is an Associate Professor and C.W. Park Faculty Fellow at the University of Illinois Urbana-Champaign. Fan Yao is an Assistant Professor in the Department of Statistics and Operations Research at UNC Chapel Hill. The authors thank Tommaso Bondi, Andrey Fradkin, Sylvia Hristakeva, and Justin Johnson for their feedback. The authors also thank participants at 2024 ISMS Marketing Science Conference, 2024 Conference on Artificial Intelligence, Machine Learning, and Business Analytics (AIML), the Artificial Intelligence in Management (AIM) 2025 Conference, the 2025 PhD student workshop at Cornell University, the Brownbag at UIUC, and AI & Marketing: New Methods and New Risks Conference at Stanford for helpful discussions. Previously circulated under the title *Platform Design When Creators Train Their AI Substitutes* (November 7, 2025; http://dx.doi.org/10.2139/ssrn.5719862).

1 Introduction

Creative platforms increasingly offer consumers a choice between purchasing creators' content and prompting an integrated generative AI system trained on their content.[1] While consumers benefit from tailoring content to their preferences by prompting AI, creators risk being displaced. In response, policymakers have called for compensating creators and allowing them to opt out of AI training (European Parliament 2026). However, compensation can change the platform's pricing incentives. Its equilibrium effects on creators' demand, participation, and payoffs are thus unclear.

We show that compensating creators can increase AI displacement and change creators' economic role from selling content to supplying training data. The result arises because compensating creators for training data weakens the platform's incentive to preserve demand for their content. Without compensation, the platform must preserve content sales to prevent creators from exiting and withdrawing their training data. Compensation relaxes this constraint: the platform can retain creators, and their training data, without preserving demand for their content.

To formalize this, we extend a Hotelling-style model of a two-sided platform to include an integrated generative AI system that competes with creators. Consumers with heterogeneous tastes are distributed between two creators that differ in quality. Consumers choose between purchasing creators' content and exerting prompting effort to generate content with AI that matches their tastes exactly. The platform sets a common price for creators' content and a separate price for AI-generated content. The quality of AI-generated content and the required prompting effort depend on the creators' content used for training.

We use this model to compare two regimes. In the regime without compensation, providing training data is a condition of platform participation, and the platform retains all revenue from AI-generated content sales. In the regime with compensation and opt-out, creators choose whether to contribute training data, and the platform shares revenue from AI-generated content sales among contributing creators.[2] In both regimes, creators receive a share of revenue from their own content sales. Creators incur a participation cost and choose whether to remain on the platform, withdrawing their training data upon exit.

The platform preserves demand for both creators' content without compensation but fully displaces the low-quality creator's demand with compensation. Without compensation, the platform sets the price of creators' content to retain the low-quality creator, whose exit would deprive the platform of her training data. With compensation, the platform sets a higher price for creators' content while retaining the low-quality creator solely as a training data supplier. In both regimes, both creators contribute training data (by choice under compensation), and the quality and price of AI-generated content are identical. Greater AI displacement thus arises not from a better or cheaper AI, but from the requirement to compensate creators changing the platform's pricing incentives.

[1] For example, consumers on the Shutterstock and Adobe Stock photography platforms can prompt generative AI trained on creators' content to generate a customized image rather than purchase one from a creator.

[2] This revenue sharing scheme is inspired by Shutterstock's Contributor Fund (Shutterstock 2026) and resembles the pro-rata revenue sharing principle used by music streaming platforms (Moreau et al. 2024).

Requiring compensation also changes how the platform responds to improvements in AI capability. In both regimes, the platform raises the price of AI-generated content identically as AI improves. Without compensation, the platform keeps the price of creators' content unchanged to preserve their demand, prevent exit, and retain their training data. With compensation, the platform instead raises the price of creators' content, progressively displacing the remaining demand for creators until AI serves the entire market. Both creators nevertheless remain as training data suppliers, earning more from their data as AI improves. Better AI thus increases displacement only when compensation allows the platform to retain training data without preserving creators' sales.

Requiring compensation increases both creators' payoffs and total surplus but reduces consumer surplus. The total surplus increase comes from reallocating demand toward AI-generated content that better matches consumers' tastes, rather than expanding market coverage or improving AI quality. Consumers lose surplus as the platform raises the price of creators' content and induces more consumers to purchase AI-generated content, from which the platform captures their full willingness to pay. As AI improves, the differences in both creators' payoffs and total surplus between compensation and no compensation increase. The platform may prefer not to compensate creators when preserving demand for their content is cheaper than sharing AI revenue. A requirement to compensate creators can thus increase total surplus when the platform would otherwise choose not to compensate, but reduces consumer surplus and platform profit.

We extend our model to examine imperfect AI "steerability" (the prompting effort required to generate content that matches consumers' tastes) (Vafa et al. 2025), heterogeneous transfer rates of creators' training data to improve AI quality, and alternative compensation rules. Compensation-induced displacement persists with moderate prompting effort, although sufficiently high prompting effort can restore demand for creators' content. It also persists when creators' data differ in their contributions to AI quality, as equal revenue sharing allows displaced creators to remain as training data suppliers even when their contributions are small. When compensation is weighted by contributions to AI quality, displacement persists under equal transfer rates, but revenue from AI-generated content sales is redistributed toward the high-quality creator. With heterogeneous transfer rates, however, quality-weighted compensation may fail to sustain the participation of displaced creators whose data contribute little to AI quality.

Our findings highlight a distinct feature of AI displacement. While automation typically reduces demand for the labor it replaces, generative AI can create a new source of demand for displaced creators: their content as training data. Compensation allows creators to earn revenue from AI-generated content sales as training data suppliers even after losing demand for their content, leaving them better off as AI improves and further displaces their sales. Compensation also increases total surplus by shifting demand toward AI-generated content, although consumers lose surplus. Compensation policies may thus shape creators' economic roles and payoffs in a future with increasingly capable generative AI.

1.1 Related Work

Our paper contributes to research on technological change and labor displacement. This literature examines how automation displaces labor from existing tasks (Autor et al. 2003, Acemoglu and Restrepo 2018, 2020, 2022), how new tasks create labor demand (Acemoglu and Restrepo 2019, Autor et al. 2024), and how displaced workers adjust (Autor and Dorn 2013, Feigenbaum and Gross 2024). The displaced creators in our model continue supplying the content they produced, but now as training data for the AI that replaces their sales. Compensation for training data allows them to benefit from AI improvements without undertaking new tasks.

We also contribute to research on AI copyright. This literature examines how copyright protection, licensing, and opt-out rights affect access to AI training data (Gans 2026, de Rassenfosse et al. 2025, Shcherbakov et al. 2025, Peukert 2025a, Cooper et al. 2025, de la Durantaye 2025, Samuelson 2026), how training on creators' content affects their subsequent contributions (Peukert et al. 2025), and how to compensate creators for AI training data (Yang and Zhang 2024, Peukert 2025b, Gao et al. 2025, Wang et al. 2024, Senftleben 2025, Ørstavik 2026, Bostoen and Monti 2025).

We contribute to this literature in three ways. First, compensation changes the platform's price for creators' content, and hence displacement, even when training data access and AI capability are unchanged. Second, creators can remain as training data suppliers and receive higher payoffs than without compensation, even when AI displaces demand for their content. Third, we show that compensation can increase total surplus while reducing consumer surplus by inducing the platform to reallocate demand toward AI, in addition to redistributing revenue to creators.

Our paper also relates to theoretical research on platform encroachment. This literature examines how platforms' competing offerings affect third-party sellers' participation and investment incentives (Jiang et al. 2011, Hagiu and Spulber 2013, Hagiu et al. 2022), and how platform pricing, channel structure, and information policies shape these incentives (Abhishek et al. 2016, Choi et al. 2025). Relatedly, Bhargava et al. (2026) show that data sharing with competitors can create strategic interdependence that softens competition. Yao et al. (2024) model competition between creators and AI trained on their content, taking data sharing as given. Taitler et al. (2025) and Keinan and Ben-Porat (2026) examine strategic data sharing with competing AI and the role of data pricing and revenue sharing. Unlike these studies, which focus on content production and data sharing, we examine competition through the platform's joint pricing of creators' and AI-generated content.

Creators in our model can earn revenue in two economic roles, by selling content to consumers or supplying training data to the platform. The platform trains AI on their content and competes with them through AI-generated content. Compensation changes which of these roles must generate revenue to keep a creator on the platform. Without compensation, the creator has to sell content; with compensation, her share of revenue from AI-generated content sales can cover her participation cost. Dependence on a creator's training data can thus sustain her participation without protecting her content from competition.

2 Model Setup

We model a platform where consumers choose between purchasing content from two quality-differentiated creators and generating customized content with AI that the platform trains on the creators' content.[3] The platform sets a common price for creators' content and a separate price for AI-generated content, while creators make participation and, under compensation, data sharing decisions. In this section, we begin by describing consumer preferences and the effects of training data on AI quality and prompting effort. We then describe creators' payoffs and decisions, followed by the platform's pricing problem. We conclude by describing the timing of the game before proceeding to the main analysis in Section 3.

2.1 Consumers

A unit mass of consumers, each with unit demand, is uniformly distributed on the Hotelling line $[0, 1]$, with creator L located at 0 and creator H at 1. Creator j produces content of quality q_j, where $0 < q_L < q_H < 1$. The platform sets a common price P for creators' content and a separate price P_{AI} for AI-generated content of quality q_{AI}. A consumer at location x incurs mismatch costs of x and $1 - x$ when purchasing from L and H, respectively, with the transportation cost normalized to one. A consumer can use AI to generate content that perfectly matches her taste but incurs prompting effort $\kappa(x)$, denoted by κ henceforth. Each consumer chooses the option that yields the highest utility, including an outside option with utility normalized to zero, such as general-purpose AI.[4]

The resulting consumer utilities are:

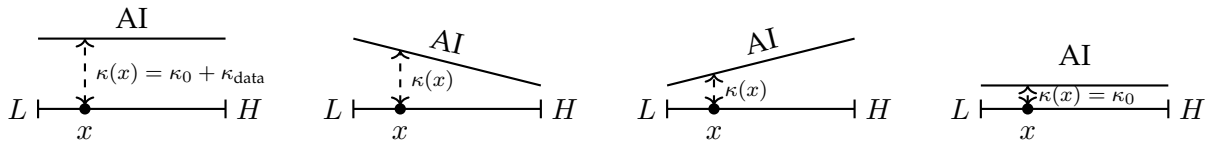
$$\begin{aligned} &U_L = q_L - P - x, && \text{if the consumer purchases from } L, \\ &U_H = q_H - P - (1 - x), && \text{if the consumer purchases from } H, \\ &U_{AI} = q_{AI} - P_{AI} - \kappa, && \text{if the consumer generates a product using AI.} \end{aligned} \tag{1}$$

Prompting effort κ. Prompting effort consists of two components. The base component $\kappa_0 \geq 0$ captures the generative AI model's inherent steerability (Vafa et al. 2025), determined by its architecture, pretraining, and post-training rather than by creators' training data. A larger κ_0 corresponds to a less steerable model. The data component $\kappa_{\text{data}} \geq 0$ captures the additional effort required when the model lacks training data close to the consumer's taste. Letting $s_j \in \{0, 1\}$

[3] We abstract away from creator-facing AI tools that lower creators' production costs, such as AI editing assistants.

[4] We adopt two tie-breaking rules for indifferent consumers: consumers favor a platform product over the outside option and creators' content over AI-generated content. When $\kappa_{\text{data}} = 1$, the second rule affects aggregate demand only when the platform trains AI on one creator's content.

Figure 1: Prompting Effort with Different Training Data



(a) Before training (b) Trained on H (c) Trained on L (d) Trained on both

Note. The prompting effort $\kappa(x)$ is the vertical distance from a consumer at x to the line representing the AI generated content, as in Eq. (2). Training on one creator's content reduces prompting effort for consumers with tastes closer to that creator's style. Training on both creators' content eliminates the taste-dependent component of prompting effort. The panels illustrate $\kappa_0 > 0$; the main analysis sets $\kappa_0 = 0$.

indicate whether the platform trains AI on creator j's content, we model prompting effort as:

$$\kappa(x; s_H, s_L) \;=\; \begin{cases} \kappa_0 + \kappa_{\text{data}}, & s_H = 0 \text{ and } s_L = 0, \\ \kappa_0 + \kappa_{\text{data}}\, x, & s_H = 0 \text{ and } s_L = 1, \\ \kappa_0 + \kappa_{\text{data}}(1 - x), & s_H = 1 \text{ and } s_L = 0, \\ \kappa_0, & s_H = 1 \text{ and } s_L = 1. \end{cases} \tag{2}$$

Training on one creator's content reduces prompting effort most for consumers whose tastes are closer to that creator's style. Figure 1 illustrates Eq. (2) visually.

For the main analysis, we assume perfect steerability ($\kappa_0 = 0$) when the platform trains AI on both creators' content and consider imperfect steerability ($\kappa_0 > 0$) in Section 4.1. We also set $\kappa_{\text{data}} = 1$ to treat prompting effort and content mismatch symmetrically.[5] When the platform trains AI on only one creator's content (as illustrated in Figures 1b and 1c), a consumer located z units from that creator incurs z units of mismatch cost when purchasing her content and z units of prompting effort when using AI to generate content that perfectly matches her taste.

Quality of AI generated content q_{AI}. Scaling studies show that AI quality improves with training data size (Kaplan et al. 2020, Hoffmann et al. 2022) and with training on specific high-quality examples (e.g., professional writing and photographs) (Zhou et al. 2023, Dai et al. 2023).

Hence, we model the quality of AI-generated content as increasing with the number of creators whose content the platform trains AI on and the quality of their content:

$$q_{AI} \;=\; q^{\text{base}} \;+\; \tau_H\, q_H\, s_H \;+\; \tau_L\, q_L\, s_L. \tag{3}$$

Here, $q^{\text{base}} \geq 0$ is the generative AI model's quality before training on creators' content, and $\tau_j \in (0, 1)$ measures how much creator j's content contributes to AI quality.

When creator j exits, the platform loses access to her data for AI training, as when a renewable

[5] $\kappa_{\text{data}} > 1$ may arise if, for example, the consumer's cost of eliminating the mismatch between the creator's content and her own taste by prompting the platform's AI is higher than that using her own tools (e.g., Adobe Photoshop).

training license expires or the flow of new content ceases. If the platform trained AI on her content, exit reduces AI quality by $\tau_j q_j$ and eliminates the associated reduction in prompting effort. For the main analysis, we set $\tau_H = \tau_L = \frac{1}{2}$, giving both creators' content equal weight in determining AI quality. We consider heterogeneous transfer rates ($\tau_H \neq \tau_L$) in Section 4.2.

Consumers can already access general-purpose AI outside the platform. To attract paying consumers, the platform's AI must offer an advantage over these alternatives. On creative content marketplaces, training on creators' content can provide this advantage by improving AI quality and reducing prompting effort. This motivates the following assumption:

Assumption 1 (Training data dependence)**.** *Without training on creators' content, the platform's AI cannot sustain positive sales revenue.*

When the platform does not train AI on either creator's content, consumers incur prompting effort $\kappa_0 + \kappa_{\text{data}} = 1$ under our main specification. Assumption 1 then requires $q^{\text{base}} \leq 1$.

2.2 Creators

Each creator $j \in \{L, H\}$ decides whether to participate on the platform, $e_j \in \{0, 1\}$. A creator who exits ($e_j = 0$) earns zero and withdraws her content and training data from the platform. A participating creator ($e_j = 1$) incurs a fixed cost ϕq_j, where $\phi > 0$, reflecting the higher cost of maintaining higher-quality content. Marginal production costs are normalized to zero.

We consider two regimes that differ in whether the platform compensates creators for training data. In both regimes, creators receive a share $\alpha \in (0, 1)$ of revenue from their own content sales. In the *no compensation regime* (NC), the platform trains AI on participating creators' content without additional payment.[6] Training data use is a condition of participation, so $s_j = e_j$. In the *compensation regime* (C), participating creators can also earn revenue by contributing their content for AI training. Each creator independently chooses whether to contribute, $s_j \in \{0, 1\}$, with $s_j \leq e_j$. The platform shares a fraction α of AI revenue equally among contributors. Creators who do not contribute can continue selling their content but receive no AI revenue.

A participating creator's gross payoff is:

$$V_j = \begin{cases} \alpha D_j P, & \text{under NC,} \\ \alpha \left(D_j P + \dfrac{s_j}{s_H + s_L} D_{AI} P_{AI} \right), & \text{under C,} \end{cases} \tag{4}$$

where D_j and D_{AI} denote demand for creator j's content and AI content, respectively.[7] A participating creator's net payoff is $V_j - \phi q_j$, while an exiting creator earns zero.

[6] Without compensation, allowing creators to opt out eliminates their incentive to share training data: doing so improves AI without increasing their direct sales revenue. Hence, we model the no compensation regime as one in which platforms obtain training rights as a condition of participation, as in Reddit and Meta (Reddit 2026, Meta 2023).

[7] We adopt the convention $0/0 = 0$. We extend the analysis to quality-weighted revenue sharing in Section 4.3.

Our focus is on whether AI can fully displace a creator who remains on the platform. We say a participating creator is *fully displaced* when $D_j = 0$. Under compensation, such a creator can remain on the platform as a training *data supplier* if her AI revenue covers her participation cost. Without compensation, a creator who loses all direct demand cannot cover her participation cost and exits. This distinction separates displacement from exit: a displaced creator may participate and continue supplying training data, whereas an exiting creator withdraws both her content and her training data from the platform.

2.3 Platform

The platform sets a common price $P \geq 0$ for creators' content (following prevailing norms on creative content marketplaces) and a price $P_{AI} \geq 0$ for AI-generated content. It retains a fraction $1 - \alpha$ of revenue from creators' content, where α is a fixed platform-specific parameter.

The platform's share of AI revenue depends on the compensation regime. Under NC, it retains all AI revenue. Under C, it pays a fraction α of AI revenue to creators who contribute training data.[8] Letting $D = D_H + D_L$ denote total demand for creators' content, the platform's profit in each regime is given by:

$$\pi = \begin{cases} (1-\alpha)DP + D_{AI}P_{AI}, & \text{under NC,} \\ (1-\alpha)(DP + D_{AI}P_{AI}), & \text{under C.} \end{cases} \tag{5}$$

When the platform does not train AI on either creator's content, AI generates no positive sales revenue under Assumption 1, so the expression in Eq. (5) under C still applies.

2.4 Timing and Equilibrium

The game proceeds in four stages, as illustrated in Figure 2. In Stage 0, the platform decides whether to integrate AI and, if so, whether to compensate creators for their data.[9] In Stage 1, the platform sets the price P for creators' content and, if AI is integrated, the price P_{AI}. In Stage 2, creators simultaneously decide whether to participate and, under C, whether to contribute their content for training. In Stage 3, consumers choose the option that maximizes their utility.

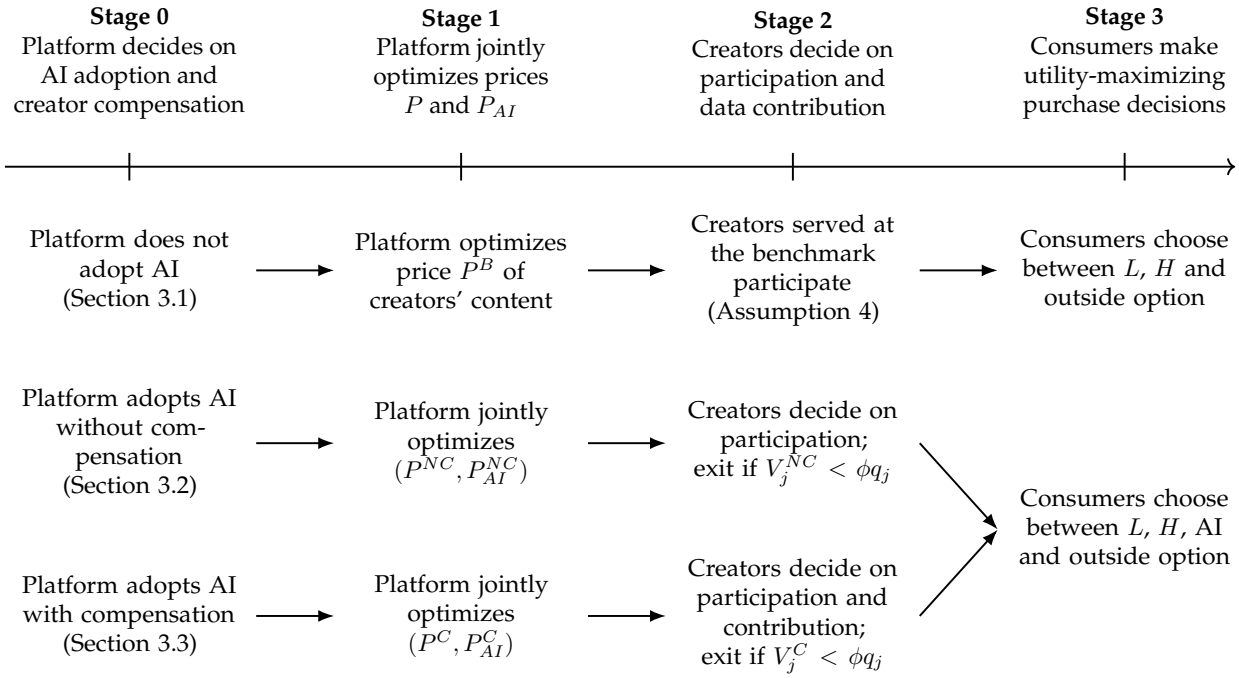
We solve for subgame-perfect equilibria by backward induction. Given the platform's prices, creators' decisions in Stage 2 form a simultaneous-move game, for which we consider pure-strategy Nash equilibria, and a mixed-strategy equilibrium where no pure-strategy equilibrium satisfies Assumption 2.[10] The platform anticipates these decisions and the resulting consumer demand

[8] We use the same revenue sharing rate for creators' content and AI-generated content to avoid introducing an additional parameter and to isolate the effect of compensating creators for their training data.

[9] Section 3.5 also considers mandated compensation when comparing private and social incentives.

[10] In the benchmark and NC regimes, creators have a pure-strategy equilibrium at every price (Appendix B.1). In the C regime, some prices have no pure-strategy equilibrium. At these prices, we consider a mixed-strategy equilibrium in which creators randomize their participation and training data contribution decisions. However, in every equilibrium, the platform chooses prices that maximize its profit and at which creators' decisions form a pure-strategy equilibrium.

Figure 2: Stages of the Game



when setting prices to maximize profit. Since creator exit withdraws training data, the platform's pricing decisions also affect AI quality and consumers' prompting effort.

The Stage 2 game requires two equilibrium selection rules. The first concerns participation when a creator's gross payoff V_j equals her participation cost ϕq_j. Since the platform sets prices before creators decide whether to participate, its optimal prices may leave a creator indifferent between participating and exiting. If an indifferent creator exits, the platform loses access to her content and training data. We resolve this indifference in favor of participation, allowing the platform to retain creators whose revenue just covers their participation costs.

Assumption 2 (Weak participation)**.** *A creator participates when she is indifferent between participating and exiting.*

The second rule concerns coordination between creators over their training data contribution decisions under compensation. At some prices, AI generates positive sales revenue only when both creators contribute. If neither contributes, a creator who contributes alone earns no additional AI revenue, so neither has an incentive to deviate from $(s_H, s_L) = (0, 0)$. However, when both contribute, AI generates revenue that is shared between them, making both creators better off. Both $(0, 0)$ and $(1, 1)$ may then be Nash equilibria. We assume that creators coordinate on the equilibrium that Pareto dominates the others when one exists.

Assumption 3 (Pareto coordination)**.** *When the Stage 2 game has multiple pure-strategy Nash equilibria*

that satisfy Assumption 2 and one Pareto dominates the others (no creator is worse off and some creator is better off), creators select that equilibrium.

We use the market before the platform integrates AI as the *benchmark*. Our interest is in whether AI displaces creators who would otherwise sell content on the platform. We focus on markets in which both creators initially participate and attract positive demand. Two factors can prevent this outcome. If the high-quality creator has a sufficiently large quality advantage, the platform may prefer to set a price at which only her content attracts demand. Even when both creators attract demand, a creator may exit if her sales revenue does not cover her participation cost. We impose two restrictions to focus on markets in which both creators are active before AI integration.

Assumption 4 (Benchmark participation)**.** *Creator qualities satisfy* $q_H/q_L < \rho^\star \equiv 1+\sqrt{2}$*, and the participation cost parameter satisfies:*

$$\phi \leq \phi^B \equiv \frac{\alpha(3q_L - q_H)(q_H + q_L)}{16q_L}.$$

The restriction $q_H/q_L < \rho^\star$ ensures that the platform's optimal price generates positive demand for both creators in the benchmark. We derive both thresholds in Proposition 1 and its proof.

3 Main Analysis

We characterize how compensation for training data changes the platform's pricing decisions, creator participation, and the allocation of consumer demand between creators' content and AI-generated content. We first establish the benchmark before AI integration and then derive the equilibria without and with compensation, before examining the platform's compensation decision and comparing welfare across the two regimes. We mark equilibrium objects with the superscripts B, NC, and C for the benchmark, the regime without compensation, and the regime with compensation, as in π^B, P^{NC}, and V_j^C. For clarity, we also summarize the notation used throughout the paper in Appendix A.

3.1 Benchmark: The Market Before AI Integration

Before the platform integrates a generative AI system, consumers choose between purchasing content from H or L and the outside option (e.g., prompting a general-purpose generative AI). The platform sets a common price $P \geq 0$, creators decide whether to participate, and the platform earns $\pi^B(P) = (1-\alpha)(D_H + D_L)P$. When both creators participate, the consumer indifferent between them is located at $x_{HL} = 1/2 - (q_H - q_L)/2$ and obtains utility $(q_H + q_L - 1)/2 - P$. The market is fully covered when $P \leq (q_H + q_L - 1)/2$ and not fully covered otherwise.

The platform faces a tradeoff between price and market coverage. Consumers near the middle of the Hotelling line incur greater mismatch costs, so attracting them requires lowering P. The platform may prefer to let some consumers choose the outside option rather than reduce the price.

When the quality advantage of H is sufficiently large, maintaining positive demand for L requires a further price reduction, making it more profitable to set a price at which only H attracts demand.

Proposition 1 (Benchmark)**.** *There is a quality ratio threshold $\rho^\star = 1 + \sqrt{2}$ such that the platform's optimal price generates positive demand for both creators under partial coverage when $q_H/q_L < \rho^\star$ and $\phi \leq \phi^B$, and positive demand only for the high-quality creator when $q_H/q_L > \rho^\star$ and $\phi \leq \alpha q_H/4$. When $q_H/q_L = \rho^\star$, the platform earns the same maximum profit whether both creators attract positive demand or only H does, provided $\phi \leq \alpha q_L/8$.*

Proof. See Appendix B.3. □

Proposition 1 distinguishes two cases according to which creators participate. When $q_H/q_L < \rho^\star$, the platform sets $P^B = (q_H+q_L)/4$, generating demands $D_L^B = (3q_L-q_H)/4$ and $D_H^B = (3q_H-q_L)/4$. A mass $1-(q_H+q_L)/2$ of consumers chooses the outside option. Platform profit is $\pi^B = (1-\alpha)(q_H+q_L)^2/8$, and each creator earns $V_j^B = \alpha D_j^B P^B$. If L's revenue covers her participation cost, H's does as well, since $V_L^B/q_L < V_H^B/q_H$. When $q_H/q_L > \rho^\star$, the platform instead sets $P^B = q_H/2$, generating demand $D_H^B = q_H/2$ and profit $\pi^B = (1-\alpha)q_H^2/4$, conditional on H's participation. The low-quality creator attracts no demand and exits.

By bounding q_H/q_L and ϕ from above, Assumption 4 restricts our attention to the first case, in which both creators participate and attract positive demand before AI integration. This allows us to examine whether AI displaces creators who would otherwise sell content and whether they remain as training data suppliers.

3.2 Without Compensation (NC)

Under the NC regime, the platform trains AI on participating creators' content and retains all AI revenue, whereas creators earn revenue only from their own content. Hence, $s_j = e_j$ and $V_j^{NC} = \alpha D_j P$. When consumers purchase AI-generated content instead of purchasing from creators, creators lose revenue and may exit, withdrawing their training data. The platform sets prices to balance AI revenue against the need to retain creators' training data. We formalize this tradeoff by first deriving consumer demand assuming both creators participate, then showing that the platform earns lower profit if a creator exits, and finally deriving its optimal prices.

We begin by deriving consumer demand at given prices when the platform trains AI on both creators' content, $(s_H, s_L) = (1,1)$. Since $q_{AI} = q^{\text{base}} + (q_H+q_L)/2$ and $\kappa = 0$, every consumer obtains $U_{AI} = q_{AI} - P_{AI}$ from the AI. The following lemma characterizes consumer demand conditional on both creators participating.

Lemma 1 (Consumer demand under NC)**.** *Suppose both creators participate, so $e_H = e_L = 1 = s_H = s_L$. Let $\widetilde{P}_{AI} \equiv P_{AI} - q^{base}$, $\delta \equiv \widetilde{P}_{AI} - P$, and $g \equiv (q_H - q_L)/2$. If $\widetilde{P}_{AI} > (q_H+q_L)/2$, consumers do not purchase AI-generated content, and demand for creators' content follows the benchmark before AI integration at the same price P (Section 3.1). If instead $\widetilde{P}_{AI} \leq (q_H+q_L)/2$, the market is fully covered, with consumer demand given by:*

Range of δ	D_L	D_{AI}	D_H
$\delta \geq 1/2$	$1/2 - g$	0	$1/2 + g$
$g < \delta < 1/2$	$\delta - g$	$1 - 2\delta$	$\delta + g$
$-g < \delta \leq g$	0	$1 - \delta - g$	$\delta + g$
$\delta \leq -g$	0	1	0

Proof. See Appendix B.4. □

We next show that the platform obtains higher profit by retaining both creators than by allowing either creator to exit. Suppose L exits (an analogous argument applies if H exits). The platform can no longer sell L's content or train AI on L's content. AI quality falls by $q_L/2$, by Equation (3). A consumer at x incurs mismatch cost $1 - x$ when purchasing from H and prompting effort $1 - x$ when purchasing AI-generated content, by Equations (1) and (2). Consumers thus choose between purchasing from H and purchasing AI-generated content based only on quality and price, which do not vary across consumers.

If the platform prices AI so that consumers obtain strictly higher utility from AI-generated content than from H's content, every consumer who would purchase from H purchases AI-generated content instead. With no demand for her content, H cannot cover the cost of participation and would prefer to exit. If neither creator participates, the platform has no content to sell and cannot train AI on any platform content. Consumers choose the outside option instead of paying for AI-generated content under Assumption 1, so the platform's profit is zero.

To retain H, the platform must set prices so that consumers obtain at least as much utility from H's content as from AI-generated content. At these prices, consumers purchase from H when indifferent between the two offerings,[4] so $D_{AI} = 0$. With demand only for H's content, the platform's profit is at most $(1 - \alpha)q_H^2/4$. This bound is strictly below the benchmark profit $\pi^B = (1 - \alpha)(q_H + q_L)^2/8$ under Assumption 4.

Thus, the platform is more profitable when it retains both creators than when either creator exits. After L's exit, the platform's profit is strictly below π^B whether H remains or exits. Analogously, the platform's profit is strictly below π^B if H exits. The platform can obtain a profit of π^B by retaining both creators, pricing their content at P^B, and pricing AI out of the market.

Retaining both creators allows the platform to additionally earn AI revenue without reducing demand for creators' content. The platform can price creators' content and AI as follows:

$$P = P^B = \frac{q_H + q_L}{4}, \qquad P_{AI} = q^{\text{base}} + \frac{q_H + q_L}{2}.$$

At these prices, consumers who previously purchased from creators continue to do so. Consumers who previously chose the outside option now purchase AI-generated content. Both creators receive the same revenue as in the benchmark before AI integration. This corresponds to row 2 of the table in Lemma 1, which we call the *three-product region*.

We now use consumer demand in the three-product region of Lemma 1 to derive how the platform prices creators' and AI-generated content to maximize profit while retaining both creators.

Proposition 2 (Equilibrium prices and demand under NC)**.** *Let $s \equiv \sqrt{q_L^2 - 4\phi q_L/\alpha}$. In every subgame-perfect equilibrium under NC (Section 2.4), both creators participate. The platform sets prices such that L's participation constraint binds, with $V_L^{NC} = \phi q_L$. H's constraint is slack, with $V_H^{NC} = \alpha(q_L + s)(2q_H - q_L - s)/4 > \phi q_H$. The platform sets prices within the three-product region, with $\delta = (q_H - s)/2 \in (g, 1/2)$. Both creators and AI-generated content attract positive demand. The equilibrium prices and demand are:*

$$P^{NC} = \frac{q_L + s}{2}, \qquad P_{AI}^{NC} = q^{base} + \frac{q_H + q_L}{2},$$
$$D_L^{NC} = \frac{q_L - s}{2}, \qquad D_H^{NC} = \frac{2q_H - q_L - s}{2}, \qquad D_{AI}^{NC} = 1 - q_H + s.$$

Proof. See Appendix B.5. □

Thus, when the platform does not compensate creators for training data, it sets the price of creators' content such that the low-quality creator just covers her participation cost, while the high-quality creator earns a positive net payoff. Although AI can perfectly match consumers' tastes, the platform prices creators' content below AI-generated content to preserve their demand and retain their training data. At the equilibrium AI price, the platform cannot raise the price of creators' content further without inducing the low-quality creator to exit and withdraw her training data.

Finally, we derive the platform's equilibrium profit in Corollary 1 by substituting the equilibrium prices and demand from Proposition 2 into the NC profit expression in Equation (5).

Corollary 1 (Platform profit under NC)**.** *The platform's profit π^{NC} is given by:*

$$\pi^{NC} = (1 - \alpha)\frac{(q_H - s)(q_L + s)}{2} + (1 - q_H + s)\left(q^{base} + \frac{q_H + q_L}{2}\right). \tag{6}$$

π^{NC} is strictly decreasing in ϕ and strictly increasing in q^{base}. Further, $\partial\pi^{NC}/\partial q^{base} = D_{AI}^{NC}$.

Proof. See Appendix B.5. □

Corollary 1 shows that higher participation costs reduce the platform's profit. To retain L, the platform lowers the price of creators' content, attracting enough additional consumers to raise her revenue and cover her higher participation cost. However, the platform receives less revenue per purchase from creators. In addition, consumers who switch from purchasing AI-generated content to purchasing from creators pay a lower price, and the platform now shares that revenue with creators. Retaining creators' training data thus becomes less profitable for the platform as their participation costs rise.

Proposition 2 and Corollary 1 also show that better AI (higher q^{base}) increases the platform's profit, but does not further displace creators' demand when they are not compensated for their

training data. As AI improves, consumers are willing to pay more for AI-generated content, but the platform must still preserve creators' demand to retain their training data. The platform maximizes profit by raising the AI price to capture the full increase in willingness to pay while charging the same price for creators' content.

Comparison with the benchmark. Relative to the benchmark before AI integration, the platform earns higher profit under NC: $\pi^{NC} > \pi^{B}$ (Corollary 4). It also prices creators' content weakly higher ($P^{NC} - P^{B} = (s-g)/2 \geq 0$). AI serves all consumers who chose the outside option before AI integration, a mass $1-(q_H+q_L)/2$, along with a mass $s-g \geq 0$ of consumers who previously purchased from creators (where $s=g$ when $\phi=\phi^B$). The low-quality creator's payoff under NC is weakly lower than in the benchmark before AI integration, whereas the high-quality creator's payoff can rise or fall depending on whether the higher price offsets the reduction in demand.

3.3 With Compensation (C)

We now consider the compensation regime (C), in which the platform shares AI revenue with creators who contribute training data. A contributing creator can remain on the platform after losing all demand for her content if her share of AI revenue covers her participation cost. The platform can thus increase demand for AI-generated content while retaining creators and their training data. Participating creators choose whether to contribute training data, so the platform may receive data from both creators, one creator, or neither creator. We begin by characterizing consumer demand for each combination of creators' contribution decisions in Lemma 2.

Lemma 2 (Consumer demand under C)**.** *Let (s_H, s_L) indicate which creators contribute training data. For each combination of these decisions, the table shows AI quality, prompting effort, and consumer demand:*

(s_H, s_L)	q_{AI}	$\kappa(x)$	Consumer demand
$(1,1)$	$q^{\text{base}} + (q_H+q_L)/2$	0	As in Lemma 1
$(1,0)$	$q^{\text{base}} + q_H/2$	$1-x$	$D_H D_{AI} = 0$
$(0,1)$	$q^{\text{base}} + q_L/2$	x	$D_L D_{AI} = 0$
$(0,0)$	q^{base}	1	$D_{AI} P_{AI} = 0$

When only H contributes training data ($(s_H, s_L) = (1,0)$), H's content and AI-generated content cannot both attract positive demand. Analogously, when only L contributes training data ($(s_H, s_L) = (0,1)$), L's content and AI-generated content cannot both attract positive demand. When neither creator contributes training data, the platform earns no revenue from AI-generated content.

Proof. See Appendix B.6. □

As in our discussion of Lemma 1, Lemma 2 helps explain why the platform obtains its highest profit when both creators contribute training data. Training on both creators' content improves

AI quality and reduces the prompting effort required to generate content matching consumers' tastes. The platform can capture consumers' greater willingness to pay by charging a higher price for AI-generated content than when only one creator contributes training data. It retains the same share of AI revenue with two contributors as with one. The platform's payments to creators for training data can cover their participation costs even when their content attracts no demand.

We now use consumer demand in Lemma 2 to derive how the platform prices creators' and AI-generated content to maximize profit while retaining both creators' training data. Without compensation, the platform preserves demand for both creators' content (Proposition 2). Proposition 3 shows that, with compensation, the platform instead sets prices at which AI fully displaces the low-quality creator. Both creators remain on the platform and contribute training data.

Proposition 3 (Equilibrium prices and demand under C)**.** *In every equilibrium, both creators participate on the platform and contribute their content for AI training, so $e_H = e_L = s_H = s_L = 1$. The following two cases give the equilibrium prices and demand depending on whether the AI model's quality before training on creators' content, q^{base}, is strictly below or greater than or equal to $(q_H - q_L)/2 = \overline{q}^{base}$:*

1. ***Partial displacement (CQ1):*** *If $q^{base} < \overline{q}^{base}$, the platform sets prices at which AI fully displaces demand for the low-quality creator's content while consumers continue purchasing from the high-quality creator. The equilibrium prices and demand are:*

$$P^C = \frac{3q_H + q_L}{4} + \frac{q^{base}}{2}, \qquad P^C_{AI} = q^{base} + \frac{q_H + q_L}{2},$$
$$D^C_L = 0, \qquad D^C_H = \frac{q_H - q_L}{4} - \frac{q^{base}}{2}, \qquad D^C_{AI} = 1 - D^C_H.$$

2. ***Full displacement (CQ2):*** *If $q^{base} \geq \overline{q}^{base}$, the platform sets prices at which AI serves the entire market and neither creator's content attracts positive demand. The equilibrium prices and demand are:*

$$P^C \geq q_H, \qquad P^C_{AI} = q^{base} + \frac{q_H + q_L}{2},$$
$$D^C_L = 0, \qquad D^C_H = 0, \qquad D^C_{AI} = 1.$$

The platform can set any price $P^C \geq q_H$ for creators' content. Neither creator earns revenue from her own content at these prices, so raising the price further leaves consumer demand, the platform's profit, and creators' payoffs unchanged.

Proof. See Appendix B.8. □

Proposition 3 formalizes why compensation allows the platform to retain creators' training data without preserving demand for their content. The platform can price creators' content high enough to fully displace the low-quality creator's demand because the low-quality creator's share of AI revenue covers her participation cost. When AI-generated content has lower quality than the

high-quality creator's content, consumers with tastes close to the high-quality creator's style pay more to purchase from the high-quality creator than to purchase AI-generated content.

As AI improves, the platform raises the AI price to capture consumers' greater willingness to pay for AI-generated content. The platform then retains more revenue from each consumer purchasing AI-generated content, making it more profitable to raise the price of creators' content and accept lower demand for the high-quality creator's content. Once AI-generated content matches the high-quality creator's content in quality, AI serves the entire market, and both creators remain as training data suppliers because their shares of AI revenue cover their participation costs. At the equilibrium prices, each creator prefers to contribute when the other contributes. A creator who opts out of AI training loses her share of AI revenue without increasing demand for her content.

Comparison with the benchmark. Relative to the benchmark before AI integration, the platform earns higher profit under C: $\pi^C > \pi^B$ (Corollary 4). It also prices creators' content strictly higher ($P^C > P^B$). AI serves all consumers who chose the outside option before AI integration, a mass $1-(q_H+q_L)/2$, along with a mass $(q_H+q_L)/2 - D_H^C > 0$ of consumers who previously purchased from creators (including at $\phi = \phi^B$). Both creators' payoffs under C are strictly higher than in the benchmark before AI integration because their shares of AI revenue more than offset the reduction in revenue from their own content.

Comparison with the regime without compensation (NC). Whether the platform earns higher profit under C or NC depends on creators' participation costs and AI quality (Corollary 4). It prices creators' content strictly higher under C ($P^C > P^{NC}$). The platform trains AI on both creators' content in both regimes ($s_H = s_L = 1$). The quality and price of AI-generated content are also identical in both regimes ($P_{AI}^C = P_{AI}^{NC} = q^{\text{base}} + (q_H+q_L)/2$).

We now examine how compensation changes both demand for creators' content and the revenue that sustains their participation on the platform. We first compare demand for creators' and AI-generated content under C and NC (Corollary 2). Under C, a creator who loses all demand for her content must cover her participation cost through her share of AI revenue. We thus examine whether displaced creators earn enough to remain on the platform and how their payoffs change as AI improves (Corollary 3).

Corollary 2 (Compensation increases AI displacement)**.** *Equilibrium demand under C and NC satisfies:*

$$D_L^C = 0 < D_L^{NC}, \qquad D_H^C < D_H^{NC}, \qquad D_{AI}^C > D_{AI}^{NC}.$$

Proof. See Appendix B.8. □

Corollary 2 shows that compensation increases AI displacement even though both creators contribute their content for AI training in both regimes, and the quality and price of AI-generated content are identical across regimes. Without compensation, the platform must preserve enough

demand for the low-quality creator's content for her revenue to cover her participation cost; otherwise, she exits and withdraws her training data. With compensation, her share of AI revenue covers her participation cost even when consumers no longer purchase from her. The platform can thus raise the price of creators' content, inducing more consumers to purchase AI-generated content while retaining creators' training data.

Displaced creators no longer earn revenue from their own content and must cover their participation costs through their share of AI revenue. We next characterize displaced creators' net payoffs and how AI improvements affect the low-quality creator's payoff in Corollary 3.

Corollary 3 (Displaced creators remain as training data suppliers)**.** *At the equilibrium prices of Proposition 3, each creator $j \in \{H, L\}$ earns strictly more than her participation cost, $V_j^C > \phi q_j$, under both partial displacement (CQ1) and full displacement (CQ2). In particular, the low-quality creator's content attracts no demand, but she earns a positive net payoff, with $D_L^C = 0$, $V_L^C > \phi q_L$, and V_L^C strictly increasing in q^{base}. Under full displacement (CQ2), neither creator's content attracts positive demand, and both creators earn only their shares of AI revenue, with $V_H^C = V_L^C = \alpha q_{AI}/2$.*

Proof. See Appendix B.8. □

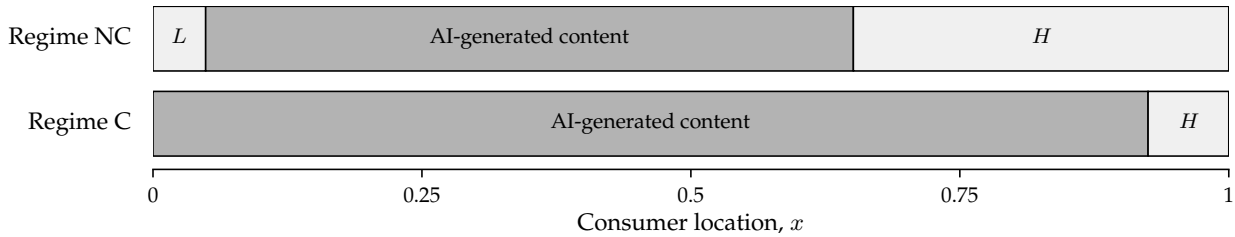
Corollary 3 shows that compensation allows creators to earn positive net payoffs from supplying training data after losing all demand for their content. Consumers no longer purchase from the low-quality creator under compensation, as illustrated in Figure 3, but the platform continues training AI on her content and sharing AI revenue with her. Her share of AI revenue exceeds her participation cost, making it profitable to remain on the platform solely as a training data supplier.

As AI improves, the low-quality creator benefits even though consumers no longer purchase from her. Consumers become willing to pay more for AI-generated content, and the platform raises the AI price. Under partial displacement (CQ1), the platform also raises the price of creators' content, inducing more consumers to purchase AI-generated content instead of purchasing from the high-quality creator. The higher AI price and greater demand for AI-generated content increase the AI revenue shared with the low-quality creator and thus her net payoff. Once AI serves the entire market under full displacement (CQ2), the high-quality creator also earns revenue solely from supplying training data. Further AI improvements allow the platform to charge more for AI-generated content, so both creators earn more from supplying training data.

3.4 The Platform's Compensation Decision

We now examine when the platform chooses to compensate creators for training data. Without compensation, the platform must preserve demand for creators' content to retain their training data. With compensation, the platform shares AI revenue with creators and can retain their training data even when consumers no longer purchase from them. Corollary 4 compares the platform's profit across the two regimes and the benchmark before AI integration.

Figure 3: Consumer Demand With and Without Compensation



Note. Equilibrium consumer demand at $q_H = 0.7$, $q_L = 0.4$, $\alpha = 0.3$, $\phi = \phi^B/2$, and $q^{\text{base}} = 0$. Segments labeled L and H show consumers purchasing from the corresponding creator. Both creators contribute training data in both regimes. Under C, consumers no longer purchase from L, who remains on the platform as a training data supplier.

Corollary 4 (The platform's compensation decision)**.** *The platform earns higher profit under both C and NC than in the benchmark before AI integration, with $\pi^C > \pi^B$ and $\pi^{NC} > \pi^B$. Moreover, π^{NC} strictly decreases in ϕ, whereas π^C is independent of ϕ. For each q^{base}, there is a threshold $\phi^\star(q^{base}) \in [0, \phi^B]$ such that the platform prefers NC when $\phi < \phi^\star$ and C when $\phi > \phi^\star$. If $0 < \phi^\star < \phi^B$, the platform is indifferent between the two regimes at $\phi = \phi^\star$, with $\pi^{NC}(\phi^\star) = \pi^C$. If $\phi^\star = 0$, the platform prefers C for every $0 < \phi \leq \phi^B$; if $\phi^\star = \phi^B$, it weakly prefers NC throughout this range.*

Proof. See Appendix B.9. □

Corollary 4 shows that the platform may prefer not to compensate creators (NC) even though compensating creators (C) allows it to raise the price of creators' content and increase demand for AI-generated content. Under C, the platform shares AI revenue with creators, reducing the revenue it retains from each consumer purchasing AI-generated content. Under NC, it retains all AI revenue but must preserve enough demand for creators' content to cover their participation costs and retain their training data. When participation costs are low, the low-quality creator can cover her cost even when few consumers purchase from her. The platform can then prefer NC because preserving the demand needed to retain creators reduces its profit by less than sharing AI revenue under C.

Higher creator participation costs can make the platform prefer C. Under NC, the platform lowers the price of creators' content so the low-quality creator can cover her higher participation cost. More consumers then purchase from creators instead of purchasing AI-generated content. They pay a lower price, and the platform retains only the share $1 - \alpha$ of that revenue, reducing its profit under NC. Under C, the platform can retain creators without changing its optimal prices or the AI revenue shared with them throughout $0 < \phi \leq \phi^B$, so its profit remains unchanged. The platform thus prefers C when $\phi > \phi^\star$. Figure 4 illustrates an example of this comparison.

The platform's compensation decision also depends on AI quality. As AI improves, the platform raises the price of AI-generated content by the same amount in both regimes. Under C, more consumers purchase AI-generated content than under NC, but the platform retains a smaller share of AI revenue. When the platform is initially indifferent between NC and C, the greater demand under C allows it to benefit more from the higher price despite sharing AI revenue with creators.

Better AI can thus make the platform prefer C at lower creator participation costs.[11]

Figure 5 illustrates when the platform chooses to compensate creators and which creators continue to attract demand. The platform prefers NC at lower creator participation costs and C at higher costs, with the solid curve marking where it is indifferent. Where the platform prefers C, the dashed line separates partial displacement (CQ1) from full displacement (CQ2).

3.5 Welfare Analysis

We now examine the welfare consequences of the greater AI displacement induced by compensating creators for training data. We first compare creators' net payoffs and consumer surplus under C and NC, then assess whether the resulting allocation increases total surplus. We determine when the platform's private compensation decision selects the regime with lower total surplus.

For each regime $r \in \{NC, C\}$, let CS^r denote consumer surplus, let $R_j^r \equiv V_j^r - \phi q_j$ denote creator j's net payoff, and define total surplus as $\mathrm{TS}^r \equiv \pi^r + R_H^r + R_L^r + \mathrm{CS}^r$.

Proposition 4 (Welfare effects of compensation)**.** *Compensation has the following effects on welfare:*

1. ***Creators' net payoffs.*** *Both creators earn strictly higher net payoffs under C than NC:*

$$R_L^{NC} = 0 < R_L^C, \qquad R_H^{NC} < R_H^C.$$

   Both differences in net payoffs increase strictly with q^{base}.

2. ***Consumer surplus.*** *Consumer surplus is strictly lower under C than NC:*

$$\mathrm{CS}^C < \mathrm{CS}^{NC}.$$

   The reduction in consumer surplus increases strictly with q^{base} for $q^{base} < \overline{q}^{base}$, where $\overline{q}^{base} = (q_H - q_L)/2$. For $q^{base} \geq \overline{q}^{base}$, AI serves the entire market under C, and the platform captures consumers' full willingness to pay, so $\mathrm{CS}^C = 0$ and the reduction remains constant.

3. ***Total surplus.*** *Total surplus is strictly higher under C than NC:*

$$\mathrm{TS}^C > \mathrm{TS}^{NC}.$$

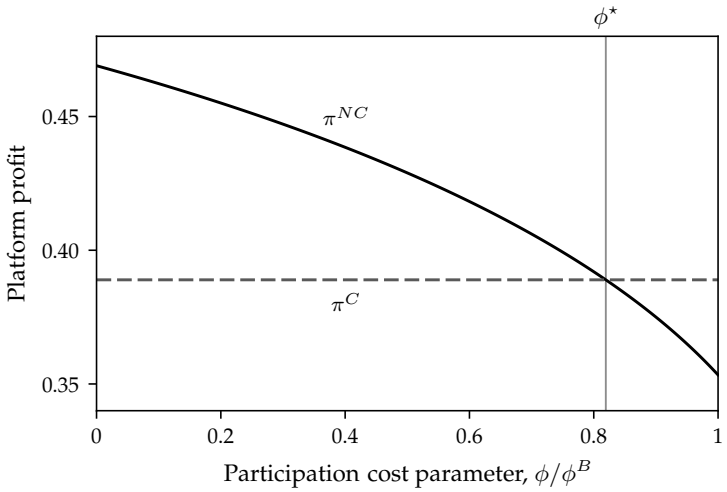
   The increase in total surplus grows strictly with q^{base}.

Proof. See Appendix B.9. □

Compensation increases both creators' net payoffs despite reducing demand for their content. Under NC, the platform sets the price of creators' content so that the low-quality creator's revenue exactly covers her participation cost, leaving her with zero net payoff. Under C, she loses all demand for her content, but her share of AI revenue covers her participation cost and leaves her

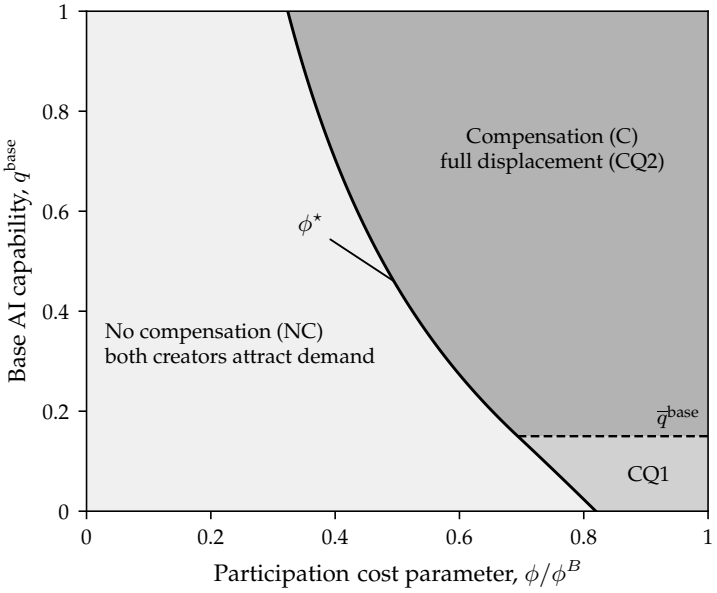
[11] Appendix B.9 establishes this result formally.

Figure 4: Platform Profit With and Without Compensation



Note. Equilibrium platform profit as ϕ/ϕ^B varies, with $q_H = 0.7$, $q_L = 0.4$, $\alpha = 0.3$, and $q^{\text{base}} = 0$. B denotes the benchmark before AI integration. The vertical line marks $\phi = \phi^\star$, where the platform earns the same profit under NC and C. The platform prefers NC when $\phi < \phi^\star$ and C when $\phi > \phi^\star$.

Figure 5: The Platform's Compensation Decision and Creator Displacement



Note. The platform's compensation decision and creator displacement as ϕ/ϕ^B and q^{base} vary, with $q_H = 0.7$, $q_L = 0.4$, and $\alpha = 0.3$. The solid curve marks $\phi = \phi^\star$, where the platform earns the same profit under NC and C. The platform prefers NC when $\phi < \phi^\star$ and C when $\phi > \phi^\star$. Under NC, both creators attract demand. Under C, the dashed line marks $q^{\text{base}} = \overline{q}^{\text{base}}$, separating partial displacement (CQ1), where only H attracts positive demand, from full displacement (CQ2), where AI serves the entire market. Both creators contribute training data in both regimes.

with a positive net payoff. The high-quality creator also earns a higher net payoff under C, as her share of AI revenue more than offsets the reduction in revenue from her own content.

Compensation reduces consumer surplus. The platform charges the same price for AI-generated content of identical quality in both regimes, but a higher price for creators' content under C. Consumers who continue purchasing from creators pay more, while some consumers who previously obtained positive surplus from creators' content instead purchase AI-generated content. The platform prices AI-generated content to capture consumers' full willingness to pay, so those who switch to AI receive no surplus. Compensation thus reduces consumer surplus through the higher price of creators' content and the resulting consumer demand (not AI quality or price).

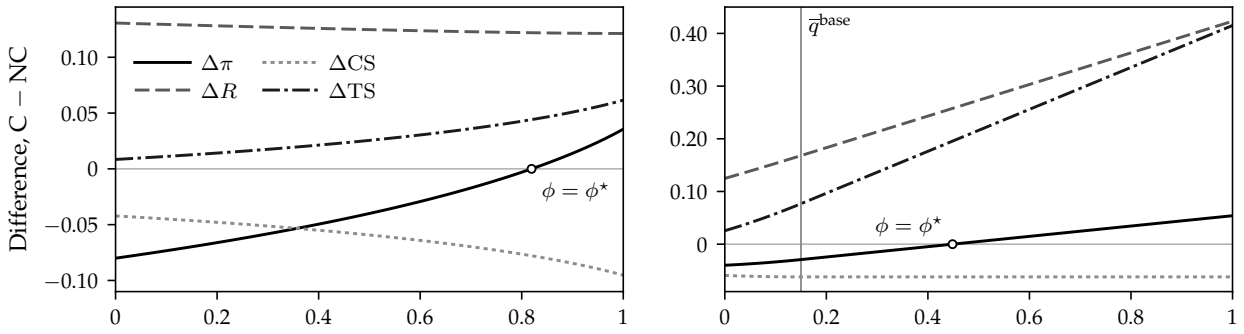
Total surplus nevertheless increases. Both regimes serve the entire market, retain both creators and their training data, and offer AI-generated content of identical quality. Since prices and revenue sharing payments are transfers, the difference in total surplus arises from which content consumers purchase. Under NC, the platform preserves demand for the low-quality creator's content to cover her participation cost. AI trained on both creators' content offers higher quality and perfectly matches those consumers' tastes. Under C, the platform can retain her training data without preserving her sales, allowing more consumers to purchase AI-generated content. Across consumers who switch from creators' content to AI-generated content, the gains in total surplus exceed the losses from reduced purchases of the high-quality creator's content.

As AI capability improves, the differences in both creators' net payoffs and total surplus between C and NC increase strictly. Under NC, the platform raises the price of AI-generated content without changing the price or demand for creators' content, leaving creators' net payoffs and consumer surplus unchanged. Under C, the platform raises the price of AI-generated content and progressively reduces demand for the high-quality creator's content, increasing the AI revenue shared with creators. More consumers purchase AI-generated content under C than NC, so improvements in AI quality contribute more to total surplus under C. Consumer surplus under C declines with demand for the high-quality creator's content and reaches zero once AI serves the entire market. The consumer surplus loss relative to NC remains constant thereafter.

The platform's privately preferred regime need not generate higher total surplus. Under C, the platform treats the AI revenue it shares with creators as a cost, whereas revenue sharing is a transfer when calculating total surplus. The platform also does not capture the full increase in total surplus from greater demand for AI-generated content. When participation costs are low, the platform can prefer NC because retaining creators through content sales is less costly than sharing AI revenue. Corollary 4 further shows that improving AI can make the platform prefer C at lower participation costs, reducing the set of participation costs for which it chooses NC.

When $\phi < \phi^\star(q^{\text{base}})$, the platform prefers NC, but a requirement to compensate creators would induce the platform to adopt C. Requiring compensation would increase both creators' net payoffs and total surplus while reducing consumer surplus and platform profit. A policymaker seeking to maximize total surplus would favor requiring compensation when the platform would otherwise choose NC, whereas a policymaker placing greater weight on consumer surplus or platform

Figure 6: Welfare Differences Between Compensation and No Compensation



(a) Participation cost parameter, ϕ/ϕ^B (b) Base AI capability, q^{base}

Note. Both panels show differences in platform profit ($\Delta\pi$), creators' net payoffs (ΔR), consumer surplus (ΔCS), and total surplus (ΔTS) between C and NC. Parameters are $q_H = 0.7$, $q_L = 0.4$, and $\alpha = 0.3$. Panel (a) varies ϕ/ϕ^B at $q^{\text{base}} = 0$. Panel (b) varies q^{base} at $\phi = \phi^B/2$, with the full displacement threshold $\bar{q}^{\text{base}}$ marked.

profit might prefer NC. Figure 6 illustrates these comparisons: compensation increases creators' net payoffs and total surplus and reduces consumer surplus throughout, while the difference in platform profit changes sign at the compensation threshold.

4 Extensions

The main analysis shows compensating creators for training data allows the platform to displace their direct sales while retaining them as data suppliers. This result relies on the AI's ability to attract consumers away from creators and on compensation being sufficient to sustain creator participation. We extend the analysis by relaxing three assumptions of the baseline model: perfect AI steerability, equal transfer rates of creators' data, and equal splitting of AI revenue among contributors. These extensions clarify when compensation continues to facilitate AI displacement and when displaced creators can remain on the platform as data suppliers.

4.1 Imperfect Steerability

The main analysis assumes perfect AI steerability ($\kappa_0 = 0$), so consumers incur no prompting effort when AI is trained on both creators' content. In this section, we relax this assumption by allowing $\kappa_0 > 0$, which captures residual prompting effort due to the model's inherent limitations, even when both creators contribute data. A higher κ_0 means a less steerable model, rendering AI less attractive than creators' content and possibly limiting the platform's ability to displace their sales.

When both creators contribute, consumer utility from AI is $U_{AI} = q^{\text{base}} + (q_H + q_L)/2 - P_{AI} - \kappa_0$. Since base AI capability and prompting effort enter utility with opposite signs, we define effective AI capability as $\theta \equiv q^{\text{base}} - \kappa_0$. For $\theta \geq 0$, the results of Section 3 still hold with θ replacing q^{base}. We therefore focus on $\theta < 0$, where prompting effort more than offsets base AI capability.

We first consider the NC regime, where creators' participation depends only on their own sales. We ask how prompting effort changes the platform's prices and the allocation of demand.

Proposition 5 (Equilibrium under NC with imperfect steerability)**.** *There exists a threshold $\widehat{\kappa}_0(\phi) > q^{base}$ such that, for $\kappa_0 < \widehat{\kappa}_0(\phi)$, the equilibrium is that of Proposition 2 with θ replacing q^{base}. In particular, the human price, consumer demands, and creators' net payoffs remain unchanged, while the AI price decreases one for one with κ_0 and the platform's profit decreases at a rate equal to AI demand D_{AI}^{NC}.*

Proof. See Appendix B.10. □

Proposition 5 shows that under NC, the platform alone bears the cost of imperfect steerability because the low-quality creator's participation constraint $\alpha D_L P \geq \phi q_L$ does not depend on AI revenue or prompting effort. The platform therefore maintains the same human price and consumer allocation as in Proposition 2, while lowering the AI price to offset the additional prompting effort. This holds when $\kappa_0 < \widehat{\kappa}_0(\phi)$, i.e., as long as AI sales remain profitable enough for the low-quality creator's participation constraint to determine the platform's optimal prices.

Under C, prompting effort also affects creators: it reduces both the AI's appeal to consumers and the AI revenue that compensates them. Our next proposition reveals whether the platform can continue to displace the low-quality creator while retaining her as a training data supplier.

Proposition 6 (Equilibrium under C with imperfect steerability)**.** *Define $\underline{\theta}_C \equiv -(1+\sqrt{2})(q_H - q_L)/2$. Let $\underline{\theta}_L(\phi) \in (-(q_H + q_L)/2, 0)$ be the value of θ at which the low-quality creator's share of AI revenue equals her participation cost ϕq_L, with compensation evaluated using the equilibrium prices and demand in Proposition 3, replacing q^{base} with θ. The equilibrium depends on effective AI capability θ as follows:*

1. ***Displacement as in Proposition 3 (CQ1 if $\theta < \overline{q}^{base}$, CQ2 otherwise):*** *If $\theta \geq \max\{\underline{\theta}_C, \underline{\theta}_L(\phi)\}$, the equilibrium is that of Proposition 3 with θ replacing q^{base}: the low-quality creator is fully displaced but remains a training data supplier, earning a strictly positive net payoff from AI compensation for every such θ except $\theta = \underline{\theta}_L(\phi)$. At $\theta = \underline{\theta}_C$, the three-product configuration (item 2) is also an equilibrium.*

2. ***Three-product configuration:*** *If $-(q_H + q_L)/2 < \theta < \underline{\theta}_C$ or $\theta = \theta_C < \theta_L(\phi)$, the platform continues to offer AI, but the low-quality creator regains positive demand for her own content.*

3. ***No AI revenue:*** *If $\kappa_0 \geq \kappa_0^* \equiv q^{base} + (q_H + q_L)/2$, i.e., $\theta \leq -(q_H + q_L)/2$, AI earns no revenue and the outcome is the benchmark of Proposition 1; for $\kappa_0 > \kappa_0^*$, AI makes no sales.*

Proof. See Appendix B.11. □

Proposition 6 shows that compensation still increases AI displacement under moderate prompting effort: for $\kappa_0 < \widehat{\kappa}_0(\phi)$, $\theta > \theta_C$, and $\theta > \theta_L(\phi)$, the inequalities in Corollary 2 still hold, and the low-quality creator remains a data supplier after losing all direct sales. When κ_0 increases, the platform lowers the AI price to offset the extra effort. Because AI sales become less profitable, it also lowers the human price, allowing the high-quality creator to retain more consumers.

Less steerable AI limits this effect in two ways. First, when effective AI capability falls below $\underline{\theta}_C$, the platform instead prefers a configuration in which all three products attract demand. The low-quality creator then regains positive direct sales while still contributing data. Second, as prompting effort increases, lower AI revenue also reduces the compensation of contributing creators. If the displaced creator's participation constraint binds before the platform reaches the three-product configuration, the prices of Proposition 3 can no longer sustain her participation, and we do not characterize the equilibrium for $\theta_C < \theta < \theta_L(\phi)$. Which occurs first depends on creator qualities and participation costs: when the quality gap is large enough, the participation constraint binds before the platform switches to the three-product configuration.

4.2 Heterogeneous Transfer Rates

The main analysis assumes equal transfer rates ($\tau_H = \tau_L = 1/2$), so each creator's contribution to AI quality is proportional to her content quality. We relax this assumption to allow creators' data to differ in how effectively they improve AI quality. Specifically, we fix $\tau_H = 1/2$ and allow $0 < \tau_L \leq 1/2$, while maintaining equal splitting of AI revenue among contributors. When both creators contribute, AI quality becomes $q_{AI} = q^{\text{base}} + \tau_H q_H + \tau_L q_L$.

A lower τ_L reduces AI quality, which can change the platform's prices, the split of demand between AI and human content, and creator compensation. We therefore ask whether the platform can still retain both creators' data while fully displacing the low-quality creator. We focus on $q_H/q_L \geq \sqrt{2}$, a sufficient condition under which the platform still prefers full displacement and equal revenue sharing still covers the low-quality creator's participation cost, even as her contribution to AI quality approaches zero.

Proposition 7 (Heterogeneous transfer rates)**.** *Suppose $q_H/q_L \geq \sqrt{2}$, $\tau_H = 1/2$, and $0 < \tau_L \leq 1/2$, and maintain equal splitting of AI revenue. Let $q_{AI} = q^{base} + \tau_H q_H + \tau_L q_L$. In equilibrium, both creators contribute, $P_{AI} = q_{AI}$, and $D_L^C = 0$. If $q_{AI} < q_H$ (CQ1), $P = (q_H + q_{AI})/2$ and $D_H^C = (q_H - q_{AI})/2$; otherwise (CQ2), $D_{AI}^C = 1$ at every human price $P \geq q_H$. In particular:*

1. *A lower τ_L reduces the equilibrium AI price and the platform's profit.*
2. *In CQ1, a lower τ_L reduces the human price and AI demand while increasing demand for the high-quality creator's content. In CQ2, AI continues to serve the entire market.*
3. *Both creators still contribute their data, and the low-quality creator remains fully displaced from direct sales ($D_L^C = 0$), earning a strictly positive net payoff from AI compensation for every such τ_L.*

Proof. See Appendix B.12. □

Proposition 7 shows that heterogeneous transfer rates change how demand and payoffs are distributed but not creators' roles: the platform still displaces the low-quality creator's sales while retaining her data, even when that data contributes little to AI quality. A lower transfer rate reduces AI quality, making AI less competitive with human content. The platform responds by lowering

the AI price. In CQ1, it also lowers the human price, shifting some consumers from AI to the high-quality creator's content. The resulting decline in AI revenue also lowers platform profit.

Both creators' payoffs decrease as τ_L falls, even though their shares of AI revenue are unchanged. Under equal splitting, each creator receives half of the creators' AI revenue, so a lower τ_L reduces the compensation received by both creators. In CQ1, the high-quality creator also gains direct demand, but this does not offset the decline in her AI compensation and the lower human price.

Despite these changes, the low-quality creator continues to participate. Although her contribution to AI quality becomes smaller, her data remains essential for AI steerability: without it, the AI attracts no consumers at the equilibrium prices. Moreover, equal splitting ties her compensation to total AI revenue rather than to her individual contribution to AI quality. This payment remains sufficient to cover her participation cost even as τ_L approaches zero (Appendix B.12).

4.3 Quality-Weighted Compensation

The main analysis assumes the creators' share of AI revenue is split equally among contributors, regardless of how much each creator's data improves AI quality. Under compensation, this rule also determines the marginal participant: since creator j participates if and only if $V_j/q_j \geq \phi$, this is the creator with the smaller ratio. Without compensation, L is the marginal participant: the platform holds her exactly at her participation constraint (Proposition 2), while H, with her larger market, has slack. The next corollary asks whether this remains true under C with equal splitting.

Corollary 5 (Reversal under equal splitting)**.** *Without compensation, $V_L^{NC}/q_L = \phi < V_H^{NC}/q_H$, whereas with compensation and equal splitting of AI revenue, $V_H^C/q_H < V_L^C/q_L$.*

Proof. See Appendix B.13. □

Corollary 5 shows that compensation reverses the order under equal splitting: both contributors receive the same share of AI revenue, but participation costs rise with quality, and H's remaining revenue from direct sales is too small to make up for the difference.

We now consider an alternative compensation rule that allocates the creators' share of AI revenue in proportion to each creator's contribution to AI quality. Specifically, we replace $s_j/(s_H + s_L)$ in Eq. (4) with $\tau_j q_j s_j/(\tau_H q_H s_H + \tau_L q_L s_L)$, and we mark equilibrium objects under this rule with the superscript QW. Under the baseline assumption of equal transfer rates ($\tau_H = \tau_L = 1/2$), creator j receives a share $q_j/(q_H + q_L)$ of the creators' AI revenue when both contribute. The next proposition examines whether this redistribution undoes the reversal in Corollary 5 and whether it also changes the equilibrium allocation.

Proposition 8 (Quality-weighted compensation)**.** *Suppose $\tau_H = \tau_L = 1/2$. Under quality-weighted compensation, the equilibrium prices, demands, platform profit, consumer surplus, and creators' total net payoff remain unchanged from Proposition 3. Both creators still contribute their data and the low-quality creator remains fully displaced from direct sales. Quality weighting redistributes AI revenue from the low-quality creator to the high-quality creator, while both continue to earn strictly positive net payoffs.*

When transfer rates are heterogeneous, suppose $q_H/q_L \geq \sqrt{2}$, $\tau_H = 1/2$, *and* $0 < \tau_L \leq 1/2$. *The outcome of Proposition 7 remains an equilibrium under quality weighting if and only if* $\tau_L \geq \tau_L^\star$, *where* $\tau_L^\star \in (0, 1/2)$ *depends on* ϕ *and* α *only through* ϕ/α. *The low-quality creator gains a strictly positive net payoff when* $\tau_L > \tau_L^\star$, *and her participation constraint binds at* $\tau_L = \tau_L^\star$.

Proof. See Appendices B.12 and B.14. □

Proposition 8 shows that quality weighting changes how creators share AI revenue but not the equilibrium allocation. Quality weighting does not alter the platform's profit under any contribution profile, since its retained share of AI revenue remains $1 - \alpha$. At the equilibrium prices of Proposition 3, neither creator can generate AI revenue alone, and both still earn positive net payoffs when contributing together. Relative to equal splitting, quality weighting transfers $\alpha(q_H - q_L)D_{AI}^C P_{AI}^C/[2(q_H + q_L)]$ from L to H, leaving creators' total net payoff unchanged.

More importantly, this redistribution undoes the reversal in Corollary 5. Under quality weighting, both creators receive the same AI compensation per unit of quality, but H also earns revenue from direct sales in CQ1. Hence, $V_H^{C,\mathrm{QW}}/q_H - V_L^{C,\mathrm{QW}}/q_L = \alpha D_H^C P^C/q_H \geq 0$, with strict inequality in CQ1. In CQ1, L is again the marginal participant; in CQ2, the two participation constraints are equally tight. Thus, the reversal is driven by equal splitting rather than compensation itself.

With heterogeneous transfer rates, quality weighting sustains the outcome of Proposition 7 only if $\tau_L \geq \tau_L^\star$, whereas equal splitting sustains it for every τ_L. As τ_L decreases, L receives a smaller share of declining AI revenue. Below $\tau_L^\star$, her compensation at the prices of Proposition 7 no longer covers her participation cost and we do not discuss the resulting equilibrium.

5 Conclusion

We study compensation for training data in a model of a creative platform where consumers choose between purchasing from creators and prompting an integrated generative AI system. We compare a regime that makes training data use a condition of platform participation with one that compensates creators and allows them to opt out of AI training. In the absence of compensation, creators cover their participation costs through their content sales, which constrains the platform to preserve their demand. With compensation, the low-quality creator chooses to remain solely as a training data supplier while AI fully displaces her demand. The platform sets a higher price for creators' content with compensation, although both creators contribute training data and the quality and price of AI-generated content are identical in the two regimes.

As AI improves, the platform raises the price of AI-generated content identically with and without compensation. Without compensation, it leaves the price of creators' content unchanged to preserve their demand and retain their training data. With compensation, it also raises the price of creators' content, progressively displacing the remaining demand for creators until AI serves the entire market. Both creators nevertheless remain on the platform, earning positive net payoffs from

supplying training data. Further AI improvements increase their revenue from training data even after AI fully displaces their demand.

Requiring compensation increases creators' net payoffs and total surplus but reduces consumer surplus. The increase in total surplus comes from reallocating demand toward AI-generated content that better matches consumers' tastes, without expanding market coverage or improving AI quality. Consumers lose surplus as the platform raises the price of creators' content and captures the willingness to pay of consumers purchasing AI-generated content. The platform's compensation decision depends on creator participation costs and AI capability: retaining creators through content sales can be more profitable than sharing revenue from AI-generated content sales when participation costs are low. Revenue sharing is a cost to the platform but a transfer when calculating total surplus. Requiring compensation in this case increases total surplus, although consumers lose surplus and the platform loses profit. As AI improves, the range of creator participation costs for which the platform prefers not to compensate can narrow.

A platform can depend on creators' training data while competing with their content if it can compensate them without preserving their sales. For policymakers, the relevant comparison must account for how the platform changes prices in response to compensation. Requiring compensation changes what consumers purchase and how revenue is divided between creators and the platform. Preserving demand for creators' content and sustaining creators' net payoffs are different objectives, and neither alone captures the consequences for consumers or total surplus. Policymakers may prefer different compensation requirements depending on how they value total surplus and the payoffs of creators, consumers, and the platform.

We extend the model to consider imperfect AI steerability, heterogeneous transfer rates, and alternative compensation rules. Compensation-induced displacement persists with moderate prompting effort, although sufficiently high prompting effort can restore demand for creators' content. Greater prompting effort also reduces AI revenue and can leave compensation below a displaced creator's participation cost at the prices derived for imperfect steerability. When compensation is weighted by contributions to AI quality, displacement persists under equal transfer rates, but the distribution of revenue between creators changes. With quality-weighted compensation and heterogeneous transfer rates, the low-quality creator may contribute too little to AI quality for her share of revenue to cover her participation cost at the prices characterized in our analysis. Equal revenue sharing can sustain her participation because the platform pays her a fixed share of total AI revenue. These extensions show that compensation-induced displacement depends on consumers' willingness to purchase AI-generated content and whether compensation covers displaced creators' participation costs.

Our work has several limitations. We assume that creators withdraw their training data when they exit the platform. This assumption is most relevant when training licenses must be renewed or creators continue producing new content. If the platform retains perpetual rights to previously supplied content, creator exit may have a weaker effect on AI quality and prompting effort, reducing the platform's incentive to preserve demand for creators' content without compensation. We also

consider two creators whose contributions affect AI quality and prompting effort, and a common price for creators' content. Future research could allow creators to choose content quality and the platform to choose separate revenue shares for creators' content and AI-generated content. A dynamic model could then examine whether compensation sustains the production of new content as AI improves. Greater AI displacement reduces revenue from direct content sales, while compensation may provide revenue from supplying new content for training. How these incentives affect future content production remains an open question.

References

Abhishek V, Jerath K, Zhang ZJ (2016) Agency selling or reselling? Channel structures in electronic retailing. *Management Sci.* 62(8):2259–2280.

Acemoglu D, Restrepo P (2018) The race between man and machine: Implications of technology for growth, factor shares, and employment. *Amer. Econom. Rev.* 108(6):1488–1542.

Acemoglu D, Restrepo P (2019) Automation and new tasks: How technology displaces and reinstates labor. *J. Econom. Perspect.* 33(2):3–30.

Acemoglu D, Restrepo P (2020) Robots and jobs: Evidence from US labor markets. *J. Political Econom.* 128(6):2188–2244.

Acemoglu D, Restrepo P (2022) Tasks, automation, and the rise in U.S. wage inequality. *Econometrica* 90(5):1973–2016.

Autor D, Chin C, Salomons A, Seegmiller B (2024) New frontiers: The origins and content of new work, 1940–2018. *Quart. J. Econom.* 139(3):1399–1465.

Autor DH, Dorn D (2013) The growth of low-skill service jobs and the polarization of the US labor market. *Amer. Econom. Rev.* 103(5):1553–1597.

Autor DH, Levy F, Murnane RJ (2003) The skill content of recent technological change: An empirical exploration. *Quart. J. Econom.* 118(4):1279–1333.

Bhargava H, Dubus A, Ronayne D, Shekhar S (2026) The strategic value of data sharing in interdependent markets. *Management Sci.* 72(2):1472–1488.

Bostoen F, Monti G (2025) Putting a price on AI training data. TILEC Discussion Paper No. 2025-10, SSRN 5840062, dated November 28, 2025, `https://doi.org/10.2139/ssrn.5840062`.

Choi JP, Kim K, Mukherjee A (2025) "Sherlocking" and platform information policy. *Management Sci.* 71(11):9812–9830.

Cooper Z, Martens B, Peukert C, Stocker V (2025) The law and economics of generative AI and copyright: A primer to core challenges for our digital future. *Rev. Network Econom.* 24(3):161–176.

Dai X, Hou J, Ma CY, Tsai S, Wang J, Wang R, Zhang P, Vandenhende S, Wang X, Dubey A, Yu M, Kadian A, Radenovic F, Mahajan D, Li K, Zhao Y, Petrovic V, Singh MK, Motwani S, Wen Y, Song Y, Sumbaly R, Ramanathan V, He Z, Vajda P, Parikh D (2023) Emu: Enhancing image generation models using photogenic needles in a haystack. Preprint, submitted September 27, `https://arxiv.org/abs/2309.15807`.

de la Durantaye K (2025) Control and compensation: A comparative analysis of copyright exceptions for training generative AI. *IIC Internat. Rev. Intellectual Property Competition Law* 56(4):737–770.

de Rassenfosse G, Jaffe AB, Waldfogel J (2025) Intellectual property and creative machines. *Entrepreneurship Innovation Policy Econom.* 4:47–79.

European Parliament (2026) Copyright and generative artificial intelligence – opportunities and challenges. Resolution of March 10, 2026, URL https://www.europarl.europa.eu/doceo/document/TA-10-2026-0066_EN.html, P10_TA(2026)0066.

Feigenbaum J, Gross DP (2024) Answering the call of automation: How the labor market adjusted to mechanizing telephone operation. *Quart. J. Econom.* 139(3):1879–1939.

Gans JS (2026) Copyright policy options for generative artificial intelligence. *J. Law Econom.* 69(1):1–19.

Gao Y, Wang Z, Kim A (2025) Fair use vs. fair pay: An economic analysis of compensation from generative AI platforms to original content creators. Working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5233915.

Hagiu A, Spulber D (2013) First-party content and coordination in two-sided markets. *Management Sci.* 59(4):933–949.

Hagiu A, Teh TH, Wright J (2022) Should platforms be allowed to sell on their own marketplaces? *RAND J. Econom.* 53(2):297–327.

Hoffmann J, Borgeaud S, Mensch A, Buchatskaya E, Cai T, Rutherford E, de Las Casas D, Hendricks LA, Welbl J, Clark A, Hennigan T, Noland E, Millican K, van den Driessche G, Damoc B, Guy A, Osindero S, Simonyan K, Elsen E, Vinyals O, Rae JW, Sifre L (2022) Training compute-optimal large language models. *Adv. Neural Inform. Processing Systems, vol. 35* (Curran Associates, Red Hook, NY), 30016–30030.

Jiang B, Jerath K, Srinivasan K (2011) Firm strategies in the "mid tail" of platform-based retailing. *Marketing Sci.* 30(5):757–775.

Kaplan J, McCandlish S, Henighan T, Brown TB, Chess B, Child R, Gray S, Radford A, Wu J, Amodei D (2020) Scaling laws for neural language models. Preprint, submitted January 23, https://arxiv.org/abs/2001.08361.

Keinan G, Ben-Porat O (2026) Strategic content creation with GenAI: To share or not to share? *Proc. ACM Web Conf. 2026* (ACM, New York), 169–180.

Meta (2023) Privacy matters: Meta's generative AI features. URL https://about.fb.com/news/2023/09/privacy-matters-metas-generative-ai-features/, originally published September 27, 2023; updated November 21, 2025.

Moreau F, Wikström P, Haampland O, Johannessen R (2024) Alternative payment models in the music streaming market: A comparative approach based on stream-level data. *Inform. Econom. Policy* 68:101103.

Ørstavik IB (2026) Copyright and generative AI: Guidance from economic incentive theory for a nuanced regime for author remuneration. *IIC Internat. Rev. Intellectual Property Competition Law*, ePub ahead of print October 29, https://doi.org/10.1007/s40319-026-01784-9.

Peukert C (2025a) Copyright and the dynamics of innovation in artificial intelligence. Working paper, https://doi.org/10.2139/ssrn.5445234.

Peukert C (2025b) The economics of data for generative AI: Trade-offs and optimal copyright policy. Working paper, September 26, 2025, https://ssrn.com/abstract=5532959.

Peukert C, Abeillon F, Haese J, Kaiser F, Staub A (2025) AI and the dynamic supply of training data. Working paper, https://doi.org/10.2139/ssrn.4807979.

Reddit (2026) User agreement. URL https://redditinc.com/policies/user-agreement, effective July 1, 2026.

Samuelson P (2026) Assessing the feasibility of collective licensing of in-copyright works as training data for generative AI systems. *Proc. Natl. Acad. Sci. USA* 123(30):e2509769122.

Senftleben M (2025) Win-win: How to remove copyright obstacles to AI training while ensuring author remuneration (and why the AI Act fails to do the magic). *Chicago-Kent Law Rev.* 100(1):7–55.

Shcherbakov V, Dalaud I, Peukert C (2025) AI needs better data than the law allows. Working paper, https://doi.org/10.2139/ssrn.5296505.

Shutterstock (2026) Shutterstock data licensing and the Contributor Fund. Updated September 18, 2026; accessed November 1, 2026, https://submit.shutterstock.com/help/en/articles/10594694-shutterstock-data-licensing-and-the-contributor-fund.

Taitler B, Madmon O, Tennenholtz M, Ben-Porat O (2025) Data sharing with a generative AI competitor. Preprint, submitted May 18, https://arxiv.org/abs/2505.12386.

Vafa K, Bentley S, Kleinberg J, Mullainathan S (2025) What's producible may not be reachable: Measuring the steerability of generative models. *Adv. Neural Inform. Processing Systems, vol. 38* (Curran Associates, Red Hook, NY), 62767–62804.

Wang JT, Deng Z, Chiba-Okabe H, Barak B, Su WJ (2024) An economic solution to copyright challenges of generative AI. Preprint, submitted April 22, https://arxiv.org/abs/2404.13964.

Yang SA, Zhang AH (2024) Generative AI and copyright: A dynamic perspective. Working paper, revised December 5, 2025, https://doi.org/10.2139/ssrn.4716233.

Yao F, Li C, Nekipelov D, Wang H, Xu H (2024) Human vs. generative AI in content creation competition: Symbiosis or conflict? *Proc. 41st Internat. Conf. Machine Learn.* (PMLR), 56885–56913.

Zhou C, Liu P, Xu P, Iyer S, Sun J, Mao Y, Ma X, Efrat A, Yu P, Yu L, Zhang S, Ghosh G, Lewis M, Zettlemoyer L, Levy O (2023) LIMA: Less is more for alignment. *Adv. Neural Inform. Processing Systems, vol. 36* (Curran Associates, Red Hook, NY), 55006–55021.

A Notation

We collect notation, case labels and thresholds of Sections 2–4 and the shorthand of Appendix B in Table 1. All these symbols are also defined where they first appear in the paper.

Table 1: Notation and Case Labels

Primitives	
$q_H,\ q_L$	Content quality of the high- and low-quality creator, $0 < q_L < q_H < 1$.
$\alpha \in (0,1)$	Creators' share of own-content revenue and, under C, contributors' share of AI revenue; the platform keeps the rest.
$\phi > 0$	Participation cost parameter: creator j's participation cost is ϕq_j.
$q^{\text{base}} \geq 0$	Base AI capability: quality of AI-generated content before training on creators' content; $q^{\text{base}} \leq 1$ by Assumption 1 when $\kappa_0 = 0$.
$\kappa_0,\ \kappa_{\text{data}}$	Base (steerability) and data components of prompting effort $\kappa(x; s_H, s_L)$, Eq. (2); $\kappa_{\text{data}} = 1$, and $\kappa_0 = 0$ except in Section 4.1.
$\tau_H, \tau_L \in (0,1)$	Transfer rates of creators' content into AI quality, Eq. (3); $\tau_H = \tau_L = \frac{1}{2}$ except in Sections 4.2–4.3, which allow $0 < \tau_L \leq \tau_H = \frac{1}{2}$.
Decisions, demands, payoffs	
$P,\ P_{AI}$	Common price of creators' content; price of AI-generated content (Stage 1).
$e_j,\ s_j \in \{0,1\}$	Participation of creator j and training of the AI on her content (Stage 2); $s_j = e_j$ under NC, while j chooses $s_j \leq e_j$ under C.
q_{AI}	Quality of AI-generated content, $q_{AI} = q^{\text{base}} + \tau_H q_H s_H + \tau_L q_L s_L$, Eq. (3).
D_H, D_L, D_{AI}	Demand for H's, L's, and AI-generated content; $D \equiv D_H + D_L$.
$V_j,\ R_j$	Gross payoff of participating creator j, Eq. (4); net payoff $R_j \equiv V_j - \phi q_j$.
$\pi,\ \text{CS},\ \text{TS}$	Platform profit, Eq. (5); consumer surplus; total surplus $\text{TS} \equiv \pi + R_H + R_L + \text{CS}$.
Shorthand	
$(x)^+$	Positive part of x: $(x)^+ \equiv \max\{x, 0\}$.
$g,\ t$	$g \equiv (q_H - q_L)/2 \in (0, \frac{1}{2})$ (Lemma 1) and $t \equiv q_H + q_L \in (0,2)$.
$\widetilde{P}_{AI},\ \delta$	Net AI price $\widetilde{P}_{AI} \equiv P_{AI} - q^{\text{base}}$ and price gap $\delta \equiv \widetilde{P}_{AI} - P$ (Lemma 1); in a covered market, demands depend on prices only through δ. Appendix B uses $\widetilde{P}_{AI} \equiv P_{AI} - \theta$, which is the same when $\theta = q^{\text{base}}$.
s	Cost index $s \equiv \sqrt{q_L^2 - 4\phi q_L/\alpha}$ (Proposition 2), distinct from s_j; it strictly decreases in ϕ and equals g at $\phi = \phi^B$, so $g \leq s < q_L$.
θ	Effective AI capability $\theta \equiv q^{\text{base}} - \kappa_0$ (Section 4.1); $\theta = q^{\text{base}}$ when $\kappa_0 = 0$. Appendix B uses the general form $\theta \equiv z - (q_H + q_L)/2$, which equals $q^{\text{base}} - \kappa_0$ when $\tau_H = \tau_L = \frac{1}{2}$.
z	Quality of AI-generated content, net of κ_0, when both creators contribute, Eq. (B.1); $z = q^{\text{base}} + (q_H + q_L)/2$ in the main analysis.

Table 1 (continued)

Symbols used only in Appendix B

$(1,1)$; $\mathcal{P}^*,\ \pi^*$	Profile in which both creators participate and contribute; a set of price vectors and the profit that every allowed Stage-2 play gives on it (Lemma B.1).
$d_k(x),\ y_k$	Distance of a consumer at x from creator k's location; quality of AI-generated content net of κ_0 when only creator k contributes, $y_k \equiv z - \tau_j q_j$ for $j \neq k$ (Lemma B.4).
$\Pi_{HL},\ \Pi_H$	Bounds on profit without the AI when both creators sell, $(1-\alpha)t^2/8$, and when at most one does, $(1-\alpha)q_H^2/4$ (Lemma B.2).
V_j^B	Creator j's revenue at the benchmark price, $\alpha(q_j - \frac{t}{4})\frac{t}{4}$.
$\delta^{NC},\ \widetilde{\pi}(\delta)$	Equilibrium price gap under NC, $(q_H - s)/2$; the platform's profit under NC at price gap δ when the AI is priced at z, Eq. (B.3).
$\pi^{NC}(\phi,\theta)$	Equilibrium profit under NC as a function of ϕ and θ, Eq. (B.2).
$D_H^C(z),\ F(z),\ G(z)$	At the displacement prices: H's demand $\frac{1}{2}(q_H - z)^+$, the platform's profit divided by $1-\alpha$, and AI revenue (Appendix B.7).
$\mathcal{P}(z)$	Displacement prices: $(\frac{q_H+z}{2}, z)$ if $z < q_H$, and (P, z) with $P \geq q_H$ otherwise.
$\Psi(\delta),\ M$	Profit divided by $1-\alpha$ under $(1,1)$ at price gap δ when the AI is priced at z, and its maximum (Lemma B.8); δ_c and δ_3 are the price gaps at its two peaks.
$V_H(z),\ A^C$	H's gross payoff at the displacement prices under equal splitting (Lemma B.10); each creator's AI payment in the equilibrium of Proposition 3.
$\Lambda(s, q^{\text{base}}),\ s^\star$	Profit difference $\pi^{NC} - \pi^C$ and its zero in s (proof of Corollary 4).
W^r	Consumers' gross utility in regime r (proof of Proposition 4).
$\widehat{\theta},\ \overline{\pi}$	$\widehat{\theta} \equiv q^{\text{base}} - \widehat{\kappa}_0(\phi)$ (proof of Proposition 5); the platform's maximal profit under $(1,1)$ (proof of Proposition 6).
$h,\ \ell,\ \eta(\tau_L)$	$h \equiv q_H/2$, $\ell \equiv \tau_L q_L$, and L's payoff under quality weighting divided by α (proof of Proposition 8).

Benchmark and regimes (superscripts on equilibrium objects)

B	Benchmark before AI integration (Section 3.1).
NC	No compensation (Section 3.2): training on a participant's content is a condition of participation, and creators earn only $\alpha D_j P$.
C	Compensation (Section 3.3): each participant chooses s_j, and contributors split the share α of AI revenue equally. The superscript QW marks quality-weighted compensation, which splits it in proportion to $\tau_j q_j s_j$ (Section 4.3).

Market structure cases

Table 1 (continued)

$\overline{q}^{\text{base}} \equiv (q_H - q_L)/2$	Full displacement threshold separating CQ1 from CQ2 (Proposition 3); when both creators contribute, $q^{\text{base}} < \overline{q}^{\text{base}}$ is equivalent to $q_{AI} < q_H$, the form Proposition 7 uses for general τ_L; Section 4.1 uses θ in place of q^{base}.
Partial displacement (CQ1)	$q^{\text{base}} < \overline{q}^{\text{base}}$: AI fully displaces L ($D_L^C = 0$), while $D_H^C > 0$.
Full displacement (CQ2)	$q^{\text{base}} \geq \overline{q}^{\text{base}}$: AI serves the entire market ($D_{AI}^C = 1$).
Three-product configuration	All three products attract demand (row 2 of the table in Lemma 1): the equilibrium configuration under NC (Proposition 2; Proposition 5 for $\kappa_0 < \widehat{\kappa}_0(\phi)$) and under C for $-(q_H + q_L)/2 < \theta < \underline{\theta}_C$ or $\theta = \theta_C < \theta_L(\phi)$ (Proposition 6).

Thresholds

$\rho^\star \equiv 1 + \sqrt{2}$	Quality ratio at which serving both creators and serving only H give the same maximal benchmark profit (Proposition 1); Assumption 4 requires $q_H/q_L < \rho^\star$.
ϕ^B	$\phi^B \equiv \alpha(3q_L - q_H)(q_H + q_L)/(16q_L)$, the largest ϕ at which both creators participate at $P^B = (q_H + q_L)/4$; Assumption 4 requires $\phi \leq \phi^B$.
$\phi^\star(q^{\text{base}}) \in [0, \phi^B]$	The platform prefers NC for $\phi < \phi^\star$ and C for $\phi > \phi^\star$ (Corollary 4).
$\widehat{\kappa}_0(\phi)$	Value of κ_0, Eq. (B.6), below which the NC equilibrium is that of Proposition 2 with θ in place of q^{base} (Proposition 5); $q^{\text{base}} < \widehat{\kappa}_0(\phi) \leq \kappa_0^*$.
$\underline{\theta}_C \equiv -(1 + \sqrt{2})g$	Under C, the equilibrium is the three-product configuration for $-(q_H + q_L)/2 < \theta < \underline{\theta}_C$ or $\theta = \theta_C < \theta_L(\phi)$, and that of Proposition 3 with θ in place of q^{base} for $\theta \geq \max\{\underline{\theta}_C, \underline{\theta}_L(\phi)\}$, with both at $\theta = \underline{\theta}_C > \theta_L(\phi)$ (Proposition 6).
$\underline{\theta}_L(\phi)$	Value of $\theta \in (-(q_H + q_L)/2, 0)$ at which L's AI payment at the prices of Proposition 3 with θ in place of q^{base} equals her participation cost ϕq_L (Proposition 6).
$\kappa_0^* \equiv q^{\text{base}} + (q_H + q_L)/2$	AI makes no sales when $\kappa_0 > \kappa_0^*$, i.e., $\theta < -(q_H + q_L)/2$ (Proposition 6).
$\tau_L^\star \in (0, \frac{1}{2})$	For $q_H/q_L \geq \sqrt{2}$, the outcome of Proposition 7 is an equilibrium under quality-weighted compensation if and only if $\tau_L \geq \tau_L^\star$; L's participation constraint binds at $\tau_L^\star$ (Proposition 8).

Note. Unless stated otherwise, we use $q_H = 0.7$, $q_L = 0.4$, $\alpha = 0.3$, $\phi = \phi^B/2$, and $q^{\text{base}} = 0$ in numerical illustrations.

B Proofs

In this section, we prove all results stated in the main paper. The equilibrium results (Propositions 1–3 and 5–8) share one logic. The platform sets prices anticipating creators' participation and data-contribution decisions. We must therefore compare its profit across everything creators may do at every price and then verify what they do at the stated prices. To formally state this logic, we first introduce additional notation and the proof strategy commonly used in these results, and

then collect the properties of creators' Stage-2 play (Appendix B.1) and of consumer demand (Appendix B.2) that the proofs share. Then, we prove these results in the order shown in the main paper, with auxiliary lemmas placed before proofs that use them.

Notation. We use two shorthands for the creators' qualities throughout:

$$t \equiv q_H + q_L \in (0,2), \qquad g \equiv \frac{q_H - q_L}{2} \in \left(0, \tfrac{1}{2}\right).$$

We also write $(x)^+ \equiv \max\{x, 0\}$ for the positive part of a number x. We cover the main analysis and the extensions of Section 4 with one argument. We thus let z denote the quality of AI-generated content, net of the base prompting effort κ_0, when the AI is trained on both creators' content:

$$z \equiv q^{\text{base}} - \kappa_0 + \tau_H q_H + \tau_L q_L. \tag{B.1}$$

By Eq. (2) and Eq. (3), the utility that a consumer at x obtains from AI-generated content depends on whose content the AI is trained on:

$$U_{AI} = \begin{cases} z - P_{AI}, & \text{both creators' content,} \\ z - \tau_j q_j - P_{AI} - d_k(x), & \text{only creator } k\text{'s content } (j \neq k), \\ q^{\text{base}} - \kappa_0 - 1 - P_{AI}, & \text{neither creator's content.} \end{cases}$$

In the second case, the AI loses creator j's contribution $\tau_j q_j$ to its quality, and the consumer incurs extra prompting effort equal to her distance from k's location, $d_L(x) = x$ or $d_H(x) = 1 - x$, as $\kappa_{\text{data}} = 1$. In the third case, $U_{AI} \leq -P_{AI}$, as Assumption 1 requires $q^{\text{base}} - \kappa_0 \leq 1$.

We further use three symbols for prices:

$$\theta \equiv z - \frac{t}{2}, \qquad \widetilde{P}_{AI} \equiv P_{AI} - \theta, \qquad \delta \equiv \widetilde{P}_{AI} - P.$$

When the AI is trained on both creators' content, $U_{AI} = \frac{t}{2} - \widetilde{P}_{AI}$, and δ is the gap between the net AI price $\widetilde{P}_{AI}$ and the price of creators' content. In the main analysis ($\kappa_0 = 0$ and $\tau_H = \tau_L = \frac{1}{2}$), $z = q^{\text{base}} + \frac{t}{2}$ and $\theta = q^{\text{base}}$, so $\widetilde{P}_{AI}$ and δ are those of Lemma 1. In Section 4.1, $\theta = q^{\text{base}} - \kappa_0$, as defined there; in Sections 4.2 and 4.3, where $\kappa_0 = 0$ and $\tau_H = \frac{1}{2}$, $\theta = q^{\text{base}} - (\frac{1}{2} - \tau_L) q_L$.

Proof strategy. We prove each equilibrium result (Propositions 1–3 and 5–8) in three steps. First, we bound the platform's profit at every price vector under every Stage-2 profile (under NC, in every pure-strategy Stage-2 equilibrium). Second, we show that the bound is attained only at the stated prices and under the stated profile. Whenever the AI sells, this is the profile in which both creators participate and contribute data. Third, we show that at these prices this profile is the only Stage-2 play consistent with Assumptions 2 and 3. Lemma B.1 turns these facts into the equilibrium characterization, and we collect in Appendix B.2 the demand properties that the three steps share.

B.1 Equilibrium Properties

We call a pure-strategy choice of e_j (it's s_j under C) by each creator a *profile*. We say that creator j *contributes* if $s_j = 1$; under NC, this means that she participates. We write $(1, 1)$ for the profile in which both creators participate and contribute if the AI is offered. As in Section 2.4, we call a Stage-2 play at a price vector *allowed* if it is a pure-strategy Nash equilibrium that satisfies Assumption 2 and that Assumption 3 selects when it applies. If no such equilibrium exists, an allowed play is a mixed-strategy Nash equilibrium, which the platform values at its expected profit.

By an *equilibrium* we mean a subgame-perfect equilibrium in which the platform chooses a price vector and creators use an allowed play at every price vector. Under C, a pure-strategy Nash equilibrium of the Stage-2 game can fail to exist at some price vectors.[12]

We use Lemma B.1 to prove every equilibrium result (Propositions 1–3 and 5–8) to pass from Stage-2 play to the platform's equilibrium prices. The lemma has three parts. Part (a) rules out mixed play in the benchmark and under NC. Part (b) reduces the search for equilibria to comparing profits across price vectors, and part (c) extends profit bounds from profiles to mixed plays.

Lemma B.1 (Stage-2 play)**.** *(a) In the benchmark and under NC (for every $\kappa_0 \geq 0$), at every price vector some pure-strategy Nash equilibrium satisfies Assumption 2, so every allowed play is pure.*

(b) Fix the benchmark or a regime (under C, with equal splitting or quality weighting). Suppose that, for a nonempty set $\mathcal{P}^*$ of price vectors, every allowed play gives profit π^* on $\mathcal{P}^*$ and strictly less elsewhere. Then the equilibria are exactly the strategy profiles in which creators use an allowed play at every price vector and the platform chooses a vector in $\mathcal{P}^*$. In particular, an equilibrium exists and yields profit π^*.

(c) At any given price vector, if every profile gives profit at most π^*, then so does every allowed play, with equality only if the play puts probability one on the profiles that give π^*.

Proof. (a) A creator sells weakly more when the other creator exits, so a creator who participates against a participating rival also participates against an absent one; this monotonicity yields a pure-strategy equilibrium. We fix prices and creators $j \neq j'$, and let $D_j(e_{j'})$ denote j's demand when she participates. If j' exits, j's content loses a competitor. Under NC, the AI also loses j''s training data, which lowers every consumer's AI utility. By Eq. (2) and Eq. (3), the fall is $\tau_{j'} q_{j'} + \kappa_{\text{data}} |x - x_j| \geq 0$ for every κ_0, where $x_L = 0$ and $x_H = 1$ are the creators' locations. As ties favor creators' content over the AI and platform products over the outside option (footnote 4), every consumer who bought from j still does, so her demand weakly rises: $D_j(0) \geq D_j(1)$.

By Assumption 2, j participates if and only if her sales revenue covers her cost: her best reply is $\beta_j(e_{j'}) = 1$ if $\alpha D_j(e_{j'}) P \geq \phi q_j$ and $\beta_j(e_{j'}) = 0$ otherwise. Hence $\beta_j(1) \leq \beta_j(0)$, and the pure-strategy equilibria satisfying Assumption 2 are the profiles with $e_j = \beta_j(e_{j'})$ for both creators. We now show that such a profile exists. If $(1, 1)$ is not one, then $\beta_j(1) = 0$ for some j. If $\beta_{j'}(0) = 1$, the profile in which j exits and j' participates is one. Otherwise $\beta_{j'}(1) \leq \beta_{j'}(0) = 0$, so j' exits whatever

[12] For example, at $(P, P_{AI}) = (0.07, 0.05)$ with $(q_H, q_L, \alpha, \phi, q^{\text{base}}, \kappa_0) = (0.9, 0.8, 0.5, 0.02, 0.8, 0)$, which satisfies Assumption 4, the creators' best replies are unique and cycle.

j does, and the profile in which j' exits and j plays $\beta_j(0)$ is one. As a pure-strategy equilibrium satisfying Assumption 2 thus exists, no mixed play is allowed there.

(b) An allowed play exists at every price vector. It is a pure-strategy equilibrium that satisfies Assumption 2 and that Assumption 3 selects whenever it applies. If there is none, it is a mixed-strategy equilibrium of the finite Stage-2 game, which exists by Nash's theorem. By definition, allowed plays are Nash equilibria of the Stage-2 subgames. A strategy profile in which creators use them is therefore an equilibrium if and only if the platform's price vector maximizes profit given those plays. By hypothesis, this holds if and only if the price vector lies in $\mathcal{P}^*$.

(c) An allowed play induces a distribution μ over the finitely many profiles of the Stage-2 game. We let $\pi(\omega)$ denote the platform's profit under profile ω. The play's expected profit $\sum_\omega \mu(\omega)\pi(\omega)$ is at most π^*, with equality only if $\pi(\omega) = \pi^*$ for every profile ω with $\mu(\omega) > 0$. □

B.2 Consumer Demand

In this section, we describe consumer demand and the profit it allows according to which creators contribute. The lemmas hold for every $\kappa_0 \geq 0$ and all transfer rates.

We start with the market in which the AI is not offered. We use Lemma B.2 to prove Proposition 1. Its profit bounds also apply whenever the AI earns no revenue (Lemma B.5).

Lemma B.2 (Demand and profit without the AI)**.** *Suppose that the platform does not offer the AI, and fix a price $P \geq 0$ for creators' content. Let $\Pi_{HL} \equiv (1-\alpha)t^2/8$ and $\Pi_H \equiv (1-\alpha)q_H^2/4$.*

(a) A participating creator j sells only to consumers with nonnegative utility from her content, a set of measure $(q_j - P)^+$. The two sets are disjoint if $P > \frac{t-1}{2}$ and cover $[0,1]$ otherwise. She sells to her whole set if she participates alone or $P > \frac{t-1}{2}$; if both participate and $P \leq \frac{t-1}{2}$, every consumer buys.

(b) Under every participation profile, profit is at most $\max\{\Pi_{HL}, \Pi_H\}$. Equality requires either that both creators participate and $P = \frac{t}{4} < q_L$, or that $P = \frac{q_H}{2}$ and only H sells.

(c) $\Pi_{HL} - \Pi_H$ has the sign of $\rho^\star - q_H/q_L$. Hence, under every participation profile, profit is at most Π_{HL} if $q_H/q_L < \rho^\star$, with equality only if both creators participate and $P = \frac{t}{4}$.

Proof. (a) By Eq. (1), $U_L \geq 0$ if and only if $x \leq q_L - P$, and $U_H \geq 0$ if and only if $x \geq 1 - q_H + P$. Each set has measure $(q_j - P)^+$ as $P \geq 0$ and $q_j < 1$. The two sets are disjoint if $q_L - P < 1 - q_H + P$, that is, if $P > \frac{t-1}{2}$. Otherwise they cover $[0,1]$. A creator who participates alone or whose set is disjoint from the other's faces no competition there and sells to all of it. If both participate and the sets cover $[0,1]$, every consumer has nonnegative utility from some content and buys.

(b) Profit is $(1-\alpha)P$ times total demand. Suppose first that both creators sell, which requires that both participate and $P < q_L$. If the market is covered ($P \leq \frac{t-1}{2}$), profit is $(1-\alpha)P \leq (1-\alpha)\frac{t-1}{2} = \Pi_{HL} - (1-\alpha)\frac{(2-t)^2}{8} < \Pi_{HL}$. If it is not, total demand is $t - 2P$ by part (a), and

$$\Pi_{HL} - (1-\alpha)P(t-2P) = 2(1-\alpha)\Big(P - \frac{t}{4}\Big)^2,$$

so profit equals Π_{HL} only at $P = \frac{t}{4}$. This price always exceeds $\frac{t-1}{2}$, as $t < 2$, and it is below q_L if

and only if $q_H < 3q_L$. Suppose next that at most one creator sells. Profit is zero if none does and, by part (a), at most $(1-\alpha)P(q_j - P)^+$ if only creator j does. For $j = H$, $\Pi_H - (1-\alpha)P(q_H - P) = (1-\alpha)(P - \frac{q_H}{2})^2$, so profit equals Π_H only at $P = \frac{q_H}{2}$; for $j = L$, profit is at most $(1-\alpha)q_L^2/4 < \Pi_H$.

(c) The sign follows from

$$\Pi_{HL} - \Pi_H = (1-\alpha)\frac{q_L^2 + 2q_H q_L - q_H^2}{8} = \frac{(1-\alpha)q_L^2}{8}\Big(\rho^\star - \frac{q_H}{q_L}\Big)\Big(\frac{q_H}{q_L} + \sqrt{2} - 1\Big).$$

If $q_H/q_L < \rho^\star$, then $\Pi_H < \Pi_{HL}$ and $q_H < 3q_L$, so part (b) gives the second claim. □

Lemma B.3 gives demand when the AI is trained on both creators' content. It is Lemma 1 of the main paper stated for a general θ. We use it for pricing under both regimes (Lemmas B.7 and B.8).

Lemma B.3 (Demand when both creators contribute)**.** *Suppose both creators participate and contribute. If $\widetilde{P}_{AI} > \frac{t}{2}$, no consumer buys AI-generated content, and consumers choose among creators' content and the outside option as they would without the AI at the same price P. If $\widetilde{P}_{AI} \le \frac{t}{2}$, the market is covered and demands are those in the table of Lemma 1; in particular, they depend on prices only through δ.*

Proof. Every consumer obtains $U_{AI} = z - P_{AI} = \frac{t}{2} - \widetilde{P}_{AI}$ from the AI. If $\widetilde{P}_{AI} > \frac{t}{2}$, this utility is negative, so no consumer buys AI-generated content. Let $\widetilde{P}_{AI} \le \frac{t}{2}$. Then $U_{AI} \ge 0$ and, as ties with the outside option go to platform products (footnote 4), every consumer buys. By Eq. (1),

$$U_L - U_{AI} = (\delta - g) - x, \qquad U_H - U_{AI} = x - (1 - g - \delta), \qquad U_L - U_H = 2(x_{HL} - x),$$

where $x_{HL} \equiv \frac{1}{2} - g$ is the consumer who is indifferent between the two creators. Ties between a creator's content and the AI go to the creator. Hence a consumer at $x \neq x_{HL}$ buys from L if and only if $x \le \delta - g$ and $x < x_{HL}$, buys from H if and only if $x \ge 1 - g - \delta$ and $x > x_{HL}$, and buys AI-generated content otherwise. With $m \equiv \min\{\delta, \frac{1}{2}\}$ and up to endpoints, L's buyers are the consumers with $x \le m - g$ and H's are those with $x \ge 1 - g - m$. As $m - g < 1$ and $1 - g - m > 0$,

$$D_L = (m - g)^+, \qquad D_H = (m + g)^+, \qquad D_{AI} = 1 - D_L - D_H,$$

which gives the four rows of the table: $m = \frac{1}{2}$ if $\delta \ge \frac{1}{2}$, and $m = \delta$ otherwise. □

Lemma B.4 covers the remaining profiles, in which at most one creator contributes. It proves the last three rows of Lemma 2. We also use it to bound the platform's profit when a creator exits under NC (Lemma B.7) or withholds her data under C (Lemma B.9).

Lemma B.4 (Demand with at most one contributor)**.** *(a) Suppose only creator k contributes, and let $j \neq k$ and $y_k \equiv z - \tau_j q_j$. Then $U_k - U_{AI} = q_k - y_k + P_{AI} - P$ at every location, so k's content and AI-generated content cannot both attract demand: $D_k D_{AI} = 0$. If the AI sells, then $D_k = 0$, $P_{AI} < y_k$, and*

$$D_{AI} P_{AI} \le \tfrac{1}{4} y_k^2, \qquad D_j P \le \tfrac{1}{4} q_j^2.$$

(b) If neither creator contributes, the platform earns no AI revenue at any price.

Proof. (a) The AI is trained on k's content, so k participates. Both $U_k = q_k - P - d_k(x)$ and $U_{AI} = y_k - P_{AI} - d_k(x)$ fall one for one with distance from k, because prompting effort mirrors the mismatch cost. Their difference is thus the same for all consumers. If it is nonnegative, every consumer weakly prefers k's content, and $D_{AI} = 0$ because ties go to creators' content. If it is negative, every consumer strictly prefers the AI, and $D_k = 0$. A buyer of AI content has $U_{AI} \geq 0$, that is, $d_k(x) \leq y_k - P_{AI}$, so $D_{AI} \leq (y_k - P_{AI})^+$. If the AI sells, then $P_{AI} < y_k$ and $D_{AI}P_{AI} \leq (y_k - P_{AI})P_{AI} \leq \frac{1}{4}y_k^2$. A buyer of j's content has $U_j \geq 0$, so that $D_j P \leq (q_j - P)^+ P \leq \frac{1}{4}q_j^2$.

(b) Then $U_{AI} \leq -P_{AI}$, so no consumer buys AI-generated content unless $P_{AI} = 0$. □

Lemma B.5 bounds the platform's profit whenever the AI earns nothing. We use it to rule out such outcomes in Lemmas B.7, B.8 and B.9, and to prove part 3 of Proposition 6.

Lemma B.5 (Profit without AI revenue)**.** *Suppose that $q_H/q_L < \rho^\star$. Consider either regime, a price vector and a profile such that AI revenue $D_{AI}P_{AI}$ is zero. The platform's profit is then at most $\Pi_{HL} = (1-\alpha)t^2/8$, which is the benchmark profit π^B under Assumption 4 (Proposition 1). Equality requires that $P = \frac{t}{4}$ and that both creators participate and sell their benchmark demands $q_j - \frac{t}{4}$.*

Proof. Without AI revenue, the platform's profit is $(1-\alpha)DP$ in both regimes by Eq. (5). A buyer of creator j's content obtains nonnegative utility from it, so D is at most the measure of consumers who obtain nonnegative utility from some creator's content. By Lemma B.2(a), this measure is total demand at price P when both creators participate and the AI is not offered. Lemma B.2(c) then gives $(1-\alpha)DP \leq \Pi_{HL}$, with equality only if $P = \frac{t}{4}$ and $D = \frac{t}{2}$. At $P = \frac{t}{4} > \frac{t-1}{2}$, the two creators' sets of potential buyers are disjoint, with measures $q_j - \frac{t}{4}$ that sum to $\frac{t}{2}$. Hence $D = \frac{t}{2}$ requires each creator to participate and sell to her entire set of potential buyers. □

We close this section with prices at which the AI is priced at its quality z, so that its buyers obtain zero utility. Every equilibrium price pair characterized below at which the AI earns revenue has this form. At such prices the AI is viable only with both creators' data, which is what ties each creator's AI payment to her own contribution. We use Lemma B.6 in the last step of the proofs of Propositions 2, 3 and 5–8, where it shows which Stage-2 play the equilibrium prices induce.

Lemma B.6 (The AI priced at its quality)**.** *Let $z > 0$ and consider prices with $P_{AI} = z$ and $P > \frac{t-1}{2}$.*

(a) In either regime, AI-generated content sells only under $(1,1)$, and under every profile a participating creator j sells $(q_j - P)^+$. Under $(1,1)$, $D_{AI} = 1 - (q_H - P)^+ - (q_L - P)^+ > 0$.

(b) Under C, with equal splitting or quality weighting, suppose that under $(1,1)$ both creators' net payoffs are nonnegative and at least one is positive. Then $(1,1)$ is the only allowed play.

Proof. (a) Under $(1,1)$, every consumer obtains $U_{AI} = z - P_{AI} = 0$ from the AI. If only creator k contributes, $U_{AI} \leq -\tau_j q_j < 0$ for $j \neq k$, and if neither contributes, $U_{AI} \leq -z < 0$. So the AI sells nothing under every profile other than $(1,1)$. As $P > \frac{t-1}{2}$, no consumer obtains nonnegative utility from both creators' content (Lemma B.2(a)). Under every profile, the consumers who obtain nonnegative utility from a participating creator j's content therefore buy it, as the AI offers

them at most zero and ties go to creators' content; these consumers have measure $(q_j - P)^+$. Under $(1,1)$, all other consumers buy AI-generated content, which gives them zero utility, so $D_{AI} = 1 - (q_H - P)^+ - (q_L - P)^+$. This is positive because $P > \frac{t-1}{2}$ and $q_H < 1$.

(b) We let R_j denote creator j's net payoff under $(1,1)$ and A_j her AI payment, which is positive by part (a) as each contributor's share of AI revenue is positive under both rules. By part (a), under any other profile a participating creator earns the same sales revenue but no AI payment. Her net payoff is then $R_j - A_j$, and a creator who exits earns zero. So no creator gains by deviating from $(1,1)$: withholding data forfeits A_j, and exit yields $0 \leq R_j$. So $(1,1)$ is a Nash equilibrium, and it satisfies Assumption 2 because both creators participate. In every other pure-strategy equilibrium, creator j earns at most $\max\{R_j - A_j, 0\} \leq R_j$, with strict inequality if $R_j > 0$. Hence $(1,1)$ Pareto dominates every other pure-strategy equilibrium, and Assumption 3 selects it. As a pure-strategy equilibrium satisfying Assumption 2 exists, no mixed play is allowed. □

B.3 Proof of Proposition 1

Proof. Lemma B.2 bounds the platform's profit by Π_{HL} when both creators sell and by Π_H when at most one does. The quality ratio decides which bound is larger. It then remains to show that the creators indeed participate at the price that attains the larger bound.

Step 1 (Stage-2 play at $P > \frac{t-1}{2}$). By Lemma B.2(a), each creator j sells $(q_j - P)^+$ whatever the other does, so her decision does not depend on the other's. She participates if and only if $\alpha(q_j - P)^+ P \geq \phi q_j$, where Assumption 2 settles equality. This is the only allowed play. Allowed plays are pure (Lemma B.1(a)), so a bound on profit under every profile also holds under every allowed play.

Step 2 (Case $q_H/q_L < \rho^\star$ and $\phi \leq \phi^B$). By Lemma B.2(c), every profile yields at most Π_{HL}, with equality only at $P^B = \frac{t}{4}$ with both creators participating. As $P^B > \frac{t-1}{2}$, Step 1 applies. Creator j earns $V_j^B = \alpha(q_j - P^B)P^B$, and

$$\frac{V_L^B}{q_L} = \frac{\alpha(3q_L - q_H)(q_H + q_L)}{16 q_L} = \phi^B, \qquad \frac{V_H^B}{q_H} - \frac{V_L^B}{q_L} = \frac{\alpha(q_H - q_L)(q_H + q_L)^2}{16 q_H q_L} > 0.$$

The low-quality creator thus has the tighter participation constraint, and ϕ^B is the largest cost at which it holds. As $\phi \leq \phi^B$, both creators participate. Every allowed play therefore yields Π_{HL} at P^B and strictly less at every other price. By Lemma B.1(b), in every equilibrium both creators participate and

$$P^B = \frac{q_H + q_L}{4}, \quad D_L^B = \frac{3q_L - q_H}{4} > 0, \quad D_H^B = \frac{3q_H - q_L}{4}, \quad \pi^B = (1-\alpha)\frac{(q_H + q_L)^2}{8},$$

while the remaining mass $1 - \frac{t}{2} > 0$ of consumers takes the outside option.

Step 3 (Case $q_H/q_L > \rho^\star$ and $\phi \leq \alpha q_H/4$). By Lemma B.2(b) and (c), every profile yields at most Π_H, with equality only at $P = \frac{q_H}{2}$ with only H selling. As $\rho^\star > 2$, $\frac{q_H}{2} > q_L > \frac{t-1}{2}$, so Step 1 applies. L would sell nothing and exits because $\phi > 0$. H sells $\frac{q_H}{2}$ and participates because $\alpha q_H^2/4 \geq \phi q_H$. By Lemma B.1(b), every equilibrium has $P^B = \frac{q_H}{2}$, $D_H^B = \frac{q_H}{2}$, $D_L^B = 0$ and $\pi^B = \Pi_H$.

Step 4 (Case $q_H/q_L = \rho^\star$ and $\phi \le \alpha q_L/8$). Here $\Pi_{HL} = \Pi_H$ by Lemma B.2(c), and $2 < \rho^\star < 3$ gives $\frac{t}{4} < q_L < \frac{q_H}{2}$. By Lemma B.2(b), profit thus equals Π_{HL} exactly at $\frac{t}{4}$ with both creators participating and at $\frac{q_H}{2}$ with only H selling. As $(2-\sqrt{2})(2+\sqrt{2}) = 2$, $\phi^B = \alpha q_L(3-\rho^\star)(1+\rho^\star)/16 = \alpha q_L/8 < \alpha q_H/4$, so $\phi \le \phi^B < \alpha q_H/4$ and the Stage-2 arguments of Steps 2 and 3 apply at $\frac{t}{4}$ and $\frac{q_H}{2}$. By Lemma B.1(b), the equilibrium prices are $\frac{t}{4}$, where both creators sell, and $\frac{q_H}{2}$, where only H sells. □

B.4 Proof of Lemma 1

Proof. In the main analysis $\theta = q^{\text{base}}$, so the lemma is Lemma B.3. □

B.5 Proof of Proposition 2 and Corollary 1

Throughout this section, we consider regime NC with equal transfer rates ($\tau_H = \tau_L = \frac{1}{2}$) and maintain Assumption 4, so $\pi^B = (1-\alpha)t^2/8$ (Proposition 1). Recall that $s \equiv \sqrt{q_L^2 - 4\phi q_L/\alpha}$. As s^2 strictly decreases in ϕ and equals g^2 at $\phi = \phi^B$, $0 < \phi \le \phi^B$ gives $g \le s < q_L$, with $s = g$ only at $\phi = \phi^B$. We let $\delta^{NC} \equiv \frac{q_H - s}{2}$, which lies in $(g, \frac{1}{2})$ because $s < q_L$ and $q_H < 1$, take P^{NC} and D_j^{NC} from Proposition 2, and define

$$\pi^{NC}(\phi, \theta) \equiv (1-\alpha)(q_H - s)\frac{q_L + s}{2} + (1 - q_H + s)\Big(\frac{q_H + q_L}{2} + \theta\Big). \tag{B.2}$$

In the main analysis, $\pi^{NC}(\phi, q^{\text{base}})$ is the profit Eq. (6) of Corollary 1.

We use Lemma B.7 on pricing under NC to prove Proposition 2 and Corollary 1. We state it for a general θ so that it also serves Proposition 5. Its message is that the platform keeps the full price of AI-generated content but only the share $1-\alpha$ of creators' revenue. Yet it must leave L enough sales to cover her participation cost, because her exit would remove her training data.

Lemma B.7 (Pricing under NC). *Let $\kappa_0 \ge 0$, so that $\theta = q^{base} - \kappa_0$ and $z = \theta + \frac{t}{2}$.*

(a) In a pure-strategy Stage-2 equilibrium in which a creator exits or the AI sells nothing, profit is at most π^B. If a creator exits, it is at most $(1-\alpha)q_H^2/4$, which is strictly below π^B.

(b) Consider a pure-strategy Stage-2 equilibrium in which both creators participate and the AI sells. Then all three products sell with $g < \delta < \frac{1}{2}$ and $\delta \ge \delta^{NC}$, and profit is at most

$$\widetilde{\pi}(\delta) \equiv \frac{t}{2} - \alpha t\delta - 2(1-\alpha)\delta^2 + (1-2\delta)\theta, \tag{B.3}$$

with equality if and only if the platform prices the AI at its quality, $P_{AI} = z$.

(c) Suppose $z > 0$. At (P^{NC}, z), the only allowed play has both creators participate. Demands are $D_j^{NC} = q_j - P^{NC}$ and $D_{AI}^{NC} = 1 - q_H + s$, L's participation constraint binds ($V_L = \phi q_L$), H's is slack ($V_H = \alpha(q_L + s)(2q_H - q_L - s)/4 > \phi q_H$), and profit is $\widetilde{\pi}(\delta^{NC}) = \pi^{NC}(\phi, \theta)$.

(d) If $\theta \ge 0$, then $\partial \pi^{NC}(\phi, \theta)/\partial s = (q_H - s) + \alpha(s - g) + \theta > 0$ and $\pi^{NC}(\phi, \theta) > \pi^B$.

Proof. (a) If neither creator participates, AI revenue is zero (Lemma B.4(b)). If only creator j participates, the AI is trained on her content alone. In equilibrium her sales cover her cost, $\alpha D_j P \ge$

$\phi q_j > 0$, so $D_j > 0$ and hence $D_{AI} = 0$ (Lemma B.4(a)). Her buyers obtain nonnegative utility from her content, so $D_j \le (q_j - P)^+$. In these two cases profit is therefore at most $(1-\alpha)P(q_j - P)^+ \le (1-\alpha)q_H^2/4 = \Pi_H$, and $\Pi_H < \Pi_{HL} = \pi^B$ by Lemma B.2(c). When both creators participate but the AI sells nothing, AI revenue is zero, and Lemma B.5 gives the bound.

(b) Both creators participate, so each covers her cost and $D_L > 0$. As $D_{AI} > 0$ as well, Lemma B.3 gives $\widetilde{P}_{AI} \le \frac{t}{2}$ and $g < \delta < \frac{1}{2}$, with $D_H = \delta + g$, $D_L = \delta - g$ and $D_{AI} = 1 - 2\delta$. Substituting $P_{AI} = P + \delta + \theta$ into Eq. (5), profit is

$$\pi(P, \delta) \equiv (1-\alpha)\, 2\delta P + (1 - 2\delta)\big(P + \delta + \theta\big).$$

Now raise both prices by the same amount, which leaves δ and hence all demands unchanged. Profit rises, as $\partial\pi/\partial P = 1 - 2\alpha\delta > 0$, and so does each creator's revenue. The only limit is that consumers must remain willing to buy AI-generated content, $P + \delta = \widetilde{P}_{AI} \le \frac{t}{2}$. Hence $\pi(P, \delta) \le \pi(\frac{t}{2} - \delta, \delta) = \widetilde{\pi}(\delta)$, with equality if and only if $P_{AI} = z$, the price at which AI buyers obtain zero utility. For the same reason, L covers her cost at (P, δ) only if she does at $(\frac{t}{2} - \delta, \delta)$, that is, only if $\alpha(\delta - g)(\frac{t}{2} - \delta) \ge \phi q_L$. Using $\phi q_L/\alpha = (q_L^2 - s^2)/4$,

$$\alpha(\delta - g)\Big(\frac{t}{2} - \delta\Big) - \phi q_L = -\alpha\Big(\delta - \frac{q_H - s}{2}\Big)\Big(\delta - \frac{q_H + s}{2}\Big),$$

so $\delta \ge \delta^{NC}$: the price gap must leave L enough buyers to cover her cost.

(c) At (P^{NC}, z), $\delta = \frac{t}{2} - P^{NC} = \delta^{NC}$, so under $(1, 1)$ Lemma B.3 gives

$$D_L = \delta^{NC} - g = \frac{q_L - s}{2}, \qquad D_H = \delta^{NC} + g = \frac{2q_H - q_L - s}{2}, \qquad D_{AI} = 1 - 2\delta^{NC} = 1 - q_H + s,$$

all positive, with $D_j = q_j - P^{NC}$. Hence $V_L = \alpha(q_L^2 - s^2)/4 = \phi q_L$ and $V_H - \phi q_H = \alpha(q_H - q_L)(q_L + s)^2/(4q_L) > 0$, and profit is $\pi(P^{NC}, \delta^{NC}) = \widetilde{\pi}(\delta^{NC})$, which equals Eq. (B.2). As $P^{NC} \ge \frac{q_L + g}{2} = \frac{t}{4} > \frac{t-1}{2}$, Lemma B.6(a) applies: each participating creator sells $q_j - P^{NC}$ whether or not the other participates. Participating thus gives L a net payoff of zero and H a positive one whatever the other creator does. The pure-strategy equilibria are therefore $(1, 1)$ and the profile in which only H participates. The latter violates Assumption 2 because the indifferent L exits. As allowed plays are pure under NC (Lemma B.1(a)), the profile $(1, 1)$ is the only allowed play.

(d) Differentiating Eq. (B.2) gives the stated derivative, which is positive because $g \le s < q_H$ and $\theta \ge 0$. So $\pi^{NC}(\phi, \theta)$ is at least its value at $s = g$, where $q_H - s = q_L + s = \frac{t}{2}$:

$$\pi^{NC}(\phi, \theta) \ge \pi^B + \Big(1 - \frac{t}{2}\Big)\Big(\theta + \frac{t}{2}\Big) > \pi^B.$$

At $s = g$, the human price and creators' demands are those of the benchmark, and the AI adds the revenue from the consumers whom the benchmark leaves unserved. □

Proof of Proposition 2. In the main analysis, $\theta = q^{\text{base}} \in [0, 1]$ and $z = q^{\text{base}} + \frac{t}{2} = P_{AI}^{NC} > 0$. We show that the platform sells as much AI-generated content as L's participation allows.

Step 1 (The platform leaves L the minimum). Consider a pure-strategy Stage-2 equilibrium in which both creators participate and the AI sells. By Lemma B.7(b), profit is at most $\widetilde{\pi}(\delta)$ and $\delta \geq \delta^{NC}$. As $\theta \geq 0$, $\widetilde{\pi}$ strictly decreases in δ,

$$\widetilde{\pi}'(\delta) = -\alpha t - 4(1-\alpha)\delta - 2\theta < 0,$$

because a wider price gap moves consumers from AI-generated content to creators' content, on which the platform earns less. The platform's profit is therefore at most $\widetilde{\pi}(\delta^{NC}) = \pi^{NC}$, with equality only if $\delta = \delta^{NC}$ and $P_{AI} = z$, that is, only at the price pair (P^{NC}, P_{AI}^{NC}).

Step 2 (Other outcomes yield less). By Lemma B.7(a), a pure-strategy Stage-2 equilibrium in which a creator exits or the AI sells nothing yields at most π^B, and $\pi^B < \pi^{NC}$ by Lemma B.7(d). So every pure-strategy Stage-2 equilibrium at any other price pair yields strictly less than π^{NC}.

Step 3 (Equilibrium). By Lemma B.7(c), at (P^{NC}, P_{AI}^{NC}) the only allowed play has both creators participate, and it yields π^{NC}. As allowed plays are pure under NC (Lemma B.1(a)), Step 2 covers all allowed plays at every other price pair. Lemma B.1(b) with $\mathcal{P}^* = \{(P^{NC}, P_{AI}^{NC})\}$ and $\pi^* = \pi^{NC}$ then shows that an equilibrium exists and that every equilibrium has these prices, with the participation decisions, demands and payoffs stated in Lemma B.7(c).

Proof of Corollary 1. The platform's equilibrium profit is $\pi^{NC}(\phi, q^{\text{base}})$, which is Eq. (6). It depends on ϕ only through s, with $ds/d\phi = -2q_L/(\alpha s) < 0$, so it strictly decreases in ϕ by Lemma B.7(d). As s does not depend on q^{base}, $\partial \pi^{NC}/\partial q^{\text{base}} = 1 - q_H + s = D_{AI}^{NC} > 0$. □

B.6 Proof of Lemma 2

Proof. The columns q_{AI} and $\kappa(x)$ follow directly from Eq. (2) and Eq. (3). Row $(1,1)$ is Lemma B.3, rows $(1,0)$ and $(0,1)$ are Lemma B.4(a), and row $(0,0)$ is Lemma B.4(b). These lemmas hold whether a creator who does not contribute her data participates or exits the platform. □

B.7 Pricing under Compensation

In this section, we characterize the platform's profit under C profile by profile, for the main analysis and the extensions at once. We maintain Assumption 4 throughout, so $\pi^B = (1-\alpha)t^2/8$ (Proposition 1). For $z \geq 0$, we define

$$D_H^C(z) \equiv \tfrac{1}{2}(q_H - z)^+, \qquad F(z) \equiv z + D_H^C(z)^2, \qquad G(z) \equiv z\left[1 - D_H^C(z)\right],$$

and we let $\mathcal{P}(z)$ denote the set of *displacement prices*: the pair $(\frac{q_H+z}{2}, z)$ if $z < q_H$, and the pairs (P, z) with $P \geq q_H$ if $z \geq q_H$. Both F and G are continuous and strictly increasing: $F'(z) = 1 - D_H^C(z) > 0$ and $G'(z) = 1 - \frac{q_H}{2} + z > 0$ for $z < q_H$, and $F' = G' = 1$ for $z > q_H$. We let $\underline{\theta}_C \equiv -(1+\sqrt{2})g$, so that $-3g < \underline{\theta}_C < -2g$; Assumption 4 gives $\underline{\theta}_C + \frac{t}{2} = \frac{(2+\sqrt{2})q_L - \sqrt{2}q_H}{2} > 0$. Lemmas B.8 and B.10 describe the displacement prices when $\theta \geq \underline{\theta}_C$ and both creators contribute. There the AI displaces L and H keeps demand $D_H^C(z)$. The platform earns $(1-\alpha)F(z)$ and AI revenue is $G(z)$.

Lemma B.8 gives the platform's most profitable prices when both creators contribute their data. This lemma underlies Proposition 3 and its extensions in Propositions 6–8.

Lemma B.8 (Profit when both creators contribute)**.** *Let $z > 0$, that is, $\theta > -\frac{t}{2}$. Under C and $(1,1)$, profit at any price pair is at most $(1-\alpha)M$, with M and the price pairs attaining the bound as follows.*

(i) Displacement. *If $\theta > \underline{\theta}_C$, then $M = F(z)$, and the bound is attained exactly at the displacement prices $\mathcal{P}(z)$, where demands are $D_L = 0$, $D_H = D_H^C(z)$ and $D_{AI} = 1 - D_H^C(z)$.*

(ii) Three products. *If $\theta < \underline{\theta}_C$, then $M = z + \frac{1}{2}\theta^2$, and the bound is attained only at the three-product prices $(\frac{t+\theta}{2}, z)$, where $D_H = g - \frac{\theta}{2}$, $D_L = -g - \frac{\theta}{2} > 0$ and $D_{AI} = 1 + \theta > 0$.*

(iii) Tie. *If $\theta = \underline{\theta}_C$, then $M = F(z) = z + \frac{1}{2}\theta^2$, and the bound is attained only at the pair in $\mathcal{P}(z)$ and at the three-product prices, with the demands stated in (i) and (ii), respectively.*
In all three cases, $M \geq F(z) \geq z$ and the bound exceeds the benchmark profit, $(1-\alpha)M > \pi^B$.

Proof. Under $(1,1)$ the platform's choice reduces to the price gap δ, which determines who buys what (Lemma B.3). We first show that the platform prices the AI at z, and then maximize over δ.

If the AI sells nothing, AI revenue is zero and profit is at most π^B (Lemma B.5), which is below the bound by the last claim of the lemma. Suppose $D_{AI} > 0$. Then $U_{AI} = z - P_{AI} \geq 0$, so $P_{AI} \leq z$. By Lemma B.3, the market is covered and demands depend on prices only through δ. Here $\delta < \frac{1}{2}$ because $D_{AI} > 0$, and $\delta = P_{AI} - \theta - P \leq z - \theta = \frac{t}{2}$ because $P \geq 0$. Substituting $P = P_{AI} - \theta - \delta$ and $D_H + D_L + D_{AI} = 1$ into Eq. (5),

$$\pi = (1-\alpha)\big[\Psi(\delta) - (z - P_{AI})\big], \qquad \Psi(\delta) \equiv (D_H + D_L)\Big(\frac{t}{2} - \delta\Big) + D_{AI}\, z.$$

So $\pi \leq (1-\alpha)\Psi(\delta)$, with equality if and only if $P_{AI} = z$. In a covered market, raising both prices by the same amount loses no buyer, so the platform raises the AI price until AI buyers obtain zero utility. The human price is then $P = \frac{t}{2} - \delta$, and by Lemma B.3,

$$\Psi(\delta) = \begin{cases} z, & \delta \leq -g \quad \text{(only the AI sells)}, \\ (\delta + g)(\frac{t}{2} - \delta) + (1 - \delta - g)z, & -g \leq \delta \leq g \quad \text{(the AI and } H \text{ sell)}, \\ 2\delta(\frac{t}{2} - \delta) + (1 - 2\delta)z, & g \leq \delta \leq \bar{\delta} \equiv \min\{\frac{1}{2}, \frac{t}{2}\} \quad \text{(all three sell)}. \end{cases}$$

Ψ is continuous, and on the second and third ranges it is strictly concave, with derivatives $-(2\delta + g + \theta)$ and $-2(2\delta + \theta)$. These vanish at $\delta_c \equiv -\frac{g+\theta}{2}$ and $\delta_3 \equiv -\frac{\theta}{2}$, where the two expressions equal $z + \frac{1}{4}(g-\theta)^2$ and $z + \frac{1}{2}\theta^2$, and

$$\Big[z + \frac{(g-\theta)^2}{4}\Big] - \Big[z + \frac{\theta^2}{2}\Big] = \frac{g^2 - 2g\theta - \theta^2}{4}. \tag{B.4}$$

The right-hand side is a downward parabola in θ with roots $\underline{\theta}_C$ and $(\sqrt{2}-1)g$, so it is positive for $\underline{\theta}_C < \theta < (\sqrt{2}-1)g$ and negative for $\theta < \underline{\theta}_C$. Four cases cover $\theta > -\frac{t}{2}$.

Case $\theta \geq g$. Both derivatives are negative for $\delta > -g$, so Ψ is maximized exactly on $\delta \leq -g$, where it equals $z = F(z)$. With $P_{AI} = z$, the condition $\delta \leq -g$ means $P \geq q_H$.

Case $-2g \le \theta < g$. Then $\delta_c \in (-g, g)$ and $\delta_3 \le g$, so Ψ decreases on $[g, \bar{\delta}]$ and is maximized only at δ_c, where it equals $z + \frac{1}{4}(g-\theta)^2 = F(z)$, as $g - \theta = q_H - z > 0$.

Case $-3g \le \theta < -2g$. Then $\delta_c \in (-g, g]$ and $\delta_3 \in (g, \frac{t}{4}) \subset (g, \bar{\delta})$, as $\theta > -\frac{t}{2}$ and $t < 2$. Ψ then has one peak on each of the last two ranges, and Eq. (B.4) ranks the two peaks. The maximum of Ψ is at δ_c alone if $\theta > \underline{\theta}_C$, at δ_3 alone if $\theta < \underline{\theta}_C$, and at both peaks if $\theta = \underline{\theta}_C$.

Case $\theta < -3g$. Then $\delta_c > g$, so Ψ increases on $[-g, g]$ and is maximized only at $\delta_3 \in (g, \frac{t}{4})$.

As $-3g < \underline{\theta}_C < -2g$, these four cases give M as stated in (i)–(iii). Recall that $P_{AI} = z$ and $P = \frac{t}{2} - \delta$. The gap δ_c lies in $(-g, g]$ whenever it is a maximizer. It gives $P = \frac{q_H + z}{2}$ and, by Lemma B.3, $D_L = 0$ and $D_H = \delta_c + g = \frac{q_H - z}{2} = D_H^C(z)$. The gap δ_3 gives $P = \frac{t+\theta}{2}$, $D_H = g - \frac{\theta}{2}$, $D_L = -g - \frac{\theta}{2}$ and $D_{AI} = 1 + \theta$, where $D_L > 0$ because $\theta < -2g$ in this case.

Finally, $M \ge F(z) \ge z$ in every case, by Eq. (B.4) when $\theta < \underline{\theta}_C$. Moreover, $\frac{t}{4} \in (g, \bar{\delta})$ because $q_H < 3q_L$, so $M \ge \Psi(\frac{t}{4}) = \frac{t^2}{8} + (1 - \frac{t}{2})z > \frac{t^2}{8}$ and $(1-\alpha)M > \pi^B$. Intuitively, the platform already earns more than in the benchmark by keeping the benchmark price $\frac{t}{4}$ for creators' content and selling AI-generated content to the consumers whom the benchmark leaves unserved. □

With a single contributor, the AI competes head-on with that creator's content. Lemma B.9 shows that the platform then earns at most the benchmark profit if the AI earns nothing, and less than $(1-\alpha)z$ otherwise. Under $(1,1)$, it can secure $(1-\alpha)z$ by letting the AI serve the whole market at price z. We use this lemma in Lemma B.11 and in the proof of Proposition 6 to show that the platform prefers a profile in which both creators contribute their data.

Lemma B.9 (Profit with at most one contributor)**.** *Consider regime C in either of two settings: equal transfer rates ($\tau_H = \tau_L = \frac{1}{2}$) with any $\kappa_0 \ge 0$, or $\kappa_0 = 0$ with $\tau_H = \frac{1}{2}$ and $0 < \tau_L \le \frac{1}{2}$. Let $z > 0$. At every price pair and under every profile with at most one contributor, the platform's profit is at most the benchmark profit π^B if AI revenue is zero, and strictly below $(1-\alpha)z$ otherwise.*

Proof. If AI revenue is zero, Lemma B.5 applies. Otherwise exactly one creator k contributes (Lemma B.4(b)) and the AI sells. With $j \ne k$, Lemma B.4(a) gives $D_k = 0$ and bounds the two remaining revenues, so profit is at most $(1-\alpha)(q_j^2 + y_k^2)/4$, where $0 < y_k = z - \tau_j q_j$. By Assumption 1, $y_k = q^{\text{base}} - \kappa_0 + \tau_k q_k < 2$, so $y_k^2/4 < y_k/2$ and

$$z - \frac{q_j^2 + y_k^2}{4} > z - \frac{q_j^2}{4} - \frac{z - \tau_j q_j}{2} = \frac{z}{2} + \frac{q_j(2\tau_j - q_j)}{4}.$$

The last expression is strictly positive if $\tau_j = \frac{1}{2}$, because $q_j < 1$. In the remaining case, $\kappa_0 = 0$ and $j = L$ with $\tau_L < \frac{1}{2}$, the expression is at least $\frac{z}{2} - \frac{q_L^2}{4} > 0$, because $z > \frac{q_H}{2} > \frac{q_L^2}{2}$. □

Lemma B.10 gives payoffs at the displacement prices. We use it to show that both creators participate (Propositions 6–8) and to compare payoffs across regimes (Corollary 3 and Proposition 4).

Lemma B.10 (Creators' payoffs at the displacement prices)**.** *Let $\theta \ge \underline{\theta}_C$ and $z > 0$, and consider $(1,1)$*

at a price pair in $\mathcal{P}(z)$. AI revenue is $G(z)$. Under equal splitting, L's gross payoff is $\alpha G(z)/2$, and H's is

$$V_H(z) \equiv \begin{cases} \alpha\big[q_H^2 + (2-q_H)z\big]/4, & z < q_H, \\ \alpha z/2, & z \geq q_H, \end{cases}$$

which is continuous and strictly increasing in z and satisfies $V_H(z) > \alpha q_H^2/4 > \phi^B q_H$. Hence H's net payoff $V_H(z) - \phi q_H$ is strictly positive for every $z > 0$ and every $\phi \leq \phi^B$.

Proof. By Lemma B.8, $D_L = 0$, $D_H = D_H^C(z)$ and $D_{AI} = 1 - D_H^C(z)$, so AI revenue is $D_{AI} z = G(z)$. As L sells nothing, her payoff is half of the creators' share of this revenue. If $z \geq q_H$, H also sells nothing and earns $\alpha z/2 \geq \alpha q_H/2$. If $z < q_H$, she sells $\frac{q_H - z}{2}$ at the price $\frac{q_H + z}{2}$, so by Eq. (4),

$$\frac{V_H(z)}{\alpha} = \frac{(q_H - z)(q_H + z)}{4} + \frac{1}{2}\Big(1 - \frac{q_H - z}{2}\Big)z = \frac{q_H^2 + (2 - q_H)z}{4}.$$

Both expressions increase in z and equal $\alpha q_H/2$ at $z = q_H$. They exceed $\alpha q_H^2/4$ because $z > 0$ and $q_H < 1$, and $\frac{\alpha q_H}{4} - \phi^B = \frac{\alpha(q_H - q_L)(q_H + 3q_L)}{16 q_L} > 0$. As $\phi \leq \phi^B$, $V_H(z) > \phi q_H$. □

In Lemma B.11, we combine Lemmas B.8 and B.9 when there is no base prompting effort, as in the main analysis and in Sections 4.2 and 4.3. It is the first step in proving Propositions 3, 7 and 8.

Lemma B.11 (Profit bound under C without base prompting effort)**.** *Let $\kappa_0 = 0$, $\tau_H = \frac{1}{2}$ and $0 < \tau_L \leq \frac{1}{2}$, and suppose $q_H/q_L \geq \sqrt{2}$ if $\tau_L < \frac{1}{2}$. Then $\theta > \underline{\theta}_C$. Under either sharing rule, the platform's profit is at most $(1-\alpha)F(z)$ at every price pair and under every profile. Equality holds if and only if the profile is $(1,1)$ and the price pair lies in $\mathcal{P}(z)$. At these pairs, $D_L = 0$, $D_H = D_H^C(z)$ and $D_{AI} = 1 - D_H^C(z)$.*

Proof. Neither consumer demand nor the platform's profit depends on how creators split their share of AI revenue. Here $z = q^{\text{base}} + \frac{q_H}{2} + \tau_L q_L > 0$ and $\theta = q^{\text{base}} - (\frac{1}{2} - \tau_L)q_L$. If $\tau_L = \frac{1}{2}$, then $\theta = q^{\text{base}} \geq 0 > \underline{\theta}_C$. If $\tau_L < \frac{1}{2}$, then $\theta > -\frac{q_L}{2}$ and

$$-\frac{q_L}{2} - \underline{\theta}_C = \frac{1 + \sqrt{2}}{2}\big(q_H - \sqrt{2}\, q_L\big) \geq 0,$$

so $\theta > \underline{\theta}_C$: when $q_H/q_L \geq \sqrt{2}$, a low transfer rate never makes the platform prefer the three-product prices. Lemma B.8(i) thus gives the claim for $(1,1)$, and $(1-\alpha)F(z) > \pi^B$. Under every other profile, profit is at most π^B or strictly below $(1-\alpha)z \leq (1-\alpha)F(z)$ by Lemma B.9. □

B.8 Proof of Proposition 3 and Corollaries 2 and 3

Proof. In the main analysis, $z = q^{\text{base}} + \frac{t}{2} = P_{AI}^C$. As $\overline{q}^{\text{base}} = g$ and $q_H = g + \frac{t}{2}$, CQ1 is the case $z < q_H$ and CQ2 the case $z \geq q_H$, and the price pairs in the statement are the displacement prices $\mathcal{P}^C \equiv \mathcal{P}(z)$. We write $D_H^C \equiv D_H^C(z)$ and $\pi^C \equiv (1-\alpha)F(z) = (1-\alpha)[P_{AI}^C + (D_H^C)^2]$.

Under compensation, a creator no longer needs direct sales to cover her participation cost, because her share of AI revenue can cover it instead. The platform can therefore choose the prices

that maximize its profit when both creators contribute, which fully displace L (Step 1), provided that the AI payment keeps both creators on the platform (Step 2).

Step 1 (Profit bound). By Lemma B.11 with $\tau_L = \frac{1}{2}$, every profile yields at most π^C at every price pair. Equality holds only under $(1,1)$ at the pairs in $\mathcal{P}^C$, where $D_L = 0$, $D_H = D_H^C$ and $D_{AI} = 1 - D_H^C$. In CQ1, the pair is (P^C, P_{AI}^C) with $P^C = \frac{q_H+z}{2} = \frac{3q_H+q_L}{4} + \frac{q^{\text{base}}}{2}$, and $D_H^C = \frac{q_H-z}{2} = \frac{g-q^{\text{base}}}{2} > 0$. In CQ2, $D_H^C = 0$ and the AI serves the whole market at every $P \geq q_H$.

Step 2 (The AI payment covers both creators' costs). Under $(1,1)$ at a pair in $\mathcal{P}^C$, each creator receives the AI payment $A^C \equiv \frac{\alpha}{2} D_{AI}^C P_{AI}^C$. As $D_{AI}^C = 1 - D_H^C \geq 1 - \frac{g}{2} > \frac{3}{4}$ and $P_{AI}^C \geq \frac{t}{2}$, $A^C > \frac{3\alpha t}{16}$, and

$$\frac{3\alpha t}{16} - \phi^B q_H = \frac{\alpha t \left[3q_L(1-q_H) + q_H^2\right]}{16\, q_L} > 0.$$

So $V_j^C \geq A^C > \phi^B q_H \geq \phi q_j$ by Assumption 4: even L, with no sales, prefers to participate.

Step 3 (Equilibrium). At every pair in $\mathcal{P}^C$, $P > q_L > \frac{t-1}{2}$, since $P^C - q_L = \frac{3g+q^{\text{base}}}{2} > 0$ in CQ1 and $P \geq q_H$ in CQ2. As both net payoffs are positive (Step 2), Lemma B.6(b) shows that $(1,1)$ is the only allowed play there, and it yields π^C. By Step 1 and Lemma B.1(c), every allowed play yields strictly less at every pair outside $\mathcal{P}^C$. Lemma B.1(b) with $\mathcal{P}^* = \mathcal{P}^C$ and $\pi^* = \pi^C$ proves the proposition: in every equilibrium the prices lie in $\mathcal{P}^C$, creators play $(1,1)$, and demands are those of Step 1. In CQ2, every $P \geq q_H$ is an equilibrium price, with the same demands, profit and payoffs.

Proof of Corollary 2. By Propositions 2 and 3, and as $s < q_L$ for $\phi > 0$, $D_L^C = 0 < \frac{q_L-s}{2} = D_L^{NC}$ and $D_H^C \leq \frac{g}{2} < 2g < 2g + \frac{q_L-s}{2} = D_H^{NC}$. As the market is covered in both regimes, the AI's demand is the remainder in each, so $D_{AI}^C - D_{AI}^{NC} = (D_H^{NC} - D_H^C) + D_L^{NC} > 0$.

Proof of Corollary 3. Step 2 above gives $V_j^C > \phi q_j$ for both creators in CQ1 and CQ2. As $D_L^C = 0$, L's payoff is her AI payment alone, $V_L^C = \frac{\alpha}{2} G(z)$ (Lemma B.10). G is continuous and strictly increasing and $z = q^{\text{base}} + \frac{t}{2}$, so V_L^C strictly increases in q^{base} on $[0,1]$. In CQ2, $D_H^C = 0$ as well, so $V_H^C = V_L^C = \frac{\alpha}{2} z = \frac{\alpha}{2} q_{AI}$, because $q_{AI} = z$ when both creators contribute. □

B.9 Proof of Corollary 4 and Proposition 4

Proof. Both results compare the equilibria of Propositions 2 and 3, in which both creators participate and contribute their data. We use the following facts from these two propositions. The quality and the price of AI-generated content are the same in the two regimes, $q_{AI} = P_{AI}^{NC} = P_{AI}^C = z = q^{\text{base}} + \frac{t}{2} > 0$. For $r \in \{NC, C\}$, we let $D^r \equiv D_H^r + D_L^r$, so that $D^r P^r$ is human-content revenue and $D_{AI}^r = 1 - D^r$. Under NC, $D^{NC} = q_H - s$, $D_H^{NC} - D_L^{NC} = 2g$ and $P^{NC} = \frac{q_L+s}{2}$, none of which depends on q^{base}. Under C, $D^C = D_H^C = \frac{1}{2}\max\{g - q^{\text{base}}, 0\} \leq \frac{g}{2}$, so $(D_H^C)^2$ is differentiable in q^{base} with derivative $-D_H^C$; in CQ1, $P^C = q_H - D_H^C$. Finally, $\phi = \alpha(q_L^2 - s^2)/(4q_L)$ strictly decreases in s and equals ϕ^B at $s = g$, so the range $\phi \in (0, \phi^B]$ corresponds to $s \in [g, q_L)$.

Proof of Corollary 4. The two regimes respond differently to ϕ and to q^{base}. A higher participation cost lowers profit only under NC, where the platform must then leave L more sales; this single crossing in ϕ defines the threshold $\phi^\star$ (Steps 1 and 2). A more capable AI raises the AI price one

for one in both regimes, and wherever the platform weakly prefers C this raises its profit by more under C than under NC; this is why the threshold falls as q^{base} rises (Step 3).

Step 1 (Both regimes beat the benchmark). By Lemma B.7(d) and Corollary 1, $\pi^{NC} > \pi^B$ and π^{NC} strictly decreases in ϕ. By Lemma B.8 with $\theta = q^{\text{base}}$, $\pi^C = (1-\alpha)F(z) = (1-\alpha)[P_{AI}^C + (D_H^C)^2] > \pi^B$, and the profit π^C does not depend on the participation cost parameter ϕ.

Step 2 (The threshold). For $s \in [g, q_L]$, we let $\Lambda(s, q^{\text{base}}) \equiv \pi^{NC} - \pi^C$, with π^{NC} given by Eq. (B.2) (at $s = q_L$, its limit as $\phi \downarrow 0$). By Lemma B.7(d), $\partial\Lambda/\partial s = (q_H - s) + \alpha(s - g) + q^{\text{base}} > 0$. Fix q^{base}. If $\Lambda(g, q^{\text{base}}) \geq 0$, we set $\phi^\star = \phi^B$; then $\pi^{NC} > \pi^C$ for $\phi < \phi^B$ and $\pi^{NC} \geq \pi^C$ at ϕ^B. If $\Lambda(q_L, q^{\text{base}}) \leq 0$, which excludes the previous case, we set $\phi^\star = 0$; then $\pi^{NC} < \pi^C$ on $(0, \phi^B]$. Otherwise $\Lambda(\cdot, q^{\text{base}})$ is negative at g and positive at q_L, so it has a unique zero $s^\star \in (g, q_L)$; we set $\phi^\star = \alpha(q_L^2 - (s^\star)^2)/(4q_L) \in (0, \phi^B)$, so that $\pi^{NC} \gtrless \pi^C$ as $\phi \lessgtr \phi^\star$, with equality at $\phi^\star$.

Step 3 (A better AI lowers the threshold). We show that $\phi^\star$ is nonincreasing in q^{base}, and strictly decreasing wherever $0 < \phi^\star < \phi^B$, as claimed in Section 3.4. The argument has three parts.

First, human-content revenue is higher under NC. On $[g, q_L]$, $D^{NC}P^{NC} = (q_H - s)(q_L + s)/2$ has s-derivative $g - s \leq 0$, so it is at least $2gq_L$, its value at $s = q_L$. In CQ1, $D^C P^C = (g - q^{\text{base}})(3q_H + q_L + 2q^{\text{base}})/8$ has q^{base}-derivative $-(t + 2q^{\text{base}})/4 < 0$, so it is at most $g(3q_H + q_L)/8$, its value at $q^{\text{base}} = 0$; in CQ2 it is zero. Hence, for all $s \in [g, q_L]$ and $q^{\text{base}} \in [0, 1]$,

$$D^{NC}P^{NC} - D^C P^C \geq 2gq_L - \frac{g(3q_H + q_L)}{8} = \frac{3g(5q_L - q_H)}{8} > 0, \tag{B.5}$$

where the last inequality holds because $q_H/q_L < \rho^\star < 5$ by Assumption 4.

Second, wherever the platform weakly prefers C, a better AI strengthens this preference: $\partial\Lambda/\partial q^{\text{base}} < 0$ wherever $\Lambda \leq 0$. In each regime, profit rises with q^{base} at the rate of the AI demand whose revenue the platform keeps: $\partial\pi^{NC}/\partial q^{\text{base}} = D_{AI}^{NC}$ by Corollary 1, and $\partial\pi^C/\partial q^{\text{base}} = (1 - \alpha)(1 - D_H^C) = (1-\alpha)D_{AI}^C$. Writing $\pi^{NC} = (1-\alpha)D^{NC}P^{NC} + D_{AI}^{NC}z$ and $\pi^C = (1-\alpha)(D^C P^C + D_{AI}^C z)$ by Eq. (5) and subtracting,

$$z\,\frac{\partial\Lambda}{\partial q^{\text{base}}} = z\big[D_{AI}^{NC} - (1-\alpha)D_{AI}^C\big] = \Lambda - (1-\alpha)\big(D^{NC}P^{NC} - D^C P^C\big),$$

which is negative when $\Lambda \leq 0$ by Eq. (B.5). In words, the platform earns more from human content under NC. It can therefore prefer C only if it keeps more AI revenue under C, and a higher AI price then benefits it more under C. In particular, $(1-\alpha)D_{AI}^C > D_{AI}^{NC}$ at indifference.

Third, let $q < q'$ be any two values of q^{base}. For each s, $\Lambda(s, q) \leq 0$ implies $\Lambda(s, q') < \Lambda(s, q)$. Otherwise $\Lambda(s, \cdot)$, which falls just to the right of q, would first return to the level $\Lambda(s, q) \leq 0$ at a point where its derivative is nonnegative. This would contradict the second part of the argument. It follows that $\Lambda(\cdot, q') < 0$ wherever $\Lambda(\cdot, q) \leq 0$. By Step 2, $\phi^\star(q) = 0$ then implies $\phi^\star(q') = 0$; $0 < \phi^\star(q) < \phi^B$ implies $\Lambda(\cdot, q') < 0$ on $[g, s^\star(q)]$, so that $\phi^\star(q') = 0$ or $s^\star(q') > s^\star(q)$, and in both cases $\phi^\star(q') < \phi^\star(q)$; and $\phi^\star(q) = \phi^B$ implies $\phi^\star(q') \leq \phi^B = \phi^\star(q)$.

Proof of Proposition 4. We now compare in turn creators' net payoffs (Step 4), consumer surplus

(Step 5) and total surplus (Step 6) across the equilibria of the two regimes.

Step 4 (Creators' net payoffs). By Proposition 2, $R_L^{NC} = 0$, so by Corollary 3, $R_L^C - R_L^{NC} = V_L^C - \phi q_L$ is positive and strictly increases in q^{base}. For H, participation costs cancel: $R_H^C - R_H^{NC} = V_H^C - V_H^{NC}$. Here $V_H^{NC} = \alpha(q_L + s)(2q_H - q_L - s)/4$ does not depend on q^{base} (Proposition 2), while $V_H^C = V_H(z)$ strictly increases in q^{base} (Lemma B.10). So $R_H^C - R_H^{NC}$ strictly increases on $[0, 1]$. It is therefore positive, because at $q^{\text{base}} = 0$, where $z = \frac{t}{2} < q_H$, the identity $(q_L + s)(2q_H - q_L - s) = q_H^2 - (q_H - q_L - s)^2$ gives

$$R_H^C - R_H^{NC} = \frac{\alpha}{8}\big[t(2 - q_H) + 2(q_H - q_L - s)^2\big] > 0.$$

Step 5 (Consumer surplus). AI buyers obtain $q_{AI} - P_{AI}^r = 0$, so consumer surplus comes only from buyers of creators' content. A buyer of j's content at distance d from j obtains $q_j - P^r - d = D_j^r - d$, since $D_j^r = q_j - P^r$ whenever $D_j^r > 0$ (Propositions 2 and 3). Integrating over $d \in [0, D_j^r]$ gives $\text{CS}^r = [(D_H^r)^2 + (D_L^r)^2]/2$, so $\text{CS}^C < \text{CS}^{NC}$ by Corollary 2. Moreover, $\text{CS}^{NC} = [(q_L - s)^2 + (2q_H - q_L - s)^2]/8$ does not depend on q^{base}, while $\text{CS}^C = (g - q^{\text{base}})^2/8$ strictly decreases in CQ1 and $\text{CS}^C = 0$ in CQ2. The reduction therefore strictly increases below $\overline{q}^{\text{base}}$ and is constant above it.

Step 6 (Total surplus). Prices are transfers: by Eq. (5) and Eq. (4), $\pi^r + V_H^r + V_L^r$ equals consumers' spending. As both creators participate, $\text{TS}^r = W^r - \phi t$, where W^r is consumers' gross utility: $q_j - d$ for a buyer of j's content at distance d from j, and q_{AI} for an AI buyer. Every consumer buys, so $D_{AI}^r = 1 - D_H^r - D_L^r$, and with $q_H - q_{AI} = g - q^{\text{base}}$ and $q_L - q_{AI} = -g - q^{\text{base}}$,

$$W^r - q_{AI} = (g - q^{\text{base}})D_H^r - (g + q^{\text{base}})D_L^r - \tfrac{1}{2}\big[(D_H^r)^2 + (D_L^r)^2\big].$$

Under C this equals $\frac{3}{2}(D_H^C)^2$, since $D_L^C = 0$ and $g - q^{\text{base}} = 2D_H^C$ in CQ1, while $D_H^C = 0$ in CQ2. Under NC, $D_H^{NC} - D_L^{NC} = 2g$ and $D_H^{NC} + D_L^{NC} = q_H - s$ turn the first two terms into $2g^2 - (q_H - s)q^{\text{base}}$ and the bracket into $4g^2 + 2D_H^{NC}D_L^{NC}$. Since q_{AI} is the same in both regimes,

$$\text{TS}^C - \text{TS}^{NC} = D_H^{NC}D_L^{NC} + (q_H - s)\, q^{\text{base}} + \tfrac{3}{2}(D_H^C)^2 > 0,$$

because $D_H^{NC}D_L^{NC} > 0$. As $(D_H^C)^2$ has derivative $-D_H^C$, the difference has q^{base}-derivative $(q_H - s) - \frac{3}{2}D_H^C > 2g - \frac{3}{4}g > 0$, where we use $s < q_L$ and $D_H^C \le \frac{g}{2}$. Hence it strictly increases on $[0, 1]$. □

B.10 Proof of Proposition 5

Proof. With base prompting effort κ_0, consumers value AI-generated content as if base capability were $\theta = q^{\text{base}} - \kappa_0$, which may be negative. Lemma B.7 describes pricing under NC for every such θ. The platform lowers the AI price one for one with κ_0 and changes nothing else as long as widening the price gap would lower its profit, which is the case for every κ_0 below $\widehat{\kappa}_0(\phi)$.

Step 1 (The threshold). Consider a pure-strategy Stage-2 equilibrium in which both creators participate and the AI sells. By Lemma B.7(b), profit is at most $\widetilde{\pi}(\delta)$, where $\delta^{NC} \le \delta < \frac{1}{2}$. The function $\widetilde{\pi}$

in Eq. (B.3) is strictly concave, and

$$\widetilde{\pi}'(\delta^{NC}) = -\alpha t - 2(1-\alpha)(q_H - s) - 2\theta = 2(\widehat{\theta} - \theta), \qquad \widehat{\theta} \equiv q^{\text{base}} - \widehat{\kappa}_0(\phi),$$

where

$$\widehat{\kappa}_0(\phi) \;\equiv\; q^{\text{base}} + \frac{\alpha(q_H + q_L)}{2} + (1-\alpha)(q_H - s) \;>\; q^{\text{base}}, \tag{B.6}$$

the inequality holding as $q_H > s$. Let $\kappa_0 < \widehat{\kappa}_0(\phi)$, that is, $\theta > \widehat{\theta}$. Then $\widetilde{\pi}$ strictly decreases on $[\delta^{NC}, \frac{1}{2})$, so profit is at most $\widetilde{\pi}(\delta^{NC}) = \pi^{NC}(\phi, \theta)$, with equality only if $\delta = \delta^{NC}$ and $P_{AI} = z$, that is, only at (P^{NC}, z). The AI price is strictly positive there: $z = \frac{t}{2} + \theta > \frac{t}{2} + \widehat{\theta} = (1-\alpha)(s-g) \geq 0$. In particular, $\frac{t}{2} + \widehat{\theta} \geq 0$, that is, $\widehat{\kappa}_0(\phi) \leq \kappa_0^*$, where $\kappa_0^* \equiv q^{\text{base}} + \frac{t}{2}$ is the level of base prompting effort in Proposition 6 at which the AI stops earning any revenue.

Step 2 (Other outcomes yield less). Using $(q_H - s)(q_L + s) = \frac{t^2}{4} - (s-g)^2$ in Eq. (B.2), the platform's gain over the benchmark is its AI revenue minus its loss on creators' content:

$$\pi^{NC}(\phi, \theta) - \pi^B = D_{AI}^{NC}\, z - (1-\alpha)\frac{(s-g)^2}{2} > (1-\alpha)(s-g)\Big(1 - q_H + \frac{s+g}{2}\Big) \geq 0,$$

where the strict inequality uses $D_{AI}^{NC} > 0$ and $z > (1-\alpha)(s-g)$ (Step 1), and the last uses $s \geq g$ and $q_H < 1$. By Lemma B.7(a), a pure-strategy Stage-2 equilibrium in which a creator exits or the AI sells nothing yields at most π^B. Together with Step 1, every pure-strategy Stage-2 equilibrium at any price pair other than (P^{NC}, z) yields strictly less than $\pi^{NC}(\phi, \theta)$.

Step 3 (Equilibrium). As $z > 0$, Lemma B.7(c) shows that at (P^{NC}, z) the only allowed play has both creators participate, and it yields $\pi^{NC}(\phi, \theta)$. Allowed plays are pure under NC (Lemma B.1(a)), so Step 2 covers every other price pair. Lemma B.1(b) with $\mathcal{P}^* = \{(P^{NC}, z)\}$ and $\pi^* = \pi^{NC}(\phi, \theta)$ then shows that every equilibrium is that of Proposition 2 with θ in place of q^{base}. Its human price, demands and net payoffs (zero for L, positive for H) do not involve κ_0. The AI price $z = q^{\text{base}} + \frac{t}{2} - \kappa_0$ falls one for one with κ_0, and $\partial \pi^{NC}(\phi, \theta)/\partial \kappa_0 = -(1 - q_H + s) = -D_{AI}^{NC}$. □

B.11 Proof of Proposition 6

Proof. Here transfer rates are equal, so $z = \theta + \frac{t}{2}$ with $\theta = q^{\text{base}} - \kappa_0$. The proposition follows what the platform does as θ falls. Lemma B.8 gives the prices it prefers when both creators contribute: the displacement prices of Proposition 3 while $\theta \geq \underline{\theta}_C$, and the three-product prices below it. We show that no other profile does better (Step 1), find when creators' payoffs support $(1,1)$ at these prices (Step 2), and conclude (Step 3). Step 4 treats the case in which the AI earns no revenue.

If $\theta \geq 0$, consumers' utilities under every profile are those of the main analysis with θ in place of q^{base}, and $\theta \leq 1$ by Assumption 1. The game is thus that of Section 3 with $\theta \in [0,1]$ in place of q^{base}. As $\underline{\theta}_C < 0$ and $\underline{\theta}_L(\phi) < 0$ (Step 2), every such θ falls under part 1. It then follows from Proposition 3 and Corollary 3, which gives $V_L^C > \phi q_L$. In Steps 1–3, let $-\frac{t}{2} < \theta < 0$, so $z \in (0, \frac{t}{2})$ and $z < q_H$. Steps 2 and 4 use creator j's revenue at the benchmark price, $V_j^B \equiv \alpha(q_j - \frac{t}{4})\frac{t}{4}$. This revenue covers her cost because $V_j^B/q_j = \alpha(1 - \frac{t}{4q_j})\frac{t}{4}$ is larger for H and $V_L^B/q_L = \phi^B \geq \phi$.

Step 1 (Profit bound). We let $\overline{\pi} \equiv (1-\alpha)F(z)$ if $\theta \geq \underline{\theta}_C$ and $\overline{\pi} \equiv (1-\alpha)(z+\frac{1}{2}\theta^2)$ if $\theta < \underline{\theta}_C$. By Lemma B.8, profit under $(1,1)$ is at most $\overline{\pi}$. If $\theta > \underline{\theta}_C$, the bound is attained only at $(P^C, P^C_{AI}) \equiv (\frac{q_H+z}{2}, z)$, the prices of Proposition 3 (CQ1) with θ in place of q^{base}. If $\theta < \underline{\theta}_C$, it is attained only at the three-product prices $(P^{3\text{prod}}, z) \equiv (\frac{t+\theta}{2}, z)$. If $\theta = \underline{\theta}_C$, it is attained at both pairs. Moreover, $\overline{\pi} > \pi^B$ and $\overline{\pi} \geq (1-\alpha)z$. By Lemma B.9, every profile with at most one contributor yields at most π^B or strictly less than $(1-\alpha)z$, hence less than $\overline{\pi}$. By Lemma B.1(c), an allowed play thus yields $\overline{\pi}$ only if it puts probability one on $(1,1)$ at a pair attaining the bound, and less otherwise.

Step 2 (Creators' payoffs under $(1,1)$). We let R_j denote creator j's net payoff under $(1,1)$.

(a) *Displacement prices.* Let $\theta \geq \underline{\theta}_C$. At (P^C, P^C_{AI}), $R_H > 0$ and L sells nothing, so she earns $V^C_L = \frac{\alpha}{2}G(z)$ (Lemma B.10). Her AI payment covers her cost if and only if the AI is capable enough. To see this, note that G strictly increases, $\frac{\alpha}{2}G(0) = 0 < \phi q_L$, and

$$\frac{\alpha}{2}G\Big(\frac{t}{2}\Big) - \phi q_L \geq \frac{\alpha}{2}G\Big(\frac{t}{2}\Big) - \phi^B q_L = \frac{\alpha t(2-q_L)}{8} > 0,$$

so a unique $\underline{\theta}_L(\phi) \in (-\frac{t}{2}, 0)$ solves $\frac{\alpha}{2}G\big(\frac{t}{2} + \underline{\theta}_L(\phi)\big) = \phi q_L$, namely

$$\underline{\theta}_L(\phi) = \tfrac{1}{2}\Big[\sqrt{(2-q_H)^2 + 16\phi q_L/\alpha} - (2+q_L)\Big],$$

and, for $\theta \geq \underline{\theta}_C$, R_L has the sign of $\theta - \underline{\theta}_L(\phi)$. This threshold exceeds $\underline{\theta}_C$, so that L's participation rather than the platform's pricing limits part 1, if and only if $\frac{\alpha}{2}G(\frac{t}{2} + \underline{\theta}_C) < \phi q_L$. Since

$$G\big(\tfrac{t}{2} + \underline{\theta}_C\big) = \Big(1 - \tfrac{(2+\sqrt{2})g}{2}\Big)\tfrac{(2+\sqrt{2})q_L - \sqrt{2}q_H}{2}$$

has positive factors (the first as $g < \frac{1}{2} < 2 - \sqrt{2}$) that decrease in q_H, this holds if and only if ϕ or q_H exceeds a cutoff given the other parameters. Either cutoff may lie outside the admissible range.

(b) *Three-product prices.* Let $\theta \leq \underline{\theta}_C$. At $(P^{3\text{prod}}, z)$, each creator earns more than in the benchmark, so $R_H > 0$ and $R_L > 0$. Each creator sells $q_j - P^{3\text{prod}}$ and receives half of the creators' share of AI revenue $(1+\theta)z$ (Lemma B.8). Substitution gives

$$V_H - V^B_H = \tfrac{\alpha}{4}z\,(2 - 2q_L + z) > 0, \qquad V_L - V^B_L = \tfrac{\alpha}{4}z\,(2 - 2q_H + z) > 0.$$

Hence $V_j > V^B_j \geq \phi q_j$ for both creators, so both net payoffs are positive.

Step 3 (Parts 1 and 2). Both candidate prices of creators' content exceed $\frac{t-1}{2}$, as $P^C - \frac{t-1}{2} = \frac{1+g+\theta}{2} > 0$ and $P^{3\text{prod}} - \frac{t-1}{2} = \frac{1+\theta}{2} > 0$. By Lemma B.6(b), $(1,1)$ is therefore the only allowed play at any candidate pair at which $R_H > 0$ and $R_L \geq 0$. If this holds at every pair attaining the bound of Step 1, Lemma B.1(b) with $\pi^* = \overline{\pi}$ then shows that an equilibrium exists and that every equilibrium has Stage-2 play $(1,1)$ and prices that attain the profit bound of Step 1.

Let $\max\{\underline{\theta}_C, \underline{\theta}_L(\phi)\} \leq \theta < 0$. The bound of Step 1 is attained at (P^C, P^C_{AI}), where Step 2(a) gives $R_H > 0$ and $R_L \geq 0$. If $\theta = \underline{\theta}_C$, it is also attained at $(P^{3\text{prod}}, z)$, where Step 2(b) gives $R_H, R_L > 0$. For $\theta > \underline{\theta}_C$, the unique equilibrium outcome is therefore that of Proposition 3 (CQ1) with θ in

place of q^{base}: $D_L^C = 0$, and $R_L > 0$ except at $\theta = \underline{\theta}_L(\phi)$. At $\theta = \underline{\theta}_C$, which requires $\underline{\theta}_C \geq \underline{\theta}_L(\phi)$, the three-product outcome of part 2 is a second equilibrium outcome.

Let $-\frac{t}{2} < \theta < \underline{\theta}_C$. The bound is attained only at $(P^{3\text{prod}}, z)$, where Step 2(b) gives $R_H, R_L > 0$ whatever $\underline{\theta}_L(\phi)$ is. The three-product configuration is therefore the unique equilibrium outcome: the AI sells $D_{AI} = 1 + \theta > 0$ at the price $z > 0$, and L regains the demand $D_L = -g - \frac{\theta}{2} > 0$.

Let $\theta = \theta_C < \theta_L(\phi)$. The bound of Step 1 is attained at (P^C, P_{AI}^C) and at $(P^{3\text{prod}}, z)$. At (P^C, P_{AI}^C), Step 2(a) gives $R_L < 0$, so L would rather exit and no allowed play puts probability one on $(1, 1)$. By Step 1, every allowed play yields less than $\bar{\pi}$ at every price pair not attaining the bound. At $(P^{3\text{prod}}, z)$, Step 2(b) gives $R_H, R_L > 0$, so $(1, 1)$ is the only allowed play by Lemma B.6(b), and it yields $\bar{\pi}$. Lemma B.1(b), with $\mathcal{P}^* = \{(P^{3\text{prod}}, z)\}$ and $\pi^* = \bar{\pi}$, then shows that the three-product configuration is again the unique equilibrium outcome, with $D_L > 0$ as $\theta_C < -2g$.

Step 4 (Part 3). Let $\theta \leq -\frac{t}{2}$, so that $z \leq 0$. Then no consumer obtains positive utility from the AI: $U_{AI} = z - P_{AI} \leq -P_{AI}$ under $(1, 1)$, and $U_{AI} < 0$ under every other profile. So AI revenue is zero at all prices and under every profile, and the AI sells nothing if $\theta < -\frac{t}{2}$. By Lemma B.5, profit is at most π^B, with equality only if $P = P^B = \frac{t}{4}$ and both creators participate and sell $q_j - P^B$. Fix a pair (P^B, P_{AI}). The AI gives no consumer positive utility, ties go to creators' content and $P^B > \frac{t-1}{2}$. Each participating creator thus sells $q_j - P^B$ whatever the other creator does (Lemma B.2(a)), and contributions affect no payoff. Her revenue V_j^B covers her cost, so both creators participate in every allowed play, L by Assumption 2 if $\phi = \phi^B$. Lemma B.1(b) and (c) with $\mathcal{P}^* = \{(P^B, P_{AI}) : P_{AI} \geq 0\}$ and $\pi^* = \pi^B$ gives the benchmark outcome, with no AI sales if $\kappa_0 > \kappa_0^*$. □

B.12 Proof of Proposition 7

Proof. Here $\kappa_0 = 0$, $\tau_H = \frac{1}{2}$ and $0 < \tau_L \leq \frac{1}{2}$, so $z = q^{\text{base}} + \frac{q_H}{2} + \tau_L q_L$ is the q_{AI} of the statement. CQ1 is the case $z < q_H$ and the price pairs of the statement are the displacement prices $\mathcal{P}(z)$. A lower transfer rate lowers z, so the platform sells a less capable AI. We show that it still sets the displacement prices, and that equal splitting still pays L enough to participate.

Step 1 (Profit bound). By Lemma B.11, every profile yields at most $(1 - \alpha)F(z)$ at any price pair. Equality requires $(1, 1)$ and a pair in $\mathcal{P}(z)$, where $D_L = 0$, $D_H = D_H^C(z)$ and $D_{AI} = 1 - D_H^C(z)$.

Step 2 (Creators' payoffs). Under $(1, 1)$ at a pair in $\mathcal{P}(z)$, H's net payoff is positive and L earns $V_L^C = \frac{\alpha}{2}G(z)$ (Lemma B.10, which applies as $\theta > \underline{\theta}_C$ by Lemma B.11). As G increases and $z > \frac{q_H}{2}$, $V_L^C > \frac{\alpha}{2}G(\frac{q_H}{2}) = \alpha q_H(4 - q_H)/16$, and

$$\frac{16}{\alpha}\Big[\frac{\alpha q_H(4 - q_H)}{16} - \phi^B q_L\Big] = q_H(4 - 5q_L) + 3q_L(q_H - q_L) > 0,$$

since $q_L \leq q_H/\sqrt{2} < \frac{4}{5}$. Thus $V_L^C > \phi^B q_L \geq \phi q_L$ for every $\tau_L \in (0, \frac{1}{2}]$: L receives half of the creators' share of AI revenue, which stays large even when her own data adds little to AI quality.

Step 3 (Equilibrium). At every pair in $\mathcal{P}(z)$, $P > \frac{q_H}{2} > \frac{t-1}{2}$. By Step 2 and Lemma B.6(b), $(1, 1)$ is the only allowed play there, and it yields $(1 - \alpha)F(z)$. By Step 1 and Lemma B.1(c), every allowed play yields strictly less at every other price pair. Lemma B.1(b) with $\mathcal{P}^* = \mathcal{P}(z)$ and $\pi^* = (1 - \alpha)F(z)$

shows that every equilibrium has prices in $\mathcal{P}(z)$, play $(1,1)$ and the demands of Step 1, and that every pair in $\mathcal{P}(z)$ is an equilibrium price pair. This proves the characterization and item 3.

Step 4 (Comparative statics). As $dz/d\tau_L = q_L > 0$ and F is strictly increasing, the AI price z and the platform's profit $(1-\alpha)F(z)$ strictly increase in τ_L (item 1). In CQ1, the human price $P = \frac{q_H+z}{2}$ and AI demand $D_{AI} = 1 - D_H^C(z)$ strictly increase in τ_L and $D_H^C(z) = \frac{q_H-z}{2}$ strictly decreases in τ_L; in CQ2, $D_{AI} = 1$ (item 2). Both creators' gross payoffs, $\frac{\alpha}{2}G(z)$ and $V_H(z)$, also strictly increase in τ_L (Lemma B.10), as claimed in the discussion that follows the proposition in Section 4.2. □

B.13 Proof of Corollary 5

Proof. Under NC, L's participation constraint binds and H's is slack (Proposition 2), so $V_L^{NC}/q_L = \phi < V_H^{NC}/q_H$. Under C, we split the gap between the two ratios into two parts: H's direct revenue per unit of quality, and the excess of L's AI payment per unit of quality over H's. We show that the second part is larger, because both creators receive the same AI payment, which is larger relative to L's lower quality. By Proposition 3, $s_H = s_L = 1$ and $D_L^C = 0$, so Eq. (4) and $1/q_L - 1/q_H = 2g/(q_H q_L)$ give

$$\frac{V_H^C}{q_H} - \frac{V_L^C}{q_L} = \frac{\alpha D_H^C P^C}{q_H} - \frac{\alpha D_{AI}^C P_{AI}^C}{2}\Big(\frac{1}{q_L} - \frac{1}{q_H}\Big) = \frac{\alpha}{q_H q_L}\big(q_L D_H^C P^C - g\, D_{AI}^C P_{AI}^C\big).$$

Also by Proposition 3, $D_H^C P^C \le \frac{g}{2} q_H$ (in CQ1, $D_H^C = \frac{g - q^{\text{base}}}{2} \le \frac{g}{2}$ and $P^C = q_H - D_H^C$; in CQ2, $D_H^C = 0$), $D_{AI}^C = 1 - D_H^C \ge 1 - \frac{g}{2}$ and $P_{AI}^C = q^{\text{base}} + \frac{t}{2} \ge \frac{t}{2}$, so

$$q_L D_H^C P^C - g\, D_{AI}^C P_{AI}^C \le \frac{g}{2}\Big[q_H q_L - \Big(1 - \frac{g}{2}\Big)t\Big] < 0,$$

because $q_H q_L < q_L < \frac{t}{2} < (1 - \frac{g}{2})t$. Hence $V_H^C/q_H < V_L^C/q_L$, as claimed. □

B.14 Proof of Proposition 8

Proof. Quality weighting changes how creators split their share of AI revenue, but not what consumers buy or what the platform earns under any profile. The platform's preferred prices are therefore unchanged, and the only question is whether L still participates when part of her AI payment goes to H. Steps 1–4 answer it for equal transfer rates (the first part of the proposition) and Steps 5 and 6 for heterogeneous transfer rates (the second part).

Step 1 (Profit bound with equal transfer rates). Let $\tau_H = \tau_L = \frac{1}{2}$, so that $z = q^{\text{base}} + \frac{t}{2} = P_{AI}^C$, and we let $\mathcal{P}^C \equiv \mathcal{P}(z)$, $D_H^C \equiv D_H^C(z)$, $D_{AI}^C \equiv 1 - D_H^C$ and $\pi^C \equiv (1-\alpha)F(z)$ be the equilibrium prices, demands and profit of Proposition 3. Lemma B.11 holds under either sharing rule. So every profile yields at most π^C at every price pair, with equality only under $(1,1)$ at the pairs in $\mathcal{P}^C$. There, demands are $D_L = 0$, $D_H = D_H^C$ and $D_{AI} = D_{AI}^C$. Prices, demands, profit and consumer surplus are thus those of Proposition 3 once we show that $(1,1)$ is the only allowed play on $\mathcal{P}^C$.

Step 2 (Redistribution). Under $(1,1)$, creator j's quality-weighted share is q_j/t: H's share rises from

$\frac{1}{2}$ to q_H/t and L's falls to q_L/t. On $\mathcal{P}^C$,

$$V_H^{C,\mathrm{QW}} - V_H^C = V_L^C - V_L^{C,\mathrm{QW}} = \frac{\alpha(q_H - q_L)}{2t}\, D_{AI}^C P_{AI}^C > 0,$$

and creators' total net payoff is unchanged. Each creator now receives the AI payment $\alpha D_{AI}^C P_{AI}^C/t$ per unit of quality, and $D_L^C = 0$, so

$$\frac{V_H^{C,\mathrm{QW}}}{q_H} - \frac{V_L^{C,\mathrm{QW}}}{q_L} = \frac{\alpha D_H^C P^C}{q_H} \;\geq\; 0, \tag{B.7}$$

with strict inequality in CQ1, where $D_H^C P^C > 0$, and equality in CQ2, where $D_H^C = 0$.
Step 3 (Positive net payoffs). By Eq. (B.7), it suffices to show $V_L^{C,\mathrm{QW}} > \phi q_L$. By Corollary 3, V_L^C increases in q^{base} and so does $V_L^{C,\mathrm{QW}} = (2q_L/t)V_L^C$. Hence $V_L^{C,\mathrm{QW}}$ is smallest at $q^{\mathrm{base}} = 0$ (CQ1), where $D_{AI}^C = 1 - \frac{q}{2}$ and $P_{AI}^C = \frac{t}{2}$. As $\phi \leq \phi^B$ (Assumption 4) and $q_L < 1$,

$$\frac{V_L^{C,\mathrm{QW}}}{q_L} - \phi \;\geq\; \frac{\alpha(4 - q_H + q_L)}{8} - \phi^B = \frac{\alpha\big[(q_H - 2q_L)^2 + q_L(8 - 5q_L)\big]}{16\, q_L} > 0. \tag{B.8}$$

Step 4 (Equilibrium with equal transfer rates). At every pair in $\mathcal{P}^C$, $P > \frac{q_H}{2} > \frac{t-1}{2}$. By Step 3 and Lemma B.6(b), $(1,1)$ is the only allowed play there, and it yields π^C. By Step 1 and Lemma B.1(c), every allowed play yields strictly less elsewhere. Lemma B.1(b) with $\mathcal{P}^* = \mathcal{P}^C$ and $\pi^* = \pi^C$ then proves the first part of the proposition, the case of equal transfer rates.
Step 5 (The threshold $\tau_L^\star$). Now let $q_H/q_L \geq \sqrt{2}$, $\tau_H = \frac{1}{2}$ and $0 < \tau_L \leq \frac{1}{2}$, and we write $h \equiv \frac{q_H}{2}$ and $\ell \equiv \tau_L q_L$, so that $z = q^{\mathrm{base}} + h + \ell$. By Lemma B.11, the profit bound $(1-\alpha)F(z)$, the price pairs $\mathcal{P}(z)$ that attain it under $(1,1)$ and the demands there are those of Proposition 7. Under $(1,1)$ on $\mathcal{P}(z)$, AI revenue is $G(z)$ (Lemma B.10). L sells nothing and receives the share $\ell/(h+\ell)$ of the creators' part of this revenue. Her payoff is thus $\alpha\eta(\tau_L)$, where $\eta(\tau_L) \equiv \frac{\ell}{h+\ell}G(z)$ does not depend on ϕ or α. A lower transfer rate now reduces her payment twice, through lower AI revenue and through her smaller share. Formally, η is continuous and strictly increasing, as $\ell/(h+\ell)$ and $G(z)$ are positive and strictly increase in ℓ. It tends to zero as $\tau_L \downarrow 0$. At $\tau_L = \frac{1}{2}$, $\ell/(h+\ell) = q_L/t$ and $z \geq \frac{t}{2}$, so $\eta(\frac{1}{2}) \geq (q_L/t)G(\frac{t}{2}) = q_L(4 - q_H + q_L)/8$, which exceeds $(\phi^B/\alpha)q_L \geq (\phi/\alpha)q_L$ by Eq. (B.8). Hence a unique $\tau_L^\star \in (0, \frac{1}{2})$ solves $\eta(\tau_L^\star) = (\phi/\alpha)q_L$. It depends on ϕ and α only through ϕ/α, and L's net payoff is negative below $\tau_L^\star$, zero at $\tau_L^\star$ and positive above it.
Step 6 (Equilibrium with heterogeneous transfer rates). H's share $h/(h+\ell)$ is at least $\frac{1}{2}$, so her payoff is at least her payoff under equal splitting, and her net payoff is positive (Lemma B.10). If $\tau_L \geq \tau_L^\star$, L's net payoff is nonnegative. As $P > \frac{q_H}{2} > \frac{t-1}{2}$ at every pair in $\mathcal{P}(z)$, Lemma B.6(b) shows that $(1,1)$ is the only allowed play there. Lemmas B.11 and B.1(b)–(c) then show that the outcome of Proposition 7 is the equilibrium outcome. If $\tau_L < \tau_L^\star$, L's net payoff under $(1,1)$ is negative at every pair in $\mathcal{P}(z)$, so she would rather exit than contribute. Then $(1,1)$ is not a Nash equilibrium at any of these pairs, and no equilibrium under quality weighting has this outcome. □