A Model of the Attention Economy: Data, Platforms, and Discovery

Daniel Chen*

February 2022, revised May 2026

Abstract

This paper develops a dynamic model of the *attention economy* featuring two interlinked markets: a *product market*, where consumers have limited awareness of firms, and an *attention market*, where platforms attract attention by investing in free services (e.g., social media, streaming, podcasts) and sell targeted ads using data on consumer preferences. The model provides a tractable framework in which platform investment, ad-driven discovery, product demand, and ad revenue are jointly determined. The effects of data and interoperability policies can reverse over time, platform quality may be inefficiently high or low, and the rise of data-rich digital advertising can coincide with a declining ad revenue-to-GDP ratio.

*Assistant Professor of Economics, Princeton University; dtchen@princeton.edu.
I especially thank Yuliy Sannikov, Andy Skrzypacz, Darrell Duffie, and Bob Wilson for their suggestions and helpful comments. I also thank Chris Tonetti for constructive feedback on an earlier draft. I have benefited from discussions with Susan Athey, Kyle Bagwell, Anirudha Balasubramanian, Martino Banchio, Markus Brunnermeier, Christoph Carnehl, Cindy Chung, Sebastian Di Tella, Ravi Jagadeesan, Chad Jones, Nadia Kotova, David Kreps, Ernest Liu, Alessandro Lizzeri, Moritz Lenel, Erik Madsen, Suraj Malladi, Mike Ostrovsky, Marzena Rostek, Karthik Sastry, Takuo Sugaya, Timur Sobolev, Laura Veldkamp, Frank Yang, Leeat Yariv, Moto Yogo, Ali Yurukoglu, Anthony Zhang, Weijie Zhong, and several seminar participants. This research was supported by the Bradley Graduate Fellowship through a grant to SIEPR. Qiyue Yu provided excellent research assistance. A previous version of this paper was titled "The Market for Attention."

> "...a wealth of information creates a poverty of attention and a need to allocate that attention efficiently..."—Herbert Simon

Simon's insight captures today's digital economy, where social media, news, and entertainment compete intensely for attention. Individuals now spend over two and a half hours per day on digital media (Evans, 2020). This attention underlies the Internet's core business model: platforms—any website or app displaying ads—collectively earn hundreds of billions of dollars each year from data-driven targeted advertising (Silk et al., 2021).

How do these platforms mediate the allocation of attention between consumers and firms, and how does this allocation shape demand, competition, and welfare over time?

A core challenge is that the attention economy involves a complex interaction between two markets. In the product market, consumers can only purchase from firms they are aware of. In the attention market, platforms provide consumers with services—e.g., social media, streaming, podcasts—in exchange for attention, and sell targeted ads that determine which firms consumers discover in the product market. Ads are allocated through auctions in which firms bid on individual impressions using platform data on consumers' preferences. Because services are often free, platforms compete through investment in service quality.

Many key issues—the welfare tradeoffs between paid products and free platform services, and the cross-market effects of policy—cannot be studied without modeling these markets jointly. This paper develops a tractable framework that does so: platforms' investment in quality drives attention, attention drives ad-funded discovery, discovery shapes product-market competition, and product-market competition in turn shapes the ad revenue that funds platform investment. To illustrate its value, I use the framework to study the dynamic effects of data and interoperability policies, the welfare properties of equilibrium, and how asymmetric data shapes competition between platforms such as traditional and digital media.

To keep the model tractable, I impose simplifying assumptions. In the baseline, platforms are monopolistically competitive: they share business models with real-world players such as Pinterest, Spotify, Meta, and YouTube, but do not interact strategically. The paper's focus is thus on market operation, *not* market concentration; a duopoly extension incorporates strategic investment under additional structure. CES preferences over products and platform services further simplify matters: on the product side, prices are fixed and firms *optimally* do not personalize them. In return, the framework treats data flexibly[1] and allows multihoming by both firms and consumers, with a key parameter—the elasticity of attention to platform quality—governing platform competition and market power.

The paper yields three sets of findings, each pointing to the importance of cross-market spillovers for economic analysis. First, the effects of data and interoperability policies can be counterintuitive and reverse over time, as interventions trade off product-market efficiency against platform quality. Second, platform investment can be inefficiently high or low, and I derive a sufficient statistic that determines the sign of the distortion. Third, though platforms can profit from access to richer data than their rivals, *total* ad revenue can *fall*.

I now describe these findings in greater detail. I begin with numerical policy experiments highlighting mechanisms related to (i) data reforms and (ii) pro-competitive policies, such as interoperability mandates, that raise the elasticity of attention to platform quality and ad load.[2] Although these policies originate in the attention market, endogenous product-market responses can overturn their initial effects.

In one experiment, giving platforms access to more informative data improves targeting and raises ad revenue in the short run. Over time, however, firms anticipate that any consumer they reach will have also been reached by better-matched competitors. This reduces a firm's willingness to pay for ads, causing long-run ad revenue to fall. Product consumption rises, but platforms invest less in quality.

In a second experiment, an interoperability mandate raises the elasticity of attention. Because consumers view ads as nuisances that reduce effective platform quality, platforms initially reduce ad load, depressing ad revenue and, in turn, investment. Lower ad load, however, weakens product discovery. Over time, firms anticipate fewer competitors for any consumer they reach, raising their willingness to pay for ads. In the long run, ad revenue and investment therefore rise. Product consumption falls even as platform quality improves.

I discuss conditions under which these patterns appear more likely to arise.

Next, I compare equilibrium to the planner's allocation. Two forces drive inefficiency. First, platforms compete for ad revenue, which creates business-

[1] Data is modeled as signals about consumer preferences, without parametric restrictions on signal structure.

[2] Such policies are intended to make platform markets more contestable by giving consumers a more "meaningful choice" among platforms. For example, a content interoperability mandate requires platforms to allow users to share content or posts across sites. As content becomes more overlapping, attention becomes more sensitive to platform-specific features such as the interface or ad load.

stealing externalities. Second, because platforms charge users zero prices, they do not internalize the consumer surplus their services directly generate. These forces can push ad display rates and platform investment either above or below efficient levels. When consumers have Cobb-Douglas utility over CES aggregates of products and platform services, I derive a simple sufficient statistic[3] that determines the sign of the investment distortion.

The policy experiments described earlier are intentionally stylized to isolate cross-market spillovers and dynamic effects, frequent policy focal points. In practice, interoperability policies often target large platforms, where strategic interaction may matter. Motivated by this, I next consider a duopoly model that preserves most of the baseline structure but takes ad display rates as exogenous, allowing platforms to interact strategically through investment. I develop a numerical algorithm to compute equilibrium and find that the mechanisms in the baseline policy experiments remain robust.

Finally, I extend the baseline model to allow two platform groups—e.g., digital and traditional media—with asymmetric data. The extension shows how data differences shape relative quality and market shares, and can generate a pattern consistent with a puzzling empirical fact: despite the rise of highly profitable digital platforms, advertising's share of GDP fell in the period from 2000 to 2018 (Silk et al., 2021). One might expect the presence of data-rich platforms to raise ad revenue: better matching creates surplus for firms, and platforms should be able to extract a share of the surplus gain. In a numerical example, however, digital platforms displace traditional media, earning a higher share of revenue due to their more informative data—yet total ad revenue declines.[4] The result hinges on product-market competition: better data improves the matching of *all* firms with consumers, so individual advantages erode and profit gains are offset, limiting the revenue platforms can extract.[5]

Beyond these main findings, the online appendix develops further extensions—network effects, platform heterogeneity, entry, and reserve prices—that illustrate the model's potential as a foundation for future work.

The paper proceeds as follows. Section 1 provides background on the attention economy and related work. Section 2 introduces a simple product market that serves as a building block for the analysis. Section 3 integrates the atten-

[3]The sufficient statistic depends on ad revenue, income, the product markup, platform substitutability, and the Cobb-Douglas weight.

[4]In the model, entry without data asymmetry cannot explain this: in fact, total ad revenue is *invariant* to the mass of platforms.

[5]In fact, with fixed prices and binding budget constraints, firms' profits gross of ad costs are constant.

tion market with the product market to develop the baseline model. Section 4 characterizes equilibrium. Section 5 presents policy experiments, Section 6 a welfare analysis, Section 7 a duopoly extension, and Section 8 a model with asymmetric data. Section 9 summarizes further extensions. Section 10 concludes.

1 Background

The term *attention economy* is now widely used to describe markets where attention is a scarce resource and a basis for economic exchange. It is often invoked in the context of digital advertising, linking it to issues of data use, platform design, and competition. These markets are inherently multisided: platforms intermediate between consumers and advertisers while shaping the allocation of attention through the quality and features of their services. Competition is shaped not only by prices—many services are offered at zero monetary cost to consumers—but also by non-price dimensions such as quality, data, and interoperability. These features pose distinctive challenges for economic analysis and policy, particularly when interventions on one side of the market have unintended consequences on others (OECD, 2018, 2021, 2022).

This paper provides a transparent benchmark for analyzing these forces. The monopolistic competition baseline clarifies mechanisms that arise from the structure of the attention economy rather than strategic interaction, and lends itself to extensions. Section 7 turns to a duopoly setting to study strategic interaction along the investment margin.

Prior work on digital advertising platforms (e.g., Bergemann and Bonatti 2023; Ambrus et al. 2016; Prat and Valletti 2021; Galperti and Perego 2022) typically models large platforms in static settings and abstracts from either a microfounded product market or content provision.[6] Similarly, Kirpalani and Philippon (2021) study matching in a large online marketplace but without content provision. Relative to this literature, my framework models the attention *economy*—bringing platform investment, ad-driven discovery, and product-market competition into a single system whose joint dynamics shape equilibrium outcomes.

A line of research in traditional advertising endogenizes free platform content provision, most notably Anderson and Coate (2005). Whereas Anderson and Coate (2005) model investment in entry of new platforms, I study

[6]Ambrus et al. (2016) and Prat and Valletti (2021) model the product market in reduced form. Bergemann and Bonatti (2023) and Galperti and Perego (2022) model a single platform and abstract from content and attention.

investment in platform quality. Both models predict potential over- or under-investment due to business-stealing externalities and platforms' inability to charge consumers. However, in their setting consumers derive no surplus from advertising and there is no role for data or targeting.

In digital advertising, ads are commonly sold via auctions at the *individual* level in *real time* as consumers engage with platforms' services. Advertisers rely on platform data to tailor bids. Recent work investigates the value of information in auctions (e.g., Board 2009; Hummel and McAfee 2016; Bergemann et al. 2021; De Corniere and De Nijs 2016), but with few exceptions, abstracts from the product market. A central result is that more informative data typically raises revenue (Board, 2009; Hummel and McAfee, 2016). In my setting, richer data can reduce ad revenue because of spillover effects on advertisers' profits in the product market.

While the role of data in the economy has been examined in general equilibrium settings before (Jones and Tonetti, 2020; Farboodi and Veldkamp, 2021), it has rarely been studied in connection with advertising.[7] A nascent literature incorporates advertising into macroeconomic models, including contemporaneous work by Cavenaile et al. (2023) and Rachel (2024), but without modeling data. Most closely, Greenwood et al. (2025) study advertising with free media goods in a static, perfectly competitive setting, where targeting distinguishes between college and noncollege consumers.

2 Simple Model of a Product Market

The baseline model has two components: a product market and an attention market. This section analyzes the product market in isolation—a simple monopolistically competitive environment isomorphic to the one in Melitz (2003). Although static, it is embedded in continuous time to serve as the backbone for the dynamic framework developed in the next section, where it is integrated with the attention market.

Time t takes values in $[0, \infty)$. A continuum of consumers $i \in \mathbb{I}$ of unit mass have heterogeneous tastes or "values" $\{v_{ij}\}$ for a continuum of products $j \in \mathbb{J}$ of mass J. Each product is produced by a distinct firm, also indexed by j for simplicity.

[7] These papers analyze data as an input into firm production. See also Demirer et al. (2024).

Consumer i has CES preferences for products:

$$C_{it} \equiv \left[\int_{\mathbb{J}} v_{ij}^{\frac{1}{\sigma}} c_{ijt}^{\frac{\sigma-1}{\sigma}} \, \mathrm{d}j \right]^{\frac{\sigma}{\sigma-1}},$$

where $\sigma > 1$ is the elasticity of substitution across products.[8] Values $\{v_{ij}\}$ are drawn independently across consumers and firms from a distribution F supported on $[0, \infty)$ with finite mean and remain constant over time.

At each time t, consumer i has income I and allocates it across products to maximize C_{it} taking firms' prices $\{p_{jt}\}$ as given. Firms set prices to maximize expected profits given a common constant marginal cost normalized to one (the numeraire). I make *no* assumptions about what firms know about consumers' values and could allow firms to personalize prices, though it turns out that firms would *optimally* choose not to do so in equilibrium.

The key friction in this otherwise standard model is that each consumer i is aware of only a subset $\Omega_{it} \subseteq \mathbb{J}$ and can consume only from this set. All that will matter for equilibrium is the mass M_t of firms in Ω_{it} and the consumer's average value $\mu_{\Omega_t} \equiv \frac{1}{M_t} \int_{\Omega_{it}} v_{ij} \, \mathrm{d}j$ for those firms. I assume both objects are common across consumers.

Proposition 1. *In the unique equilibrium of the product market:*

1. *Firm j sets price*

$$p_{jt} = \frac{\sigma}{\sigma - 1}. \tag{1}$$

2. *Firm j's flow profit from selling to consumer i is $\pi_{\mathbb{J}t} v_{ij}$ where*

$$\pi_{\mathbb{J}t} = \frac{I}{\sigma M_t \mu_{\Omega_t}}. \tag{2}$$

3. *Consumer i's aggregate consumption is*

$$C_{it} = I (M_t \mu_{\Omega_t})^{\frac{1}{\sigma-1}}. \tag{3}$$

Intuition is provided in the Appendix. As seen from (1), CES preferences imply a constant markup independent of consideration sets or information firms hold about consumers' values.[9] This property simplifies integration with the attention market because prices are fixed. Moreover, flow profits (2) and aggregate consumption (3) depend only on a single sufficient statistic: the cumulative value $M_t \mu_{\Omega_t}$ of firms in consideration sets. The attention market will endogenize the evolution of M_t and μ_{Ω_t}.

[8] Values enter with exponent $1/\sigma$ as a normalization for notational convenience.

[9] The optimality of a constant markup is a feature of CES demand. The Appendix explains why it still holds here with heterogeneous values.

3 Baseline Model of the Attention Economy

This section integrates the product market with the attention market to form the baseline model of the attention economy. In what follows, I treat the microfounded relationships (1)–(3) as primitives.

The attention market consists of a continuum of platforms $k \in \mathbb{K}$ of mass K. Like product firms, platforms are monopolistically competitive: each offers a distinct service (e.g., streaming, social media, podcasts) and thus has market power but is atomistic. Services are free but require attention to consume. Platforms provide these services to attract attention for ad display and invest to maintain their quality.

If a consumer allocates attention x_{ikt} to platform k, she views ads at Poisson intensity $a_{kt}x_{ikt}$, where a_{kt} is the platform's chosen *ad display rate*. Each ad opportunity is sold via a second-price auction the instant it arises. Due to latency, only N firms, drawn uniformly at random from $\Omega_{it}^c \equiv \mathbb{J} \setminus \Omega_{it}$ (outside the consideration set), submit bids within the instant.[10] The highest bidder wins and pays the second-highest bid. The winning firm enters Ω_{it} at the end of the instant and exits (i.e., is "forgotten") at an independent exponential time with rate λ_f, ensuring that search frictions persist in the long run.

Firms bid based on signals of consumers' values for their products. These signals, interpreted as *platform-supplied data*, induce posterior expectations $\{\hat{v}_{ij}\}$ of consumers' values $\{v_{ij}\}$. Expectations are drawn independently from a continuous cdf G, are constant over time, and do not depend on platform identity, since all platforms provide the same data.[11] Section 8 relaxes this last assumption.

I begin by describing how consideration sets evolve, taking attention, ad display rates, and bidding strategies as given. I then characterize agents' decision problems. Because consumers allocate attention symmetrically in equilibrium, I suppress the index i on x_{ikt} when convenient.

Evolution of Consideration Sets. The key state variables are the mass M_t and cdf H_t of *expected* values for firms in Ω_{it}, both common across consumers. By the law of large numbers,[12] the mean μ_{H_t} coincides with the mean μ_{Ω_t} of *actual* values in Ω_{it}. Since $M_t\mu_{\Omega_t}$ is a sufficient statistic for firms' profits and

[10] Firms only target consumers they do not already serve. A firm does not benefit from displaying an ad to a consumer whose consideration set it is already in.

[11] By Blackwell (1953), G is a mean-preserving contraction of the prior F over values but is otherwise unrestricted.

[12] Throughout, I assume the law of large numbers. This can be formalized using methods from Duffie et al. (2025).

consumption, we do not need to track the true value distribution.

Figure 1 summarizes the evolution of the state variables. The cdf H_t^c corresponds to expected values in Ω_{it}^c and is pinned down given H_t via an accounting identity:[13]

$$M_t H_t(\hat{v}) + (J - M_t) H_t^c(\hat{v}) = JG(\hat{v}), \quad \hat{v} \in [0, \infty). \tag{4}$$

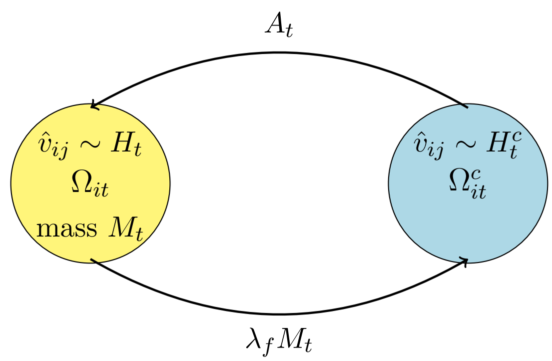


Figure 1: Evolution of consideration sets

The rate of inflow into Ω_{it} is:

$$A_t \equiv \int_{\mathbb{K}} a_{kt} x_{kt} \, \mathrm{d}k. \tag{5}$$

In equilibrium, the winning firm has the highest expected value, so the cdf of inflowing firms is $(H_t^c)^N$. Because firms exit at random, outflow occurs at rate $\lambda_f M_t$ with cdf H_t. Thus:

$$\frac{\mathrm{d}}{\mathrm{d}t}[M_t H_t(\hat{v})] = A_t H_t^c(\hat{v})^N - \lambda_f M_t H_t(\hat{v}), \quad \hat{v} \in [0, \infty). \tag{6}$$

Taking $\hat{v} \to \infty$ yields:[14]

$$\dot{M}_t = A_t - \lambda_f M_t. \tag{7}$$

Together, (4) and (6) fully characterize the dynamics of the key state variables.

[13]The mass of firms with expected values less than $\hat{v}$ in both Ω_{it} and Ω_{it}^c must coincide with that of the whole economy.

[14]Equations (6) and (7) apply whenever $M_t < J$ and I impose parameter conditions in Section 4 so this is always so. The conditions for interchanging the derivative and limit hold here (but (7) can alternatively be derived from aggregate flow balance).

Firms' Bidding Problem. Anticipating M_t and H_t, firms choose bids to maximize discounted profits. The formal problem is stated in Online Appendix B; here I describe it recursively.

Because auctions are second price, firm j's optimal bid for a consumer i with $\hat{v}_{ij} = \hat{v}$ equals the *gain* in continuation value from entering Ω_{it}:

$$B_t(\hat{v}) = V_t^{\text{In}}(\hat{v}) - V_t^{\text{Out}}(\hat{v}). \tag{8}$$

V_t^{In} and V_t^{Out} are the present values of profits from selling to consumer i when the firm is currently inside and outside Ω_{it}, respectively. While outside, firm j enters Ω_{it} with Poisson intensity $\lambda_{et}(\hat{v})$, characterized shortly. Given discount rate $\rho > 0$, V_t^{In} and V_t^{Out} satisfy the Hamilton–Jacobi–Bellman (HJB) equations:

$$\dot{V}_t^{\text{In}}(\hat{v}) = \rho V_t^{\text{In}}(\hat{v}) - \lambda_f \underbrace{\left[V_t^{\text{Out}}(\hat{v}) - V_t^{\text{In}}(\hat{v})\right]}_{\text{loss from exit}} - \underbrace{\pi_{\mathbb{J}t}\hat{v}}_{\text{expected flow profit}} \tag{9}$$

$$\dot{V}_t^{\text{Out}}(\hat{v}) = \rho V_t^{\text{Out}}(\hat{v}) - \lambda_{et}(\hat{v}) \left(\underbrace{V_t^{\text{In}}(\hat{v}) - V_t^{\text{Out}}(\hat{v})}_{\text{gain from entry}} - \underbrace{\mathbb{E}\left[B_t^{(1)} \middle| B_t(\hat{v}) > B_t^{(1)}\right]}_{\text{expected ad payment}} \right). \tag{10}$$

Here, $B_t^{(1)}$ is the highest bid among the $N-1$ other bidders. If B_t is increasing, then $B_t^{(1)} \overset{d}{=} B_t(\hat{v}^{(1)})$ where $\hat{v}^{(1)}$ is the maximum of $N-1$ draws from H_t^c.

In that case, accounting implies[15]

$$\lambda_{et}(\hat{v}) = \underbrace{\frac{NA_t}{J - M_t}}_{\text{auction entry intensity}} \underbrace{H_t^c(\hat{v})^{N-1}}_{\text{win probability}}. \tag{11}$$

Platforms' Ad and Investment Decisions. Platforms choose ad display and investment to maximize discounted profits. Given B_t, the expected ad price is $\pi_{\mathbb{K}t} = \mathbb{E}[B_t(\hat{v}_t^{(2)})]$, where $\hat{v}_t^{(2)}$ is the second-highest of N independent draws from H_t^c. Anticipating $\pi_{\mathbb{K}t}$, platform k solves:

$$\max_{\{a_{kt}\},\{\ell_{kt}\}} \int_0^\infty e^{-\rho t} \Big(\pi_{\mathbb{K}t} a_{kt} x_{kt}(a_{kt}, q_{kt}) - \ell_{kt} \Big) \, \mathrm{d}t \tag{12}$$

[15]Each ad display triggers an auction with N invited bidders, so invitations are sent at total rate NA_t. The first term in (11) divides this rate by the mass of firms outside Ω_{it}.

subject to the law of motion for quality:

$$\dot{q}_{kt} = \ell_{kt}^{\varphi} - \delta q_{kt}, \quad q_{k0} = q_0, \tag{13}$$

where $\varphi < 1$ ensures decreasing returns and $\delta > 0$ is the depreciation rate. The notation $x_{kt}(a_{kt}, q_{kt})$ indicates that the platform internalizes how its choices affect consumer attention.

Consumers' Attention Choices. Consumers derive utility from both products and platform services:

$$U_i \equiv \int_0^\infty e^{-\rho t} u(C_{it}, X_{it})\, \mathrm{d}t,$$

where $u(\cdot)$ is increasing in both arguments. Platform services enter utility through a CES aggregator:

$$X_{it} \equiv \left[\int_{\mathbb{K}} \left(\nu(a_{kt}) q_{kt} x_{ikt} \right)^{\frac{\epsilon-1}{\epsilon}} \mathrm{d}k \right]^{\frac{\epsilon}{\epsilon-1}},$$

where $\nu(\cdot)$ is positive and decreasing so ads are nuisances, and $\epsilon > 1$ is the elasticity of substitution across platforms.

At each t, consumer i has one unit of attention to allocate across platforms. For simplicity, she does *not* internalize how attention affects $\{C_{is}\}_{s \geq t}$ through the evolution of her consideration set. In equilibrium, this myopia is without loss of optimality: because all platforms set the same ad display rate, how the consumer divides attention among them is inconsequential. She thus solves:

$$\max_{\{x_{ikt}\}} X_{it} \quad \text{s.t.} \quad \int_{\mathbb{K}} x_{ikt}\, \mathrm{d}k \leq 1. \tag{14}$$

Attention is therefore allocated independently of the utility function u.

Equilibrium Concept.

Definition 1. *Taking the product market relationships* (1)–(3) *as primitives, an* equilibrium *of the economy with initial condition* (M_0, H_0, q_0) *is a collection of processes*

$$\big\{ \{M_t\}, \{H_t\}, \{q_{kt}\}, \{B_t(\cdot)\}, \{a_{kt}\}, \{\ell_{kt}\}, \{x_{ikt}(\cdot)\} \big\}$$

such that: (i) firms, platforms, and consumers solve their respective problems; and (ii) $\{M_t\}, \{H_t\}$, *and* $\{q_{kt}\}$ *satisfy their laws of motion.*

3.1 Modeling Remarks

The baseline model adopts simplifying assumptions to provide a transparent benchmark. Several are relaxed in later sections or online appendices, as summarized in the introduction. Beyond these, one can show that the results extend to any standard auction format by revenue equivalence. Moreover, if firms can infer preferences from purchase histories, the environment converges to one with perfect information (i.e., $\hat{v}_{ij} = v_{ij}$ for all i, j).

A key modeling choice is to allocate ads through dynamic auctions. This both reflects the microstructure of digital advertising and ensures tractability under multihoming: in static settings, symmetric monotone bidding equilibria break down when a firm can enter multiple auctions for the same consumer. Spreading competition over time guarantees that each firm participates in at most one auction per consumer at a time, supporting equilibrium existence. Similar difficulties arise in static Walrasian formulations when platforms have asymmetric data. Online Appendix G discusses these issues in more detail.

Beyond these technical benefits, the dynamic structure separates short-run from long-run effects, captures investment incentives more faithfully, and delivers steady-state predictions that are independent of initial conditions.

4 Equilibrium of the Baseline Model

The equilibrium derivation proceeds in five steps. Proofs are in the Appendix.

First, solving the attention allocation problem (14) yields:

$$x_{kt} = \frac{[\nu(a_{kt})q_{kt}]^{\epsilon-1}}{\int_{\mathbb{K}}[\nu(a_{zt})q_{zt}]^{\epsilon-1}\,\mathrm{d}z}, \quad k \in \mathbb{K}. \tag{15}$$

Thus, attention depends on effective quality—reflecting both service quality and ad load—relative to competitors.

Second, given (15), each platform chooses its ad rate to maximize flow profit:

$$A = \arg\max_{a_{kt}} \pi_{\mathbb{K}t} a_{kt} \frac{[\nu(a_{kt})q_{kt}]^{\epsilon-1}}{\int_{\mathbb{K}}[\nu(a_{zt})q_{zt}]^{\epsilon-1}\mathrm{d}z} = \arg\max_{a} a\nu(a)^{\epsilon-1} \tag{16}$$

This choice is *static* because no individual platform has measurable effects on consideration sets. CES implies *constant ad rates* determined only by primitives $\nu(\cdot)$ and ϵ.

Third, setting $A_t = A$, I solve (4) and (6) *analytically* for M_t and H_t, under symmetric and monotone bidding strategies.[16]

Fourth, given *arbitrary paths* for M_t and H_t, I solve (8)–(10) *explicitly* for bid functions $B_t(\cdot)$, verifying monotonicity and uniqueness.

Fifth, from $B_t(\cdot)$, I compute the average ad price $\pi_{\mathbb{K}t}$ and solve (12), yielding a simple boundary value problem that characterizes investment and quality paths.

Together, these steps deliver a tractable characterization of the full equilibrium transition path. The full system is reported in the Appendix, with most variables expressed analytically. Theorem 1 summarizes steady-state properties. In equilibrium, attention-market outcomes and product-market outcomes are jointly determined through consideration sets that link the two markets. In particular, ad revenue (Part 4), investment and quality (Part 5), and aggregate consumption (Part 7) reflect this interaction. Capturing this joint determination allows policy and welfare to be evaluated within the model.

Theorem 1. *Suppose A uniquely solves $\max_a a\nu(a)^{\epsilon-1}$. If $A/\lambda_f < J$ and $\epsilon - 1 < 1/\varphi$, then a unique equilibrium exists for any feasible initial condition (M_0, H_0, q_0), characterized by equations (29)–(33) in the Appendix.*[17] *The equilibrium converges to a steady state in which:*

1. *Each platform displays ads at rate A.*

2. *The mass M of products in Ω_{it} is A/λ_f.*

3. *The cdf H satisfies $H(\cdot) = H^c(\cdot)^N$, where H^c solves:*

$$MH^c(\hat{v})^N + (J - M)H^c(\hat{v}) = JG(\hat{v}), \quad \hat{v} \in [0, \infty).$$

4. *Total ad revenue is $\pi_{\mathbb{K}}A$, where $\pi_{\mathbb{K}}$ is the average ad price:*

$$\pi_{\mathbb{K}} = \pi_{\mathbb{J}} \int_0^\infty \frac{1 - NH^c(\hat{v})^{N-1} + (N-1)H^c(\hat{v})^N}{\rho + \lambda_f + \lambda_e(\hat{v})} \,\mathrm{d}\hat{v}. \tag{17}$$

   Here, $\lambda_e(\hat{v})$ is the entry intensity into consideration sets (see (11)), and $\pi_{\mathbb{J}}$ scales firms' flow profits (see (2), with $\mu_{\Omega_t} = \mu_H$).

[16] To compute M_t and H_t, it is not necessary to first solve for attention allocation or platform quality. This simplicity relies on platform symmetry. With asymmetric data (Section 8), the evolution of H_t depends on how attention is allocated, introducing an additional fixed point into the equilibrium characterization.

[17] An initial condition is *feasible* if $M_0 \,\mathrm{d}H_0 \leq J \,\mathrm{d}G$, since there cannot be more firms with a given expectation in Ω_{it} than exist in $\mathbb{J}$.

5. *Each platform invests at rate:*

$$\ell_{\mathbb{K}} = \frac{\varphi\delta\pi_{\mathbb{K}}A(\epsilon-1)}{K(\rho+\delta)}, \tag{18}$$

and has quality $q = (1/\delta)\ell_{\mathbb{K}}^{1/\varphi}$.

6. *Consumers allocate attention as in* (15).

7. *Aggregate platform consumption is* $X_{it} = K^{1/(\epsilon-1)}\nu(A)q$, *and aggregate product consumption is* $C_{it} = I(M\mu_H)^{1/(\sigma-1)}$.

The condition $\epsilon - 1 < 1/\varphi$ ensures concavity of platforms' investment problem; under the reverse inequality, no equilibrium exists. Similarly, $A/\lambda_f < J$ ensures consideration sets remain incomplete so that (6) applies.

Three implications of Theorem 1 are especially relevant for the analysis that follows, highlighting the structural sources of platform market power and the role of data.

Remarks

1. **Elasticity of attention.** Parts 5–6 show that the key parameter governing platform market power in attracting attention is ϵ. As $\epsilon \to 1$, platforms capture equal shares of attention regardless of quality or ad load. Holding all else constant, as ϵ falls, investment declines (Part 5) while the ad display rate $A = \max_a a\nu(a)^{\epsilon-1}$ rises (Part 1).

2. **Search frictions.** Part 4 shows that search frictions generate market power in the ad market: as the entry intensity $\lambda_e(\cdot)$ rises pointwise, ad revenue falls and eventually vanishes. As $A/\lambda_f \to J$, the economy converges to a classical frictionless benchmark in which consideration sets contain all firms, platforms earn no rents, and therefore do not invest in quality.

3. **Role of data.** Part 3 shows that data shapes the composition H of consideration sets through G. Through its effect on H, data reallocates demand toward higher-valuation firms, thereby altering ad revenue (Part 4), investment incentives (Part 5), aggregate consumption (Part 7), and ultimately welfare.

5 Policy Experiments

Data policies and interoperability mandates are central to current debates on digital markets. This section shows that their effects can be *counterintuitive and time-dependent*: static predictions can reverse in transition, and interventions often trade off product-market efficiency against platform investment and quality. The analysis is intentionally stylized to highlight mechanisms and provide intuition for market conditions when these patterns arise. Throughout, formal results apply on the parameter domain where a unique equilibrium exists.

5.1 Data

I first consider data policies, analyzing steady state before turning to transition dynamics. Such policies change the informativeness of the data platforms supply to firms for bidding in ad auctions. An example is a privacy law that restricts the granularity of the data platforms can collect on users.

Data informativeness ranges from uninformative—when all firms' posterior expectations equal the prior mean μ_F—to perfectly informative, when posterior expectations coincide with consumers' true values ($G = F$). Intermediate cases are ordered by the mean-preserving spread order (Blackwell, 1953).[18]

Steady state Proposition 2 shows that banning data use (or ad targeting) *raises* ad revenue and platform quality but lowers product consumption in steady state.[19]

Proposition 2. *In steady state, as data becomes uninformative (i.e., along any sequence of continuous distributions G converging pointwise to the degenerate distribution at μ_F), ad revenue and platform quality converge to their suprema, while aggregate product consumption converges to its infimum, over all continuous distributions G. Moreover, aggregate product consumption is monotone in informativeness.*

Surprisingly, ad revenue is *highest* when data is *uninformative*. A data ban therefore *raises* platform *profits*.[20] Three forces drive the result:

[18] A cdf G is a mean-preserving spread of a cdf $\hat{G}$ if $\int_{-\infty}^{x} G(t)\,\mathrm{d}t \le \int_{-\infty}^{x} \hat{G}(t)\,\mathrm{d}t, \forall x$ with equality when $x = \infty$.

[19] Because G is assumed continuous to rule out ties, the case of uninformative data is formalized in a limiting sense in Proposition 2.

[20] Platform profits net of investment are monotone in ad revenue.

1. **Fixed total firm profits.** Under CES demand, prices are fixed and budgets bind, so total firm profits equal I/σ *regardless* of G.

2. **Full extraction.** With uninformative data, Bertrand competition in auctions extracts these profits fully.

3. **Information rents.** Informative data creates bidder information rents that reduce ad revenue.

The fixed-profit force is a stark implication of CES demand. The broader mechanism, however, is product-market competition: when data improves, each firm internalizes that it faces rivals who are likewise better-matched for any given consumer. In the CES setting this force makes total firm profits exactly invariant to G; more generally, competition reduces the surplus firms retain from better targeting.[21] Provided this effect is sufficiently strong, the finding that less informative data may raise ad revenue does not hinge on profits literally being fixed.

Beyond the case of uninformative data, analytical monotone comparative statics are difficult to establish. Nevertheless, numerical results indicate that more informative data typically reduces ad revenue.

The intuition for this finding is as follows. Although more informative data raises the average value of winning firms—an effect that tends to increase ad revenue since bids are monotone—three forces push in the opposite direction as seen from the expression for ad revenue in Part 4 of Theorem 1.

1. **Higher match rates for high-value firms.** More informative data raises the entry rates $\lambda_e(\hat{v})$ of firms with relatively higher $\hat{v}$.

2. **Lower average value outside consideration sets.** The average value μ_{H^c} of firms bidding in auctions falls.

3. **Lower profit coefficient.** As μ_H rises, the coefficient $\pi_{\mathbb{J}}$ on firms' profits falls.

In steady state these forces often dominate, causing ad revenue to decline. This is so, particularly because of the third effect. To see why, note that a

[21] Advertising's share of GDP has remained remarkably constant despite major changes in advertising technologies since 1925. As Chemi (2014) notes, "the pie is not growing...the easiest way to make more money is to steal larger slices of the pie." Silk et al. (2021) find that advertising's share of GDP declined slightly during the period 2000-2018 when digital advertising grew in dominance. I revisit this phenomenon in Section 8.

firm's expected flow profit from selling to a consumer is

$$\pi_{\mathbb{J}}\hat{v} = \frac{I\hat{v}}{\sigma M \mu_H}.$$

More informative data raises the $\hat{v}$ of winning firms, but it also raises μ_H, the average $\hat{v}$ in consideration sets. In steady state these coincide, so improvements in match quality do not increase winners' expected profits. This is another expression of the fixed-profits force.

Because auctions are second price, ad revenue depends on the expected profit of the *second*-highest bidder rather than that of the winner. Specifically, it depends on the ratio of the second-highest $\hat{v}$ in auctions to μ_H. This ratio is maximized at one when data is uninformative and typically—though not universally—declines as data becomes more informative.

Transition dynamics The steady-state analysis suggests that transition dynamics matter because some key forces unfold only gradually as the composition of firms in consideration sets adjusts. Following an increase in data informativeness, two mechanisms—the decline in the average expected value of bidders outside consideration sets, μ_{H^c}, and the fall in the profit coefficient $\pi_{\mathbb{J}}$—occur over time. As a result, ad revenue may rise on impact even though it ultimately falls in the long run.

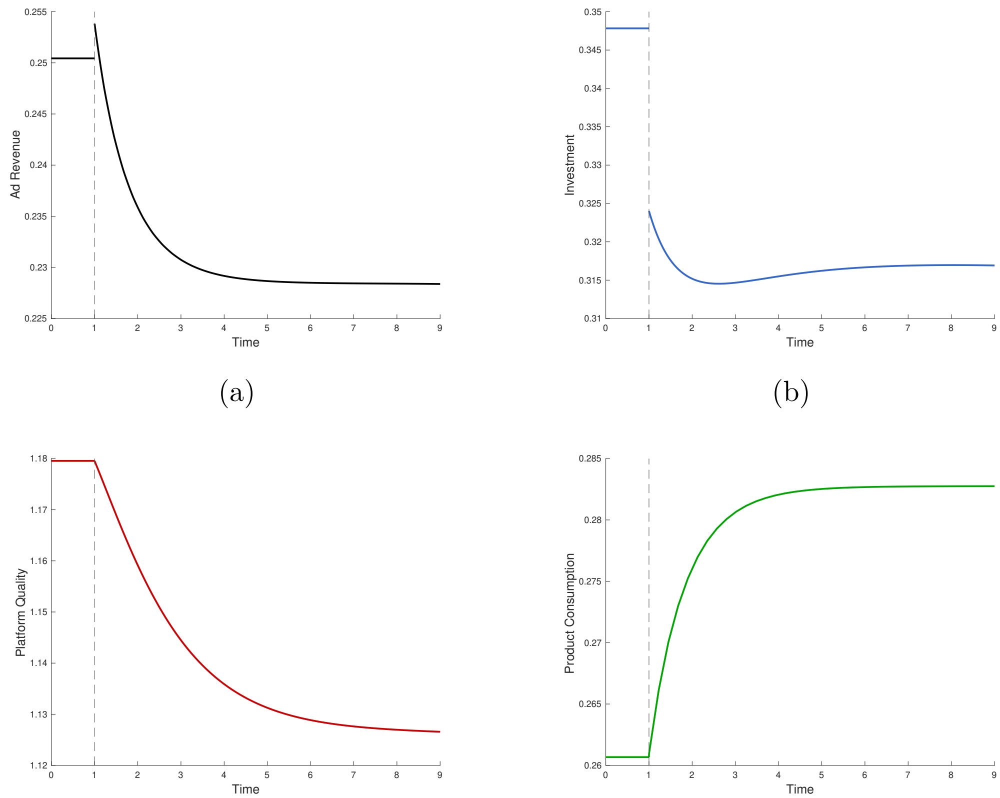
Figure 2 illustrates this pattern following an unanticipated increase in data informativeness at $t = 1$, where G shifts from U$[.2, .8]$ to U$[0, 1]$ as the economy transitions between steady states.[22] This change may represent, for example, a relaxation of privacy laws that allows platforms to target ads with more granular data.

Ad revenue *initially rises* but eventually falls below its starting level. Investment and quality decline steadily, while product consumption rises monotonically. Investment falls despite the temporary increase in ad revenue because the increase is short-lived and platforms are forward-looking.

In this example, data liberalization can appear beneficial for platforms *on impact* but ultimately reduces platform profits *in the long run.* Because platform quality and product consumption move in opposite directions, the welfare effect is ambiguous and depends on $u(\cdot)$.

I close with three remarks.

[22]Figure 6 in the Appendix shows the specific garbling of new data into old data corresponding to this change in G. The details of the garbling—i.e., the joint distribution of old and new expectations—do not affect the steady state, but they matter for transition dynamics because they determine the initial distribution of expectations in consideration sets after the shock.



(a) (b)

(c) (d)

Figure 2: Transition dynamics of increased data informativeness

Notes: Transition between steady states following a data shock at $t = 1$, where the signal distribution changes from $G = \mathrm{U}[.2, .8]$ to $G = \mathrm{U}[0, 1]$ (corresponding to the garbling in Figure 6). Parameters: $\rho = 0.1$, $\epsilon = 1.33$, $\lambda_f = 1$, $K = 0.1$, $J = 1$, $\sigma = 3$, $N = 5$, $I = 1$, $\varphi = 0.5$, $\delta = 0.5$, and $\nu(a) = 1 - 7.5a$.

Remarks

1. **When these patterns arise.** The patterns in Figure 2 tend to arise in numerical examples when firms are less patient and platforms are more patient. Less patient firms put less weight on the future decline in profits caused by better-matched competitors, so bids respond more strongly to the immediate improvement in targeting. More patient platforms put more weight on the eventual decline in ad revenue, so investment falls despite the short-run increase in ad revenue.

2. **Private vs. collective incentives.** Though platforms may collectively

be worse off when data improves for *all* platforms, *individual* platforms typically benefit from private access to more informative data. I provide an illustrative example in Section 8.

3. **Settings with personalized pricing.** In alternative models where prices are personalized in equilibrium, total firm profits may vary with data informativeness. Whether the qualitative comparative statics here are reinforced or mitigated in such settings is an open question.[23]

5.2 Interoperability

Interoperability mandates aim to reduce user lock-in and promote platform competition by making attention more *elastic* to platform quality. For example, consider a content-interoperability rule requiring platforms to allow content sharing across services. Because content becomes more similar across platforms, differences in interface quality or ad load play a more prominent role in attracting attention.

In practice, interoperability mandates may involve technological changes to platforms that alter the aggregator X or even the utility function u on a case-by-case basis. The analysis below remains agnostic about these details and assumes only that *attention* retains its CES shape and becomes *more elastic*—captured by a higher ϵ—after the policy.[24] Interpreted broadly, the analysis is therefore relevant for any pro-competitive intervention that raises the elasticity of attention.

Steady state The following proposition characterizes the steady-state effects of an increase in ϵ.

Proposition 3. *An increase in ϵ raises ad revenue and platform quality, but reduces product consumption, in steady state.*

The mechanism is as follows.

1. **Lower ad load.** More elastic attention induces platforms to reduce ad display rates A and invest more in quality q.

2. **Smaller consideration sets.** Lower A shrinks consideration sets ($M = A/\lambda_f$), reducing product consumption C.

[23] Rhodes and Zhou (2022) show that under "full market coverage," personalized pricing can reduce firm profits, reinforcing the direction of the steady-state results here.

[24] For instance, the aggregator X may be scaled or translated following the mandate.

3. **Higher ad prices.** The ad price $\pi_{\mathbb{K}}$ rises sufficiently that total ad revenue $\pi_{\mathbb{K}} A$ increases despite the decline in A. As seen from (32) in Part 4 of Theorem 1, $\pi_{\mathbb{K}}$ rises because (i) the firm profit coefficient $\pi_{\mathbb{J}}$ scales inversely with A via $M = A/\lambda_f$, and (ii) lower A reduces match rates $\lambda_e(\cdot)$.

   Although the cdfs H_t and H_t^c also vary with A through its effect on M (Part 3 of Theorem 1), I show in Online Appendix C that $\pi_{\mathbb{K}} A$ nevertheless increases.

The increase in ad prices occurs gradually, operating through the dynamics of M_t, H_t, and H_t^c, whereas the effect on the ad rate A is immediate. As a result, ad revenue may *decline* initially, making transition dynamics important.

Transition dynamics A one-time 1% increase in ϵ at $t = 1$ produces the transition path shown in Figure 3.

Interoperability depresses ad revenue and investment *at first*, yet raises them *in steady state*. In the long run product consumption declines while platform quality improves. Thus, as with an increase in data informativeness, short-run and long-run effects differ and the welfare effect is ambiguous and depends on $u(\cdot)$.

Remarks

1. **When these patterns arise.** The patterns in Figure 3 tend to arise in numerical examples when nuisance costs are steeper and firms and platforms are less patient. Steeper nuisance costs make the drop in ad load larger on impact. Less patient firms place less weight on the future decline in consideration sets, so bids do not rise enough initially to offset the lower ad load. As a result, ad revenue falls more sharply in the short run. Less patient platforms, in turn, put less weight on the eventual rebound in ad revenue when choosing investment.

2. **Alternative entry channels.** If there is an alternative entry channel into consideration sets, higher ϵ might *reduce* steady-state ad revenue because M may no longer shrink proportionally with A.[25]

[25] This logic does not apply to word-of-mouth referrals if initial discoveries must originate from advertising. In that case, word-of-mouth is a downstream consequence of advertising rather than a truly separate entry channel.

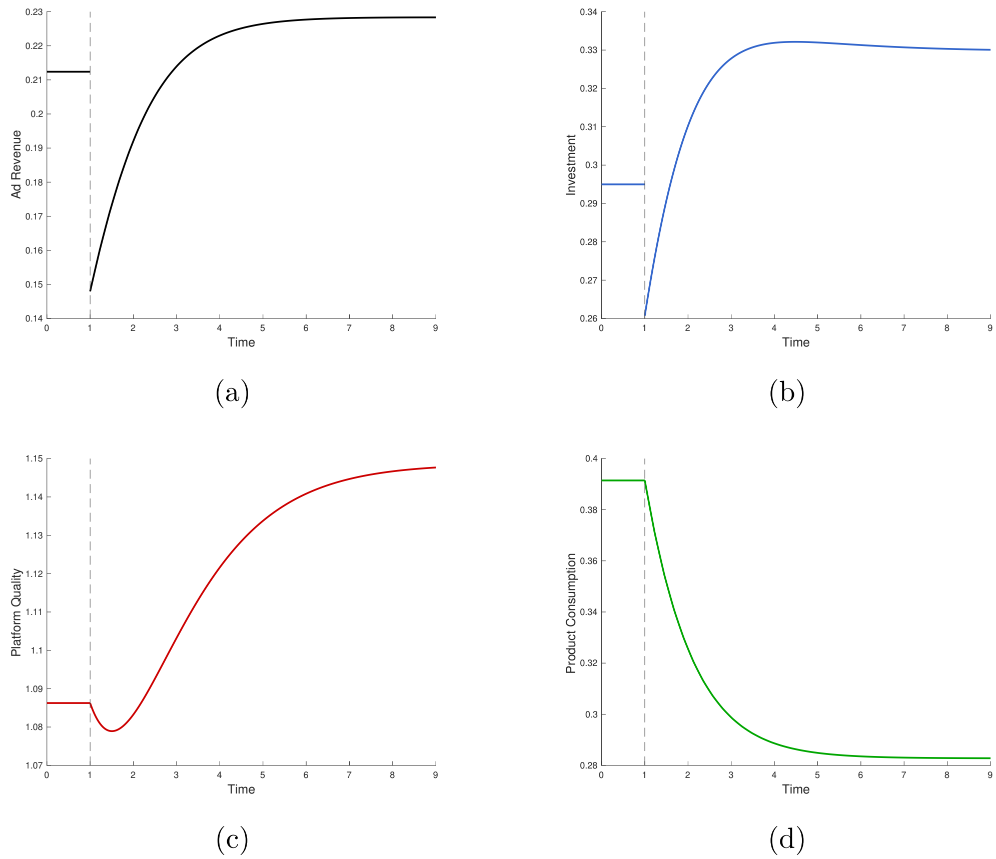


Figure 3: Transition dynamics of increased attention elasticity

Notes: Transition between steady states following an unanticipated shock to ϵ from 1.33 to 1.33(1.01) at $t = 1$. Parameters: $\lambda_f = 1$, $K = .1$, $J = 1$, $\rho = .1$, $\sigma = 3$, $N = 5$, $G = \mathrm{U}[0, 1]$, $I = 1$, $\varphi = .5$, $\delta = .5$, $\nu(a) = (1 - .8395a^{.01})^{63.1472}$.

6 Welfare Analysis in General Equilibrium

So far, I have analyzed policy interventions without characterizing the underlying sources of inefficiency. This section closes the economy so that welfare corresponds to consumer surplus, and then compares the decentralized equilibrium with the planner's allocation. Specializing to Cobb–Douglas utility u, I show that equilibrium ad display and investment rates may lie on either side of the efficient levels, and I derive a sufficient statistic that signs the investment distortion. Although this statistic relies on monopolistic competition among platforms, the planner's solution extends directly to oligopoly settings.

6.1 General Equilibrium

To close the baseline model, I endogenize consumers' incomes. Each consumer supplies L units of labor inelastically at each date. Labor is the only input: platforms use it for investment, and firms use it for production. Thus ℓ_{kt} now denotes the labor hired by platform k for investment. One unit of output requires one unit of labor, so firms' marginal cost equals the wage, which I normalize to one.

Labor market clearing at time t requires

$$L = \int_{\mathbb{K}} \ell_{kt} \, \mathrm{d}k + \int_{\mathbb{J}} \ell_{jt} \, \mathrm{d}j,$$

where

$$\ell_{jt} = \int_{\mathbb{I}} c_{ijt} \, \mathbb{1}_{\{j \in \Omega_{it}\}} \, \mathrm{d}i$$

is firm j's total labor demand.

Consumers collectively own all firms and platforms in equal shares. Since platforms earn only ad revenue and all other costs are wages, income I_t equals aggregate firm revenue:

$$I_t = \int_{\mathbb{J}} p_{jt} \ell_{jt} \, \mathrm{d}j.$$

This identity also follows from product-market clearing.

An equilibrium is defined as in the baseline model, augmented with the income process $\{I_t\}$ and the requirement that labor and product markets clear at all dates.

Proposition 4 provides the general-equilibrium counterpart to Theorem 1.

Proposition 4. *Theorem 1 applies as stated, except with income $I_t = \frac{\sigma}{\sigma-1}(L - K\ell_{\mathbb{K}t})$ in place of I. The equilibrium converges to a steady state in which each platform invests at rate*

$$\ell_{\mathbb{K}} = \frac{\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathbb{K}}A(\epsilon-1)}{\rho+\delta+\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathbb{K}}A(\epsilon-1)} \frac{L}{K}, \tag{19}$$

where $\hat{\pi}_{\mathbb{K}} \equiv \pi_{\mathbb{K}}/I$ denotes the average ad price per unit of income.

The comparative statics from Propositions 2 and 3 also extend to this general equilibrium setting, with $\hat{\pi}_{\mathbb{K}}A$ replacing $\pi_{\mathbb{K}}A$ where relevant (Online Appendix D.2).

I next characterize the planner's allocation before comparing it with the decentralized equilibrium.

6.2 Planner's Problem

The planner chooses ad display and investment rates to maximize welfare, taking as given that consumers choose demands to maximize flow utility and firms set prices and bids to maximize discounted profits. The planner is constrained to treat platforms symmetrically, observes the same consumer data as platforms, and cannot impose technological changes such as interoperability. Even so, this benchmark is valuable for evaluating data and interoperability policies, whose welfare effects depend on the wedge between the decentralized and planner allocations (see Corollary 2.2). Although the planner could also set firms' prices or bids, these choices are already efficient.[26]

Given (M_0, H_0, q_0), the planner solves

$$\max_{\{\ell_{\mathbb{K}t}\},\{A_t\}} \int_0^\infty e^{-\rho t} u(C_t, X_t)\,\mathrm{d}t \tag{20}$$

subject to

$$\begin{aligned} C_t &= (L - K\ell_{\mathbb{K}t})(M_t \mu_{H_t})^{\frac{1}{\sigma-1}}, \\ X_t &= K^{\frac{1}{\epsilon-1}} \nu(A_t) q_{\mathbb{K}t}, \end{aligned}$$

with M_t and H_t following (4)–(6), and $q_{\mathbb{K}t}$ following (13) given $\ell_{\mathbb{K}t}$.[27]

I focus on steady-state solutions.[28]

Definition 2. *A* steady-state solution *of the planner's problem consists of initial conditions (M^*, H^*, q^*) and constants A^* and $\ell_{\mathbb{K}}^*$ such that the planner sets $A_t = A^*$ and $\ell_{\mathbb{K}t} = \ell_{\mathbb{K}}^*$, with $M_t = M^*$ and $H_t = H^*$ for all t.*

For clarity, I assume

$$u(C_t, X_t) = C_t^{1-\tau} X_t^{\tau}, \tag{21}$$

with $\tau \in (0, 1)$ for the remainder of this section.

Theorem 2, proven in Online Appendix D, characterizes steady-state investment for arbitrary ρ and steady-state ad display in the limit as $\rho \to 0$.[29]

[26] Firms charge markups, but all produce identical quantities, which is efficient by concavity of the CES aggregator and ex-ante symmetry. Bids are efficient because they are monotone.

[27] In C_t, the term $L - K\ell_{\mathbb{K}t}$ is labor available for production.

[28] I conjecture that, under standard regularity conditions, any solution converges to the steady state, but do not pursue this here.

[29] The ad display rate for general ρ can also be derived using Pontryagin's Maximum Principle but is less transparent to state.

Theorem 2. *Let u be given by* (21)*. Any steady-state solution has investment*

$$\ell_{\mathbb{K}}^{*} = \frac{\varphi\delta\frac{\tau}{1-\tau}}{\rho+\delta+\varphi\delta\frac{\tau}{1-\tau}}\frac{L}{K}. \tag{22}$$

Suppose a steady-state solution exists for all ρ near zero. Let $A^*(\rho)$ denote the corresponding ad display rate. Then

$$\lim_{\rho\to 0} A^*(\rho) = \arg\max_{a}\,[a\mu_H(a)]^{\frac{1-\tau}{\sigma-1}}\,\nu(a)^{\tau}, \tag{23}$$

whenever the right-hand side is well defined. Here $\mu_H(a)$ is the steady-state average quality of firms in consideration sets when all platforms use ad display rate a.

Corollary 2.1. *If the planner controls investment but not ad display, the unique steady-state choice of investment remains* (22)*.*

6.3 Planner vs Laissez-Faire

I now compare the planner's allocation with the decentralized equilibrium.

Investment Comparing (19) with (22) yields a simple sufficient statistic.

Proposition 5. *For a given τ, the deviation $\ell_{\mathbb{K}} - \ell_{\mathbb{K}}^{*}$ is increasing in*

$$\frac{\sigma}{\sigma-1}\,\hat{\pi}_{\mathbb{K}}A(\epsilon-1) - \frac{\tau}{1-\tau}. \tag{24}$$

Equilibrium investment is too high (too low) when (24) *is positive (negative). Either case can arise, since $\hat{\pi}_{\mathbb{K}}$ does not depend on τ (see* (40)*).*

This statistic depends on four interpretable objects: (*i*) the product-market markup $\sigma/(\sigma-1)$; (*ii*) ad revenue per unit of income $\hat{\pi}_{\mathbb{K}}A$; (*iii*) the elasticity $\epsilon-1$ of attention with respect to platform quality; and (*iv*) the utility weight τ on platform consumption.[30]

These reflect distinct distortions. Markups capture market power, which shifts labor toward platform investment. Ad revenue reflects business-stealing motives whose strength is governed by the elasticity $\epsilon-1$. The utility weight appears because platforms, which charge zero prices, do not internalize the consumer-surplus gains from higher quality whereas the planner does.

A proportional tax or subsidy on ad revenue can correct these distortions.

[30]Since τ appears in the sufficient statistic (24), its empirical counterpart matters for evaluating investment distortions. In principle, one could estimate τ and ϵ jointly by an experiment that charges users for access to platform services.

Corollary 2.2. *A proportional tax on platform ad revenue at rate*

$$1 - \frac{\tau}{1-\tau}\,\frac{\sigma-1}{\sigma}\,\frac{1}{\hat{\pi}_{\mathbb{K}} A(\epsilon-1)},$$

rebated lump-sum to consumers, restores $\ell_{\mathbb{K}}$ to the efficient level (22). *A negative rate corresponds to a subsidy financed by consumers.*

Remark Condition (24) also helps evaluate policy changes. The comparative statics for $\pi_{\mathbb{K}} A$ in Propositions 2 and 3 extend to $\hat{\pi}_{\mathbb{K}} A$ in general equilibrium. Thus a ban on data use reduces investment efficiency and welfare whenever investment is initially too high. Analogous logic applies to changes in ϵ, such as those induced by interoperability.

Ad Display Rate From (16) and (23) the equilibrium ad rate maximizes

$$a\,\nu(a)^{\epsilon-1},$$

whereas, as $\rho \to 0$, the planner's rate maximizes

$$[a\mu_H(a)]^{\frac{1-\tau}{\sigma-1}}\nu(a)^{\tau}.$$

Hence the equilibrium rate may be above or below the planner's. As an illustration, suppose ν is continuous and satisfies $\nu(\overline{a}) = 0$ for some $\overline{a} > 0$.[31] Then, as $\epsilon \to 1$, the equilibrium rate converges to $\overline{a}$ whereas the planner's rate does not, so the equilibrium rate is inefficiently high.

Platforms internalize the nuisance cost of ads only through its effect on attention, since zero prices prevent platforms from capturing the consumer-surplus gains from lower ad loads directly. When ϵ is small and attention is inelastic, this internalization is weak, leading to excessive ads.

Conversely, the equilibrium rate may be too low. As $\tau \to 0$, the planner's optimal rate converges to $\overline{a}$, while the equilibrium rate remains unchanged. In this case the planner places little weight on platform consumption and primarily values product consumption, increasing the ad display rate to expose consumers to more products. Platforms do not internalize this channel and may choose a lower rate than the planner.

[31]The baseline assumption $\nu > 0$ was made for simplicity, since demand (15) is undefined if almost all platforms choose a with $\nu(a) = 0$. Theorem 1 extends to the case in which ν vanishes above $\overline{a} > 0$, provided we restrict attention to equilibria in which each platform sets ad display below $\overline{a}$.

7 Duopoly

This section shows that the framework can also accommodate strategic behavior in a duopoly setting and develops a simple numerical algorithm for computing Markov equilibrium. For tractability, ad display rates are taken as exogenous. I use the duopoly model to assess the robustness of Section 5's findings, particularly since interoperability mandates often target large platforms. I remark on the case of endogenous ad rates at the end of the section.

7.1 Setup

There are two platforms, $k \in \mathbb{K} \equiv \{1, 2\}$. Platform k's quality evolves according to

$$\mathrm{d}q_{kt} = (\ell_{kt}^{\varphi} - \delta q_{kt})\,\mathrm{d}t + \eta q_{kt}\,\mathrm{d}B_{kt}, \qquad q_{k0} = q_0, \tag{25}$$

where B_{1t} and B_{2t} are independent Brownian motions. The law of motion mirrors the baseline model, but with the addition of noise. A small amount of noise is useful for numerical stability and, as in related dynamic duopoly models, appears to help ensure existence of a Markov equilibrium.[32] In the numerical examples, I set $\eta = 0.12$.

To isolate strategic effects along the investment margin, both platforms display ads at a common exogenous rate A. For the particular choice of A given by (16), this assumption admits an interpretation in which platforms set ad display *as if* they were individually small on this margin, as in the baseline. All other elements of the baseline environment are retained.

7.2 Markov Equilibrium

I focus on Markov equilibrium where strategies depend only on the two quality levels and time. Let $V_k(q_1, q_2, t)$ denote platform k's value function. Suppressing arguments, value functions satisfy the coupled HJB system

$$\rho V_k = \sup_{\ell_k} \mathcal{L}^{(\ell_1,\ell_2)} V_k + \pi_{\mathbb{K}t} A \frac{q_k^{\epsilon-1}}{q_1^{\epsilon-1} + q_2^{\epsilon-1}} - \ell_k, \quad k \in \mathbb{K} \tag{26}$$

[32] See Budd et al. (1993) and Harris et al. (2010). In such settings, deterministic formulations often fail to admit Markov equilibrium because, under certain Markov strategy profiles, an arbitrarily small lead can generate large payoff gains for a player. Introducing noise can restore equilibrium existence by softening these leader–follower dynamics. Consistent with this, Budd et al. (1993), who analyze a similar model, claim in their Section 3 that any nontrivial amount of noise ensures equilibrium existence. Given the close connection between our models (see Online Appendix E), I conjecture the same mechanism applies here.

where $\mathcal{L}^{(\ell_1,\ell_2)}$ is the infinitesimal generator of the joint quality process: for any sufficiently smooth $F : \mathbb{R}^3 \to \mathbb{R}$,

$$\mathcal{L}^{(\ell_1,\ell_2)}F = F_t + (\ell_1^{\varphi} - \delta q_1)F_{q_1} + (\ell_2^{\varphi} - \delta q_2)F_{q_2} + \tfrac{1}{2}\eta^2 q_1^2 F_{q_1 q_1} + \tfrac{1}{2}\eta^2 q_2^2 F_{q_2 q_2}.$$

Since all equilibrium objects other than investment coincide with those in the baseline model, solving for a Markov equilibrium reduces to finding value functions and investment policies that satisfy this system.

Algorithm Overview I solve the HJB system using an implicit upwind finite-difference scheme. In steady state, value functions and policies depend on (q_1, q_2) but not on t.

1. **Steady state:** Fix the steady-state ad price from Theorem 1. From an initial guess of value functions, iterate (26) backward in time, updating value functions and controls until convergence to steady state.

2. **Transition path:** Using the steady-state value functions as the terminal condition, iterate (26) backward from a large terminal time to compute value functions and controls along the transition path.

Details and theoretical foundations are discussed in Online Appendix E.

7.3 Policy Experiments Under Duopoly

I revisit the policy experiments from Section 5 to assess the robustness of the results to strategic interaction in investment. I initialize both platforms' quality levels at the point where their drifts are zero in steady-state and plot the transition following a parameter shock along the path where *realized* Brownian shocks are zero. Because the equilibrium is symmetric, the two platforms share a common quality trajectory along this path. Although transition dynamics may vary across realized Brownian paths, I plot this trajectory because it is representative given the modest volatility parameter $\eta = 0.12$.

Figure 4 presents two experiments. Panels (a)–(b) examine a data shock from $G = \mathrm{U}[.2, .8]$ to $G = \mathrm{U}[0, 1]$. Panels (c)–(d) examine an interoperability shock that raises ϵ (from 1.33 to 1.33(1.01)) and lowers the ad rate A (from .2 to .1). If platforms set ad display as in the baseline model (by ignoring strategic effects or the effects on consideration sets), this change in A would arise endogenously from the change in ϵ. The experiment thus isolates the *incremental effects* of strategic investment relative to the baseline model. I plot quality and investment paths. Consumption and ad revenue follow the

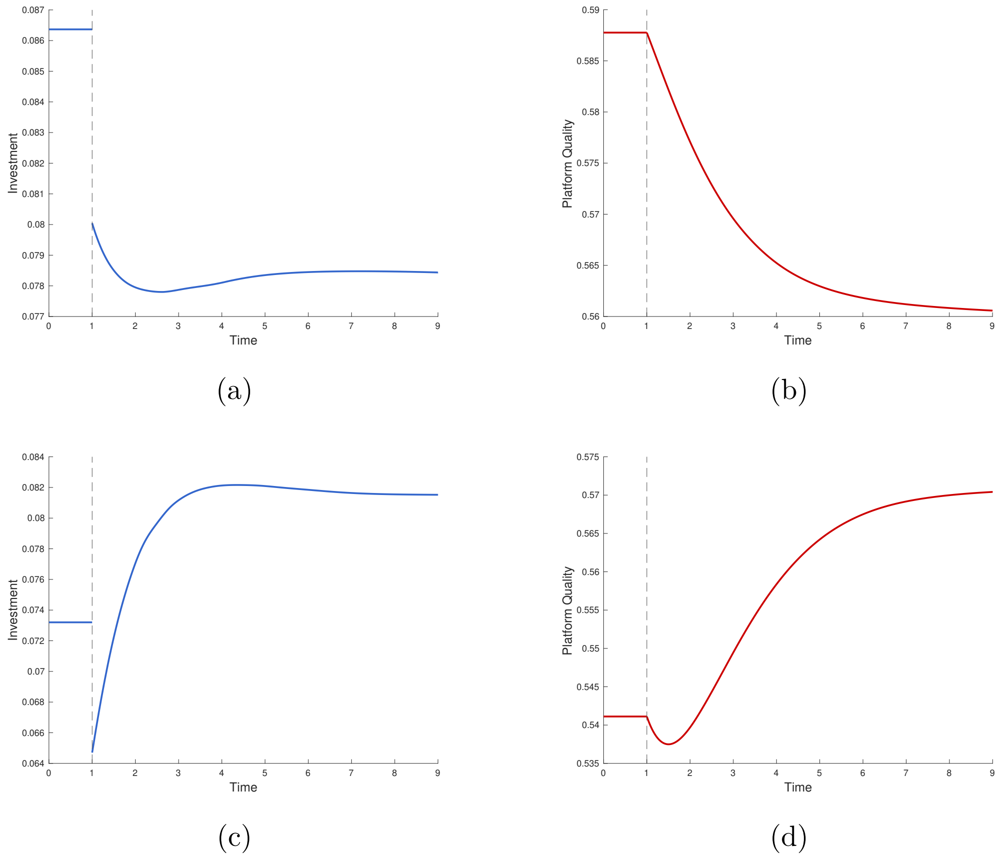


Figure 4: Transition dynamics under duopoly

Notes: Panels (a)–(b) show responses to a shock to G from U$[.2, .8]$ to U$[0, 1]$, using parameters from Figure 2 except with $I = 10$. Panels (c)–(d) show responses to simultaneous shocks to ϵ from 1.33 to 1.33(1.01) and to A from .2 to .1, using parameters from Figure 3 except with $I = 10$.

same paths as in the baseline up to a constant scaling factor and are therefore omitted for brevity.

Across both experiments, the qualitative comparative statics mirror those in the monopolistically competitive case, showing that the key economic mechanisms are robust to strategic investment.

7.4 Comment on Endogenous Ad Display Rates

The analysis treats the ad display rate A as exogenous. Endogenizing ad display requires tracking the entire distribution H_t as a state variable, which complicates the equilibrium computation. A possible route is to instead model a continuum of duopolistic platform markets (for example via a nested CES structure). In that environment, ad display remains a static choice and no individual platform affects the deterministic law of motion of consideration sets. At the same time, platforms remain large under a narrower market definition, with strategic interaction in both investment and ad display. The algorithm developed here appears extendable to this setting particularly for steady state.[33]

8 Data Asymmetry

So far, platforms have shared the same data, and we found that they may *collectively* prefer coarser information. I now extend the framework to allow data asymmetry between two groups of platforms, linking data informativeness to relative market shares and quality levels. In this environment, *individual* platforms often benefit from more granular data even when, collectively, platforms do not. Building on this, I offer a candidate mechanism for the puzzling decline in advertising's share of GDP during 2000–2018, precisely when *data-rich* digital platforms rose to prominence (Silk et al., 2021).

The dynamic auction structure is crucial for keeping the analysis here tractable, as discussed in Subsection 3.1. To my knowledge, relatively little work studies platform competition with asymmetric data.[34]

8.1 Setup

There are two groups of platforms indexed by $z \in \{1, 2\}$, with masses m_z. As before, v_{ij} denotes consumer i's value for firm j. When bidding on a platform in group z, firm j receives signal ζ_{zij}. I assume

$$(v_{ij}, \zeta_{1ij}, \zeta_{2ij}) \sim Q$$

on $[0, \infty) \times \mathbb{R}^2$, independently across i and j. Let G be the joint cdf of $(\zeta_{1ij}, \zeta_{2ij})$ with continuous density g. All other aspects of the baseline model are retained.

[33] In such a setting, the interoperability counterfactual corresponds to raising the elasticity in each market, capturing the aggregate effect of many mandates (or a mandate applied across markets).

[34] See exceptions Ichihashi (2021), Greenwood et al. (2025), and Bergemann and Bonatti (2011). The latter two analyze perfect competition between traditional and digital media. In Greenwood et al. (2025), targeting distinguishes college from noncollege consumers; in Bergemann and Bonatti (2011), targeting ability is captured through the joint distribution of preferences and media use.

8.2 Equilibrium

With asymmetric data, a state variable is the *joint distribution* H_t of signals $(\zeta_{1ij}, \zeta_{2ij})$ in consideration sets. Its law of motion depends on the attention shares allocated across the two platform groups, which depend on platform investment, which depends on ad revenue, which depends on H_t—introducing an additional fixed point relative to the baseline model. Bidding strategies also differ by platform group. Despite this added structure, the model remains parsimonious, and I provide a simple algorithm to compute steady-state equilibrium.[35]

Algorithm Overview A steady-state equilibrium is computed as follows:

1. Guess an attention share x_1 allocated to group 1 platforms.
2. Fixing x_1, iterate the law of motion (45) for H_t forward until convergence.
3. Compute steady-state bid functions and average ad prices via a contraction map (48).
4. Given ad prices, platform qualities and the implied attention share x_1 follow explicitly (50).
5. If the implied x_1 matches the guess, a steady-state equilibrium is found; otherwise update the guess and repeat.

By iterating over a grid of guesses for x_1 on $[0, 1]$, one can identify all steady-state equilibria and verify uniqueness.

8.3 A Puzzling Empirical Pattern

The model can reproduce the following stylized fact:

> *"Perhaps the most puzzling feature…is that the rapid growth of digital advertising has occurred over a period during which the share of U.S. economic activity (as measured by GDP) represented by total advertising expenditures has been in decline."* (Silk et al., 2021)

[35] Details are in Online Appendix F.

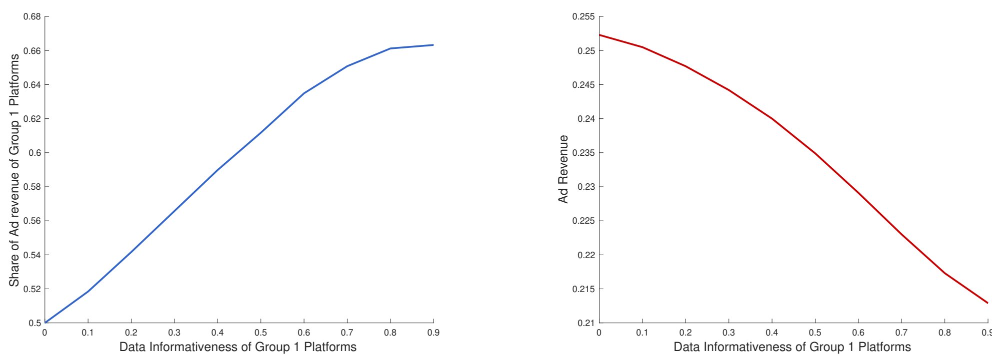
The mechanism mirrors Section 5: More informative data raises the average expected value of firms in consideration sets, μ_{H_t}. Because firms' profits gross of advertising costs remain fixed, the surplus platforms can extract is limited. With greater information rents in auctions, *total* ad revenue can fall even as data-rich platforms gain market share. Though other explanations may exist, the model suggests that the data richness of digital platforms may itself have contributed to the decline in ad revenue.

To illustrate, suppose $v_{ij} = e^{Z_{ij}}$ with $Z_{ij} \sim N(0, \sigma_Z^2)$. There are two platform groups of equal mass. Firm j's signals are

$$\zeta_{1ij} = Z_{ij} + \Delta u, \qquad \zeta_{2ij} = Z_{ij} + u,$$

where $u \sim N(0, \sigma_u^2)$ and $0 \leq \Delta \leq 1$.[36] When $\Delta < 1$, group 1 platforms are data rich (digital), and group 2 platforms are data poor (traditional).

Figure 5 plots the steady-state share of ad revenue accruing to group 1 as a function of its data advantage $1-\Delta$. The left-most point ($\Delta = 1$) corresponds



(a) (b)

Figure 5: Group 1 share of total ad revenue and total ad revenue

Notes: The group 1 share of ad revenue and total ad revenue are plotted as functions of the data advantage of group 1 platforms, measured by $1-\Delta$, for parameter values $\sigma_Z^2 = .5$, $\sigma_u^2 = 2$, $I = 1$, $A = .01$, $\lambda_f = 1$, $F = .1$, $N = 20$, $\rho = 1.6$, $\sigma = 3$, $\varphi = .75$, $\nu(a) = 1 - 62.5a$. Results hold for any platform mass K.

to the case in which both groups are traditional media. Moving rightward, each point $\Delta < 1$ represents a steady state in which data-rich digital platforms have entered alongside traditional media. Both total ad revenue and its split across groups are invariant to the total mass of platforms, so entry alone *cannot*

[36] Firm j never observes both signals simultaneously. When bidding on platform z, the firm observes only ζ_{zij} and does not know the realization of the other signal.

generate the patterns in the figure. Thus, the comparison between $\Delta = 1$ and any $\Delta < 1$ is consistent with *any* pattern of entry or exit, provided the new steady state contains equal masses of each group.

Consistent with the stylized fact, group 1 captures a rising share of ad revenue as its data advantage grows, yet *total* ad revenue declines. In this example, the higher share reflects both higher ad prices and higher attention shares. Thus, individual platforms gain from more informative data even when platforms as a whole do not.

9 Further Extensions

I summarize five extensions of the baseline model to illustrate its ability to accommodate a wide range of market features. Although presented separately, these extensions can—in steady state—be combined within a unified framework. All five admit analysis in general equilibrium, where welfare has a natural interpretation, and Extensions 1–3 also appear tractable in the duopoly and asymmetric-data environments.

1) Network Effects In the first extension, found in Online Appendix H, I introduce network effects by assuming that a platform's effective quality is $\eta(x_{kt})\nu(a_{kt})q_{kt}$, where $\eta(\cdot)$ increases in total attention $x_{kt} = \int_{\mathbb{I}} x_{ikt}\,\mathrm{d}i$. Analogous to (15), equilibrium attention satisfies

$$x_{kt} = \frac{[\eta(x_{kt})\nu(a_{kt})q_{kt}]^{\epsilon-1}}{\int_{\mathbb{K}}[\eta(x_{zt})\nu(a_{zt})q_{zt}]^{\epsilon-1}\mathrm{d}z}, \quad k \in \mathbb{K}.$$

Since x_{kt} appears on both sides, to obtain explicit solutions I assume $\eta(x) = x^{\zeta}$ with $\zeta > 0$.

For any set $E_t \subset \mathbb{K}$ of positive mass, there is a solution of the form

$$x_{kt} = \frac{[\nu(a_{kt})q_{kt}]^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}}}{\int_{E_t}[\nu(a_{zt})q_{zt}]^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}}\mathrm{d}z}\mathbb{1}_{\{k\in E_t\}}, \quad k \in \mathbb{K}.$$

Thus network effects generate *equilibrium multiplicity*. Under the refinement that $E_t = \mathbb{K}$ at each t, a unique equilibrium exists. The equilibrium mirrors that of the baseline model but with a higher elasticity of attention to quality equal to $(\epsilon-1)/[1-\zeta(\epsilon-1)] > \epsilon-1$. This higher elasticity reduces equilibrium ad display rates and increases long-run investment. The refinement is natural when $\mathbb{K}$ is interpreted as a stable set of active platforms.

2) Heterogeneous Platform Productivity The second extension, found in Online Appendix I, allows platforms to differ in the productivity of their investments. I characterize full equilibrium dynamics. By varying the productivity distribution, the model can generate a broad class of non-atomic platform share distributions, enhancing its suitability for quantitative work and empirical applications.

3) Zero Prices In the third extension, found in Online Appendix J, I provide intuition for why zero prices for platform services may emerge endogenously in some environments. I allow platforms to charge only nonnegative prices. Negative prices may be infeasible if platforms cannot reliably distinguish bots from humans (Corrao et al., 2023).

The intuition is straightforward: if the marginal utility of product consumption is much higher than that of platform consumption, attention becomes highly sensitive to platform prices. In such environments, platforms optimally rely solely on advertising revenue and set prices to zero.

4) Firm and Platform Entry The fourth extension, found in Online Appendix K, introduces entry by both firms and platforms subject to upfront costs. I characterize the unique steady-state equilibrium and show that more informative data can induce firm entry and platform exit, while greater platform substitutability can trigger exit by both firms and platforms.

5) Reserve Prices The fifth extension, found in Online Appendix L, allows platforms to set reserve prices in ad auctions. I show that candidate steady-state equilibria typically feature positive reserve prices—even with a continuum of platforms—because search frictions in the ad market inhibit competition. As in the baseline model, steady-state ad revenues are maximized when data is uninformative.

10 Concluding Discussion

This paper proposes a framework for the modern attention economy in which platforms offer free services to attract attention and monetize that attention through targeted advertising. The framework captures the interaction between attention and product markets, platforms' dynamic investment incentives, and welfare tradeoffs between paid products and free platform services. These features are salient in ongoing policy debates but rarely analyzed together in formal models.

The paper provides a tool for investigating mechanisms that arise from these features. The results suggest a common theme: the importance of accounting for cross-market spillovers and dynamic effects when analyzing the attention economy. First, I study data and interoperability policies and find that their effects can be counterintuitive and change sign over time, trading off product consumption against platform quality. In practice, interoperability policies are often aimed at large platforms where strategic effects may be important; evidence from the duopoly variant suggests that similar patterns can arise in such settings.

Second, I investigate welfare in general equilibrium and find that investment and advertising can be too high or too low. With Cobb-Douglas utility, I derive a simple sufficient statistic—based on a few interpretable quantities: the utility weight, ad revenue, income, elasticity of attention, and product markup—that determines the sign of the investment distortion.

Third, I extend the model to capture data asymmetry between two groups of platforms and highlight a channel through which product-market spillovers may help reconcile the rise of data-rich digital platforms with the decline in advertising's share of GDP in the period 2000–2018.

In addition to these results, the framework can flexibly accommodate several salient features of digital markets. I consider extensions with network effects, platform heterogeneity, reserve prices, and endogenous entry, and examine when zero pricing for platform services can arise endogenously.

There are many promising directions for future work. Endogenizing ad display rates in a duopoly setting would allow a more complete analysis of strategic interaction.[37] Although CES preferences preclude personalized pricing, a topic of growing policy attention, the model can be extended to study product steering, a related phenomenon. To study robustness of the findings to other modes of discovery, the framework could also be extended to give consumers an outside option to purchase products without first viewing ads. A sharper characterization of when more informative data reduces ad revenue would further enrich the analysis of data policies.[38] Finally, modifying the investment technology to accommodate long-run growth could connect the framework to policy discussions around innovation.

Much remains to be explored. I hope the framework offered here can serve as a useful step toward a deeper understanding of the attention economy.

[37] See Subsection 7.4, which outlines an approach for the case of duopolistic competition.

[38] Toward this end, methods from Bergemann et al. (2022), Ganuza and Penalva (2010), and the literature on information in auctions may be useful.

References

Ambrus, A., Calvano, E., and Reisinger, M. (2016). Either or both competition: A "two-sided" theory of advertising with overlapping viewerships. *American Economic Journal: Microeconomics*, 8(3):189–222.

Anderson, S. P. and Coate, S. (2005). Market provision of broadcasting: A welfare analysis. *The review of Economic studies*, 72(4):947–972.

Barles, G. and Souganidis, P. E. (1991). Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic analysis*, 4(3):271–283.

Bergemann, D. and Bonatti, A. (2011). Targeting in advertising markets: implications for offline versus online media. *The RAND Journal of Economics*, 42(3):417–443.

Bergemann, D. and Bonatti, A. (2023). Data, competition, and digital platforms. *arXiv preprint arXiv:2304.07653*.

Bergemann, D., Bonatti, A., and Gan, T. (2022). The economics of social data. *The RAND Journal of Economics*.

Bergemann, D., Heumann, T., Morris, S., Sorokin, C., and Winter, E. (2021). Optimal information disclosure in auctions. *AER: Insights*.

Blackwell, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, pages 265–272.

Board, S. (2009). Revealing information in auctions: the allocation effect. *Economic Theory*, 38(1):125–135.

Brunnermeier, M. K. and Sannikov, Y. (2017). Macro, money and finance: A continuous-time approach. *Handbook of Macroeconomics*, 2B:1497–1546.

Budd, C., Harris, C., and Vickers, J. (1993). A model of the evolution of duopoly: Does the asymmetry between firms tend to increase or decrease? *The Review of Economic Studies*, 60(3):543–573.

Cavenaile, L., Celik, M. A., Perla, J., and Roldan-Blanco, P. (2023). A theory of dynamic product awareness and targeted advertising.

Chemi, E. (2014). Advertising's century of flat-line growth. *Bloomberg News*. Available at `https://www.bloomberg.com/news/articles/2014-03-03/advertisings-century-of-flat-line-growth`.

Corrao, R., Flynn, J. P., and Sastry, K. A. (2023). Nonlinear pricing with underutilization: A theory of multi-part tariffs. *American Economic Review*, 113(3):836–860.

De Corniere, A. and De Nijs, R. (2016). Online advertising and privacy. *The RAND Journal of Economics*, 47(1):48–72.

Demirer, M., Hernández, D. J. J., Li, D., and Peng, S. (2024). Data, privacy laws and firm production: Evidence from the gdpr. Technical report, National Bureau of Economic Research.

Duffie, D., Qiao, L., and Sun, Y. (2025). Continuous time random matching. *The Annals of Applied Probability*, 35(3):1755–1790.

Evans, D. S. (2020). The economics of attention markets. *Available at SSRN 3044858*.

Farboodi, M. and Veldkamp, L. (2021). A growth model of the data economy. Technical report, National Bureau of Economic Research.

Galperti, S. and Perego, J. (2022). Competitive markets for personal data. *Available at SSRN 4309966*.

Ganuza, J.-J. and Penalva, J. S. (2010). Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica*, 78(3):1007–1030.

Greenwood, J., Ma, Y., and Yorukoglu, M. (2025). "you will:" a macroeconomic analysis of digital advertising. *Review of Economic Studies*, 92(3):1837–1881.

Harris, C., Howison, S., and Sircar, R. (2010). Games with exhaustible resources. *SIAM Journal on Applied Mathematics*, 70(7):2556–2581.

Hummel, P. and McAfee, R. P. (2016). When does improved targeting increase revenue? *ACM Transactions on Economics and Computation (TEAC)*, 5(1):1–29.

Ichihashi, S. (2021). Competing data intermediaries. *The RAND Journal of Economics*, 52(3):515–537.

Jones, C. I. and Tonetti, C. (2020). Nonrivalry and the economics of data. *American Economic Review*, 110(9):2819–58.

Kirpalani, R. and Philippon, T. (2021). Data sharing and market power with two-sided platforms. Technical report.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

OECD (2018). Quality considerations in digital zero-price markets.

OECD (2021). Competition in digital advertising markets. OECD Roundtables on Competition Policy Papers 259, OECD Publishing, Paris.

OECD, O. (2022). Handbook on competition policy in the digital age.

Prat, A. and Valletti, T. M. (2021). Attention oligopoly. *American Economic Journal: Microeconomics, Forthcoming*.

Rachel, L. (2024). Leisure-enhancing technological change. Available at: `https://sites.google.com/site/lukaszrachel/home?authuser=0`.

Rhodes, A. and Zhou, J. (2022). Personalized pricing and competition. *Available at SSRN 4103763*.

Silk, A. J., Berndt, E. R., et al. (2021). Aggregate advertising expenditure in the us economy: Measurement and growth issues in the digital era. *Foundations and Trends® in Marketing*, 15(1):1–85.

A Appendix

A.1 Intuition for Proposition 1

Consumer i's CES preferences imply a demand for product $j \in \Omega_{it}$ of

$$c_{ijt} = \frac{I v_{ij}}{\int_{\Omega_{it}} v_{iz} p_{zt}^{1-\sigma} \mathrm{d}z} p_{jt}^{-\sigma}. \tag{27}$$

Given the demand in (27), firm j's flow profit from selling to consumer i is

$$\frac{I v_{ij}}{\int_{\Omega_{it}} v_{iz} p_{zt}^{1-\sigma} \mathrm{d}z} p_{jt}^{-\sigma} (p_{jt} - 1). \tag{28}$$

Since v_{ij} and Ω_{it} appear in a term that scales demand by the same factor for any given price, they are irrelevant to the firm's pricing decision. Firm j's optimal price is therefore a constant markup irrespective of its information about the consumer's preferences.

A.2 Equilibrium Properties Along the Transition Path

Theorem 1 characterizes the equilibrium in steady state. Along the transition path, the equilibrium has the following properties.

- Ad display rates and attention are as in Parts 1 and 6 of Theorem 1.
- The mass of firms in Ω_{it} is

$$M_t = \frac{A}{\lambda_f} - \left(\frac{A}{\lambda_f} - M_0\right) e^{-\lambda_f t}. \tag{29}$$

- The cdf H_t^c of expected values in Ω_{it}^c satisfies

$$\int_{H_0^c(\hat{v})}^{H_t^c(\hat{v})} \frac{1}{JG(\hat{v}) - Mu^N - (J-M)u} \,\mathrm{d}u = \frac{\ln\left[M - M_0 + (J-M)e^{\lambda_f t}\right]}{J-M} \tag{30}$$

  at each $\hat{v} \in [0, \infty)$. Given H_t^c, H_t is found from (4).

- Firm j's bid in an auction for a consumer with expected value $\hat{v}$ is

$$B_t(\hat{v}) = \int_0^{\hat{v}} \int_t^{\infty} \pi_{\mathbb{J}s} e^{-\int_t^s [\rho + \lambda_f + \lambda_{ez}(y)]\,\mathrm{d}z} \,\mathrm{d}s\,\mathrm{d}y, \quad \hat{v} \in [0, \infty). \tag{31}$$

- The average ad price in an auction is

$$\pi_{\mathbb{K}t} = \int_0^{\infty} \int_t^{\infty} \pi_{\mathbb{J}s} e^{-\int_t^s [\rho + \lambda_f + \lambda_{ez}(\hat{v})]\,\mathrm{d}z} \,\mathrm{d}s\,[1 - O_t(\hat{v})]\,\mathrm{d}\hat{v}, \tag{32}$$

  where $O_t = (H_t^c)^N + N(H_t^c)^{N-1}(1 - H_t^c)$ denotes the cdf of the second-highest expected value among firms in an auction.

- Platform investment ℓ_{kt} and quality q_{kt} solve the boundary value problem

$$\begin{aligned}
\dot{\ell}_{kt} &= \frac{\rho + \delta}{1 - \varphi}\ell_{kt} - \frac{\varphi}{1-\varphi}\frac{\pi_{\mathbb{K}t}A(\epsilon - 1)}{Kq_{kt}}\ell_{kt}^{\varphi}, \\
\dot{q}_{kt} &= \ell_{kt}^{\varphi} - \delta q_{kt}, \\
\lim_{t\to\infty} \ell_{kt} &= \ell_{\mathbb{K}}, \\
q_{k0} &= q_0,
\end{aligned} \tag{33}$$

  where $\ell_{\mathbb{K}}$ is steady-state investment from Part 5 of Theorem 1. Numerical solutions to (33) are computable in seconds via the shooting method.

A.3 Proof of Theorem 1

Proof. The proof follows the steps outlined in Section 4. I begin with Step 3, since Step 1 is standard and Step 2 was completed in the main text. Steps 1 and 2 yield Parts 1 and 6 of Theorem 1.

Step 3. Calculate M_t, H_t, H_t^c. Because the ODE (7) is linear, it can be solved directly to yield (29). Taking $t \to \infty$ yields the steady-state value of M in Part 2 of Theorem 1.

To derive H_t and H_t^c, I conjecture (and later verify) that bidding strategies are monotone, so that (6) applies:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(M_t H_t(\cdot)\right) = A H_t^c(\cdot)^N - \lambda_f M_t H_t(\cdot).$$

Using the accounting identity (4) to express this in terms of H_t^c alone gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[(J - M_t) H_t^c(\cdot)\right] = \lambda_f [JG(\cdot) - (J - M_t) H_t^c(\cdot)] - A H_t^c(\cdot)^N.$$

Expanding the left-hand side and substituting $\dot{M}_t = A - \lambda_f M_t$ from (7),

$$\begin{aligned}(J - M_t)\dot{H}_t^c(\cdot) &= \lambda_f [JG(\cdot) - (J - M_t) H_t^c(\cdot)] - A H_t^c(\cdot)^N + (A - \lambda_f M_t) H_t^c(\cdot) \\ &= \lambda_f [JG(\cdot) - M H_t^c(\cdot)^N - (J - M) H_t^c(\cdot)],\end{aligned}$$

where $M = A/\lambda_f$ is the steady-state value of M_t. Setting the right-hand side to zero gives the steady-state cdf H^c in Part 3 of the theorem.

Since M on the right-hand side is constant, the ODE is separable and can be solved analytically, yielding (30). As $t \to \infty$, the right-hand side of (30) diverges, so the integral on the left-hand side must diverge as well. The integrand is bounded except where its denominator vanishes, so $H_t^c(\hat{v})$ must approach a zero of the denominator. That zero is the steady-state value $H^c(\hat{v})$ identified above, so H_t^c (and therefore H_t) converges to its steady-state level.

Step 4. Firms' Bidding Strategies. Recall from Section 3 that the rate at which a firm outside Ω_{it} enters Ω_{it}, conditional on $\hat{v}_{ij} = \hat{v}$, is $\lambda_{et}(\hat{v}) = \lambda_{at} W_t(\hat{v})$, where $\lambda_{at} \equiv N A_t/(J - M_t)$ is the rate at which the firm is invited to bid and $W_t(\hat{v})$ is the probability of winning conditional on being invited. The value functions V_t^{In} and V_t^{Out} satisfy the HJB equations

$$\dot{V}_t^{\mathrm{In}}(\hat{v}) = \rho V_t^{\mathrm{In}}(\hat{v}) - \lambda_f \left[V_t^{\mathrm{Out}}(\hat{v}) - V_t^{\mathrm{In}}(\hat{v})\right] - \pi_{\mathbb{J}t}\hat{v}$$

and

$$\dot{V}_t^{\text{Out}}(\hat{v}) = \rho V_t^{\text{Out}}(\hat{v}) - \underbrace{\lambda_{at} W_t(\hat{v})}_{=\lambda_{et}(\hat{v})} \left(V_t^{\text{In}}(\hat{v}) - V_t^{\text{Out}}(\hat{v}) - \mathbb{E}\left[B_t^{(1)} \,\middle|\, B_t(\hat{v}) > B_t^{(1)} \right] \right),$$

where $B_t^{(1)}$ is the maximum of the $N-1$ other bids in that auction. If B_t is increasing, then $W_t = (H_t^c)^{N-1}$ and $B_t^{(1)} \overset{d}{=} B_t(\hat{v}^{(1)})$, where $\hat{v}^{(1)} \sim W_t$.

In any equilibrium, because the auction is second price, $B_t = V_t^{\text{In}} - V_t^{\text{Out}}$. Subtracting the two HJB equations gives

$$\dot{B}_t(\hat{v}) = [\rho + \lambda_{at} W_t(\hat{v}) + \lambda_f]\, B_t(\hat{v}) - \lambda_{at} \int_0^{B_t(\hat{v})} s \,\mathrm{d}\tilde{W}_t(s) - \pi_{\mathbb{J}t} \hat{v},$$

where $\tilde{W}_t$ is the cdf of $B_t^{(1)}$, so $W_t(\hat{v}) = \tilde{W}_t(B_t(\hat{v}))$ by definition. Differentiating in $\hat{v}$ yields

$$\dot{B}_t'(\hat{v}) = [\rho + \lambda_{at} W_t(\hat{v}) + \lambda_f]\, B_t'(\hat{v}) - \pi_{\mathbb{J}t},$$

which has solution

$$B_t'(\hat{v}) = e^{\int_0^t [\rho + \lambda_f + \lambda_{az} W_s(\hat{v})] \,\mathrm{d}s} \left(B_0'(\hat{v}) - \int_0^t \pi_{\mathbb{J}s} e^{-\int_0^s [\rho + \lambda_f + \lambda_{az} W_z(\hat{v})] \,\mathrm{d}z} \,\mathrm{d}s \right).$$

Since B_t is Lipschitz (Lemma 1 in Online Appendix B), the term in parentheses must vanish as $t \to \infty$; otherwise B_t' would diverge. Therefore

$$B_t'(\hat{v}) = e^{\int_0^t [\rho + \lambda_f + \lambda_{az} W_s(\hat{v})] \,\mathrm{d}s} \int_t^\infty \pi_{\mathbb{J}s} e^{-\int_0^s [\rho + \lambda_f + \lambda_{az} W_z(\hat{v})] \,\mathrm{d}z} \,\mathrm{d}s,$$

which is positive, so the bidding function is increasing as conjectured. This holds for any path of H_t, confirming that the assumption underlying (6) is valid.

The bidding function (31) then follows by substituting $H_t^c(\hat{v})^N = W_t(\hat{v})$. A direct calculation yields the average ad price (32) and its steady-state counterpart in Part 4 of Theorem 1.

Step 5. Investment and Quality. Given the average ad price and consumers' attention choices, the Hamiltonian for platform k's problem is

$$\mathcal{H}(t, q_{kt}, \ell_{kt}, \lambda_t) = \pi_{\mathbb{K}t} A \frac{q_{kt}^{\epsilon-1}}{\int_{\mathbb{K}} q_{lt}^{\epsilon-1} \,\mathrm{d}l} - \ell_{kt} + \lambda_t \left(\ell_{kt}^{\varphi} - \delta q_{kt} \right),$$

where the costate λ_t satisfies

$$\rho \lambda_t - \dot{\lambda}_t = \pi_{\mathbb{K}t} A (\epsilon - 1) \frac{q_{kt}^{\epsilon-2}}{\int_{\mathbb{K}} q_{lt}^{\epsilon-1} \,\mathrm{d}l} - \lambda_t \delta.$$

Lemma 2 in Online Appendix B shows that the platform's problem (12) has a unique solution when $\epsilon - 1 < 1/\varphi$, so all platforms employ the same investment strategy in equilibrium. The costate equation therefore simplifies to

$$\rho\lambda_t - \dot{\lambda}_t = \frac{\pi_{\mathbb{K}t} A(\epsilon - 1)}{K q_{kt}} - \lambda_t \delta.$$

The first-order condition for the Hamiltonian gives

$$\lambda_t = \frac{1}{\varphi} \ell_{kt}^{1-\varphi},$$

and differentiating in t yields

$$\dot{\lambda}_t = \frac{1-\varphi}{\varphi} \ell_{kt}^{-\varphi} \dot{\ell}_{kt}.$$

Substituting into the costate equation gives

$$\dot{\ell}_{kt} = \frac{\rho + \delta}{1-\varphi} \ell_{kt} - \frac{\varphi}{1-\varphi} \frac{\pi_{\mathbb{K}t} A(\epsilon - 1)}{K q_{kt}} \ell_{kt}^{\varphi}, \tag{34}$$

where recall that

$$\dot{q}_{kt} = \ell_{kt}^{\varphi} - \delta q_{kt}, \qquad q_{k0} = q_0. \tag{35}$$

The steady-state values of ℓ_{kt} and q_{kt} follow from (34)–(35), yielding Part 5 of the theorem.

It remains to establish existence and uniqueness of an equilibrium trajectory satisfying (34)–(35). Three lemmas, proved in Online Appendix B, deliver this. Lemma 3 shows that a solution exists for any initial conditions. Lemma 4 shows that any such solution either vanishes, diverges, or converges to steady state. Only the last is consistent with equilibrium, so the equilibrium trajectory solves the boundary value problem (33). Lemma 5 establishes that this boundary value problem has a unique solution.

The final step verifies that the Hamiltonian approach is valid—that is, that platform k optimizes by investing according to the solution to (33), provided its rivals do the same. When $\epsilon < 2$, the Hamiltonian is jointly concave in state and control, so the Mangasarian sufficient conditions apply directly. In the more general case $\epsilon - 1 < 1/\varphi$, joint concavity may fail. I instead apply the calculus of variations: the Gateaux derivative of platform k's objective with respect to the control vanishes in all directions when investment satisfies (33), and the second-order condition holds because the objective is concave in the control (Lemma 2). Details are omitted for brevity.

Finally, Part 7 of the theorem, which records steady-state product and platform consumption aggregates, follows immediately from the earlier parts. $\square$

A.4 Data Shock: Garbling

Figure 6 shows how $\tilde{v} \sim \mathrm{U}[0,1]$ is garbled into $\hat{v} \sim \mathrm{U}[.2,.8]$.

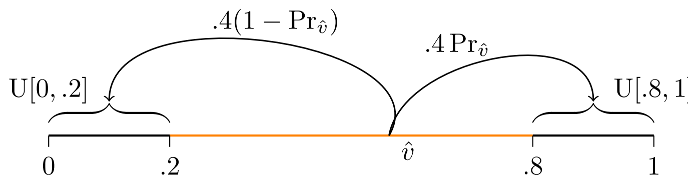


Figure 6: Joint distribution of pre- and post-shock expectations

Notes: Let $\hat{v}$ be any point in the orange region $[.2,.8]$. Following the shock, $\hat{v}$ stays put in that $\tilde{v} = \hat{v}$ with probability .6. Otherwise, with the residual probability .4, it jumps to one of the black regions $[0,.2] \cup [.8,1]$. Conditional on jumping up (down), $\tilde{v}$ is distributed uniformly across the upper (lower) black region. The probability $\Pr_{\hat{v}}$ that $\hat{v}$ jumps up is such that the martingale property holds: $\mathbb{E}[\tilde{v}|\hat{v}] = \hat{v}$.

Online Appendix

B Supplemental Material for Section 4

Appendix B.1 gives the formal statement of a firm's bidding problem. Appendix B.2 proves the auxiliary lemmas used in the proof of Theorem 1.

B.1 Formal Statement of Firms' Bidding Problem

Let τ_z denote the zth time of entry into an auction for consumer i. That is, τ_z is the zth arrival of a Poisson process that ticks at rate $\lambda_{at} + \lambda_f \mathbb{1}_{\{j \in \Omega_{it}\}}$. Taking the bidding strategy $B_t(\cdot)$ of its rivals as given, firm j sets bids to maximize the present value of flow profits (including the costs of advertising) from selling to a typical consumer i:

$$\Pi_{\mathbb{J}} = \max_{\{b_z\}} \mathbb{E}\left[\int_0^\infty e^{-\rho s}\pi_{\mathbb{J}s}(\hat{v}_{ij})\mathbb{1}_{\{j\in\Omega_{is}\}}\,\mathrm{d}s - \sum_{z=1}^{\infty} e^{-\rho\tau_z} B_{\tau_z}\left(\hat{v}_z^{(1)}\right)\mathbb{1}_{\left\{b_z > B_{\tau_z}\left(\hat{v}_z^{(1)}\right)\right\}}\right]$$

where $\hat{v}_z^{(1)}$ is the highest expectation of the $N-1$ other bidders in the zth auction: $\hat{v}_z^{(1)} \sim \left(H_{\tau_z}^c\right)^{N-1}$ conditional on τ_z. Above, the bid b_z in the zth auction is a measurable function of the expectation $\hat{v}_{ij}$ and time t.

B.2 Auxiliary Lemmas for Theorem 1

This subsection proves Lemmas 2–5, used in the proof of Theorem 1.

Lemma 1. *In an equilibrium, the bidding function $B_t(\cdot)$ is Lipschitz at each t.*

Proof. Since $B_t = V_t^{\mathrm{In}} - V_t^{\mathrm{Out}}$, it suffices to show that V_t^{In} and V_t^{Out} are Lipschitz. I argue for V_t^{In}; the argument for V_t^{Out} is analogous.

Fix two expected values $\hat{v} > \tilde{v}$. Consider a firm with expected value $\tilde{v}$ that deviates by bidding as if its expected value were $\hat{v}$. The deviating firm's bidding induces the same dynamics for $\mathbb{1}_{\{j\in\Omega_{is}\}}$ as a firm actually with expected value $\hat{v}$, but its true flow profit at time s is $\pi_{\mathbb{J}s}\tilde{v}$ rather than $\pi_{\mathbb{J}s}\hat{v}$. Letting $\hat{V}_t^{\mathrm{In}}(\tilde{v})$ denote the value of this deviation,

$$V_t^{\mathrm{In}}(\hat{v}) - \hat{V}_t^{\mathrm{In}}(\tilde{v}) = (\hat{v} - \tilde{v})\,\mathbb{E}\left[\int_t^\infty e^{-\rho(s-t)}\pi_{\mathbb{J}s}\,\mathbb{1}_{\{j\in\Omega_{is}\}}\,\mathrm{d}s\right],$$

where the expectation is taken under the $\hat{v}$-bidding law. Note that ad costs cancel since both firms bid identically.

Optimality of $V_t^{\text{In}}(\tilde{v})$ implies $V_t^{\text{In}}(\tilde{v}) \geq \hat{V}_t^{\text{In}}(\tilde{v})$, so

$$V_t^{\text{In}}(\hat{v}) - V_t^{\text{In}}(\tilde{v}) \leq (\hat{v} - \tilde{v})\, \mathbb{E}\left[\int_t^{\infty} e^{-\rho(s-t)} \pi_{\mathbb{J}s}\, \mathbb{1}_{\{j \in \Omega_{is}\}}\, \mathrm{d}s\right] \leq \frac{\sup_{s \geq t} \pi_{\mathbb{J}s}}{\rho}\, (\hat{v} - \tilde{v}).$$

Since $V_t^{\text{In}}(\hat{v}) \geq V_t^{\text{In}}(\tilde{v})$ by monotonicity (a firm with higher expected value can always replicate the strategy of a firm with lower expected value), the difference is non-negative, and

$$\left|V_t^{\text{In}}(\hat{v}) - V_t^{\text{In}}(\tilde{v})\right| \leq \frac{\sup_{s \geq t} \pi_{\mathbb{J}s}}{\rho}\, |\hat{v} - \tilde{v}|.$$

Since $\sup_{s \geq t} \pi_{\mathbb{J}s} < \infty$, V_t^{In} is Lipschitz. □

Lemma 2. *If $\epsilon - 1 < 1/\varphi$, there exists a unique solution to the platform problem 12 which is strictly concave. If $\epsilon - 1 > 1/\varphi$, the platform's problem does not have a solution.*

Proof. Solving out the ODE for q_t yields

$$q_{kt} = e^{-\delta t} \int_0^t e^{\delta s} \ell_{ks}^{\varphi}\, \mathrm{d}s + e^{-\delta t} q_0.$$

The flow utility of platform k is then

$$\pi_{\mathbb{K}t} \frac{\left(e^{-\delta t} \int_0^t e^{\delta s} \ell_{ks}^{\varphi}\, \mathrm{d}s + e^{-\delta t} q_0\right)^{\epsilon - 1}}{\int_{\mathbb{K}} q_{zt}^{\epsilon-1}\, \mathrm{d}z} - \ell_{kt}.$$

Suppose that $\epsilon - 1 < 1/\varphi$. Then we observe that the flow utility must be concave in $\{\ell_{kt}\}$—The second term is linear; the first term's concavity is determined by the numerator, which is a CES aggregator with share weights determined by the exponential function. It is well known that this aggregator is strictly concave as long as $\epsilon - 1 < 1/\varphi$.

Since the flow utility is concave in $\{\ell_{kt}\}$ at each t, the objective function must also be concave in $\{\ell_{kt}\}$. As a result, any solution to platform k's problem must be unique.

Now suppose that $\epsilon - 1 > 1/\varphi$. I claim that there can not exist an equilibrium. Fix an arbitrary strategy $\{\ell_{kt}\}$ and then consider scaling up by some factor χ. For large χ, flow utility is determined primarily by the term

$$\chi^{\varphi(\epsilon-1)}\pi_{\mathbb{K}t}\frac{\left(e^{-\delta t}\int_0^t e^{\delta s}\ell_{ks}^{\varphi}\,\mathrm{d}s\right)^{\epsilon-1}}{\int_{\mathbb{K}} q_{zt}^{\epsilon-1}\,\mathrm{d}z} - \chi\ell_{kt}.$$

Thus if $\varphi(\epsilon-1) > 1$, it is possible for the platform to acheive an arbitrarily high value for the objective simply be scaling up χ. □

Lemma 3. *There exists a unique solution to the ODE system*

$$\begin{aligned}\dot{\ell}_{kt} &= \frac{\rho+\delta}{1-\varphi}\ell_{kt} - \frac{\varphi}{1-\varphi}\frac{\pi_{\mathbb{K}t}A(\epsilon-1)}{Kq_{kt}}\ell_{kt}^{\varphi} \\ \dot{q}_{kt} &= \ell_{kt}^{\varphi} - \delta q_{kt}\end{aligned} \tag{36}$$

for the whole domain $t \in [0,\infty)$ for any positive initial conditions (ℓ_{k0}, q_{k0}). Moreover, the solution for investment $\{\ell_{kt}\}$ is increasing and continuous in its initial condition ℓ_{k0}.

Proof. Since the system can be written in standard form as $[\dot{\ell}_{kt}\,\dot{q}_{kt}] = f(t, [\ell_{kt}\,q_{kt}])$ where f is continuous and locally Lipschitz in $[\ell_{kt}\,q_{kt}]$ the existence and uniqueness of local solutions follows from Picard-Lindelof. These properties also imply that solutions are continuous in initial conditions. Observe that the ODE for ℓ_{kt} has an absorbing point at zero and by Gronwall's inequality

$$\ell_{kt} \leq e^{\frac{\rho+\delta}{1-\varphi}t}$$

whenever a solution exists. Similarly from the ODE for q_{kt} we have

$$q_{kt} = e^{-\delta t}q_0 + \int_0^t e^{-\delta(t-s)}\ell_{ks}^{\varphi}\,\mathrm{d}s \leq e^{-\delta t}q_0 + \int_0^t e^{-\delta(t-s)}e^{\varphi\frac{\rho+\delta}{1-\varphi}t}\,\mathrm{d}s.$$

But then the solution of the ODE system can not explode in finite time or become negative and so the maximal domain of existence must be all of $t \in [0,\infty)$.

To see that $\{\ell_{kt}\}$ is monotone in ℓ_{k0} divide both sides of the ODE for ℓ_{kt} by ℓ_{kt}. Then we have

$$\frac{\dot{\ell}_{kt}}{\ell_{kt}} = \frac{\rho+\delta}{1-\varphi} - \frac{\varphi}{1-\varphi}\ell_{kt}^{\varphi-1}\frac{\pi_{\mathbb{K}t}A(\epsilon-1)}{Kq_{kt}}.$$

Let $f_t = \ln(\ell_{kt})$. It suffices to show $\{f_t\}$ is monotone in f_0. We have

$$\dot{f}_t = \frac{\rho+\delta}{1-\varphi} - \frac{\varphi}{1-\varphi}(e^{f_t})^{\varphi-1}\frac{\pi_{\mathbb{K}t}A(\epsilon-1)}{Kq_{kt}}$$

with

$$\dot{q}_{kt} = e^{\varphi f_t} - \delta q_{kt}.$$

Now I observe that $\dot{f}$ is increasing in both f_t and q_{kt} with the former being true since $\varphi < 1$. Moreover, any solution for q_{kt} is monotone in $\{f_s, s \leq t\}$. Let f_{10} and f_{20} be two initial conditions for f_t. Suppose $f_{10} > f_{20}$. Let $\{f_{1t}\}$ and $\{f_{2t}\}$ be the corresponding solutions. It follows that $\dot{f}_{1t} > \dot{f}_{2t}$ for all t and so $f_{1t} - f_{2t} > f_{10} - f_{20}$ for all t. □

Lemma 4. *Any solution of the ODE system* (36) *for investment ℓ_{kt} for any initial conditions either diverges, vanishes, or converges to steady state.*

Proof. In what follows, let (ℓ_{SS}, q_{SS}) denote the unique steady state of the ODE system (36).

I first show that if $q_{kt} \to q_{SS}$, then $\ell_{kt} \to \ell_{SS}$. To see why, suppose $q_{kt} \to q_{SS}$ and fix $\alpha > 0$. There exists $\zeta > 0$ sufficiently small such that

$$\frac{\rho + \delta}{1 - \varphi} - \frac{\varphi}{1 - \varphi} \frac{(\pi_{\mathbb{K}} + \zeta)\, A(\epsilon - 1)}{K\,(q_{SS} - \zeta)} (\ell_{SS} + \alpha)^{\varphi - 1} > 0. \tag{37}$$

Note that if α and ζ were zero, the left-hand side above would be zero since (ℓ_{SS}, q_{SS}) are by definition steady state solution to (36).

Since $\pi_{\mathbb{K}t} \to \pi_{\mathbb{K}}$ and $q_{kt} \to q_{SS}$, for all t sufficiently large, we have

$$\begin{aligned} \frac{\dot{\ell}_{kt}}{\ell_{kt}} &= \frac{\rho + \delta}{1 - \varphi} - \frac{\varphi}{1 - \varphi} \frac{\pi_{\mathbb{K}t} A(\epsilon - 1)}{K q_{kt}} \ell_{kt}^{\varphi - 1} \\ &> \frac{\rho + \delta}{1 - \varphi} - \frac{\varphi}{1 - \varphi} \frac{(\pi_{\mathbb{K}t} + \zeta)\, A(\epsilon - 1)}{K\,(q_{SS} - \zeta)} \ell_{kt}^{\varphi - 1} \\ &> 0. \end{aligned}$$

Therefore, for some t sufficiently large, if ever $\ell_{kt} > \ell_{SS} + \alpha$, then $\dot{\ell}_{ks}/\ell_{ks}$ will be bounded below by (37) for all $s \geq t$ and thus ℓ_{ks} must diverge. But then clearly, $q_{kt} \to q_{SS}$ can not hold. We can follow an analogous argument to show that if ever $\ell_{kt} < \ell_{SS} - \alpha$ for some t sufficiently large then ℓ_{kt} must eventually vanish and so $q_{kt} \to q_{SS}$ also can not hold. Since $\alpha > 0$ was arbitrary it follows that if $q_{kt} \to q_{SS}$, then $\ell_{kt} \to \ell_{SS}$.

Therefore, suppose that $q_{kt} \nrightarrow q_{SS}$. Then there exists $\alpha > 0$ such that for any T, there exists $t > T$ such that $|q_{kt} - q_{SS}| > \alpha$. The proof is complete if we can show that ℓ_{kt} must either diverge or vanish.

Let t be sufficiently large so that, $|q_{kt} - q_{SS}| > \alpha$ and $|\pi_{\mathbb{K}t} - \pi_{\mathbb{K}}| < \zeta$ both hold. There are two cases to consider.

1. In the first case, $q_{kt} \geq q_{SS} + \alpha$. We may without loss assume that $\ell_{kt} \geq \ell_{SS}$ because after a long enough time, there must be t such that $\ell_{kt} \geq \ell_{SS}$ since otherwise q_{kt} could not have been reached. Therefore $\dot{\ell}_{ks} > 0$, if ζ was chosen sufficiently small,

$$\begin{aligned}\frac{\dot{\ell}_{kt}}{\ell_{kt}} &= \frac{\rho+\delta}{1-\varphi} - \frac{\varphi}{1-\varphi}\frac{\pi_{\mathbb{K}t}A(\epsilon-1)}{Kq_{kt}}\ell_{kt}^{\varphi-1} \\ &> \frac{\rho+\delta}{1-\varphi} - \frac{\varphi}{1-\varphi}\frac{(\pi_{\mathbb{K}t}+\zeta)A(\epsilon-1)}{K(q_{SS}+\alpha)}\ell_{SS}^{\varphi-1} \\ &> 0.\end{aligned}$$

   This implies that $\ell_{kt} \to \infty$.

2. In the second case, $q_{kt} \leq q_{SS} - \alpha$. Analogous logic to case 1 shows that $\ell_{kt} \to 0$.

□

Lemma 5. *There exists a unique solution to the boundary-value problem* (33).

Proof. I first prove uniqueness. Suppose for contradiction that there are two solutions ℓ_{kt} and $\hat{\ell}_{kt}$ which respectively have initial conditions ℓ_{k0} and $\hat{\ell}_{k0}$ and suppose that $\hat{\ell}_{k0} > \ell_{k0}$. In the last part of the proof of Lemma 3 we showed that

$$\ln(\hat{\ell}_{kt}) - \ln(\ell_{kt}) \geq \ln(\hat{\ell}_{k0}) - \ln(\ell_{k0})$$

for all t. But then it can not be the case that both satisfy the boundary condition, a contradiction.

I now prove existence. To do this, consider a version of (33) where the boundary at ∞ is instead a boundary at some finite $T > 0$. If a solution exists to this boundary problem with boundary at T then it will be unique by an analogous argument to the one we just gave for boundary at ∞. It is easy to see from (36) that by increasing ℓ_{k0} we can get ℓ_{kT} as high as we would like and by decreasing ℓ_{k0} we can get ℓ_{kT} as close to zero as we would like—the RHS of (36) diverges as ℓ_{kt} diverges and becomes negative for all ℓ_{kt} in a neighborhood of zero.

Thus, by continuity (established in Lemma 3), there must be some initial condition for which $\ell_{kT} = \ell_{SS}$ which I denote by $\ell_{k0}(T)$. I.e., the solution to the initial value problem with this initial condition is the solution to the boundary value problem with boundary at T.

Now consider a sequence of times $\{T_n\}$ such that $\lim_{n\to\infty} T_n = \infty$. Consider the corresponding sequence of solutions of the boundary value problem with each solution extended out to infinity. By Lemma 4 we can categorize these solutions based on their tail behavior. Namely, there must be an infinite number of solutions that diverge or an infinite number that vanish.

Suppose that the former is true. The other case can be dealt with analogously. Consider ℓ_{k0}^* defined as the infimum of the set of initial conditions of the diverging solutions: $\ell_{k0}^* = \inf\{\ell_{k0}(T_n), n \in 1, 2, \ldots\}$. Let the solution for investment with this initial condition be denoted by ℓ_{kt}^*. We will argue that ℓ_{kt}^* solves the boundary value problem. The proof proceeds by contradiction. There are two cases to consider.

1. Suppose for contradiction that ℓ_{kt}^* diverges. Let T^* denote the last time that $\ell_{kt}^* = \ell_{SS}$: $T^* = \sup\{t|\ell_{kt}^* = \ell_{SS}\}$. As long as the set $\{t|\ell_{kt}^* = \ell_{SS}\}$ is nonempty, T^* must be finite. Suppose for contradiction $\{t|\ell_{kt}^* = \ell_{SS}\}$ is empty. Then any solution with initial condition $\ell_{k0} \geq \ell_{k0}^*$ likewise never hits ℓ_{SS} by monotonicity (Lemma 3) a contradiction of the definition of ℓ_{k0}^*.

   Then ℓ_{kt}^* is the solution to the boundary value problem with boundary at T^*. Next consider $T_n > T^*$ where n is chosen such that the corresponding solution to boundary value problem with boundary at T_n diverges. Then it must be that $\ell_{k0}(T_n) < \ell_{k0}^*$ since solutions are monotone in initial conditions and this solution hits at a later time than T^*. But then we have a contradiction of the definition of ℓ_{k0}^*.

2. Now suppose for contradiction that ℓ_{kt}^* eventually vanishes. By inspecting (36) I see that there exists $\underline{\ell} > 0$ and $\underline{q} > 0$ such that if at any point in time $\ell_{kt} < \underline{\ell}$ and $q_{kt} < \underline{q}$ then ℓ_{kt} must vanish.

   Let T^* now be defined as the first time that $\ell_{kt}^* < \underline{\ell} - \alpha$ and $q_{kt}^* < \underline{q} - \alpha$ for $\alpha > 0$. But then if we perturb ℓ_{k0}^* up by an arbitrarily small amount, the solution for this perturbed initial condition at T^* must be such that ℓ_{kT^*} and q_{kT^*} must move up by at least α. This is so since the solution must diverge by the definition of ℓ_{k0}^*. This contradicts the continuity of the solutions in initial conditions established in Lemma 3.

Therefore, it follows that ℓ_{kt}^* must converge to steady state and thus solves the boundary-value problem. $\square$

C Supplemental Material for Section 5

Appendix C.1 proves Proposition 2. Appendix C.2 proves Proposition 3. Appendix C.3 presents additional comparative statics omitted from the main text.

C.1 Proof of Proposition 2

Proof of Proposition 2. The proof proceeds via two Lemmas.

Lemma 6 proves that product consumption is monotone in data informativeness.

Lemma 6. *An increase in G in the mean-preserving spread order leads to an increase in the steady state value of $M\mu_H$ and product consumption C and a decrease in the cdf H^c in second-order stochastic dominance.*

Proof. Suppose that G increases in the mean-preserving spread order to $\hat{G}$. Let $\hat{H}^c$ denote the steady state cdf under $\hat{G}$. Define $\gamma : [0,\infty) \to [-1,1]$ and $\nu : [0,\infty) \to [-1,1]$ such that

$$\hat{G}(y) = G(y) + \gamma(y) \text{ and } \hat{H}^c(y) = H^c(y) + \nu(y)$$

for all $y \in [0,\infty)$. Then by Part 3 of Theorem 1, it follows that

$$J\left[G(y) + \gamma(y)\right] = M\left[H^c(y) + \nu(y)\right]^N + (J - M)\left[H^c(y) + \nu(y)\right]$$

and

$$JG(y) = MH^c(y)^N + (J - M)H^c(y).$$

Substracting the bottom equation from the top equation gives

$$\gamma(y) = \nu(y)\left(\frac{J - M}{J} + \frac{M}{J}[H^c(y) + \nu(y)]^{N-1}\right).$$

Integrating both sides from 0 to $s \in [0,\infty)$ we derive

$$\int_0^s \nu(y)\,\mathrm{d}y\left[\frac{J - M}{J} + \frac{M}{J}\hat{H}^c(s)^{N-1}\right] - \frac{M}{J}\int_0^s\int_0^y \nu(l)\,\mathrm{d}y\,\mathrm{d}\hat{H}^c(y)^{N-1} \geq 0. \tag{38}$$

Above, I have used integration by parts and the fact that $\hat{G}$ is a mean-preserving spread of G implies that $\int_0^s \gamma(y)\,\mathrm{d}y \geq 0$ for each $s \in [0,\infty)$. I now argue that $\int_0^s \nu(y)\,\mathrm{d}y \geq 0$ for all $s \in [0,\infty)$ with strict inequality at some point $s \in [0,\infty)$. This implies both that H^c dominates $\hat{H}^c$ in second-order stochastic dominance and so $\mu_{\hat{H}} > \mu_H$. Suppose for contradiction that there exists a point $s \in [0, \infty)$ such that $\int_0^s \nu(y)\,\mathrm{d}y < 0$. Let

$$l^* = \inf \left\{ l | \int_0^l \nu(y)\,\mathrm{d}y < 0, l > 0 \right\}.$$

If $l^* > 0$, then (38) is violated at l^* which is a contradiction. Then it must be that $l^* = 0$. But by inspecting (38), we see that $\int_0^s \nu(y)\,\mathrm{d}y$ must be increasing in s when it fist departs from 0 as otherwise (38) is violated for s close to the point of departure. Thus $l^* \neq 0$, a contradiction. It follows that $\int_0^s \nu(y)\,\mathrm{d}y \leq 0$ for each $s \in [0, \infty)$. Strict inequality must occur at a some point since $\hat{G}$ is a mean-preserving spread of G.

□

Next, Lemma 2 shows that ad revenue is maximal in a limiting sense when data is uninformative. From Part 5 of Theorem 1, platform quality is then also maximal. To make the notation explicit, write $\pi_{\mathbb{K}}(G)$ for the average ad price as a function of the cdf G of expected values.

Lemma 7. *Let $\{G_n\}$ be a sequence of continuous cdfs converging pointwise to the Heaviside function centered at μ_F: $\lim_{n\to\infty} G_n(\hat{v}) = \mathbb{1}_{\{\hat{v}\geq\mu_F\}}$ for each $\hat{v} \in [0, \infty)$. Then $\lim_{n\to\infty} \pi_{\mathbb{K}}(G_n) = \sup_G \pi_{\mathbb{K}}(G)$ where the supremum is over all continuous cdfs G supported on $[0, \infty)$.*

Proof. Let G be arbitrary. We have

$$\begin{aligned}
\pi_{\mathbb{K}}(G) &= \mathbb{E}[B(\hat{v}^{(2)})] \\
&= \mathbb{E}\left[\pi_{\mathbb{J}} \int_0^{\hat{v}^{(2)}} \frac{1}{\rho + \lambda_f + \lambda_e(s)}\,\mathrm{d}s\right] \\
&\leq \frac{I\mathbb{E}[\hat{v}^{(2)}]}{\sigma M \mu_H(\rho + \lambda_f)} \\
&\leq \frac{I\mathbb{E}[\hat{v}^{(1)}]}{\sigma M \mu_H(\rho + \lambda_f)} \\
&= \frac{I}{\sigma M(\rho + \lambda_f)}
\end{aligned}$$

where above $\hat{v}^{(2)} \sim (H^c)^N + N\,(H^c)^{(N-1)}\,(1 - H^c)$ and $\hat{v}^{(1)} \sim (H^c)^N$. In the fourth line we use the fact that in steady state $H = (H^c)^N$. The notation has supressed the dependency of H^c and H on G.

Using (17) in Theorem 1 we have

$$\begin{aligned}
\lim_{n\to\infty} \pi_{\mathbb{K}}(G_n) &= \lim_{n\to\infty} \frac{I}{\sigma M \int_0^\infty [1-H(s,G_n)]\,\mathrm{d}s} \\
&\quad \times \int_0^\infty \frac{1 - NH^c(s,G_n)^{N-1} + (N-1)H^c(s,G_n)^N}{\rho + \lambda_f + \lambda_a H^c(s,G_n)^N}\,\mathrm{d}s \\
&= \frac{I}{\sigma M \mu_F} \int_0^\infty \frac{1 - N\mathbb{1}_{s\geq\mu_F} + (N-1)\mathbb{1}_{s\geq\mu_F}}{\rho + \lambda_f + \lambda_a \mathbb{1}_{s\geq\mu_F}}\,\mathrm{d}s \\
&= \frac{I}{\sigma M \mu_F} \int_0^{\mu_F} \frac{1}{\rho+\lambda_f}\,\mathrm{d}s \\
&= \frac{I}{\sigma M(\rho+\lambda_f)}
\end{aligned}$$

where in the second equality I have used the dominated convergence theorem to pass the limit through the integral. Above I have made explicit the dependency of H^c on G_n in the notation. □

The proof of Proposition 2 is complete. □

C.2 Proof of Proposition 3

Proof of Proposition 3. The proof proceeds via a series of lemmas.

Lemma 8. *An increase in ϵ leads to a decrease in the ad display rate A.*

Proof. By Theorem 1, A is equal to

$$\arg\max_a a\nu(a)^{\epsilon-1} = \arg\max \ln(a) + (\epsilon-1)\ln(\nu(a)). \tag{39}$$

Pick $\epsilon_2 > \epsilon_1$, and let a_1 solve (39) when $\epsilon = \epsilon_1$ and define a_2 analogously . By unique optimality,

$$\ln a_1 + (\epsilon_1 - 1)\ln\nu(a_1) > \ln a_2 + (\epsilon_1 - 1)\ln\nu(a_2)$$

and

$$\ln a_2 + (\epsilon_2 - 1)\ln\nu(a_2) > \ln a_1 + (\epsilon_2 - 1)\ln\nu(a_1).$$

Add these two equations and cancel terms:

$$(\epsilon_2 - \epsilon_1)\big(\ln\nu(a_2) - \ln\nu(a_1)\big) > 0.$$

Since $\epsilon_2 - \epsilon_1 > 0$ and $\ln\nu(\cdot)$ is decreasing, this implies $a_2 < a_1$. □

Combining Lemma 8 and the following shows that an increase in ϵ leads to a decrease in product consumption.

Lemma 9. *An increase in A leads to a decrease in H_t and H_t^c in first-order stochastic dominance and an increase in $M_t\mu_{H_t}$ and C_{it} at all t.*

Proof. Recall that in Step 4 of Section 4 we saw that

$$\begin{aligned}\dot{H}_t^c(\cdot) &= \lambda_f \frac{[JG - MH_t^c(\cdot)^N - (J-M)H_t^c(\cdot)]}{J - M_t} \\ &= \lambda_f \frac{J\,[G - H_t^c(\cdot)] + M\left[H_t^c(\cdot) - H_t^c(\cdot)^N\right]}{J - M_t}.\end{aligned}$$

When A goes up, both M and M_t increase as seen from (29) and so $\dot{H}_t^c$ is higher holding fixed the value of H_t^c at time t.

By a standard comparison argument for differential equations, H_t^c must increase pointwise when A increases. That is, H_t^c decreases in the sense of first-order stochastic dominance. This in turn implies that $M_t\mu_{H_t} = K\mu_G - (J - M_t)\mu_{H_t^c}$ must increase. By Part 7 of Theorem 1, $C_{it} = I(M_t\mu_{H_t})^{\frac{1}{\sigma-1}}$ so it also must increase.

To show that H_t decreases in first-order stochastic dominance recall that

$$\begin{aligned}&(\dot{M_tH_t}) = A(H_t^c)^N - \lambda_f M_t H_t \\ &\Rightarrow (A - \lambda_f M_t)H_t + \dot{H}_t M_t = A(H_t^c)^N - \lambda_f M_t H_t \\ &\Rightarrow \dot{H}_t = \frac{A}{M_t}[(H_t^c)^N - H_t] \\ &\Rightarrow \dot{H}_t = \frac{A}{\frac{A}{\lambda_f} - \left(\frac{A}{\lambda_f} - M_0\right)e^{-\lambda_f t}}[(H_t^c)^N - H_t].\end{aligned}$$

In the last line, we see that an increase in A leads to an increase in $\dot{H}_t$ holding fixed the value of H_t. Again, using standard comparison arguments for differential equations, it is easy to see that this implies H_t^c must increase pointwise when A increases. $\square$

Combining Lemma 8 and the following lemma shows that an increase in ϵ leads to an increase in ad revenue $\pi_{\mathbb{K}}A$.

Lemma 10. *An increase in A leads to a decrease in steady state ad revenue $\pi_{\mathbb{K}}A$.*

Proof. From Part 4 of Theorem 1, we have

$$\pi_{\mathbb{K}} A = \frac{\lambda_f}{\sigma \int_0^\infty 1 - H^c(s)^N \, \mathrm{d}s} \int_0^\infty \frac{1 - NH^c(s)^{N-1} + (N-1)H^c(s)^N}{\rho + \lambda_f + \lambda_e(s)} \, \mathrm{d}s$$

where $\lambda_e(s) = \lambda_a H^c(s)^{N-1}$ for each $s \in [0, \infty)$.

To prove that ad revenue is decreasing, it suffices to prove the ratio of the integrand in the numerator to the integrand in the denominator is decreasing at each point s.

By Lemma 9, an increase in A leads to a decrease in H^c in first-order stochastic dominance. Thus, since λ_e increases pointwise, it suffices to show that

$$\frac{1 - NH^c(s)^{N-1} + (N-1)H^c(s)^N}{1 - H^c(s)^N} = N\frac{1 - H^c(s)^{N-1}}{1 - H^c(s)} - (N-1)$$

is decreasing in $H^c(s)$ which can be done simply by computing the derivative. I omit this step. □

This completes the proof of the proposition. □

C.3 Additional Comparative Statics Omitted From the Main Text

Lemma 11. *An increase in J leads to an increase in H_t^c and H_t in first-order stochastic dominance and thus an increase in $M_t \mu_{H_t}$ and C_{it} at all t.*

Proof. From the proof of Lemma 9, we have

$$\dot{H}_t^c(\cdot) = \lambda_f \frac{K\left[G - H_t^c(\cdot)\right] + M\left[H_t^c(\cdot) - H_t^c(\cdot)^N\right]}{J - M_t}.$$

Holding fixed H_t^c, the right-hand side is decreasing in F. By standard comparison arguments for differential equations it follows that H_t^c must decrease pointwise and thus increase in the sense of first-order stochastic dominance.

Also, from Lemma 9, we have

$$\dot{H}_t(\cdot) = \frac{A}{M_t}[H_t^c(\cdot)^N - H_t(\cdot)].$$

Since M_t is unaffected and H_t^c is lower pointwise when F increases, H_t must also be lower pointwise. Thus H_t increases in the sense of first-order stochastic dominance. It follows immediately that $M_t \mu_{H_t}$ increases. By Part 8 of Theorem 1 $C_{it} = I(M_t \mu_{H_t})^{\frac{1}{\sigma - 1}}$ so it must increase as well. □

Lemma 12. *An increase in J leads to an increase in steady state ad revenue $\pi_{\mathbb{K}}A$.*

Proof. By Lemma 11, an increase in J leads to an increase in H^c in first-order stochastic dominance and then following the same steps as in Lemma 10, we see that this leads to an increase in steady state ad revenue $\pi_{\mathbb{K}}A$. □

Lemma 13. *An increase in ϵ leads to a decrease in C_{it} at each point in time and an increase in steady state ad revenue $\pi_{\mathbb{K}}A$ and if $K \leq 1$, steady state platform consumption X.*

Proof. An increase in ϵ leads to a decrease in A since (16) is submodular in ϵ and a_{kt}. By Lemma 10 this leads to an increase in steady state ad revenue $\pi_{\mathbb{K}}A$ and in turn platform investment (18) and thus platform quality. Since $X = K^{\frac{1}{\epsilon-1}}\nu(A)q_t$ and q_t increased while A decreased, if $K \leq 1$ then X must increase. □

Lemma 14. *An increase in K has no effects on ad revenue $\pi_{\mathbb{K}t}A$ at any time t and leads to an increase in steady state platform consumption X.*

Proof. As seen from (31), the equilibrium ad revenue $\pi_{\mathbb{K}t}A$ does not depend on K. In steady state, using (18)

$$X = K^{\frac{1}{\epsilon-1}}\frac{\ell_{\mathbb{K}}^{\varphi}}{\delta} = \frac{1}{\delta}K^{\frac{1}{\epsilon-1}-\varphi}\left(\frac{\varphi\delta\pi_{\mathbb{K}}A(\epsilon-1)}{(\delta+\rho)}\right)^{\varphi}.$$

This is increasing in K if the coefficient $\epsilon - 1 \leq 1/\varphi$. This is a necessary condition for an equilibrium to exist as shown in Lemma 2. □

D Supplemental Material for Section 6

Appendix D.1 presents proofs of the results in Section 6. Appendix D.2 presents general equilibrium analogs of Propositions 2 and 3.

D.1 Proofs of main results

Proof of Proposition 4. Product market clearing implies that $I_t = \frac{\sigma}{\sigma-1}(L - K\ell_{\mathbb{K}t})$. That is, income is equal to the markup times the quantity of labor left over for production.

To compute (19), let $\hat{\pi}_{\mathbb{K}} = \pi_{\mathbb{K}}/I$ be the average ad price per unit of income. From (17),

$$\hat{\pi}_{\mathbb{K}} = \frac{1}{\sigma M\mu_H}\int_0^{\infty}\frac{1 - NH^c(s)^{N-1} + (N-1)H^c(s)^N}{\rho+\lambda_f+\lambda_e(s)}\,\mathrm{d}s. \tag{40}$$

Then, rewriting (18), we have

$$\ell_{\mathbb{K}} = \frac{\varphi\delta\hat{\pi}_{\mathbb{K}}A(\epsilon-1)}{K(\rho+\delta)}I. \tag{41}$$

Given investment,

$$I = \frac{\sigma}{\sigma-1}(L-K\ell_{\mathbb{K}}).$$

Solving this linear system for $\ell_{\mathbb{K}}$ yields (19). The rest of the proof is analogous to that of Theorem 1. □

Proof of Theorem 2. To characterize the planner's steady state investment, it suffices to take as given steady-state ad display A^* and suppose that we have already reached steady state for H_t. The Hamiltonian for the planner's problem for investment is

$$\begin{aligned}\mathcal{H}(t,q_t,\lambda_t,\ell_{\mathbb{K}t}) = \left[(L-K\ell_{\mathbb{K}t})\left(M\mu_H\right)^{\frac{1}{\sigma-1}}\right]^{1-\tau}\left[K^{\frac{1}{\epsilon-1}}\nu(A^*)q_t\right]^{\tau} \\ + \lambda_t\left(\ell_{\mathbb{K}t}^{\varphi}-\delta q_t\right)\end{aligned}$$

where λ_t satisfies

$$\rho\lambda_t - \dot{\lambda}_t = \left[(L-K\ell_{\mathbb{K}t})\left(M\mu_H\right)^{\frac{1}{\sigma-1}}\right]^{1-\tau}\tau K^{\frac{\tau}{\epsilon-1}}\nu(A^*)^{\tau}q_t^{\tau-1} - \delta\lambda_t.$$

Maximizing the Hamiltonian with respect to the control yields

$$\lambda_t\varphi\ell_{\mathbb{K}t}^{\varphi-1} = (1-\tau)K[L-K\ell_{\mathbb{K}t}]^{-\tau}\left[\left(M\mu_H\right)^{\frac{1}{\sigma-1}}\right]^{1-\tau}\left[K^{\frac{1}{\epsilon-1}}\nu(A^*)q_t\right]^{\tau}.$$

In steady state, λ_t must therefore be a constant. We have

$$\lambda_t = \frac{\left[(L-K\ell_{\mathbb{K}t})\left(M\mu_H\right)^{\frac{1}{\sigma-1}}\right]^{1-\tau}\tau K^{\frac{\tau}{\epsilon-1}}\nu(A^*)^{\tau}q_t^{\tau-1}}{\delta+\rho}.$$

Substituting ths into the previous equation, we have

$$\tau\frac{(L-K\ell_{\mathbb{K}t})\,\varphi\ell_{\mathbb{K}t}^{\varphi-1}}{\delta+\rho} = (1-\tau)Kq_t.$$

In steady state, $q_t = \ell_{\mathbb{K}t}^{\varphi}/\delta$ so then

$$\tau\frac{\varphi\,(L-K\ell_{\mathbb{K}t})}{\delta+r} = (1-\tau)K\frac{\ell_{\mathbb{K}t}}{\delta}.$$

Rearranging gives

$$\ell_{\mathbb{K}t} = \frac{\varphi\delta\frac{\tau}{1-\tau}}{\rho+\delta+\varphi\delta\frac{\tau}{1-\tau}}\frac{L}{K}.$$

The Hamiltonian is concave in both the state and the control and therefore satisfies the Mangasarian sufficient conditions for an optimal control.

In the limit as $\rho \to 0$, the steady state ad rate chosen by the planner must maximize the flow utility of consumers which amounts to (23). □

Proofs of Proposition 5 and Corollary 2.2. These results follow immediately from Proposition 4 and Theorem 2. □

D.2 Comparative Statics Omitted from the Main Text

Proposition 6. *The conclusions of Proposition 2 extend to the general equilibrium setting of Section 6.*

Proof. I first show that ad revenue is maximal in the limiting sense as data becomes uninformative. By an analogous argument to Lemma 2, $\hat{\pi}_{\mathbb{K}}$ is maximal as data becomes uninformative. Total ad revenue is $\pi_{\mathbb{K}}A = I\hat{\pi}_{\mathbb{K}}A$. Suppose for contradiction that ad revenue is not maximal in this limit; then under some informative data, I must be high enough that ad revenue rises. From the proof of Proposition 4, $I = \frac{\sigma}{\sigma-1}(L-K\ell_{\mathbb{K}})$, so a higher I requires lower investment $\ell_{\mathbb{K}}$. But (41) implies that lower investment requires lower ad revenue—a contradiction.

For product consumption, recall $C = I(M\mu_H)^{1/(\sigma-1)}$. As data becomes uninformative, μ_H is minimal (Lemma 6). Ad revenue is maximal in this limit, so by (41) investment is maximal, which makes $I = \frac{\sigma}{\sigma-1}(L-K\ell_{\mathbb{K}})$ minimal. All three factors push C to its infimum.

Platform quality is maximal in the limit by Part 5 of Theorem 1, since investment is maximal. □

Proposition 7. *The conclusions of Proposition 3 on platform quality and product consumption extend to the general equilibrium setting, with $\hat{\pi}_{\mathbb{K}}A$ replacing $\pi_{\mathbb{K}}A$.*

Proof. By an analogous argument to that in the proof of Proposition 3, $\hat{\pi}_{\mathbb{K}}A$ rises with ϵ.

I next show that platform quality rises with ϵ. Suppose for contradiction that platform quality (and hence investment $\ell_{\mathbb{K}}$) falls when ϵ rises. By (41), lower investment implies lower ad revenue. Total ad revenue is $\pi_{\mathbb{K}}A = I\hat{\pi}_{\mathbb{K}}A$, so since $\hat{\pi}_{\mathbb{K}} A$ has risen, lower ad revenue requires lower I. But $I = \frac{\sigma}{\sigma-1}(L - K\ell_{\mathbb{K}})$, so lower I requires higher $\ell_{\mathbb{K}}$—a contradiction.

For product consumption, recall $C = I(M\mu_H)^{1/(\sigma-1)}$. In general equilibrium, ϵ affects C only through I, since M and μ_H do not depend on ϵ. Since platform quality rises, investment $\ell_{\mathbb{K}}$ rises as well, so $I = \frac{\sigma}{\sigma-1}(L - K\ell_{\mathbb{K}})$ falls. Therefore C falls. □

E Supplemental Material for Section 7

This appendix presents supplemental material for the duopoly model of Section 7. Appendix E.1 provides details of the algorithm to compute an equilibrium. Appendix E.2 describes a reformulation in which there is a single state variable and uses it to reveal the connection between my model and that of Budd et al. (1993).

E.1 Details of the Algorithm to Compute Equilibrium

I use an implicit upwind finite difference scheme to solve (26). Here, I provide details of this finite-difference scheme. In particular, I describe the iterative procedure (used in Steps 1 and 2 outlined in Section 7) to calculate the value function at an earlier time step given the value function at a later time step. At the end of this subsection, I discuss properties of the algorithm and mathematical foundations.

Discretize the State Space I first compactify and then discretize each dimension of the state space in a standard way using evenly spaced grids. The two quality dimensions share a common grid with gap Δ_q, and the time dimension uses gap Δ_t between grid points.

Finite Difference Scheme Throughout, I supress the index k on the value function V_k for ease of notation. In the remaining parts of this document, let

$$V_{q_1}^{+}(q_1, q_2, t) \equiv \frac{V(q_1 + \Delta_q, q_2, t) - V(q_1, q_2, t)}{\Delta_q}$$

denote the forward difference and

$$V_{q_1}^{-}(q_1, q_2, t) \equiv \frac{V(q_1, q_2, t) - V(q_1 - \Delta_q, q_2, t)}{\Delta_q}$$

the backward difference approximation of $\partial V/\partial q_1$. Define V_t^+ analogously as the forward difference approximation of $\partial V/\partial t$.

Let

$$V^c_{q_1,q_1}(q_1,q_2,t) = \frac{V(q_1+\Delta_q,q_2,t) - 2V(q_1,q_2,t) + V(q_1-\Delta_q,q_2,t)}{\Delta_q^2}$$

denote the central difference approximation of $\partial^2 V/\partial q_1^2$. Define $V^c_{q_2,q_2}$ analogously.

To implement the (semi-)implicit upwind scheme, I work with the following discretized version of (26):

$$\begin{aligned}\rho V(q_1,q_2,t) = \sup_{\ell_{k1}} & V_t^+(q_1,q_2,t) + V_{q_1}^+(q_1,q_2,t)(\ell_{k1}^{\varphi} - \delta q_1)^+ \\ & + V_{q_1}^-(q_1,q_2,t)(\ell_{k1}^{\varphi} - \delta q_1)^- + V_{q_2}^+(q_1,q_2,t)\,(\ell_{k2}^{\varphi} - \delta q_2)^+ \\ & + V_{q_2}^-(q_1,q_2,t)\,(\ell_{k2}^{\varphi} - \delta q_2)^- - \ell_{k1} + \pi_{Kt} A \frac{q_1^{\epsilon-1}}{q_1^{\epsilon-1}+q_2^{\epsilon-1}}. \\ & + \frac{1}{2}\nu^2 q_1^2 \tilde{V}^c_{q_1,q_1}(q_1,q_2) + \frac{1}{2}\eta^2 q_2^2 \tilde{V}^c_{q_2,q_2}(q_1,q_2).\end{aligned} \tag{42}$$

Thus, the finite-difference scheme is *upwind* in the sense that it uses the forward difference approximation of $\partial V/\partial q_1$ ($\partial V/\partial q_2$) whenever the drift of q_1 (q_2) is positive and the backward difference whenever the drift is negative. The scheme is *implicit* in the sense that it uses the forward-difference approximation of $\partial V/\partial t$. Note that (42) is a relationship between $V(\cdot, t+\Delta)$ and $V(\cdot,t)$ and thus we can use it to calculate $V(\cdot,t)$ given $V(\cdot,t+\Delta)$. The rest of this subsection details how this is done. I note that because the scheme is both upwind and implicit, $V(\cdot,t)$ depends *monotonically* on $V(\cdot,t+\Delta)$, which is essential for the algorithm's *stability*. Monotonicity and stability are critical for convergence of the numerical scheme to the solution (Barles and Souganidis, 1991).

Optimal Controls Using (42) to compute the value functions backward requires solving the optimization problem on the right-hand side. Using first-order conditions, we can show that there are only two candidates for the optimal controls: $\max\{\ell_k^+, \ell_k^c\}$ and $\min\{\ell_k^-, \ell_k^c\}$, where ℓ_k^+, ℓ_k^-, and ℓ_k^c are defined as follows.

Define ℓ_k^+ by

$$V_{q_1}^+(q_1,q_2,t)\varphi(\ell_k^+)^{\varphi-1} = 1,$$

ℓ_{k1}^- by

$$V_{q_1}^-(q_1,q_2,t)\varphi(\ell_k^-)^{\varphi-1} = 1,$$

and ℓ_k^c by

$$(\ell_k^c)^\varphi = \delta q_1.$$

Construct the M matrix Consider the following term which appears on the right-hand side of (26):

$$\begin{aligned} V_{q_1}^+(q_1,q_2,t)(\ell_{k1}^\varphi - \delta q_1)^+ + V_{q_1}^-(q_1,q_2,t)(\ell_{k1}^\varphi - \delta q_1)^- \\ + V_{q_2}^+(q_1,q_2,t)\,(\ell_{k2}^\varphi - \delta q_2)^+ + V_{q_2}^-(q_1,q_2,t)\,(\ell_{k2}^\varphi - \delta q_2)^- \\ + \frac{1}{2}\eta^2 q_1^2 V_{q_1,q_1}(q_1,q_2) + \frac{1}{2}\eta^2 q_2^2 V_{q_2,q_2}(q_1,q_2). \end{aligned} \tag{43}$$

I express the operator above in terms of a matrix M. Specifically, let N denote the number grid points for both quality dimensions. The expression (26) is to be expressed as MV where V is a $N^2 \times 1$ vector representing the value function at each quality (q_1, q_2) grid point and M is a $N^2 \times N^2$ matrix.

The rows of V are partitioned into N groups of N rows. Within each group, the state q_1 is the same and only q_2 varies.

More specifically, if[39]

$$V(q_1,q_2,t) \mathrel{\widehat{=}} k$$

where k is the index of the vector V, then the corresponding indices of the other value function terms in (43) are:

$$\begin{aligned} V(q_1,q_2+\Delta_q,t) &\mathrel{\widehat{=}} k+1, \\ V(q_1,q_2-\Delta_q,t) &\mathrel{\widehat{=}} k-1, \\ V(q_1+\Delta_q,q_2,t) &\mathrel{\widehat{=}} k+N, \\ V(q_1-\Delta_q,q_2,t) &\mathrel{\widehat{=}} k-N. \end{aligned}$$

Given the vector V, the rows and columns of M are organized accordingly. Namely, M is populated along the main diagonal (which contains the coefficients in (43) on $V(q_1,q_2,t)$), the +1 and -1 diagonals (which contain the coefficients on $V(q_1,q_2+\Delta_q,t)$ and $V(q_1,q_2-\Delta_q,t)$ respectively), and the $+N$ and $-N$ diagonals (contain the coefficients on $V(q_1+\Delta_q,q_2,t)$ and $V(q_1-\Delta_q,q_2,t)$) respectively. All other entries in M are zero. Thus, M is a *sparse matrix* with only 5 populated diagonals. This sparsity is *critical* for computational speed and a key advantage of the continuous-time formulation. I take advantage of it in the MATLAB code using the `spdiags` function.

[39]Recall that $\widehat{=}$ means "corresponds."

Calculate Value Functions at Previous Time Step Using the M matrix (constructed using the *optimal controls at time t* in (43)) together with (42), we can succintly express $V(\cdot, t-\Delta_t)$ in terms of $V(\cdot, t)$ via the equation

$$V(\cdot, t-\Delta_t) = \left[I\left(1+\rho\Delta_t\right) - \Delta_t M_t\right]^{-1}\left(\Delta_t g_t + V(\cdot, t)\right)$$

where I is the identity matrix and $g_t = -\ell^*_{kt} + \pi_{Kt} A \frac{q_1^{\epsilon-1}}{q_1^{\epsilon-1}+q_2^{\epsilon-1}}$ is the flow profit from the optimal control ℓ^*_{kt}. For clarity, I have explicitly indexed the M matrix and optimal control by time t.

Discussion of the Algorithm Though I have not proven convergence theoretically, the algorithm is informed by established mathematical foundations. The upwind finite-difference scheme used here falls within the class of monotone, stable, and consistent schemes that Barles and Souganidis (1991) prove converge to viscosity solutions of elliptic PDEs under suitable conditions. While their result does not directly apply to *coupled* HJB systems like the one studied here, the structural properties of the scheme offer reassurance about the reliability of the numerical method. The computation is also fast—even though there are two state variables, with minimal optimization and no parallelization, on a 2021 base model 16 inch M1 Macbook pro, the MATLAB code calculates steady state in roughly 2 minutes and the full equilibrium dynamics in roughly 10 minutes for reasonably fine time and quality grids.

Furthermore, the approach taken here has precedent in economics. For example, Brunnermeier and Sannikov (2017) use an analogous procedure to solve a coupled HJB system in a macrofinance model. Brunnermeier and Sannikov (2017) consider a setting with a single state variable and atomistic agents, whereas I study a model with two state variables and atomic players. However, in both cases, computing equilibrium involves solving a coupled HJB system and the numerical algorithms we use are conceptually similar.

E.2 Model Variant with Single Spatial State Variable

Here I describe a slight reformulation of the duopoly model that reduces the number of state variables: instead of tracking the quality levels of both platforms separately, it suffices for each platform to track only the share of attention received by platform 1. In the process, I remark on the close connection between the duopoly model and that of Budd et al. (1993). As discussed in Section 7, this connection (together with the comment in Footnote 32) motivates the conjecture that a nontrivial amount of noise is sufficient for equilibrium existence in my setting. This is because Budd et al. 1993 show in their similar model that any nontrivial amount of noise ensures equilibrium existence (provided some standard and permissive conditions on other parameters are met).

Suppose that the effect of investment on quality scaled with quality. That is, suppose instead of (25) we have

$$\mathrm{d}q_{kt} = \left(\ell_{kt}^{\varphi} - \delta\right) q_{kt}\,\mathrm{d}t + \eta q_{kt}\,\mathrm{d}B_{kt}.$$

Then one can show that the share of attention x_{kt} that platform k receives is a Markov process. Namely, using Itô's formula,

$$\begin{aligned}\mathrm{d}x_{1t} = x_{1t}(1-x_{1t})&\left[(\epsilon-1)(\ell_{1t}^{\varphi}-\ell_{2t}^{\varphi}) + \frac{1}{2}(\epsilon-1)^2\eta^2(1-2x_{1t})\right]\mathrm{d}t\\ &+ (\epsilon-1)\eta x_{1t}(1-x_{1t})\left(\mathrm{d}B_{1t}-\mathrm{d}B_{2t}\right).\end{aligned}$$

where recall that $x_{1t} = q_{1t}^{\epsilon-1}/(q_{1t}^{\epsilon-1}+q_{2t}^{\epsilon-1})$.

Therefore we can express

$$\begin{aligned}\mathrm{d}x_{1t} = x_{1t}(1-x_{1t})&\left[(\epsilon-1)(\ell_{1t}^{\varphi}-\ell_{2t}^{\varphi}) + \frac{1}{2}(\epsilon-1)^2\eta^2(1-2x_{1t})\right]\mathrm{d}t\\ &+ (\epsilon-1)\eta x_{1t}(1-x_{1t})\sqrt{2}\,\mathrm{d}B_t.\end{aligned} \tag{44}$$

where $B \equiv \left(1/\sqrt{2}\right)(B_1 - B_2)$ is a standard Brownian motion.

Thus, in this setup it is logical to look for a steady state equilibrium that is Markov with only a single state variable: x_{1t}. That is, the two quality state variables collapse onto a single state variable.

The model in Budd et al. (1993) consists of a duopoly setting where profits depend on market share x_{kt} as in my model through some exogenously given function $\pi(\cdot)$. They do not focus on a specific microfoundation for market share and profits in terms of quality as I do but rather assume x_{kt} is directly controlled by the investment of the two players via the law of motion

$$\mathrm{d}x_{kt} = \left(\ell_{kt} - \ell_{-kt}\right)\mathrm{d}t + \sigma\,\mathrm{d}B_t$$

where σ is a positive constant in the interior of the state space $[0,1]$ and zero at the boundaries. Budd et al. (1993) also assume that there is some given flow cost $c(\cdot)$ of investment for both players.

Thus our models are qualitatively similar except for the fact that in (44) there is an extra factor $x_{1t}(1-x_{1t})$ that multiplies the drift and volatility terms and there is also the second term in brackets in the drift term.[40]

[40] Other differences are merely comestic such as the fact that investment is raised to the power φ in my model but not in Budd et al. (1993)—these differences disappear by appropriately relabeling variables.

Note that at the boundaries of the state space, the volatility disappears in both of our models.

F Supplemental Material for Section 8

I present a characterization of the steady state equilibrium and a sketch of its derivation.

F.1 Steady State Equilibrium Characterization

The following Theorem 3 summarizes the steady state equilibrium properties.

Theorem 3. *Suppose that $\epsilon - 1 < 1/\varphi$ and that A is the unique solution of $\max_a a\nu(a)^{\epsilon-1}$. In any steady state equilibrium with increasing bidding functions the following hold:*

1. *Consumer i's demands for products are as in* (27) *and her demands for platforms are as in* (15)*.*
2. *Firm j sets prices as in* (1)*.*
3. *Platform k displays ads at rate A as in* (16)*.*
4. *The size of consideration sets is $M = A/\lambda_f$.*
5. *Firm j's expected flow profits from sales are as in* (2)*.*
6. *Bidding functions $\boldsymbol{B} = (B_1, B_2)$ for the two groups are the fixed point of the operator $\boldsymbol{\Lambda}$ defined by* (48) *which is a contraction map whenever values are bounded.*
7. *The attention shares received by the two groups are given by* (50)*.*
8. *The rates of investment by the two groups are given by* (49)*.*

Proof. I now sketch the proof, i.e., the procedure to solve for the steady state equilibrium. Much of the analysis of the baseline model ports over to this extended setup. Demands, prices, ad rates, and the size of consideration sets will be as in equations (27), (15), (16), and (29). However, we now must keep track of the joint distribution of the two signals inside and outside of consideration sets.

Step 1. Law of Motion of H_t and H_t^c: Let H_t^c denote the joint cdf of signals outside of consideration sets at time t. Let H_t denote the joint cdf of signals inside of consideration sets at time t. Let h_t^c and h_t be their corresponding pdfs. Let $H_t^c(\zeta_1, \infty)$ denote $\lim_{\zeta \to \infty} H_t^c(\zeta_1, \zeta)$ and let $H_t^c(\infty, \zeta_2)$ be defined analogously. Suppose that the winner in an auction on a platform in group z is the firm with the highest group z signal. Then the law of motion of h_t must satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[M_t h_t(\boldsymbol{\zeta})\right] = A\left[x_{1t} N H_t^c(\zeta_1, \infty)^{N-1} h_t^c(\boldsymbol{\zeta}) + x_{2t} N H_t^c(\infty, \zeta_2)^{N-1} h_t^c(\boldsymbol{\zeta}) - h_t(\boldsymbol{\zeta})\right]. \quad (45)$$

In (45), with abuse of notation, x_{zt} denotes the total share of attention devoted to group z platforms. The first two terms in the brackets represents the inflow coming from the winners in the auctions on the two platform groups. The third term in the bracket represents the outflow as the consumer forgets about products.

To derive the steady state h, first fix an initial guess of x_1, the steady state level of x_{1t}. Then set $x_{1t} = x_1$, $x_{2t} = 1 - x_1$, $M_t = M$ and use the accounting identing $M_t h_t + (J - M_t) h_t^c = Fg$ to iterate (45) forward to convergence at each point $\boldsymbol{\zeta}$ in a fine grid on a region that contains almost all of G's mass.

Step 2. Calculate Bidding Strategies: Given M and h, we next compute equilibrium bidding strategies. To do so let

$$\mu_H = \int_{\mathbb{R}^2} \mathbb{E}[v_{ij} | \boldsymbol{\zeta}] h(\boldsymbol{\zeta}) \, \mathrm{d}\boldsymbol{\zeta},$$

$$\pi_{\mathbb{J}} = \frac{I}{\sigma M \mu_H},$$

$$O_1(\cdot) = H^c(\cdot, \infty)^{N-1},$$

and

$$O_2(\cdot) = H^c(\infty, \cdot)^{N-1}.$$

Above, μ_H is the average value of the firms in consideration sets, $\pi_{\mathbb{J}}$ is the coefficient of firms' flow profits, O_1 determines the probability that a firm wins an auction if it takes place on a platform in group 1, and O_2 determines the probability that a firm wins an auction if it takes place on a platform in group 2.

In a steady state, bidding strategies correspond to a pair of functions $\boldsymbol{B} = (B_1, B_2)$. Here, $B_z : \mathbb{R} \mapsto [0, \infty)$ maps firm j's group z signal ζ_{zij} to its bid $B_z(\zeta_{zij})$ in a group z auction for consumer i. To derive $\boldsymbol{B}$, let $\boldsymbol{\zeta}_{ij} = (\zeta_{1ij}, \zeta_{2ij})$ and let $V(\boldsymbol{\zeta}_{ij})$ be firm j's continuation value from selling to consumer i at the time of auction entry if it knows $\boldsymbol{\zeta}_{ij}$ but does not know which platform hosts the auction.

More precisely, V satisfies the recursive equation

$$\begin{aligned} V(\boldsymbol{\zeta}_{ij}) = \sum_{z=1}^{2} x_z O_z(\zeta_{zij}) \left[\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} \mathbb{E}[v_{ij} | \boldsymbol{\zeta}_{ij}] + \frac{\lambda_f}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \right] \\ - x_l \left[\int_{-\infty}^{\zeta_{zij}} B_z(s) \, \mathrm{d}O_z(s) + \left[1 - O_z(\zeta_{zij})\right] \frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \right] \end{aligned} \quad (46)$$

The first term in brackets is the discounted expected flow profit that firm j earns from entering Ω_{it}. It exits at rate λ_f and subsequently enters another auction at rate λ_a which corresponds to the second term. On the second line, the first term in brackets is the expected payment in a group z auction. The last term is the continuation value in the event that firm j loses the auction, weighted by the probability that this happens.

Since auctions are second-price, in each auction, firm j simply bids the gain in its continuation value from winning the auction. Then,

$$\begin{aligned} B_z(\zeta_{lij}) &= \mathbb{E} \left[\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} v_{ij} + \frac{\lambda_f}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \middle| \zeta_{zij}, j \in \Omega_{it}^c \right] \\ &\quad - \mathbb{E} \left[\frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \middle| \zeta_{zij}, j \in \Omega_{it}^c \right] \\ &= \mathbb{E} \left[\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} v_{ij} - \frac{\rho}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \middle| \zeta_{zij}, j \in \Omega_{it}^c \right]. \end{aligned} \quad (47)$$

Above, $\lambda_a = NA/(J - M)$ is the rate of auction entry. The expectation is conditional on only the group z signal and the fact that the firm is outside the consideration set since this is all that the firm knows when it bids.

Using (47) and (46), we can show that $\boldsymbol{B}$ is the fixed point of an operator $\boldsymbol{\Lambda} := (\Lambda_1, \Lambda_2)$ where $\Lambda_z : C^+(\mathbb{R})^2 \Rightarrow C^+(\mathbb{R})$ takes in a pair of functions $\boldsymbol{f} = (f_1, f_2)$ and outputs another function[41]

$$\Lambda_z(\boldsymbol{f})(\cdot) = \mathbb{E} \left[\frac{\pi_{\mathbb{J}} v_{ij} + \lambda_a \sum_{l=1}^{2} x_z \int_{-\infty}^{\zeta_{lij}} f_l(s) \, \mathrm{d}O_l(s)}{\lambda_f + \rho + \lambda_a \sum_{l=1}^{2} x_l O_l(\zeta_{zij})} \middle| \zeta_{zij} = \cdot, j \in \Omega_{it}^c \right]. \quad (48)$$

[41] $C^+(\mathbb{R})$ denotes the set of nonnegative continuous functions on $\mathbb{R}$.

The map $\mathbf{\Lambda}$ is a contraction with modulus $\lambda_a/(\lambda_a + \lambda_f + \rho)$ with respect to the sup-norm whenever values v_{ij} are bounded above by some level $\overline{v}$. For numerical purposes, this will always be the case. Even without this bounded support assumption, we see that $\mathbf{\Lambda}$ is increasing. Thus, starting from an initial $\boldsymbol{B}$ such that $B_z > \Lambda_z(\boldsymbol{B})$ for each $z \in \{1, 2\}$ it follows that $\{\mathbf{\Lambda}^n(\boldsymbol{B})\}_{n=1}^{\infty}$ is a decreasing sequence which converges to the fixed point. Thus we can compute the equilibrium bidding strategies by iterating on (48).

Step 3. Calculate Investment: In a steady state equilibrium, a platform in group $z \in \{1, 2\}$ invests a constant level $\ell_{\mathbb{K}z}$ to mantain quality level $q_z = \ell_{\mathbb{K}z}^{\varphi}/\delta$.

Let platform k belong to group z. The Hamiltonian for platform k's optimization problem is

$$\mathcal{H}(t, q_{kt}, \lambda_t, \ell_{kt}) = \pi_{\mathbb{K}z} A \frac{q_{kt}^{\epsilon-1}}{m_z q_z^{\epsilon-1} + m_{-z} q_{-z}^{\epsilon-1}} - \ell_{kt} + \lambda_t \left(\ell_{kt}^{\varphi} - \delta q_{kt}\right)$$

where λ_t, the costate variable, evolves according to

$$\rho\lambda_t - \dot{\lambda}_t = \pi_{\mathbb{K}z} A \left(\epsilon - 1\right) \frac{q_{kt}^{\epsilon-2}}{m_z q_z^{\epsilon-1} + m_{-z} q_{-z}^{\epsilon-1}} - \lambda_t \delta.$$

By the Maximum Principle, a necessary condition for optimality is that the control ℓ_{kt} maximizes the Hamiltonian along the optimal trajectory:

$$\lambda_t \varphi \ell_{kt}^{\varphi-1} = 1.$$

Under the conjectured stationary strategy then

$$\lambda_t \varphi \ell_{\mathbb{K}z}^{\varphi-1} = 1.$$

This implies that λ_t must be a constant λ. By the costate evolution equation,

$$\lambda = \frac{\pi_{\mathbb{K}z} A \left(\epsilon - 1\right)}{\rho + \delta} \frac{q_z^{\epsilon-2}}{m_z q_z^{\epsilon-1} + m_{-z} q_{-z}^{\epsilon-1}}.$$

Substituting, we have

$$\frac{\pi_{\mathbb{K}z} A \left(\epsilon - 1\right)}{\rho + \delta} \varphi \ell_{\mathbb{K}z}^{\varphi-1} = m_z q_z + m_{-z} \left(\frac{q_{-z}}{q_z}\right)^{\epsilon-2} q_{-z}.$$

This implies that

$$\frac{\pi_{\mathbb{K}z} A \left(\epsilon - 1\right)}{\rho + \delta} \varphi \ell_{\mathbb{K}z}^{\varphi-1} = m_z \frac{\ell_{\mathbb{K}z}^{\varphi}}{\delta} + m_{-z} \left(\frac{\ell_{\mathbb{K}-z}}{\ell_{\mathbb{K}z}}\right)^{\varphi(\epsilon-2)} \frac{\ell_{\mathbb{K}-z}^{\varphi}}{\delta}.$$

Dividing both sides by $\ell_{\mathbb{K}z}^{\varphi}/\delta$ we arrive at

$$\frac{\delta\pi_{\mathbb{K}z}A(\epsilon-1)}{\rho+\delta}\varphi\ell_{\mathbb{K}z}^{-1}=m_z+m_{-z}\left(\frac{\ell_{\mathbb{K}-z}}{\ell_{\mathbb{K}z}}\right)^{\varphi(\epsilon-1)}.$$

By symmetry by considering the problem of a platform k in group $-z$,

$$\frac{\delta\pi_{\mathbb{K}-z}A(\epsilon-1)}{\rho+\delta}\varphi\ell_{\mathbb{K}-z}^{-1}=m_{-z}+m_z\left(\frac{\ell_{\mathbb{K}z}}{\ell_{\mathbb{K}-z}}\right)^{\varphi(\epsilon-1)}.$$

Let $y:=\ell_{\mathbb{K}z}/\ell_{\mathbb{K}-z}$. Using the above two equations, I derive

$$\frac{\pi_{\mathbb{K}z}}{\pi_{\mathbb{K}-z}}\frac{1}{y}=\frac{m_z+m_{-z}y^{-\varphi(\epsilon-1)}}{m_{-z}+m_z y^{\varphi(\epsilon-1)}}.$$

Equivalently,

$$y=\left(\frac{\pi_{\mathbb{K}z}}{\pi_{\mathbb{K}-z}}\right)^{\frac{1}{1-\varphi(\epsilon-1)}}.$$

Thus,

$$\ell_{\mathbb{K}z}=\frac{\varphi\delta\pi_{\mathbb{K}z}A(\epsilon-1)}{\rho+\delta}\frac{1}{m_z+m_{-z}\left(\frac{\pi_{\mathbb{K}-z}}{\pi_{\mathbb{K}z}}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}}}\tag{49}$$

Note that if $\epsilon\leq 2$ then the Hamiltonian is jointly concave in the state and control and so I have identified the optimal control. If $\epsilon-1<\frac{1}{\varphi}$ one can verify using the same approach outline in Step 5 for the proof of Theorem 1.

Step 4. Calculate Quality Levels and Attention Shares: From (49) we also have the quality level

$$q_z=\frac{\ell_{\mathbb{K}z}^{\varphi}}{\delta}$$

and attention share

$$x_z=\frac{m_z q_z^{\epsilon-1}}{m_z q_z^{\epsilon-1}+m_{-z}q_{-z}^{\epsilon-1}}=\frac{m_z\left(\frac{\pi_{\mathbb{K}z}}{\pi_{\mathbb{K}-z}}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}}}{m_z\left(\frac{\pi_{\mathbb{K}z}}{\pi_{\mathbb{K}-z}}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}}+m_{-z}}\tag{50}$$

of group z platforms.

Step 5. Calculate Surpluses Consumer surplus is simply $u(C, X)/r$ where

$$C = I(M\mu_H)^{\frac{1}{\sigma-1}}$$

and

$$X = \left(\sum_{z=1}^{2} m_z q_z^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}}$$

are product and platform consumption respectively. □

F.2 Summary of the Algorithm

For ease of viewing, I summarize the algorithm once more here in the appendix.

1. Guess a value of x_1.
2. Iterate (45) forward to compute h.
3. Iterate (48) to compute per-unit income bid functions and average ad prices.
4. Check whether the guess of x_1 aligns with (50).
5. If yes, done. If not, repeat with a revised guess.

All other equilibrium objects are characterized in closed form in terms of the output of this algorithm and primitives. Though inefficient, one can simply run steps 2-4 for each guess of x_1 in a fine grid on $[0, 1]$. This is relatively fast and allows one to solve for *all* steady state equilibria and in particular, check uniqueness.

G Discussion of Dynamics

In this appendix, I describe some methodological advantages of a dynamic analysis.

Consider an alternative one period model in which firms participate in multiple auctions for a given consumer. In such a model, there will not exist a bidding equilibrium in symmetric strategies (and it is unclear if there are other asymmetric equilibria). It is easiest to show this when values are supported on a $[0, \overline{v}]$ where $\overline{v} < \infty$. Suppose that there was a symmetric bidding equilibrium with increasing bidding strategies. Then a firm with value $\overline{v}$ would have a profitable deviation to bidding a zero amount in one of the auctions it participates on. This is because the firm is guaranteed to win all of the auctions it participates on if it follows the equilibrium strategy. But, the firm has only unit demand for displaying an ad. By deviating in this way, the firm can reduce its cost while still displaying an ad. By spreading the auction competition out over time I am able to avoid this issue. Thus, dynamics allows us to have an auction analysis in which consumers multi-home and there is interplatform competition in the sale of ads.

Of course, we could consider a one period model where there is no interplatform competition in the sale of ads and each firm participates in one auction each. There would be some matching of firms to auctions which we would have to take a stance on. If N or A are sufficiently large, then regardless of the matching some firms *must* participate on multiple auctions—there aren't enough distinct bidders to be allocated to fill up the auctions. In other words, comparative statics on N or A would have to be limited to a certain range where this is not the case. This is unattractive for the model especially since A is endogenous. You could of course assume N and A are sufficiently small so that so that not all auctions are filled. But, these additional modeling assumptions, in my opinion, seem unnatural. An attempt to explain them would probably appeal to unmodeled frictions such as the fact that it takes time for firms to locate auctions for a given consumer. The dynamic modeling simply makes this intuition formal.

Consider an alternative one period formulation with competitive pricing in the ad market rather than auctions. The baseline model can be solved in a competitive pricing environment. However, in that model, ads would be sold consumer by consumer and it would be as though the platforms inform firms about the *identity* i of the consumer when they make their purchases. In reality, firms only see signals and they do not know if they correspond to the same individual just as they do in the baseline model of Section 3. Of course, a disadvantage of switching to competitive pricing is that future work can not explore the model using ad auction level data. Moreover, its not clear how to solve the extended version of the model where platforms may have different data with competitive pricing. In this model, we can not let platforms sell ads consumer by consumer as if the platforms know the consumers' identities because then the firms could combine the data they receive from the different platforms. Thus, suppose each firm sees only signals of the consumers' valuations but not their identities when choosing which ads to purchase. Firms would have to do some inference about the likelihood they will also purchase an ad for the same consumer on the other platforms. This inference effect leads to complicated purchasing strategies that are nontrivial to characterize.

H Extension: Network Effects

I extend the baseline model to allow for network effects.

H.1 Setup

I redefine the CES aggregate for platform consumption to be

$$X_{it} = \left[\int_{\mathbb{K}} [\eta(x_{kt})\nu(a_{kt})q_{kt}x_{ikt}]^{\frac{\epsilon-1}{\epsilon}} \, \mathrm{d}k \right]^{\frac{\epsilon}{\epsilon-1}}$$

where $\eta(x) = x^{\zeta}$ where $\zeta > 0$. I retain all other aspects of the baseline model of Section 3.

H.2 Equilibrium Characterization

Theorem 4. *Suppose that $\hat{A}$ is the unique solution of $\max_a a\nu(a)^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}}$. If $\hat{A}/\lambda_f < J$ and $(\epsilon-1)/[1-\zeta(\epsilon-1)] < 1/\varphi$, then there exists a unique equilibrium in which each platform $k \in \mathbb{K}$ receives a positive amount of attention $x_{kt} > 0$ at all times t for any feasible initial conditions M_0, H_0, and q_0. The equilibrium converges to a steady state and has the following properties:*

1. *Consumer i's demands for products are as in* (27) *and her demands for platforms are as in* (53).
2. *Firm j sets prices as in* (1).
3. *Platform k displays ads at rate $\hat{A}$.*
4. *The size of consideration sets is given by* (29) *and the cdfs of the expected values of firms inside and outside of them are characterized by* (30) *and* (4) *with $\hat{A}$ in place of A.*
5. *Firm j's expected flow profits from sales are as in* (2) *and the rates at which firm j matches with consumers are as in* (11).
6. *Firm j bids according to* (31).
7. *Platform k's quality and investment solve the boundary-value problem* (33) *except with $(\epsilon-1)/[1-\zeta(\epsilon-1)]$ in place of $\epsilon-1$.*
8. *Total consumer, firm, and platform surplus are as in Step 8 of Section 4 except*
$$X_{it} = K^{\frac{1}{\epsilon-1}-\zeta}\nu(\hat{A})q_t.$$

Moreover the sufficient conditions are almost necessary: if either $\hat{A}/\lambda_f \geq J$ or $(\epsilon-1)/[1-\zeta(\epsilon-1)] > 1/\varphi$ then there does not exist an equilibrium.

Proof. As discussed in Section 9, the attention received by platform k solves

$$x_{kt}^{\zeta} = \frac{x_{kt}^{\zeta(\epsilon-1)}[\nu(a_{kt})q_{kt}]^{\epsilon-1}}{Y} \tag{51}$$

where

$$Y = \int_{\mathbb{K}} x_{kt}^{\zeta(\epsilon-1)}[\nu(a_{kt})q_{kt}]^{\epsilon-1}\,\mathrm{d}k.$$

Solving (51) for x_{kt} yields two possibilities:

$$x_{kt} = \frac{[\nu(a_{kt})q_{kt}]^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}}}{Y^{\frac{1}{1-\zeta(\epsilon-1)}}} \tag{52}$$

or $x_{kt} = 0$. Under the equilibrium refinement, all platforms must receive positive attention share (52). Integrating both sides of (52) over $\mathbb{K}$ yields

$$Y^{\frac{1}{1-\zeta(\epsilon-1)}} = \int_{\mathbb{K}} [\nu(a_{kt})q_{kt}]^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}}\,\mathrm{d}k.$$

Then substituting into (52) gives

$$x_{kt} = \frac{[\nu(a_{kt})q_{kt}]^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}}}{\int_{\mathbb{K}}[\nu(a_{kt})q_{kt}]^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}}\,\mathrm{d}k}. \tag{53}$$

Thus, the only change relative to the baseline model is that the elasticity of attention with respect to platform quality is now higher. The rest of the equilibrium derivation follows the same steps as in Section 4. □

I Extension: Heterogeneous Platform Productivity

I solve an extension of the baseline model in which platforms may differ in the productivity of their investments.

I.1 Setup

Platform k now solves

$$\max_{\{\ell_{kt}\}} \int_0^{\infty} e^{-\rho t}\left(\pi_{\mathbb{K}t} A \frac{q_{kt}^{\epsilon-1}}{\int_{\mathbb{K}} q_{zt}^{\epsilon-1}\,\mathrm{d}z} - \alpha_k \ell_{kt}\right)\mathrm{d}t$$

where

$$\dot{q}_{kt} = \ell_{kt}^{\varphi} - \delta q_{kt}.$$

The only difference relative to the baseline model is the parameter $\alpha_k > 0$ which controls the productivity of platform k. Let P denote the frequency distribution of $\alpha_k, k \in \mathbb{K}$. I retain all other aspects of the baseline model.

I.2 Equilibrium Characterization

Theorem 5. *Suppose that A is the unique solution of $\max_a a\nu(a)^{\epsilon-1}$ and $\epsilon - 1 < 1/\varphi$. Then there exists a unique equilibrium where platform k's investment is*

$$\ell_{kt} = \left(\frac{1}{\alpha}\right)^{\frac{1}{1-\varphi(\epsilon-1)}} \ell_t$$

and quality level is

$$q_{kt} = \left(\frac{1}{\alpha}\right)^{\frac{\varphi}{1-\varphi(\epsilon-1)}} q_t$$

where ℓ_t and q_t solve the ODE system

$$\begin{aligned}
(\rho+\delta)\frac{1}{\varphi}\ell_t - \frac{1-\varphi}{\varphi}\dot{\ell}_t &= \frac{\pi_{\mathbb{K}t} A(\epsilon-1)}{\int \left(\frac{1}{\alpha}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}} \, \mathrm{d}P(\alpha)} \frac{\ell_t^{\varphi}}{q_t} \\
\dot{q}_t &= \ell_t^{\varphi} - \delta q_t
\end{aligned}$$

with given initial condition q_0 and boundary at infinity

$$\lim_{t\to\infty} \ell_t = \frac{\delta \pi_{\mathbb{K}t} A(\epsilon-1)}{(\rho+\delta)} \frac{\varphi}{\alpha_k} \left[\int \left(\frac{1}{\alpha}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}} \mathrm{d}P(\alpha)\right]^{-1}.$$

Proof. All equilibrium objects besides investment are derived as in the proof of Theorem 1.

The Hamiltonian for platform k's problem is

$$\mathcal{H}(t, q_{kt}, \ell_{kt}, \lambda_{kt}) = \pi_{\mathbb{K}t} A \frac{q_{kt}^{\epsilon-1}}{\int_{\mathbb{K}} q_{zt}^{\epsilon-1} \, \mathrm{d}z} - \alpha_k \ell_{kt} + \lambda_t \left(\ell_{kt}^{\varphi} - \delta q_{kt}\right)$$

where the costate variable λ_t solves

$$\rho\lambda_{kt} - \dot{\lambda}_{kt} = \pi_{\mathbb{K}t} A(\epsilon-1) \frac{q_{kt}^{\epsilon-2}}{\int_{\mathbb{K}} q_{lt}^{\epsilon-1} \, \mathrm{d}l} - \lambda_{kt}\delta.$$

The FOC for maximizing the Hamiltonian yields

$$\lambda_{kt} = \frac{\alpha_k}{\varphi}\ell_{kt}^{1-\varphi}.$$

Differentiating both sides with respect to time yields

$$\dot{\lambda}_{kt} = \alpha_k \frac{1-\varphi}{\varphi}\ell_{kt}^{-\varphi}\dot{\ell}_{kt}.$$

Then we have

$$(\rho+\delta)\frac{\alpha_k}{\varphi}\ell_{kt}^{1-\varphi} - \alpha_k\frac{1-\varphi}{\varphi}\ell_{kt}^{-\varphi}\dot{\ell}_{kt} = \pi_{\mathbb{K}t}A(\epsilon-1)\frac{q_{kt}^{\epsilon-2}}{Q_t} \tag{54}$$

where

$$Q_t = \int_{\mathbb{K}} q_{zt}^{\epsilon-1}\,\mathrm{d}z.$$

Step 1. Steady State: I first derive the steady-state equilibrium before returning to solve for the full dynamics. In steady state,

$$(\rho+\delta)\frac{\alpha_k}{\varphi}\ell_{kt}^{1-\varphi} = \pi_{\mathbb{K}t}A(\epsilon-1)\frac{q_{kt}^{\epsilon-2}}{Q_t}$$

and

$$q_k = \frac{\ell_k^{\varphi}}{\delta}.$$

I obtain

$$(\rho+\delta)\frac{\alpha_k}{\varphi}\ell_{kt}^{1-\varphi} = \pi_{\mathbb{K}t}A(\epsilon-1)\frac{\ell_k^{(\epsilon-2)\varphi}}{\delta^{\epsilon-2}Q_t}$$

which can be solved to yield

$$\ell_k = \left(\frac{\pi_{\mathbb{K}t}A(\epsilon-1)}{\delta^{\epsilon-2}Q_t(\rho+\delta)}\frac{\varphi}{\alpha_k}\right)^{\frac{1}{1-\varphi(\epsilon-1)}}.$$

This implies that in steady state,

$$q_{kt} = \frac{1}{\delta}\left(\frac{\pi_{\mathbb{K}t}A(\epsilon-1)}{\delta^{\epsilon-2}Q_t(\rho+\delta)}\frac{\varphi}{\alpha_k}\right)^{\frac{\varphi}{1-\varphi(\epsilon-1)}}$$

Then, using the definition of Q_t we have

$$\frac{1}{\delta^{\epsilon-1}}\int\left(\frac{\pi_{\mathbb{K}t}A(\epsilon-1)}{\delta^{\epsilon-2}Q_t(\rho+\delta)}\frac{\varphi}{\alpha}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}}\mathrm{d}P(\alpha) = Q_t$$

which can be solved to yield

$$Q_t = \left[\frac{1}{\delta^{\epsilon-1}} \int \left(\frac{\pi_{\mathbb{K}t} A(\epsilon-1)}{\delta^{\epsilon-2}(\rho+\delta)} \frac{\varphi}{\alpha} \right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}} \mathrm{d}P(\alpha) \right]^{1-\varphi(\epsilon-1)}.$$

Therefore, the steady-state level of investment is

$$\ell_k = \left(\frac{\pi_{\mathbb{K}t} A(\epsilon-1)}{\delta^{\epsilon-2}(\rho+\delta)} \frac{\varphi}{\alpha_k} \right)^{\frac{1}{1-\varphi(\epsilon-1)}} \left[\frac{1}{\delta^{\epsilon-1}} \int \left(\frac{\pi_{\mathbb{K}t} A(\epsilon-1)}{\delta^{\epsilon-2}(\rho+\delta)} \frac{\varphi}{\alpha} \right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}} \mathrm{d}P(\alpha) \right]^{-1}$$

which simplifies to

$$\ell_k = \frac{\delta \pi_{\mathbb{K}t} A(\epsilon-1)}{(\rho+\delta)} \frac{\varphi}{\alpha_k} \left[\int \left(\frac{1}{\alpha} \right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}} \mathrm{d}P(\alpha) \right]^{-1}.$$

Step 2. Transition Path: Now I return to solve for dynamics away from steady state. Let me denote by ℓ_t the investment and q_t the quality of a platform that has $\alpha = 1$. I conjecture that

$$\ell_{kt} = \left(\frac{1}{\alpha} \right)^{\frac{1}{1-\varphi(\epsilon-1)}} \ell_t.$$

This of course implies then that

$$q_{kt} = \left(\frac{1}{\alpha} \right)^{\frac{\varphi}{1-\varphi(\epsilon-1)}} q_t.$$

Then we have from (54)

$$(\rho+\delta) \frac{1}{\varphi} \ell_t^{1-\varphi} - \frac{1-\varphi}{\varphi} \ell_t^{-\varphi} \dot{\ell}_t = \pi_{\mathbb{K}t} A(\epsilon-1) \frac{q_t^{\epsilon-2}}{Q_t}.$$

By the conjecture, we have

$$Q_t = \int \left(\frac{1}{\alpha} \right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}} \mathrm{d}P(\alpha) q_t^{\epsilon-1}.$$

Therefore

$$(\rho+\delta) \frac{1}{\varphi} \ell_t^{1-\varphi} - \frac{1-\varphi}{\varphi} \ell_t^{-\varphi} \dot{\ell}_t = \frac{\pi_{\mathbb{K}t} A(\epsilon-1)}{\int \left(\frac{1}{\alpha} \right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}} \mathrm{d}P(\alpha)} \frac{1}{q_t}.$$

From here, the same method as in the proof of Theorem 1 can be used to verify the conjecture and complete the proof. I omit these steps for brevity. □

J Extension: Zero Prices

In this section, I give an informal argument that zero prices can arise, for some parameter conditions, in an equilibrium of a variant of the baseline model where platforms can charge nonnegative prices. I rule out negative prices on the grounds that it is too difficult for platforms to verify human usage as opposed to usage by bots. Under this premise, charging a negative price is unsustainable for a platform.

It is easiest to make the point when there are an integer number K of atomic platforms and when consumers have Cobb-Douglas utility as in Section 6. However, these assumptions are not central to the logic of the argument. The core of the argument is simply that if consumers enjoy product consumption much more than platform consumption, then the attention spent on a platform will decrease quickly in the price set by the platform. This is because, if the consumer spends more attention on the platform, the consumer can spend less income on consuming products. When the elasticity of attention with respect to price is sufficiently high, it is better for a platform to rely solely on advertising to earn revenue.

Suppose that all platforms but platform 1 charge a zero price and have quality level q. Let q_1 denote platform 1's quality level. Let $p \geq 0$ denote the price charged by platform 1. Consumer i chooses how much attention x_1 to allocate to platform 1 to maximize flow utility which amounts to maximizing

$$(I - px_1)^{1-\tau} \left(M\mu_H\right)^{\frac{1-\tau}{\sigma-1}} \left[(x_1 q_1)^{\frac{\epsilon-1}{\epsilon}} + (K-1)\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\tau\epsilon}{\epsilon-1}} \nu(A)^{\tau}.$$

The first two terms comprise product consumption and the second two terms comprise platform consumption. Above, I have used the fact that the consumer will want to allocate attention evenly across the $K-1$ remaining platforms. After spending px_1 units of income on consuming platform 1, the consumer has only $I - px_1$ left to spend on products.

The first order condition for consumer i's problem is

$$\begin{aligned} p(1-\tau)(I-px_1)^{-\tau} &\left[(x_1 q_1)^{\frac{\epsilon-1}{\epsilon}} + (K-1)\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\tau\epsilon}{\epsilon-1}} = \\ &(I-px_1)^{1-\tau}\tau \left[(x_1 q_1)^{\frac{\epsilon-1}{\epsilon}} + (K-1)\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\tau\epsilon}{\epsilon-1}-1} \\ &\qquad \times \left[q_1 (x_1 q_1)^{\frac{\epsilon-1}{\epsilon}-1} - q\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}-1}\right]. \end{aligned}$$

Canceling out some terms and rearranging yields

$$p\left[(x_1q_1)^{\frac{\epsilon-1}{\epsilon}}+(K-1)\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}}\right]=$$
$$\frac{\tau}{1-\tau}\left[q_1(x_1q_1)^{\frac{\epsilon-1}{\epsilon}-1}-q\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}-1}\right](I-px_1)$$

which is linear in the price p.

Solving for p yields the inverse demand curve:[42]

$$p(x_1)=$$
$$\frac{I\left[q_1(x_1q_1)^{\frac{\epsilon-1}{\epsilon}-1}-q\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}-1}\right]}{\frac{1-\tau}{\tau}\left[(x_1q_1)^{\frac{\epsilon-1}{\epsilon}}+(K-1)\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}}\right]+\left[q_1(x_1q_1)^{\frac{\epsilon-1}{\epsilon}-1}-q\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}-1}\right]x_1}.$$

Since $I-px_1$ must be positive and p must be nonnegative, the domain of the inverse demand curve is the set of demands x_1 such that the numerator is nonnegative. That is, the domain is

$$x_1\in\left[0,\frac{q_1^{\epsilon-1}}{q_1^{\epsilon-1}+(K-1)q^{\epsilon-1}}\right].$$

This is intuitive: the domain consists of attention levels that are less than that which would arise if platform 1 also charged a price of zero.

We can now formulate platform 1's pricing problem which is to choose x_1 in this domain to maximize flow profit:

$$p(x_1)x_1+\pi_{\mathbb{K}}Ax_1.$$

We are interested in parameter conditions for when

$$x_1=\frac{q_1^{\epsilon-1}}{q_1^{\epsilon-1}+(K-1)q^{\epsilon-1}}$$

is optimal which corresponds to setting a zero price. This amounts to looking for parameter conditions such that

$$\frac{\mathrm{d}[p(x_1)x_1]}{\mathrm{d}x_1}+\pi_{\mathbb{K}}A=p'(x_1)x_1+p(x_1)+\pi_{\mathbb{K}}A>0 \tag{55}$$

[42]One can see that the inverse demand curve is monotone since the objective is submodular in (p,x_1).

for all x_1 in the domain. This will happen if $p(\cdot)$ does not decrease too fast. Intuitively this will be the case when τ is close to zero so that the consumer cares little about platform use and so attention will be very sensitive to the price set by platform 1.

By inspection, if τ is sufficiently close to 0,

$$p(x_1) \approx I \frac{\tau}{1-\tau} \frac{q_1 (x_1 q_1)^{\frac{\epsilon-1}{\epsilon}-1} - q \left(\frac{1-x_1}{K-1} q\right)^{\frac{\epsilon-1}{\epsilon}-1}}{(x_1 q_1)^{\frac{\epsilon-1}{\epsilon}} + (K-1) \left(\frac{1-x_1}{K-1} q\right)^{\frac{\epsilon-1}{\epsilon}}}.$$

I show later that we can bound

$$\frac{\mathrm{d}[p(x_1) x_1]}{\mathrm{d} x_1}$$

from *below* for all x_1 in the domain by an amount that can be made arbitrarily close to 0 by making τ sufficiently close to 0. Thus, when τ is sufficiently close to 0, it follows that (55) holds for all x_1 in the domain and a price of zero is optimal. I will take this fact as given now and show it formally later.

I have therefore shown that platform 1 does not have a profitable deviation to charging a positive price when τ is close to 0. In principle the platform could deviate both in its investment strategy and in its pricing strategy. But, so far, our analysis has fixed an arbitrary quality level for q_1. For some value of τ, say $\tau(q_1)$, which depends on q_1 we have shown it is not profitable to charge a positive price. Let $\overline{q}$ be relatively large and $\underline{q}$ be relatively small and consider

$$\tau^* = \inf_{q_1 \in [\underline{q}, \overline{q}]} \tau(q_1) < 1.$$

Then for parameter $\tau = \tau^*$ it is never optimal for a platform to deviate to any quality $q \in [0, \overline{q}]$. By setting $\overline{q}$ sufficiently high, investment costs are sufficiently high that it is obviously not optimal to deviate in terms of investment to end up at any quality $q \geq \overline{q}$. Similarly if $q \leq \underline{q}$ for some $\underline{q}$ sufficiently low deviating by cutting back on investment is not profitable because profits are too low at any positive price that the platform can set.

The analysis so far has assumed that the second order condition is satisfied. By inspection the second order condition is also satisifed since the objective is concave in the relevant domain

$$x_1 \in \left[0, \frac{q_1^{\epsilon-1}}{q_1^{\epsilon-1} + (K-1) q^{\epsilon-1}}\right]$$

provided $p \geq 0$ as we have assumed. We only need concavity over this domain since we know any higher choice of attention is always suboptimal.

I will now show that we can bound the derivative of

$$p(x_1)x_1 = \frac{I\left[q_1(x_1q_1)^{\frac{\epsilon-1}{\epsilon}-1} - q\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}-1}\right]x_1}{\frac{1-\tau}{\tau}\left[(x_1q_1)^{\frac{\epsilon-1}{\epsilon}} + (K-1)\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}}\right] + \left[q_1(x_1q_1)^{\frac{\epsilon-1}{\epsilon}-1} - q\left(\frac{1-x_1}{K-1}q\right)^{\frac{\epsilon-1}{\epsilon}-1}\right]x_1}$$

from below by an amount arbitrarily close to 0 as claimed.

Define f such that

$$p(x_1)x_1 = I\frac{f'(x_1)x_1}{\frac{1-\tau}{\tau}\frac{\epsilon}{\epsilon-1}f(x_1) + f'(x_1)x_1}.$$

Then

$$\begin{aligned}\frac{1}{I}\frac{\mathrm{d}[p(x_1)x_1]}{\mathrm{d}x_1} &= \frac{f''(x_1)x_1 + f'(x_1)}{\frac{1-\tau}{\tau}\frac{\epsilon}{\epsilon-1}f(x_1) + f'(x_1)x_1} \\ &\quad - \frac{f'(x_1)x_1}{\left[\frac{1-\tau}{\tau}\frac{\epsilon}{\epsilon-1}f(x_1) + f'(x_1)x_1\right]^2}\left[\left(\frac{1-\tau}{\tau}\frac{\epsilon}{\epsilon-1}+1\right)f'(x_1) + f''(x_1)x_1\right].\end{aligned}$$

Can we bound this from below? Consider taking the limit as τ tends to 0. Then the above, for a fixed x_1, converges to 0. However, we need to make sure that the infimum of the above over the entire range converges to 0. This will necessarily be the case if we can bound $f'(x_1)$, $f''(x_1)$, $f(x_1)$, and $f'(x_1)x_1$. The concern is about points x_1 near 0 since some of these terms explode there. Suppose we know that for any τ close to 0 that it is not optimal to set $x_1 < \epsilon$ for some $\epsilon > 0$, then we are done since $f'(x_1)$, $f''(x_1)$, $f(x_1)$, and $f'(x_1)x_1$ are bounded on $\left[\epsilon, \frac{q_1^{\epsilon-1}}{q_1^{\epsilon-1}+(K-1)q^{\epsilon-1}}\right]$

To show this, note that

$$p(x)x + \pi_{\mathbb{K}}Ax \leq p(x)x + \pi_{\mathbb{K}}A\epsilon$$

whenever $x \leq \epsilon$. But

$$p(x)x \leq I\frac{f'(x)x}{\frac{\epsilon}{\epsilon-1}f(x) + f'(x)x}$$

for all $\tau \leq 1/2$. Then

$$p(x)x + \pi_{\mathbb{K}}Ax \leq \sup_{x\in[0,\epsilon]} I\frac{f'(x)x}{\frac{\epsilon}{\epsilon-1}f(x) + f'(x)x} + \pi_{\mathbb{K}}A\epsilon$$

and the upper bound holds uniformly over all $\tau \leq \frac{1}{2}$. For some $\epsilon > 0$ this right hand side is always less than

$$\pi_{\mathbb{K}} A \frac{q_1}{q_1 + (K-1)q}.$$

Thus, there is some $\epsilon > 0$ lower bound such that its not optimal to set $x_1 < \epsilon$ for any parameter $\tau \leq 1/2$.

K Extension: Firm and Platform Entry

I extend the baseline model to allow for entry of firms and platforms.

K.1 Setup

To enter the market, a firm must pay a cost $e_{\mathbb{J}} > 0$ and a platform must pay a cost $e_{\mathbb{K}} > 0$. I retain all other aspects of the baseline model of Section 3 except that in equilibrium, the measure of firms F and the measure of platforms P are such that firms and platforms earn zero profits net of entry costs.

K.2 Steady-State Equilibrium Characterization

For the notion of steady state equilibrium, I assume that each platform enters with the steady state quality level to keep the analysis simple. One might consider other conventions such as having a given platform enter with some given quality level q_0 that may differ from steady state and then characterize transition dynamics for that platform while restricting all other equilibrium properties in steady state. I will not explore that here. Under either convention, the measure of firms in the steady state will be the same.

Theorem 6. *Suppose that A is the unique solution of $\max_a a\nu(a)^{\epsilon-1}$. If $A/\lambda_f < J$ and $\epsilon \leq 1/\varphi$ then there exists a steady state equilibrium. If one exists, it is unique and has the same steady-state properties as in Theorem 1 for a given J and K which satisfy*

$$K = \frac{\pi_{\mathbb{K}} A}{e_{\mathbb{K}}} \left(1 - \frac{\varphi\delta(\epsilon - 1)}{\rho + (1-\alpha)\delta}\right), \tag{56}$$

and

$$J = \frac{\frac{I}{\sigma} - \pi_{\mathbb{K}} A}{e_{\mathbb{J}}}. \tag{57}$$

Proof. Equations 56 and (57) are zero profit conditions. Note that in (57), $\pi_{\mathbb{K}}$ depends on J. Thus, to prove uniqueness I must prove there is a unique solution for J in (57). This follows by Lemma 12 which shows that $\pi_{\mathbb{K}}$ is increasing in J. The other parts of the theorem follows the same roadmap as in the proof of Theorem 1. □

It is straightforward to extend most of the comparative statics for steady state in Appendix C to this setting.

L Extension: Reserve Prices

I extend the baseline model to allow platforms to set reserve prices.

L.1 Setup

Each platform k sets a reserve price to maximize the expected revenue in each auction taking as given the reserve prices chosen by its rivals. All other aspects of the model are as in the baseline model of Section 3.

L.2 Steady State Equilibrium Characterization

Theorem 7. *Suppose that $\epsilon - 1 < 1/\varphi$ and that A is the unique solution to $\max_a a\nu(a)^{\epsilon-1}$. In a steady state equilibrium, the following hold:*

1. *Consumer i's demands for products are as in (27) and her demands for platforms are as in (15).*

2. *Firm j sets prices as in (1).*

3. *Platform k displays ads at rate A.*

4. *The size of consideration sets is $M = A/\lambda_f$.*

5. *Firm j's expected flow profits from sales are as in (2).*

6. *Firm j sets reserve price $R = \frac{\pi_{\mathbb{J}}}{\lambda_f+\rho}Y$ where Y solves*

$$Y = \frac{1 - H^c(Y)}{h^c(Y)}$$

where

$$H^c(Y) = \frac{K}{J - M}G(Y)$$

and

$$h^c(Y) = \frac{\frac{K}{M} g(Y)[1 - H^c(Y)^N]}{NH^c(Y)^{N-1} + \frac{J-M}{M}\left[1 - H^c(Y)^N\right]}.$$

7. *Firm j bids according to*

$$B(\hat{v}_{ij}) = \pi_{\mathbb{J}} \int_Y^{\hat{v}_{ij}} \frac{1}{\rho + \lambda_f + \lambda_a H^c(s)^{N-1}} \, \mathrm{d}s + R$$

whenever $\hat{v}_{ij} \geq Y$.

8. *The cdf of the expected values of a consumer for firms outside of her consideration sets solves*

$$H^c(s)^N - \left(\frac{K}{J-M}\right)^N G(Y)^N$$
$$= \left[\frac{K}{M} G(s) - \frac{J-M}{M} H^c(s)\right] \left(1 - \left[\frac{K}{J-M}\right)^N G(Y)^N\right]$$

for $s \geq Y$.

9. *Each platform k invests at rate* (18).

Proof. It is clear that consumers' demands and firms' flow profits and prices will be the same as in the baseline model of Section 3.

Each platform k sets the rate it displays ads to consumers to maximize flow utility:

$$A = \arg\max_{a_{kt}} \pi_{\mathbb{K}} \frac{a_{kt}}{1 - H^c(Y)^N} \nu(a_{kt})^{\epsilon - 1} = \arg\max a_{kt} \nu(a_{kt})^{\epsilon - 1}$$

as before. Here $\pi_{\mathbb{K}}$ denotes the average ad price in auction. If a_{kt} is the rate that ads are displayed, then $a_{kt}/[1 - H^c(Y)^N]$ is the rate that auctions are held since an ad is only displayed if one of the bidders has bid above the reserve.

Thus, in a steady state equilibrium, the rate that a firm enters an auction is now

$$\lambda_a = \frac{NA}{(J-M)[1 - H^c(Y)^N]}.$$

In a second-price auction, a firm bids the gain its continuation value from winning the auction. Thus

$$B(\hat{v}_{ij}) = \frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} \hat{v}_{ij} + \frac{\lambda_f}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\hat{v}_{ij}) - \frac{\lambda_a}{\lambda_a + \rho} V(\hat{v}_{ij}) \tag{58}$$

where $V(\hat{v}_{ij})$ is the continuation value from selling to consumer i at the time of entry into an auction. It is defined recursively by the equation

$$\begin{aligned}V(\hat{v}_{ij}) = \left[1 - H(\hat{v}_{ij})^{N-1}\right] \frac{\lambda_a}{\lambda_a + \rho} V(\hat{v}_{ij}) + \\ H^c(\hat{v}_{ij})^{N-1} \Bigg(\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} \hat{v}_{ij} + \frac{\lambda_f}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\hat{v}_{ij}) \\ - \mathbb{E}\left[\max\{B(\hat{v}^{(1)}), R\} | \hat{v}_{ij} > \hat{v}^{(1)}\right] \Bigg) \quad (59)\end{aligned}$$

whenever $\hat{v}_{ij} \geq Y$. In this equation, $\hat{v}^{(1)} \sim (H^c)^{N-1}$ represents the highest expected value of the other bidders in an auction.

Since the cutoff bidder must bid its value, given the reserve price R, it follows that

$$Y = R\frac{\lambda_f + \rho}{\pi_{\mathbb{J}}}.$$

To ease notation, let $O(\hat{v}_{ij}) = H^c(\hat{v}_{ij})^{N-1}$. Then

$$O(\hat{v}_{ij})\mathbb{E}\left[\max\{B(\hat{v}^{(1)}), R\} | \hat{v}_{ij} > \hat{v}^{(1)}\right] = RO(R) + \int_Y^{\hat{v}_{ij}} B(s)O'(s)\, \mathrm{d}s.$$

Substituting into (59) yields

$$V(\hat{v}_{ij}) \left(1 - \frac{\lambda_a}{\lambda_a + \rho}\right) = O(\hat{v}_{ij})B(\hat{v}_{ij}) - RO(R) - \int_Y^{\hat{v}_{ij}} B(s)O'(s)\, \mathrm{d}s$$

for $\hat{v}_{ij} \geq Y$. Then substituting into (58)

$$\begin{aligned}B(\hat{v}_{ij}) = \frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} \hat{v}_{ij} \\ - \frac{\rho}{\lambda_f + \rho} \frac{\lambda_a}{\rho} \left[O(\hat{v}_{ij})B(\hat{v}_{ij}) - RO(R) - \int_Y^{\hat{v}_{ij}} B(s)O'(s)\, \mathrm{d}s\right].\end{aligned}$$

Differentiating with respect to $\hat{v}_{ij}$, I solve explicitly for $B'(\hat{v}_{ij})$. Using the boundary condition $B(Y) = R$, we find that

$$B(\hat{v}_{ij}) = \pi_{\mathbb{J}} \int_Y^{\hat{v}_{ij}} \frac{1}{\rho + \lambda_f + \lambda_a H^c(s)^{N-1}}\, \mathrm{d}s + R$$

for $\hat{v}_{ij} \geq Y$. Note that, any bidder with a value $\hat{v}_{ij} < Y$ optimizes by bidding any amount less than or equal to Y. However, in the event that a platform deviates to a lower, reserve price, these bidders must bid their intrinsic values for the ad in that

$$B(\hat{v}_{ij}) = \frac{\pi_{\mathbb{J}}}{\lambda_f + \rho}$$

when $\hat{v}_{ij} \leq Y$ since their continuation values are 0.

Next, we derive the steady state H and H^c. Matching inflows with outflows gives,

$$\frac{NH^c(s)^{N-1}h^c(s)}{1 - H^c\left(Y\right)^N} = h(s). \tag{60}$$

for all $s \geq Y$. On the left we have the pdf of the highest expected values of the firms in the ad auctions. On the right we have the pdf of the expected values of firms who are forgotten uniformly at random.

Then, integrating we have

$$H^c(s)^N - H^c(Y)^N = H(s)\left(1 - H^c(Y)^N\right)$$

for $s \geq Y$. Recall the accounting identity $MH + (J - M)H^c = FG$. Then

$$H^c(s)^N - H^c(Y)^N = \left[\frac{K}{M}G(s) - \frac{J-M}{M}H^c(s)\right]\left[1 - H^c(Y)^N\right]$$

for $s \geq Y$. Using the fact that $H(Y) = 0$, the accounting identity gives

$$H^c(Y) = \frac{K}{J-M}G(Y). \tag{61}$$

Substituting into the above equation, we find that

$$\begin{aligned} H^c(s)^N - \left(\frac{K}{J-M}\right)^N G(Y)^N \\ = \left[\frac{K}{M}G(s) - \frac{J-M}{M}H^c(s)\right]\left(1 - \left[\frac{K}{J-M}\right)^N G(Y)^N\right] \end{aligned}$$

for $s \geq Y$. Thus, given Y, this equation can be used to compute H^c.

It will also be useful to derive $h^c(Y)$ which we can do by using the equation (60) which implues

$$NH^c(Y)^{N-1}h^c(Y) = h(Y)\left(1 - H^c\left(Y\right)^N\right).$$

Then using an accounting identity, we have

$$NH^c(Y)^{N-1}h^c(Y) = \left[\frac{K}{M}g(Y) - \frac{J-M}{M}h^c(Y)\right]\left[1 - (H^c(Y)^N\right]$$

which rearranges to

$$h^c(Y) = \frac{\frac{K}{M} g(Y)[1 - H^c(Y)^N]}{NH^c(Y)^{N-1} + \frac{J-M}{M}\left[1 - H^c(Y)^N\right]}. \tag{62}$$

I will now write down the optimality condition for the reserve price. Suppose that platform k sets a reserve price that induces cutoff $\hat{Y}$. Then platform k's expected profit in an auction is:

$$\int_{\hat{Y}}^{\infty} B(s)[N(N-1)H^c(s)^{N-2}(1 - H^c(s))h^c(s)]\,\mathrm{d}s \\ + \hat{Y}\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} N\left[1 - H^c(\hat{Y})\right] H^c(\hat{Y})^{N-1}. \tag{63}$$

Platform k sets $\hat{Y} = Y$ to maximize the above expression. The necessary first order condition for optimality is

$$\begin{aligned} &- B(Y)[N(N-1)H^c(Y)^{N-2}\left(1 - H^c(Y)\right)h^c(Y)] \\ &\qquad + Y\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} N\left[(N-1)H^c(Y)^{N-2} - NH^c(Y)^{N-1}\right] h^c(Y) \\ &\qquad\qquad + \frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} N[1 - H^c(Y)]H^c(Y)^{N-1} = 0. \end{aligned}$$

Simplifying, and using the fact that $B(Y) = Y$, we arrive at the familiar equation

$$Y = \frac{1 - H^c(Y)}{h^c(Y)}.$$

Y is the solution to this simple equation but recall that H^c and h^c are themselves functions of Y given by (61) and (62) respectively.

From here, given the average ad price $\pi_{\mathbb{K}}$ which coincides with (63) evaluated at $\hat{Y} = Y$, platforms' investment rates are determined as in the baseline model. The condition that $\epsilon - 1 < 1/\varphi$ is used in this step to ensure that all platforms follow the same investment strategy. As seen from Lemma 2 it is almost a necessary condition for equilibrium existence.

□