A Model of the Attention Economy

Daniel Chen*

February 2022, revised August 2025

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 advertisements using data on consumers. The model provides a tractable and extensible framework that endogenizes dynamics of outcomes in both markets. The effects of data and interoperability policies may reverse over time, the rise of data-rich digital advertising can coincide with declining ad revenue-to-GDP ratio, and platform quality may be inefficiently high or low.

JEL: D40, D44, D83, D60, L15

*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 Chis 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 significant part of the modern economy revolves around capturing and monetizing human attention. The average person spends more than two and a half hours consuming digital media each day (Evans, 2020). Across the web, platforms collectively earn hundreds of billions of dollars each year from targeted advertising (Silk et al., 2021). By a platform, I mean any website or app that displays ads.

This paper develops a dynamic model of the attention economy, in which two markets interact. In the product market, consumers are aware of only a fraction of available firms. In the attention market, platforms sell targeted ad space to product firms. To attract the attention needed to sell ads, platforms invest in the quality of their free services (e.g., social media, streaming, podcasts). Ads are allocated through auctions at the individual level where firms bid using platform data on consumers' preferences. As consumers see ads, they discover new firms; over time, they forget some they once knew.

The main contribution of this paper is a tractable framework that endogenizes both the product and attention markets. The model treats data flexibly, accounts for welfare from both paid products and free services, and accommodates multihoming by firms and consumers.[1] A key parameter—the elasticity of attention to platform quality—governs interplatform competition and market power. Because the model is dynamic, it captures investment incentives and intertemporal spillovers. To illustrate its value, I apply it to study the dynamic effects of policies, competition between platforms with asymmetric data (e.g., digital versus traditional media), and welfare.

For tractability, I model platforms as nonatomic monopolistic competitors. This focuses attention on the operation of the market rather than market concentration. Platforms in the model share business models with real-world players—ranging from smaller sites like Pinterest, Spotify, or Reddit to large platforms such as Meta or YouTube—but do not interact strategically. A duopoly variant shows the framework can accommodate strategic interaction under more restrictive assumptions. CES preferences for both products and platform services further simplify the analysis: on the product side, prices are constant, determined solely by the elasticity parameter, and firms *optimally* do not personalize them—a stark but convenient simplification.

I consider three applications.

The first presents numerical policy experiments highlighting mechanisms that may arise under data reforms and interoperability mandates. The latter aim to promote platform competition by increasing the elasticity of attention

[1] By "flexibly," I mean that data is represented nonparametrically, as information about consumers' preferences.

to platform quality and ad load.[2] In the model, ads impose nuisance costs that reduce a platform's effective quality. The analysis reveals that cross-market spillovers can lead to counterintuitive effects that reverse course over time, and that policies often entail tradeoffs between product-market efficiency and platform quality. For instance, granting platforms access to more granular data can raise ad revenue in the short run but *reduce it* in the long run. In the end, product consumption improves, but platforms invest less in quality. By contrast, interoperability may *depress* investment at first but raise it later. In the end, because platforms cut back on ad display rates, product consumption is worse even as platform quality improves. I discuss conditions when these patterns appear more likely to arise.

The analysis is intentionally stylized to isolate cross-market spillovers and dynamic effects, frequent policy focal points. In practice, interoperability polices often target the market power of large platforms where strategic effects may matter. I provide supporting evidence my findings are robust in a duopoly variant that retains most of the model's architecture but treats ad display rates as exogenous. I develop a numerical algorithm to compute equilibrium and outline how future work may endogenize ad rates in a setting with a continuum of duopolistic platform markets.

The second application extends the model to allow two groups of platforms—e.g., digital and traditional media—with asymmetric data. This extension shows how differences in data shape relative quality and market shares, and can generate patterns consistent with an empirical puzzle: despite the rise of data-rich platforms, advertising's share of GDP declined during 2000–2018 (Silk et al., 2021). In a numerical example, "digital platforms" displace "traditional media" without expanding the "overall pie," despite their superior data. This result hinges on product market competition: better data improves the matching of *all* firms with consumers, so individual advantages are eroded and profit gains are offset, limiting the ad revenue platforms can extract.[3]

The third application compares equilibrium to a planner's allocation. Two forces drive inefficiency. First, platforms compete for ad revenue, which creates business-stealing externalities. Second, because platforms charge users zero prices, they fail to account for the consumer surplus their services directly generate. Together, these forces can push ad display rates and investment in platform quality either above or below efficient levels. When consumers

[2] 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.

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

have Cobb–Douglas utility over the CES product aggregate and the CES platform aggregate, I derive a sufficient statistic[4]—based on a few interpretable quantities—that signs the efficiency of equilibrium investment.

Finally, I show in online appendices that the model extends naturally to settings with network effects, platform heterogeneity, entry, and reserve prices.

The purpose of these results is to demonstrate the model's potential as a framework for analyzing the attention economy.

The rest of the paper is organized as follows. Section 1 provides background on the attention economy and related research. Section 2 introduces and analyzes the product market in isolation. Section 3 develops the baseline model by integrating the attention market with the product market. Section 4 then characterizes the equilibrium. Sections 5–8 turn to applications beginning with policy experiments, followed by the duopoly model, data asymmetry, and welfare. Section 9 briefly summarizes further extensions and Section 10 concludes with a recap and directions for future work.

1 Background

The term *attention economy* is now widely used to describe markets where human attention is a scarce resource and the basis for economic exchange. It is often invoked in the context of digital advertising, linking it directly 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 design, particularly when interventions on one side of the market have unintended consequences on others (OECD, 2018, 2021, 2022).

The monopolistic competition setup in this paper provides a transparent benchmark for understanding their implications. While market dominance is a salient topic of policy discussion and strategic effects are important in that context, starting with this benchmark clarifies the economic mechanisms that arise purely from the operation of markets. A natural approach is to begin here to build intuition, and then turn to a numerically intensive model to examine how results depend on strategic interaction.

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

In this way, my paper complements prior work on digital advertising platforms (e.g., Bergemann and Bonatti 2023; Ambrus et al. 2016; Prat and Valletti 2021; Galperti and Perego 2022) which typically examines large platforms in static settings, and abstracts from either a microfounded product market or content provision.[5] In a similar vein, Kirpalani and Philippon (2021) study data-driven matching in a large online marketplace without content provision.

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 investment in platform quality. Both of our models predict potential over- or under-investment due to business-stealing externalities and platforms' inability to charge consumers. However, in theirs, consumers derive no surplus from advertising and there is no role for data or ad 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 their 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 a few exceptions, abstracts from the product market. A central result is that more informative data typically raises revenue (Board, 2009; Hummel and McAfee, 2016). I show in my setting that richer data can often reduce ad revenue due to spillover effects on advertisers' profits in the product market.

While the role of data in the economy has been explored before in general equilibrium models (Jones and Tonetti, 2020; Farboodi and Veldkamp, 2021), it has rarely been studied in the context of advertising.[6] A nascent literature incorporates advertising into macroeconomic models, including contemporaneous papers 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.

[5] 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.

[6] These papers study data's role as part of firm production. See also Demirer et al. (2024).

2 Simple Model of a Product Market

Our starting point is a simple monopolistically competitive product market, isomorphic to the one in Melitz (2003).[7] Though the product market is static, I embed it in continuous time because it serves as the backbone for the dynamic baseline model introduced in the next section.

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.

Consumer i has CES preferences for products given by

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

where $\sigma > 1$ is the elasticity of substitution across products. As a convenient normalization for notation, values appear with an exponent $1/\sigma$.

Values $\{v_{ij}\}$ are drawn independently across consumers and firms from a cumulative distribution function (cdf) F supported on $[0, \infty)$ with finite mean and remain constant over time.

At each time t, consumer i has I units of income that she allocates across products to maximize C_{it} taking firms' prices $\{p_{jt}\}$ as given. Firms set prices $\{p_{jt}\}$ to maximize expected profits taking as given their common constant marginal cost of production, which I normalize to 1 (the economy's numeraire). I make *no* assumptions about what firms know about consumers' values and could allow firms to set personalized prices though it turns out that firms would choose *not* to do so in equilibrium.

I now introduce the key friction into this otherwise standard model of product market: each consumer i is aware of only a subset $\Omega_{it} \subseteq \mathbb{J}$ of the firms in the economy and can consume positive amounts from this set only. 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 that these objects are common to all consumers.

Equilibrium is defined in the usual way with firms and consumers solving their respective problems and markets clearing. Proposition 1 characterizes the key properties of the equilibrium and is obtained by standard methods.

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

[7] See Footnote 7 in Melitz (2003).

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 in the Appendix. As seen from (1), consumers' CES preferences imply the optimality of a constant markup and this does not depend on consideration sets or any information firms know about consumers' values.[8] This immediately implies that if firms could charge different consumers different prices, they would not benefit from doing so. This property, though restrictive in shutting down any potential effects of personalized pricing, is ideal from the viewpoint of tractability, when I later integrate the attention market because there is no need to consider effects on prices.

Also ideal is that firms' flow profit (2) and consumer's aggregate consumption (3) depend only on the cumulative value $M_t \mu_{\Omega_t}$ of firms in consideration sets. The rest of this paper is concerned with endogenizing the evolution of M_t and μ_{Ω_t} in the attention market.

3 Baseline Model of The Attention Economy

We now build directly on the product market by integrating it with the attention market. Together, these two markets comprise the *attention economy*. In what follows, we take the microfounded relationships (1)–(3) as primitives.

The central players in the attention market are a continuum of platforms $k \in \mathbb{K}$ of mass K. Like product firms, platforms are monopolistically competitive: each provides a distinct service (e.g., streaming, social media, or podcasts) and thus has market power but is atomistic. Services are free but require attention to consume. Indeed, platforms provide them *in order to* attract attention to be able to display ads. Critically, these services are *not*

[8] The optimality of a constant markup is a well-known feature of CES demand. The Appendix explains why it still holds here with heterogeneous values for products.

advertising, but rather media through which ads can be displayed and maintaining their quality requires platform investment.

If she spends attention x_{ikt} on platform k, consumer i views ads on that platform at a Poisson intensity $a_{kt}x_{ikt}$ where a_{kt} is the platform's chosen *ad display rate*. Each ad opportunity is allocated via auction the instant it arises. Due to latency, only a fixed number N of firms, drawn uniformly at random from $\Omega_{it}^c \equiv \mathbb{J} \setminus \Omega_{it}$ (i.e., from outside the consideration set), submit bids within the instant.[9] Each auction is second price: the highest bidder wins and pays the second-highest bid to the platform. The winning firm then displays its ad, thus entering Ω_{it} at the "end of the instant." Once inside, each firm exits or is "forgotten" at an independent exponential time with rate λ_f.

When bidding, firms are endowed with independent signals of consumers' values for their own products. We can interpret these signals or "data" as being supplied by the platforms in which case all platforms have the same data. These signals give rise to posterior expectations $\{\hat{v}_{ij}\}$ of consumers' values $\{v_{ij}\}$ that are independent draws from some cdf G.[10] Expectations are *constant* across time. To avoid ties in the auctions, I assume G is continuous.

I now describe how consideration sets evolve taking as given several key objects: attention, ad rates, and bidding strategies. I then describe how these objects—together with platforms' quality levels—are endogenously determined by the decisions of consumers, firms, and platforms. Finally, I conclude with modeling remarks. Because consumers will have identical preferences for platforms and thus allocate attention in the same way, in what follows, I drop the index i on x_{ikt} whenever convenient.

Evolution of Consideration Sets The key state variables of the economy are the mass M_t and cdf H_t of expected values, of firms in Ω_{it} which are common to all consumers. Because expectations are unbiased and independent, the mean μ_{H_t} of H_t must coincide with the average value μ_{Ω_t} of firms in Ω_{it}, by the law of large numbers (LLN).[11] Because $M_t\mu_{\Omega_t}$ is a sufficient statistic for firms' flow profits (2) and aggregate product consumption (3), there is no need to track the true value distribution.

Figure 1 below summarizes how these state variables evolve. The cdf H_t^c corresponds to expected values in Ω_{it}^c and is pinned down given H_t—because the mass of firms with expected values less than $\hat{v}$ in both Ω_{it} and Ω_{it}^c must

[9]Firms only search for consumers that they are not already selling to. In the model, there is no benefit to a firm of displaying an ad to a consumer i if it is already in Ω_{it}.

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

[11]I assume the LLN applies but expect this can be formalized as in Duffie et al. (2025).

coincide with that of the whole economy,

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

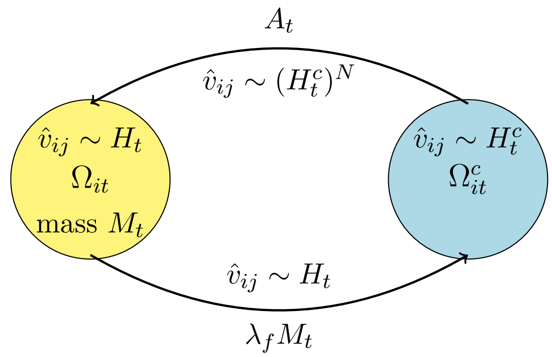


Figure 1: Evolution of Consideration Sets

As depicted in Figure 1, the rate of inflow into Ω_{it} is

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

after aggregating across all platforms. Suppose, as will be the case in equilibrium, that the winning firm in each auction has the highest expected value for the consumer. Then the cdf of expected values of inflowing firms is $(H_t^c)^N$ where $(H_t^c)^N(\cdot) \equiv H_t^c(\cdot)^N$. Because firms are forgotten "at random," the outflow has rate $\lambda_f M_t$ and cdf H_t. In sum,

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

All of this assumes the LLN. Taking limits as $\hat{v} \to \infty$ yields[12]

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

Together, (4) and (6) fully specify the dynamics of the state variables. We turn now to agents' decision problems.

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

Firms' Bidding Problems Anticipating the future paths of M_t and H_t, firms bid in the auctions to maximize the present-discounted value (PDV) of their flow profits, taking as given the bidding strategies of their rivals. A formal description of the bidding problem is in the Online Appendix A. Here, I give an informal recursive description.

Because auctions are second price, a standard dominance argument implies that a firm j's optimal bid in an auction for a consumer i with $\hat{v}_{ij} = \hat{v}$ must be the *gain* in continuation value from entering Ω_{it}:

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

Here, V_t^{In} (V_t^{Out}) represents the PDV of flow profits from selling to the consumer in the future if the firm is currently in Ω_{it} (Ω_{it}^c). While outside Ω_{it}, the firm enters an auction for consumer i at a Poisson intensity $\lambda_{et}(\hat{v})$ to be described shortly. Given discount rate $\rho > 0$, V_t^{In} and V_t^{Out} must satisfy the Hamilton-Jacobi-Bellman (HJB) equations

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

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

Above, $B_t^{(1)}$ denotes the highest bid among the $N-1$ other bidders in an auction which determines the ad price in the event that firm j wins. If B_t is increasing, then $B_t^{(1)} \overset{d}{=} B_t(\hat{v}^{(1)})$ where $\hat{v}^{(1)} \sim (H_t^c)^{N-1}$.

In that case,

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

where the first term follows from accounting: the total rate of auction entry NA_t is split among the mass $J - M_t$ of firms outside Ω_{it}.

Platforms' Ad and Investment Decisions Firms' bids determine ad prices and thus platforms' incentives. Given B_t, the expected ad price in any auction is $\pi_{\mathbb{K}t} \equiv \mathbb{E}\left[B_t\left(v_t^{(2)}\right)\right]$ where $v_t^{(2)}$ is the second highest of N independent draws from the cdf H_t^c. Anticipating the future path of $\pi_{\mathbb{K}t}$, platform k sets ad display a_{kt} and investment ℓ_{kt} to solve

$$\Pi_{\mathbb{K}} \equiv \max_{\{a_{kt}\},\{\ell_{kt}\}} \int_0^\infty e^{-\rho t}\left(\pi_{\mathbb{K}t} \underbrace{a_{kt}x_{kt}(a_{kt}, q_{kt})}_{\text{total ad display rate}} - \ell_{kt}\right) \mathrm{d}t \tag{12}$$

subject to the law of motion of quality

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

which starts at an initial level, common to all platforms, of $q_{k0} = q_0$. I assume $\varphi < 1$ so there are decreasing returns. Absent investment, quality depreciates at rate δ as the platform's content grows stale or less relevant over time. In (12), I have explicitly indicated the dependence of consumer attention on the platform's quality and ad display rate. To complete the model, I now turn to consumers' preferences and attention choices.

Consumers' Attention Choices Consumer i's lifetime utility from both products and platforms is

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

where $u(\cdot)$ is increasing and

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

is a CES aggregate of platforms' services. Above, $\nu(\cdot)$ is positive and decreasing so ads are nuisances that reduce effective platform quality. The parameter $\epsilon > 1$ is the elasticity of substitution across platforms.

At each time t, the consumer has a unit of attention to spend on platforms. She does *not* internalize how platform use affects $\{C_{it}\}$ through the evolution of her consideration set. In equilibrium, this myopia is without loss of optimality: because all platforms set the same ad display rates, how she divides attention among them has no meaningful effect. She thus maximizes flow utility $u(C_{it}, X_{it})$, which, since platforms charge zero prices, reduces to

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

Since u does not appear in (14), the marginal utilities for products and for platforms are decoupled in equilibrium, a key source of inefficiency analyzed in Section 8.

Equilibrium Concept Having laid out the baseline model, I now define equilibrium. Recall that the product market relationships (1)–(3) are treated as primitives.

Definition 1. *An* equilibrium *of the economy with initial condition* (M_0, H_0, q_0) *is a collection of processes* $(\ \{M_t\},\ \{H_t\},\ \{q_{kt}\},\ \{B_t(\cdot)\},\ \{a_{kt}\},\ \{\ell_{kt}\},\ \{x_{ikt}(\cdot)\ \})$ *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 makes a number of simplifying assumptions for a transparent benchmark—several are relaxed in later sections or in online appendices, as summarized in the introduction. Beyond these, one can show that results apply to any standard auction format by revenue equivalence, and if firms can infer preferences from purchase histories, the environment converges to one with perfect information (i.e., $\hat{v}_{ij} = v_{ij}\ \forall\ i, j$).

An important modeling choice is to allocate ads through dynamic auctions. This both reflects the microstructure of digital markets and keeps the model solvable 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 ensures that each firm participates in at most one auction per consumer at a time, supporting equilibrium existence. Similar difficulties arise in static Walrasian formulations, especially with asymmetric data (see Online Appendix F). Moreover, the dynamic structure is not merely a technical fix: it allows the analysis to separate the short from long run, capture investment incentives more faithfully, and deliver long run predictions that do not depend on a choice of initial conditions.

4 Equilibrium of the Baseline Model

I briefly sketch the derivation of the equilibrium and then present its main properties, characterized almost explicitly.

The derivation, located in Online Appendix A, proceeds as follows. First, I solve consumers' attention allocation problem (14), which 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, a platform's attention depends on its effective quality–which reflects both service quality and ad load–relative to that of its competitors.

Second, given (15), I solve for each platform k's ad rate which is chosen to maximize flow profit:

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

The choice of ad rate is a *static problem* because no individual platform has a measurable impact on consideration sets. Similar to product prices, 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.[13]

Fourth, given *arbitrary paths* for M_t and H_t, I solve (8)–(10) *explicitly* for optimal bid functions $B_t(\cdot)$. Since each $B_t(\cdot)$ is increasing, this verifies that (6), used in the previous step, is indeed valid and uniquely so.

Fifth, from $B_t(\cdot)$, I calculate the average ad price $\pi_{\mathbb{K}t}$ and then use the maximum principle to characterize platforms' investment and quality paths via a simple boundary value problem, solvable numerically in seconds.

Theorem 1 characterizes the resulting equilibrium, linking attention, ad display, and quality to product market outcomes. For brevity, I report only steady state properties; transition dynamics are in the Appendix.

Theorem 1. *Suppose A is the unique solution to $\max_a a\nu(a)^{\epsilon-1}$. If $A/\lambda_f < J$ and $\epsilon - 1 < 1/\varphi$, then a unique equilibrium exists for any feasible initial conditions (M_0, H_0, q_0), characterized by Appendix equations (29)—(33).[14] 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 of expected values of firms in Ω_{it} satisfies $H(\cdot) = H^c(\cdot)^N$ where H^c uniquely solves*

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

[13] That is, we do not need to first derive attention and platforms' quality levels. This simplicity relies on platform symmetry. E.g., if platforms have asymmetric data, as studied in Section 7, the evolution of H_t depends on how attention is allocated which introduces another fixed point into the equilibrium derivation.

[14] 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 $\mathbb{J}$ where with abuse of notation, H_0 and G denote the associated probability laws rather than the cdfs.

4. *The total ad revenue is* $\pi_{\mathbb{K}} A$*, where*

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

is the average ad price. Above, $\lambda_e(\cdot)$ *is the intensity of entry into* Ω_{it} *(from (11)), and* $\pi_{\mathbb{J}}$ *is the coefficient on firms' flow profits (from (2) evaluated at* $\mu_{\Omega_t} = \mu_H$*).*

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. *Each consumer allocates attention (15) as a function of platforms' ad rates and quality levels.*

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 flipped inequality an equilibrium does not exist. Similarly, the condition $A/\lambda_f < J$ ensures consideration sets do not grow to contain all firms, so that (6) applies.

To preview mechanisms central to the next section on policy experiments, I highlight three implications of Theorem 1 related to platform market power and data.

Remark 1. *Part 6 shows that the key parameter governing platform market power in providing services is* ϵ*. As* $\epsilon \to 1$ *platforms hold captive an equal share of attention regardless of quality or ad display rate. Moreover, as* ϵ *decreases, investment decreases holding all else constant (Part 5), and the ad display rate* $A = \max_a a\nu(a)^{\epsilon-1}$ *increases (Part 1).*

Remark 2. *Part 4 shows that search frictions are a source of platform market power in selling ads: holding all else constant, as the intensity* $\lambda_e(\cdot)$ *of firms' entry into consideration sets rises pointwise, ad revenue falls and eventually vanishes.*

Remark 3. *Part 3 shows that data shapes the composition* H *of consideration sets through* G *(the cdf of posterior expectations in all of* $\mathbb{J}$*). Through* H*, data in turn affects product consumption (Part 7), ad revenue (Part 4), investment and platform quality (Part 6), and aggregate platform consumption (Part 7).*

With the equilibrium characterized, the rest of the paper turns to various applications, beginning with policy experiments.

5 Policy Experiments

I study two regulations often discussed in debates on the attention economy: data policies and interoperability mandates. This section shows how cross-market spillovers can overturn intuitive predictions, produce short-run versus long-run reversals, and create tradeoffs between product consumption and platform quality. The stylized analysis here is intended to highlight mechanisms and yields valuable intuition for market conditions when these patterns are more likely to arise. All formal results below apply on the parameter domain where a unique equilibrium exists.

5.1 Data

I begin with data policies, considering first steady state and then later transition dynamics. These policies affect the informativeness of the data platforms have access to. Data can range from uninformative—all firms share the prior mean μ_F—to perfectly informative, where posteriors coincide with true values ($G = F$). Between these two extremes, informativeness can be ranked by how dispersed posteriors are; formally, by the mean-preserving spread order (Blackwell, 1953).[15]

Steady State Proposition 2 implies that banning data use or equivalently, ad targeting, always *increases* ad revenue and platform quality, but at the cost of lower product consumption.[16]

Proposition 2. *In steady state, as data becomes uninformative (i.e., G converges 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, relative to all G dominated by F in the mean-preserving spread order. Moreover, aggregate product consumption is monotone in informativeness.*

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

[16] Because G is assumed continuous to rule out ties, "uninformative data" is understood in the limiting sense.

Surprisingly, ad revenue is *highest* when data is uninformative. The intuition rests on three points.

1. Total firm profits are constant at I/σ regardless of G because firms charge fixed markups and budget constraints bind.

2. With uninformative data, Bertrand competition in ad auctions extracts these profits fully.

3. Informative data creates *information rents*, reducing ad revenue.

Point 1 is an artifact of CES preferences, but there is suggestive evidence that the "pie" is in effect, fixed: advertising's share of GDP has remained remarkably constant across evolving advertising technologies since 1925.[17] There is also suggestive evidence more informative data can reduce ad revenue: advertising's share of GDP declined during the period when digital advertising grew in dominance (Silk et al., 2021). I revisit this second phenomenon in Section 7.

Although global monotone comparative statics are hard to prove, more informative data *typically* reduces ad revenue. Intuition can be gleaned from the expression for ad revenue in Part 4 of Theorem 1. Three effects drive the result:

1. Higher match rates for firms with high expected values ($\lambda_e(\hat{v})$ rises).

2. Lower average value of firms outside consideration sets who bid ($\mu_{H_t^c}$ falls).

3. A reduction in firms' profit coefficient $\pi_{\mathbb{J}}$ as μ_H rises.

These offset the effect of a higher average value of winning firms on ad revenue. In steady state, they tend to dominate, particularly because of the third effect. To see this, consider a firm's expected flow profit from selling to a consumer:

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

More informative data raises the average $\hat{v}$ of winning bidders, but it also raises μ_H, the average $\hat{v}$ in consideration sets. In *steady state*, these must coincide. Thus, better match quality has *no net effect* on winning firms' expected profits. Because auctions are second price, ad revenue instead depends

[17] As Chemi (2014) puts it,"the pie is not growing...the easiest way to make more money is to steal larger slices of the pie."

on the expected profit of the second-highest bidder—formally, the ratio of the average second-highest $\hat{v}$ in auctions to μ_H. This ratio reaches its maximum of one when data is uninformative and typically, *though not always*, falls as data becomes more informative.

Transition Dynamics The steady-state analysis suggests that transition dynamics may be important since the second and third effects that depress ad revenue unfold only *gradually* as the average value of firms in consideration sets evolves. Figure 2 shows the path between steady states after an unanticipated increase in data informativeness at $t = 1$. Specifically, G shifts from U$[.2, .8]$ to U$[0, 1]$, representing, for instance, a relaxation of privacy laws that allows platforms to target with more granular data. Figure 6 in the Appendix explicitly constructs the garbling of new data into old data that corresponds to this change in G.[18]

As shown in Figure 2, ad revenue initially *rises* but then falls below its starting level. Platforms may appear better off at first, but in the long run they are *worse off* (as profits decline despite lower investment). Since the rise is short-lived and platforms are forward-looking, investment weakens and quality declines steadily, while consumption rises monotonically. Hence, the welfare effect is ambiguous and depends on u.

Three remarks are in order.

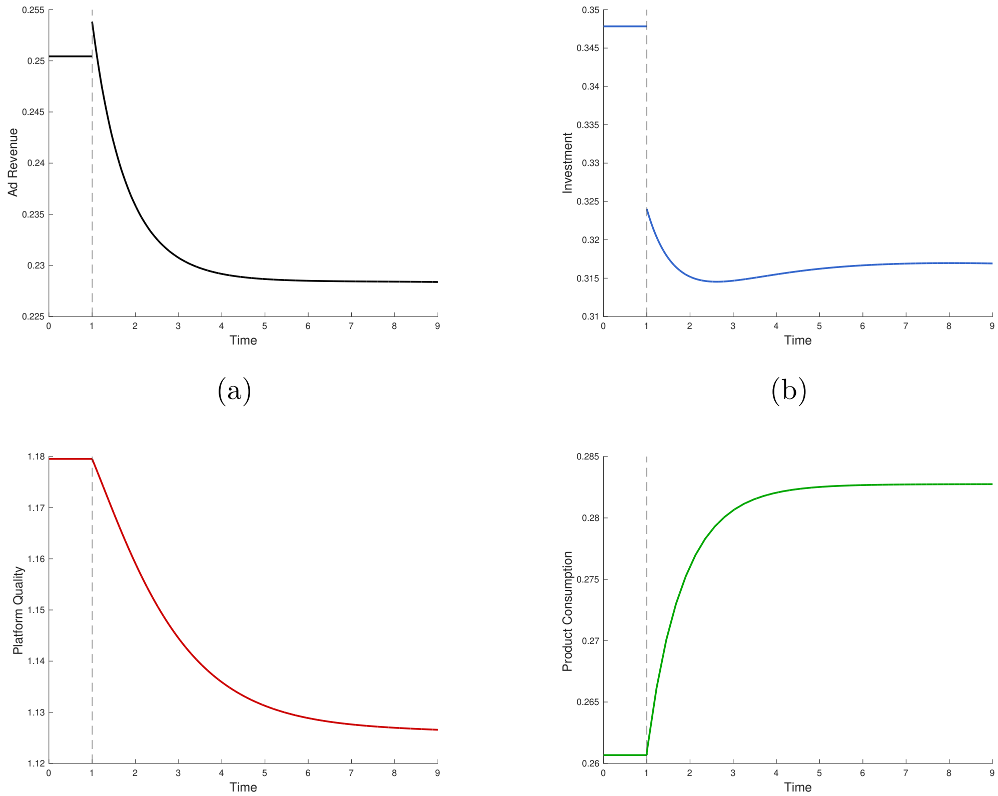
Remark 4. *Patterns like those in Figure 2 appear more likely when firms are less patient, platforms are more patient, and the data change is large.*

Remark 5. *In other settings where prices may adjust, total firm profits (gross of advertising costs) may vary with data informativeness but the comparative statics here may still carry over. For instance, Rhodes and Zhou (2022) show that under "full market coverage," personalized pricing reduces firm profits—reinforcing the direction of the steady-state results.*

Remark 6. *Individual platforms still typically benefit privately from more informative data, even though they are collectively worse off. I provide an illustrative example in Section 7.*

Having examined how changes in data affect market outcomes, I now turn to interoperability, which operates through a different channel.

[18]The details of the garbling—i.e., the joint distribution of old and new expectations—do not affect the steady state, but they do matter for transition dynamics since they determine the initial cdf of expectations in consideration sets following the shock.



(a) (b)

(c) (d)

Figure 2: Transition dynamics of a shock to G

Notes: I plot the transition between steady states following an unanticipated change in G from $\mathrm{U}[.2, 8]$ to $\mathrm{U}[0, 1]$ at $t = 1$ for the garbling in Figure 6. Other parameters are $\rho = .1$; $\epsilon = 1.33$; $\lambda_f = 1$; $K = .1$; $J = 1$; $\sigma = 3$; $N = 5$; $I = 1$; $\varphi = .5$; $\delta = .5$; $\nu(a) = 1 - 7.5a$.

5.2 Interoperability

Interoperability mandates aim to reduce user lock-in and promote competition by making attention more sensitive to platform quality. An example is a content-interoperability rule requiring platforms to allow content sharing across services. In such cases, with content more similar across platforms, differences in interface quality or ad load play a larger role in attracting attention.

In practice, interoperability mandates involve technological changes that could modify the aggregator X or even the utility function u on a case-by-case basis. The results below do not take a stand on how such changes manifest and assumes only that *attention* retains its CES shape and is *more elastic*, as captured by higher ϵ, following the policy.[19]

Steady State Proposition 3 presents the model's predictions for the 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. Higher ϵ makes it easier for platforms to attract attention from rivals, lowering the ad display rate A and promoting investment in quality q.

2. Lower A shrinks consideration sets reducing consumption C.

3. Despite lower A, higher ad prices $\pi_{\mathbb{K}}$ raises ad revenue $\pi_{\mathbb{K}}A$, further reinforcing investment in q.

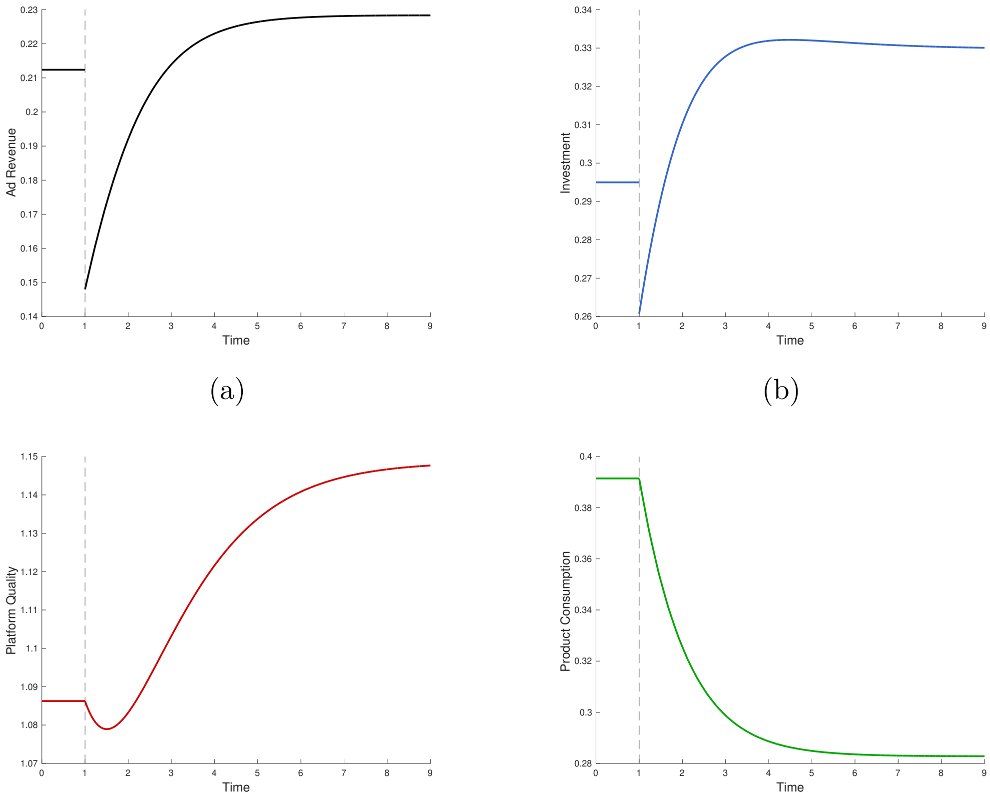
Parts 1 and 2 are intuitive but Part 3 is less so. To offer intuition, consider the expression for the ad price $\pi_{\mathbb{K}}$ in Part 4 of Theorem 1. Holding fixed H_t, the coefficient $\pi_{\mathbb{J}} \propto 1/A$ because the total mass of ads satisfies $M = A/\lambda_f$ (Part 2 of Theorem 1)—firms' profits are higher when consideration sets shrink, raising ad prices and offsetting the direct effect of lower A on ad revenue $\pi_{\mathbb{K}}A$. In addition, lower A reduces match rates $\lambda_e(\cdot)$, further supporting higher ad prices. Although the cdfs H_t and H_t^c are also affected by a change in A (Part 3 of Theorem 1), I nevertheless verify via a brute-force calculation in Online Appendix B that ad revenue increases.

The increase in $\pi_{\mathbb{K}}$ is gradual, as it operates through the dynamics of M_t and H_t/H_t^c, whereas the effect on the ad rate A is immediate. This contrast suggests that ad revenue may *decline* at first, making transition dynamics important.

Transition Dynamics A one-time 1% increase in ϵ at $t = 1$ from its initial value of 1.33 generates the patterns in Figure 3: ad revenue and quality initially fall as A drops, then recover in the long run; consumption falls persistently. Thus, as with data, short-run and long-run effects differ and the welfare impact is ambiguous and depends in u.

I conclude this section with two remarks.

[19] For example, the aggregator X may be scaled or translated following the change.



(a) (b) (c) (d)

Figure 3: Transition dynamics of a shock to ϵ

Notes: I plot the transition between steady states following an an unanticipated increase in ϵ from 1.33 to 1.33(1.01) at $t = 1$ for parameters $\lambda_f = 1$; $K = .1$; $J = 1$; $\rho = .1$; $\sigma = 3$; $N = 5$; $G = \text{U}[0, 1]$; $I = 1$; $\varphi = .5$; $\delta = .5$; $\nu(a) = (1 - .8395a^{.01})^{63.1472}$.

Remark 7. *The patterns in Figure 3 appear more likely to arise when firms and platforms are less patient and nuisance costs are steeper.*

Remark 8. *If there is an alternative entry channel into consideration sets (e.g., traditional media), higher ϵ might reduce steady-state platform ad revenue, since the mass of consideration sets no longer shrinks in proportion to platforms' ad display rate. This logic does not apply when the alternative entry channel is word-of-mouth if we assume word-of-mouth must first originate from an ad.*

6 Duopoly

Our analysis has abstracted from strategic effects to focus on cross-market spillovers and dynamics. Here, I show the framework can accomodate strategic interaction in a variant with duopolistic platforms and develop a simple algorithm to compute Markov equilibrium. The model takes ad display rates as exogenous, which keeps the focus on investment incentives while preserving tractability. I use the duopoly model to test the robustness of the previous section's findings. This is especially relevant since interoperability mandates often target the market power of large platforms. I discuss the case of endogenous ad rates at the end of the section.

6.1 Setup

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

$$\mathrm{d}q_{kt} = (\ell_{kt}^{\varphi} - \delta q_{kt})\,\mathrm{d}t + \eta q_{kt}\,\mathrm{d}B_{kt}, \quad q_{k0} = q_0, \tag{19}$$

where B_{1t} and B_{2t} are independent standard Brownian motions. This law of motion is identical to that of the baseline except for the addition of noise. Some noise is needed for numerical stability and, I suspect, for existence of a Markov equilibrium. I conjecture any nontrivial amount ($\eta \neq 0$) suffices; in the numerical examples below, I set $\eta = 0.12$, a modest level.[20] Injecting noise to ensure stability or restore existence also appears in other related dynamic duopoly models (e.g., Budd et al. 1993, Harris et al. 2010).[21]

To simplify, I assume both platforms display ads at a common exogenous rate A. Except for the above changes, I retain all other aspects of the baseline model.

6.2 Markov Equilibrium

I seek an equilibrium that is Markov in the quality levels of the two platforms and time. In that case, platform k's value function, $V_k(\cdot)$, takes q_1, q_2 and t as arguments.

[20] Budd et al. (1993), who analyze a similar model, claim in their Section 3 that any nontrivial amount of noise ensures equilibrium existence. I elaborate on the connection between our models in Online Appendix C.

[21] In such settings, deterministic formulations often fail to admit Markov equilibrium because, under certain Markov strategy profiles, an arbitrarily small lead can yield large payoff gains for a player. Introducing noise can restore equilibrium existence by softening these leader-follower dynamics. I suspect the same mechanism operates in my model.

Suppressing arguments for notational simplicity, value functions satisfy the coupled HJB system

$$\rho V_k = \sup_{\ell_k} \mathcal{L}^{(\ell_1,\ell_2)} V_k + \pi_{\mathbb{K}t} A \frac{q_k^{\epsilon-1}}{q_1^{\epsilon-1} + q_2^{\epsilon-1}} - \ell_k, \quad k \in \mathbb{K} \tag{20}$$

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

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

Because all equilibrium objects other than investment coincide with the baseline model, solving for a Markov equilibrium reduces to finding value functions and investment policies that jointly satisfy (20).

Algorithm Overview I solve (20) using an implicit upwind finite-difference scheme. There are two steps.

1. Compute steady-state value functions and controls by fixing the steady-state ad price (in Theorem 1) and iterating the discretization of (20) backward until convergence from a terminal guess of value functions.

2. Compute value functions and controls along the transition path by iterating backward from steady state.

The algorithm is efficient, stable, and generalizable. Details and mathematical foundations are discussed in Online Appendix C.

6.3 Policy Experiments Under Duopoly

I conduct an analysis similar to that of Section 5, to investigate robustness to strategic interaction along the investment margin. I focus on the case where *realized* Brownian shocks are all zero, which is representative given the modest value of $\eta = .12$. As in the baseline, steady state is defined by constant investment and quality. I plot transition dynamics for investment and quality between steady states following parameter shocks.[22]

In Figure 4, panels (a)–(b) show the effects of a data policy where G is shocked from U$[.2, 8]$ to U$[0, 1]$. Panels (c)–(d) show the effects of an interoperability policy that simultaneously shocks the elasticity ϵ (from 1.33 to 1.33(1.01)) and the ad display rate A (from .2 to .1). This shock to A would

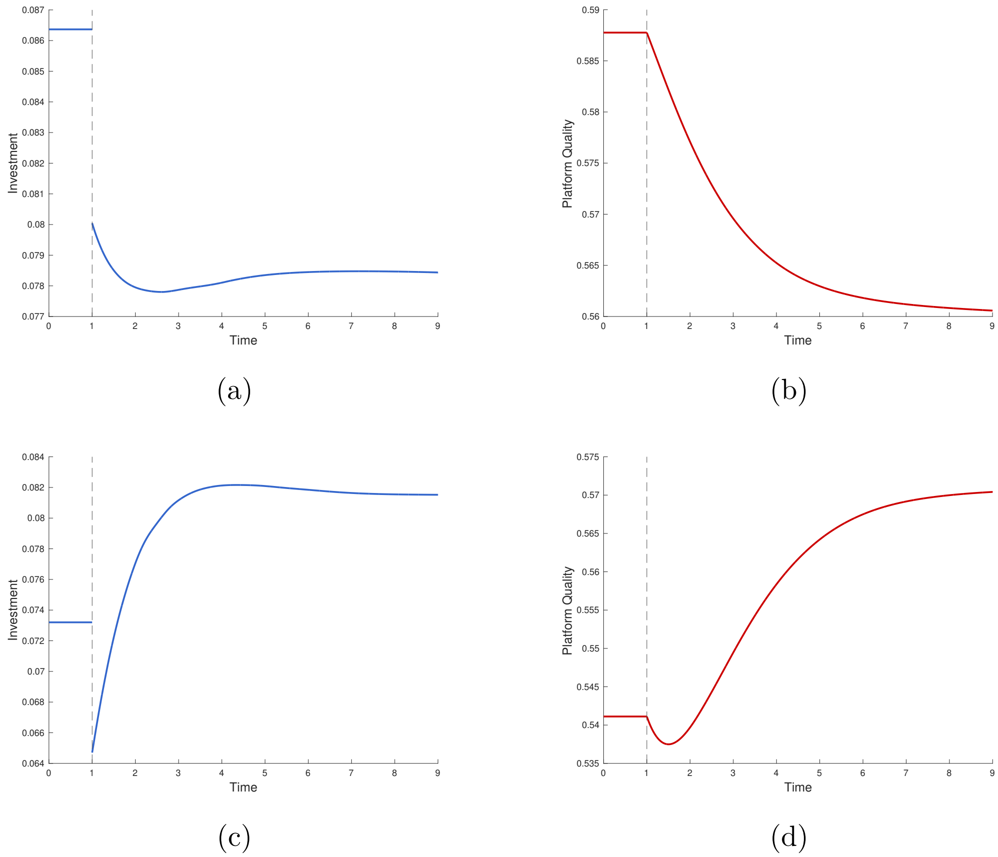


Figure 4: Transition dynamics in the duopoly model

Notes: Panels (a)-(b) depict dynamics of a shock to G from U$[.2, 8]$ to U$[0, 1]$ for parameters as in Figure 2 except with $I = 10$. Panel (c)-(d) does so for a simultaneous shock to ϵ from 1.33 to 1.33(1.01) and A from .2 to 1 for parameters as in Figure 3 except with $I = 10$.

arise endogenously if platforms behave as in the baseline model. This lets the analysis isolate strategic effects from investment.

Across both experiments, the qualitative comparative statics closely mirror those in the monopolistically competitive case, suggesting that the key economic mechanisms remain robust.

[22] I omit consumption and ad revenue, whose paths are identical to those in the baseline model up to a constant scaling factor.

6.4 Comment on Endogenous Ad Display Rates

A limitation of the analysis is that the ad rate A is exogenous. A natural extension would be to endogenize it. One plausible assumption is that platforms commit to ad rates at $t = 0$, adjusting only in response to policy shocks. This seems reasonable since users typically observe ad loads only after engaging with a platform, implying *some* commitment power.

Under this assumption, a positive shock to the elasticity ϵ should induce each platform to lower its ad rate once (at the time of the shock) to attract attention—just as in the baseline model. Strategic retaliation may dampen this effect but seems unlikely to eliminate it in which case equilibrium outcomes remain broadly similar for an appropriate choice of $\nu(\cdot)$.

Formally solving for endogenous ad rates is difficult because it would require keeping the entire cdf H_t as a state variable.

This challenge can be avoided by modeling a continuum of *duopolistic* platform markets, e.g., with a nested CES structure. Then ad display remains a static problem and no individual platform affects the (deterministic) law of motion of consideration sets. Yet, there is strategic interaction in both investment and ad display and platforms are large under a narrower market definition. The algorithm here seems readily extendable, at least for steady state. In this setting, the analog of the interoperability counterfactual is to raise the elasticity in each market, reflecting the cumulative effect of many mandates.

7 Data Asymmetry

So far, platforms have shared the same data, and perhaps surprisingly, we found that in some cases they may *collectively* prefer coarser data. Here, I extend the baseline to allow asymmetry between two groups of platforms, linking data informativeness to relative market shares and quality levels. In this setting, *individual* platforms often gain from more granular data even when, collectively, they do not. Building on this, I offer a mechanism for the puzzling decline in advertising's share of GDP during 2010–2018, when *data-rich* digital platforms rose to prominence (Silk et al., 2021).

The dynamic auction structure plays a crucial role in keeping the analysis here tractable as discussed in Subsection 3.1. To my knowledge, relatively little work has studied platform competition with asymmetric data.[23]

[23] Exceptions include Ichihashi (2021), Greenwood et al. (2025), and Bergemann and Bonatti (2011). The latter two study perfect competition between traditional and digital media.

7.1 Setup

There are two groups of platforms indexed by $z \in \{1, 2\}$. The mass of platforms in group z is 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 that $(v_{ij}, \zeta_{1ij}, \zeta_{2ij}) \sim Q$ defined on $[0, \infty) \times \mathbb{R}^2$, independently across i and j. Let G be the joint cdf of ζ_{1ij} and ζ_{2ij} derived from Q. I assume G has a continuous density g. I retain all other aspects of the baseline model.

7.2 Equilibrium

The equilibrium derivation is in Online Appendix D. Now, one must track the *joint distribution* H_t of signals in consideration sets. Unlike before, the law of motion of H_t depends on the attention shares allocated across the two platform groups—shares that themselves depend on platform investment, which in turn depends on ad revenue, which depends on H_t. A firm's bidding strategy now also depends on which platform group hosts the auction. Nevertheless, the model remains parsimonious: I provide a simple algorithm to compute steady-state equilibrium efficiently.

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

1. Guess a value for the attention share x_1 allocated to group 1 platforms.
2. Fixing x_1, iterate the law of motion for H_t (Online Appendix eq. (43)) forward until convergence to steady state.
3. Compute steady-state bid functions and average ad prices by iterating a contraction map (Online Appendix eq. (46)).
4. Given the ad prices, platform quality levels and the attention share x_1 are explicit (Online Appendix eq. (48)).
5. If x_1 matches the initial guess, a steady-state equilibrium is found. Otherwise, revise the guess and repeat.

Iterating over a grid of guesses for x_1 is a relatively fast way to solve for all steady-state equilibria and verify uniqueness.

In Greenwood et al. (2025) targeting data distinguishes college from noncollege consumers in "separate spheres of economic activity." In Bergemann and Bonatti (2011) targeting ability is captured quite differently by the joint distribution of preferences and media use.

7.3 Puzzling Empirical Pattern

I will now show that the model can generate patterns consistent with 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)

The model's ability to do this hinges on the effect of digital platforms' more informative data on competition in the product market (i.e., the effect of a rise in μ_{H_t} on firms' profits) which limits the surplus platforms can extract from firms. It thus provides further support for the mechanism identified in Section 5 where it was shown that more informative data can reduce ad revenue. Though there may be other explanations, this "puzzling" stylized fact emerges as a natural consequence in the model.

To illustrate, suppose consumer i's value for product j is log-normal: $v_{ij} = e^{Z_{ij}}$ where $Z_{ij} \sim N(0, \sigma_Z^2)$. There are two groups of platforms of equal mass. Firm j sees signals

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

and

$$\zeta_{2ij} = Z_{ij} + u$$

where $0 \leq \Delta \leq 1$ and $u \sim N(0, \sigma_u^2)$ is independent of all other model primitives. Thus, whenever $\Delta < 1$, group 1 platforms are data rich and intended to represent modern digital platforms while group 2 platforms are data poor and represent traditional platforms with limited targeting ability.

Figure 5 plots the steady-state share of ad revenue accruing to group 1 platforms as a function of their data advantage, measured by $1 - \Delta$, for the parameters listed below the figure. The left-most point on the x-axis ($\Delta = 1$) corresponds to the case when both platform groups comprise only traditional media. I note that both total ad revenue and its split between the two groups are invariant to the total mass of platforms in the economy. Thus, entry, without a difference in data, can *not* generate the patterns in the figure. Moreover, when comparing $\Delta = 1$ with another point, the figure is *consistent with any pattern of entry of digital platforms and exit of traditional platforms*, provided the new steady state contains equal shares of each. Consistent with the stylized fact, group 1 platforms capture a higher share of the ad revenue and this share is increasing in their data advantage yet *total* ad revenue decreases. In

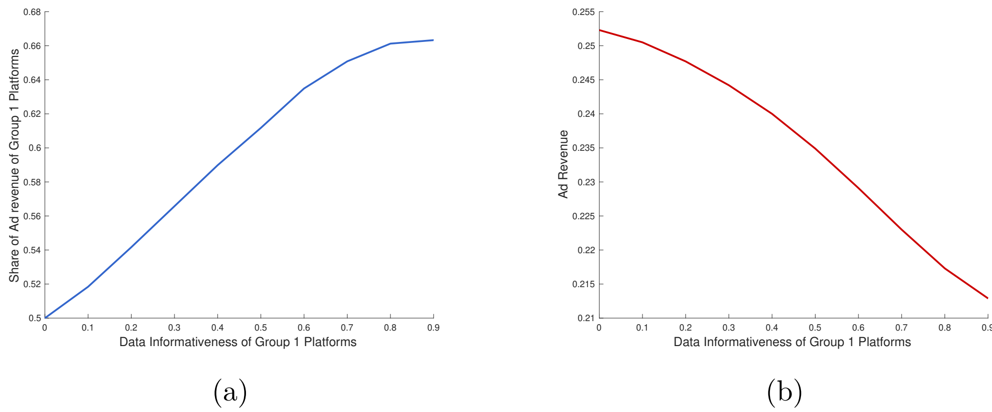


Figure 5: Group 1 share of total ad revenue and total ad revenue

Notes: The group 1 share of ad revenue and total ad revenue is plotted as a function of the informativeness of group 1 platforms' data (measured by $1-\Delta$) for parameter values $\sigma_Z^2 = .5$; $\sigma_u^2 = 2$; $I = 1$; $A = .01$; $\lambda_f = 1$; $F = .1$; $N = 20$; $\rho = 1.6$; $\sigma = 3$; $\varphi = .75$; $\nu(a) = 1 - 62.5a$. The figure applies for any measure K of platforms.

this example, the higher share of ad revenue results from both higher ad prices and shares of attention. This shows that individual platforms gain from more informative data even when, collectively, platforms do not.

8 Welfare Analysis in General Equilibrium

Thus far, we have studied policies but left open the question of which inefficiencies are there to begin with. Here, I show how to close the economy so welfare is simply consumer surplus. Then, I compare the equilibrium to the planner's economy to address this question. Specializing to Cobb-Douglas utility u, I find that equilibrium ad display and investment rates may be either too high or too low and derive a sufficient statistic that can sign the investment distortion. While this statistic relies on monopolistically competitive platforms, the planner's solution extends readily to oligopoly settings.

8.1 General Equilibrium

To close the baseline model, I endogenize consumers' incomes. Each consumer supplies L units of labor inelastically at each time. Labor is the only productive input, used by platforms for investment and by firms for production. Accordingly, ℓ_{kt} now denotes the labor hired by platform k for investment.

Each unit of output requires one unit of labor so a firm's marginal cost of production is the wage, which I normalize to one.

The labor market clears at time t if

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

where $\ell_{jt} = \int_{\mathbb{I}} c_{ijt} \mathbb{1}_{\{j \in \Omega_{it}\}} \, \mathrm{d}i$ is the total labor hired by firm j.

Consumers are assumed to own equal shares of all firms and platforms in the economy. Therefore, since platforms only extract ad revenue from firms and all other costs are wages, income I_t must equal total 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 with the addition of a process $\{I_t\}$ for income and the condition that the labor and product markets clear at all times.

Proposition 4 is the general equilibrium analog of Theorem 1, but for brevity reports only steady-state investment.

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*

$$\ell_{\mathbb{K}} = \frac{\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathbb{K}} A(\epsilon - 1)}{\rho + \delta + \varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathbb{K}} A(\epsilon - 1)} \frac{L}{K} \tag{21}$$

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

In what follows, I compare equilibrium ad and investment rates to those chosen by the planner.

8.2 Planner's Problem

The planner sets platforms' ad display and investment rates to maximize welfare taking as given that consumers set their demands to maximize flow utility and firms set prices and bid to maximize the PDV of flow profits. I assume that the planner is constrained to treat platforms symmetrically, has the same consumer data as platforms, and cannot enforce technological changes such as interoperability. Nonetheless, within the model, this benchmark remains valuable for evaluating changes in data or interoperability, whose welfare effects often depend on how the equilibrium deviates from the planner's outcome ex ante (see Remark 9). While I could allow the planner to set firms' prices or bids, it turns out that these are already efficient.[24]

Formally, given initial conditions (M_0, H_0, q_0), the planner solves

$$\max_{\{\ell_{\mathbb{K}t}\},\{A_t\}} \int_0^\infty e^{-\rho t} u(C_t, X_t)\, \mathrm{d}t \tag{22}$$

subject to

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

with M_t and H_t evolving according to (4)–(6), and $q_{\mathbb{K}t}$ following (13) given $\ell_{\mathbb{K}t}$.[25]

In what follows, I restrict attention to the steady-state solution.[26]

Definition 2. *A* steady-state solution *of the planner's problem is defined by initial conditions (M^*, H^*, q^*) and constants A^*, $\ell_{\mathbb{K}}^*$ such that the planner solves (22) by setting $A_t = A^*$ and $\ell_{\mathbb{K}t} = \ell_{\mathbb{K}}^*$, with $M_t = M^*$ and $H_t = H^*$ for all t.*

To obtain a sharp characterization, I assume

$$u(C_t, X_t) = C_t^{1-\tau} X_t^{\tau} \tag{23}$$

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

Theorem 2, proven in Online Appendix E, characterizes the planner's steady state investment for an arbitrary discount rate ρ and her ad display rate in the limit as $\rho \to 0$.[27]

Theorem 2. *Let u be as in (23). Then, any steady-state solution of the planner's problem has investment*

$$\ell_{\mathbb{K}}^* = \frac{\varphi\delta\frac{\tau}{1-\tau}}{\rho + \delta + \varphi\delta\frac{\tau}{1-\tau}} \frac{L}{K}. \tag{24}$$

[24]Though firms charge markups, they each produce the same amount which is efficient given the concavity of the CES aggregator and ex-ante symmetry. Further, bids are efficient because they are monotone.

[25]In the expression for C_t, $L - K\ell_{\mathbb{K}t}$ is the labor left over for production.

[26]I conjecture that under technical conditions, any solution to the planner's problem converges to the steady-state solution but do not investigate this here.

[27]The solution for the ad display rate for arbitrary ρ can be characterized using Pointryagin's Maximum Principle but is less clean to state.

Suppose that a steady-state solution exists for all ρ in a neighborhood of zero. Let $A^*(\rho)$ denote the ad display rate in the steady-state solution when the discount rate is ρ. Then

$$\lim_{\rho\to 0} A^*(\rho) = \arg\max_a \left[a\mu_H(a)\right]^{\frac{1-\tau}{\sigma-1}} \nu(a)^\tau \tag{25}$$

whenever the right-hand side is well defined. Here $\mu_H(a)$ denotes the steady-state average value of firms in consideration sets when all platforms display ads at constant rate a.

Corollary 2.1. *If the planner can set platforms' investment but not their ad display, then the unique steady-state solution for the planner's choice of investment remains as in (24).*

8.3 Planner vs Laissez-Faire

I now compare the planner's solution with the decentralized equilibrium, beginning with investment.

Investment Comparing equations (21) and (24) yields a simple sufficient statistic for assessing efficiency.

Proposition 5. *For a given utility weight τ, the deviation $\ell_{\mathbb{K}} - \ell^*_{\mathbb{K}}$ between the decentralized and planner's steady-state investment is increasing in*

$$\frac{\sigma}{\sigma-1}\,\hat{\pi}_{\mathbb{K}} A(\epsilon-1) - \frac{\tau}{1-\tau}. \tag{26}$$

When (26) is positive (negative), equilibrium investment is too high (too low) relative to the planner's benchmark. Either case can arise, as can be seen by varying τ since $\hat{\pi}_{\mathbb{K}}$ does not depend on τ (see Online Appendix equation (49)).

The statistic in (26) depends on just four interpretable quantities: (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.[28]

Each captures a distinct source of potential distortion. The markup captures product market power, which reduces production and reallocates labor

[28] One way to estimate ϵ and τ jointly would be to run an experiment in which users are charged prices for accessing platform services.

toward platform investment. Ad revenue per unit of income reflects business-stealing: platforms invest simply to capture ad revenue from rivals. The elasticity $\epsilon - 1$ determines how effectively platforms can do so by raising quality and capturing attention.

Critically, none of these motives for investment internalize the effects on consumer surplus. Because platforms charge zero prices, they do not appropriate the value of their own quality improvements. By contrast, the planner internalizes the effects of investment on consumers. Hence $\tau/(1-\tau)$ which captures consumers' marginal value for platform use, appears in (26).

These inefficiencies can be corrected with a proportional tax or subsidy on platforms' ad revenues.

Corollary 2.2. *A proportional tax (or subsidy) on platform ad revenues of*

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

redistributed to (or funded by) consumers restores steady-state investment to the efficient level $\ell_{\mathbb{K}}^*$ in (24).

Remark 9. *In the model, condition (26) can also help sign the welfare effect of policy changes. In particular, the comparative statics for $\pi_{\mathbb{K}}A$ in Propositions 2 and 3 also apply to $\hat{\pi}_{\mathbb{K}}A$ in general equilibrium. Thus, a ban on data use worsens investment efficiency and overall welfare in steady state if investment is initially too high. Similar logic applies to changes in ϵ, arising, for example, from an interoperability mandate.*

Ad Display Rate The equilibrium ad rate maximizes

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

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

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

By inspection, the equilibrium ad rate can be either too high or too low relative to the planner's. For example, suppose ν is continuous and satisfies $\nu(\overline{a}) = 0$ for some $\overline{a} > 0$.[29] Then, as $\epsilon \to 1$, the equilibrium ad rate converges to $\overline{a}$,

[29]In the baseline model, I assumed $\nu > 0$ to simplify exposition, as demand (15) is undefined if almost all platforms choose an ad rate a such that $\nu(a) = 0$. A version of Theorem 1 continues to apply when ν vanishes beyond some $\overline{a} > 0$.

while the planner's choice is unaffected. Thus the equilibrium ad rate can be too high.

This is possible because platforms internalize the nuisance cost of ads only to the extent that it affects their ability to divert attention from rivals. When ϵ is small—i.e., attention is inelastic—this internalization is weak, and platforms over-serve ads. The distortion also stems from the fact that platforms charge zero prices and thus cannot capture the consumer surplus.

On the other hand, the equilibrium ad rate can also be too low. As τ tends to zero, the planner's optimal ad rate converges to $\overline{a}$, while the equilibrium ad rate remains the same. In this case, the planner assigns little weight to platform consumption and instead prioritizes product consumption. To expose consumers to as many products as possible, the planner increases the ad rate. Platforms, by contrast, do not internalize the effect of ad exposure on product consumption and therefore may set a lower ad rate than the planner.

9 Summary of Further Extensions

I summarize five extensions of the baseline model to highlight its ability to accomodate a range of market features. While the extensions are analyzed separately, they could also be combined—at least in steady state—within a unified model. All five can be analyzed in general equilibrium where there is a simple welfare criterion. Extensions 1–3 also appear tractable in the duopoly and asymmetric-data settings.

1) Network Effects In the first extension, found in Online Appendix G, I introduce network effects; I assume the effective quality of a platform k is instead $\eta(x_{kt})\nu(a_{kt})q_{kt}$ where $\eta(\cdot)$ is increasing in the total attention $x_{kt} = \int_{\mathbb{I}} x_{ikt}\,\mathrm{d}i$ spent on the platform. Analogous to (15) of the baseline model, 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} now appears on both sides above, to obtain explicit solutions, I assume $\eta(x) = x^{\zeta}$ where $\zeta > 0$ controls the strength of the network effects.

For each set $E_t \subset \mathbb{K}$ of positive mass, there is a solution with

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

Thus, network effects lead to equilibrium multiplicity. Under the refinement that $E_t = \mathbb{K}$ at each t, there is a unique equilibrium, characterized as in the baseline model, except with a higher elasticity of attention to quality equal to $(\epsilon - 1)/[1 - \zeta(\epsilon - 1)] > \epsilon - 1$. This higher elasticity leads platforms to display ads at a lower rate and invest more in the long run. The refinement is natural if $\mathbb{K}$ is interpreted as a stable set of active platforms.

2) Heterogeneous Platform Productivity In the second extension, found in Online Appendix H, platforms differ in the productivity of their investments. I characterize full equilibrium dynamics. By varying the distribution of productivities, the model can generate a wide range of non-atomic distributions of platform market shares, enhancing its potential for quantitative analysis and empirical relevance.

3) Zero Prices In the third extension, found in Online Appendix I, I provide an informal argument that zero prices for platforms' services can emerge endogenously for some parameters. I consider a variant of the model where platforms may charge nonnegative prices. Negative prices are often unsustainable, as platforms cannot easily distinguish bots from humans (Corrao et al., 2023).

The intuition is simple: if a consumer's marginal utility for product consumption is much higher than for platform consumption, attention will be highly sensitive to platforms' prices. In such settings, it is optimal for platforms to rely entirely on advertising revenue, i.e., to set zero prices.

4) Firm and Platform Entry In the fourth extension, found in Online Appendix J, I allow both firms and platforms to enter the market by paying an upfront cost. I characterize the unique steady-state equilibrium and show that more informative data can often induce firm entry and platform exit, while greater platform substitutability leads to exit of both firms and platforms.

5) Reserve Prices In the fifth extension, found in Online Appendix K, platforms may set reserve prices in ad auctions. I find 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. Moreover, as in the baseline model, steady-state ad revenues are maximal when data is uninformative.

10 Concluding Discussion

This paper proposes a framework for the modern attention economy in which platforms attract attention with free services and monetize it through targeted advertising. The framework captures the interaction between the 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 around data, platforms, and digital advertising, yet are rarely analyzed together in formal models.

The paper provides a tool for investigating mechanisms that arise from these structural features, which seem relevant across a range of settings—the baseline model assumes monopolistically competitive platforms, however, much of the framework can also be analyzed in a duopoly variant. I show this in a model with exogenous ad display rates and outline how one might endogenize them in a setting with a continuum of *duopolistic* platform markets (Subsection 6.4).

Three applications illustrate the model's ability to yield new insights. First, I study data and interoperability policies and show that policy effects may reverse over time, cross-market spillovers can overturn intuitive comparative statics, and interventions often involve tradeoffs between product consumption and platform investment. The analysis yields intuition for when these patterns are more likely to arise. In practice, interoperability polices 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 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 platforms with the decline in advertising's share of GDP.

Third, I analyze welfare in general equilibrium, identify the sources of inefficiency and, in the case of Cobb-Douglas utility, provide a sufficient statistic that signs whether platform investment is inefficiently high or low.

In addition to these applications, I study extensions with network effects, platform heterogeneity, endogenous entry, and auction reserve prices, and examine when zero pricing for platform services can emerge endogenously.

There are several promising directions for future work. First, although the assumption of CES preferences rules out personalized pricing, which has received growing policy attention (Rhodes and Zhou, 2022), the model can be extended to study product steering, a related phenomenon.

Second, it would enrich the study of data policies to allow platforms to decide how much data to disclose and to derive sharper conditions for when more informative data reduces ad revenue. Toward this end, methods from Bergemann et al. (2022), Ganuza and Penalva (2010), and the classical auction literature may prove useful.

Third, endogenizing ad display rates in a duopoly model—following the approach outlined in Subsection 6.4—would allow a more complete analysis of strategic interaction.

Fourth, modifying the platform investment technology to accomodate long-run growth may allow the framework to speak to policy discussions around innovation.

Fifth, future work could endogenize platform data by linking it directly to consumer attention to study the economic implications of that channel.

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.

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.

Appendix

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 is

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

from selling to consumer i. 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.

Equilibrium Properties Along Transition Path I record equilibrium properties, referred to in Theorem 1, that hold at each time $t \in [0, \infty)$.

- 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

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

  It is easy to conjecture and verify that (31) satisfies $B_t = V_t^{\text{In}} - V_t^{\text{Out}}$ since, under the conjecture, (9) and (10) become decoupled first-order linear ODEs for which there are explicit solutions.

- 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 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 are computable in seconds via the shooting method.

Data Shock: Garbling Figure 4 shows how $\tilde{v} \sim \mathrm{U}[0,1]$ is garbled into $\hat{v} \sim \mathrm{U}[.2,.8]$.

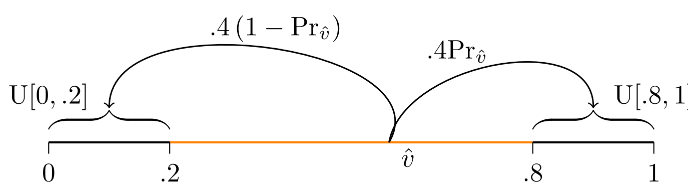


Figure 6: Joint distribution of pre and post shock expectations

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 $\mathrm{Pr}_{\hat{v}}$ that $\hat{v}$ jumps up is such that the martingale property holds: $\mathbb{E}[\tilde{v}|\hat{v}] = \hat{v}$.

Online Appendix

A Supplemental Material for Section 4

Appendix A.1 completes the proof of Theorem 1. Appendix A.2 gives the formal statement of a firm's bidding problem.

A.1 Proof of Theorem 1

Proof of Theorem 1. 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. Note that Steps 1 and 2 yield Parts 1 and 6 of the theorem.

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

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

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

I express this in terms of only H_t^c by using the accounting identity (4), yielding

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left[(J - M_t)H_t^c(\cdot)\right] &= \lambda_f[JG(\cdot) - (J - M_t)H_t^c(\cdot)] - AH_t^c(\cdot)^N \\ &= \lambda_f[JG - MH_t^c(\cdot)^N - (J - M)H_t^c(\cdot)].\end{aligned}$$

$$\Leftrightarrow$$

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

The steady state H^c is such that the right-hand side above is zero yielding Part 3 of the theorem.

Since M on the right-hand side is the steady state level and thus constant, the ODE is separable and can be solved analytically for H_t^c. This yields (30). By inspecting (30), we see that H_t^c (and therefore H_t) must eventually converge to its steady state level: the right-hand side of (30) diverges which implies the denominator of the integrand on the left-hand side of (30) must vanish.

Step 4. Firms' Bidding Strategies: Recall, from Section 3, that V_t^{In} and V_t^{Out} satisfy the HJB equations

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

and

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

Above, $W_t(\hat{v})$ denotes the probability that firm j wins an auction at time t for consumer i conditional on $\hat{v}_{ij} = \hat{v}$ and $B_t^{(1)}$ is equal in distribution to the maximum of the $N-1$ other bids in an auction at time t. If B_t is increasing then $W_t = (H_t^c)^{N-1}$ and $B_t^{(1)} \sim B_t(\hat{v}^{(1)})$ where $\hat{v}^{(1)} \sim W_t$.

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

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

where $\tilde{W}_t$ denotes the cdf of $B_t^{(1)}$. By definition, $W_t(\hat{v}) = \tilde{W}_t(B_t(\hat{v}))$.

Differentiating both sides with respect to $\hat{v}$ gives

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

The solution to this ODE is

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

The term in parenthesis must vanish as $t \to \infty$. Otherwise B_t' diverges which can not happen because B_t is bounded:

$$0 \le B_t(\hat{v}) = V_t^{\text{In}}(\hat{v}) - V_t^{\text{Out}}(\hat{v}) \le V_t^{\text{In}} \le \frac{\sup_{s \in [0,\infty)} \pi_{\mathbb{J}s}\hat{v}}{\rho}$$

for all t and for all $\hat{v}$. The last term above is the flow profit from being in the consideration set at *all* points in time in the future for an upper bound on the ad price. This upper bound is finite, for an arbitrary path of H_t, because, given the positive ad display rate A, M_t is bounded from below and thus so is μ_{Ω_t} and so $\pi_{\mathbb{J}s}$ is bounded from above (see (2)).

We therefore have

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

and so the bidding function is increasing as conjectured. Importantly, this is so for an *arbitrary* path of H_t which implies that our assumption that (6) applies did not entail any loss of generality.

The bidding function (31) then follows immediately by subbing in $H_t^c(\hat{v})^N = W_t(\hat{v})$. From there, a straightforward calculation yields the average ad price in (32) and in turn its steady state counterpart reported in Part 4 of the theorem.

Step 5. Investment and Quality: Given the average ad price and attention choice, 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 variable λ_t solves

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

Lemma 1 below shows that any solution to the platform's problem is unique (12)when $\epsilon - 1 < 1/\varphi$. Thus, each platform must employ the same investment strategy. Hence, along an equilibrium trajectory,

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

The first-order condition for maximizing the Hamiltonian yields

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

Differentiating both sides with respect to time yields

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

Then, from the costate equation we arrive at

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

where recall that

$$\dot{q}_{kt} = \ell_{kt}^{\varphi} - \delta q_{kt} \tag{35}$$

starting from $q_{k0} = q_0$. The steady state values of ℓ_{kt} and q_{kt} are derived from (34) and (35), yielding Part 5 of the theorem.

To prove existence and uniqueness of an equilibrium, we must prove there exists a solution to (34)—(35) that is consistent with equilibrium. Lemma 2 below proves that a solution exists for any initial conditions. Lemma 3 establishes that any solution for an arbitrary initial condition, is such that investment either vanishes, diverges, or converges to steady state. From here, it is clear that the only possibility that is consistent with an equilibrium is convergence to steady state. This yields the boundary value problem (33). Lemma 4 proves that there exists a unique solution to (33).

The last step is to verify the Hamiltonian approach is "valid." When $\epsilon < 2$, the Hamiltonian is jointly concave in state and control and thus the Mangasarian sufficient conditions are satisfied implying that each platform k does indeed optimize by investing according to $\{\ell_{kt}\}$ that solves (33) provided its rivals do the same.

In the more general case when $\epsilon - 1 < 1/\varphi$ the Hamiltonian is no longer necessarily jointly concave in the state and control. To verify that each platform k is optimizing I use a brute-force application of the calculus of variations. I first show that if investment satisfies (33), then the Gateaux derivative of platform k's objective with respect to the control is zero in all directions. I then show that the second order condition is satisfied in all directions which is straightforward since platform k's objective is concave in the control as shown in Lemma 1 below. I omit the details of these steps for brevity.

Lastly, I note that Part 7 of the theorem, which records steady-state product and platform consumption aggregates, is an immediate consequence of earlier parts of the theorem. □

The rest of this subsection, proves Lemmas 1—4 used in Step 5 above.

Lemma 1. *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 2. *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

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

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

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

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 3. *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

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

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 4. *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 2 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 2), 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 3 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 2) 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 2.

Therefore, it follows that ℓ^*_{kt} must converge to steady state and thus solves the boundary-value problem. □

A.2 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 PDV 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: $\hat{v}_z^{(1)} \sim \left(H^c_{\tau_z}\right)^{N-1}$ 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 Supplemental Material for Section 5

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

B.1 Proof of Proposition 2

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

Lemma 5 proves that product consumption is monotone in data informativeness.

Lemma 5. *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

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

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. It follows immediately that platform quality is also maximal from Part 5 of Theorem 1.

Lemma 6. *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

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

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

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

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. □

B.2 Proof of Proposition 3

Proof of Proposition 3. The proof proceeds via a series of lemmas. Some of these lemmas are stronger than necessary because they are of independent interest.

I first derive the relationship beween ϵ and A and then later the relationship between A and C (Lemma 8 as well as A and $\pi_{\mathbb{K}} A$ (Lemma 9). Together these results prove the proposition.

Lemma 7. *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 $a_2 < a_1$. $\square$

Combining Lemma 7 and the following shows that an increase in ϵ leads to a decrease in product consumption.

Lemma 8. *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 8 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

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

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

Lemma 9. *An increase in A leads to an 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 8, 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. □

B.3 Additional Comparative Statics Omitted From the Main Text

Lemma 10. *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 8, 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 8, 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. $\square$

Lemma 11. *An increase in J leads to an increase in steady state ad revenue $\pi_{\mathbb{K}} A$.*

Proof. By Lemma 10, 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 9, we see that this leads to an increase in steady state ad revenue $\pi_{\mathbb{K}} A$. $\square$

Lemma 12. *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 9 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. $\square$

Lemma 13. *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 1. $\square$

C Supplemental Material for Section 6

This appendix presents supplemental material for the duopoly model of Section 6. Appendix C.1 provides details of the algorithm to compute an equilibrium. Appendix C.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).

C.1 Details of the Algorithm to Compute Equilibrium

I use an implicit upwind finite difference scheme to solve (20). 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 6) 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

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

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

To implement the (semi-)implicit upwind scheme, I work with the following discretized version of (20):

$$\begin{aligned}
\rho V(q_1, q_2, t) = \sup_{\ell_{k1}} & V_t^{+}(q_1, q_2, t) + V_{q_1}^{+}(q_1, q_2, t)(\ell_{k1}^{\varphi} - \delta q_1)^{+} \\
& + V_{q_1}^{-}(q_1, q_2, t)(\ell_{k1}^{\varphi} - \delta q_1)^{-} + V_{q_2}^{+}(q_1, q_2, t)\,(\ell_{k2}^{\varphi} - \delta q_2)^{+} \\
& + V_{q_2}^{-}(q_1, q_2, t)\,(\ell_{k2}^{\varphi} - \delta q_2)^{-} - \ell_{k1} + \pi_{Kt} A \frac{q_1^{\epsilon-1}}{q_1^{\epsilon-1} + q_2^{\epsilon-1}}. \\
& + \frac{1}{2}\nu^2 q_1^2 \tilde{V}_{q_1,q_1}^{c}(q_1, q_2) + \frac{1}{2}\eta^2 q_2^2 \tilde{V}_{q_2,q_2}^{c}(q_1, q_2).
\end{aligned} \tag{40}$$

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 (40) 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 (40) 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 (20):

$$\begin{aligned} V_{q_1}^+(q_1, q_2, t)(\ell_{k1}^{\varphi} - \delta q_1)^+ + V_{q_1}^-(q_1, q_2, t)(\ell_{k1}^{\varphi} - \delta q_1)^- \\ + V_{q_2}^+(q_1, q_2, t)\,(\ell_{k2}^{\varphi} - \delta q_2)^+ + V_{q_2}^-(q_1, q_2, t)\,(\ell_{k2}^{\varphi} - \delta q_2)^- \\ + \frac{1}{2}\eta^2 q_1^2 V_{q_1,q_1}(q_1, q_2) + \frac{1}{2}\eta^2 q_2^2 V_{q_2,q_2}(q_1, q_2). \end{aligned} \tag{41}$$

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 (20) 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[30]

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

where k is the index of the vector V, then the corresponding indices of the other value function terms in (41) are:

$$V(q_1, q_2 + \Delta_q, t) \mathrel{\widehat{=}} k + 1,$$

$$V(q_1, q_2 - \Delta_q, t) \mathrel{\widehat{=}} k - 1,$$

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

$$V(q_1 - \Delta_q, q_2, t) \mathrel{\widehat{=}} k - N.$$

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 (41) 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.

Calculate Value Functions at Previous Time Step Using the M matrix (constructed using the *optimal controls at time* t in (41)) together with (40), 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

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

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.

C.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 6, this connection (together with the comment in Footnote 21) 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 (19) we have

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

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

$$\begin{aligned} \mathrm{d}x_{1t} = x_{1t}(1-x_{1t})\left[(\epsilon-1)(\ell_{1t}^{\varphi}-\ell_{2t}^{\varphi})+\frac{1}{2}(\epsilon-1)^2\eta^2(1-2x_{1t})\right]\mathrm{d}t \\ +(\epsilon-1)\eta x_{1t}(1-x_{1t})\sqrt{2}\,\mathrm{d}B_t. \end{aligned} \tag{42}$$

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

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

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 (42) 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.[31]

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

D Supplemental Material for Section 7

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

D.1 Steady State Equilibrium Characterization

The following Theorem 3 summarizes the steady state equilibrium properties.

[31]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.

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 (46) which is a contraction map whenever values are bounded.*
7. *The attention shares received by the two groups are given by (48).*
8. *The rates of investment by the two groups are given by (47).*

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[M_t h_t(\boldsymbol{\zeta})\right] = A\left[x_{1t} N H_t^c(\zeta_1, \infty)^{N-1} h_t^c(\boldsymbol{\zeta}) + x_{2t} N H_t^c(\infty, \zeta_2)^{N-1} h_t^c(\boldsymbol{\zeta}) - h_t(\boldsymbol{\zeta})\right]. \quad (43)$$

In (43), 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 (43) 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

$$V(\boldsymbol{\zeta}_{ij}) = \sum_{z=1}^{2} x_z O_z(\zeta_{zij}) \left[\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} \mathbb{E}[v_{ij}|\boldsymbol{\zeta}_{ij}] + \frac{\lambda_f}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \right]$$
$$- x_l \left[\int_{-\infty}^{\zeta_{zij}} B_z(s) \, \mathrm{d}O_z(s) + \left[1 - O_z(\zeta_{zij})\right] \frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \right] \quad (44)$$

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,

$$\begin{aligned} B_z(\zeta_{lij}) &= \mathbb{E}\left[\frac{\pi_{\mathbb{J}}}{\lambda_f+\rho}v_{ij} + \frac{\lambda_f}{\lambda_f+\rho}\frac{\lambda_a}{\lambda_a+\rho}V(\boldsymbol{\zeta}_{ij})\middle|\zeta_{zij}, j \in \Omega_{it}^c\right] \\ &\quad - \mathbb{E}\left[\frac{\lambda_a}{\lambda_a+\rho}V(\boldsymbol{\zeta}_{ij})\middle|\zeta_{zij}, j \in \Omega_{it}^c\right] \\ &= \mathbb{E}\left[\frac{\pi_{\mathbb{J}}}{\lambda_f+\rho}v_{ij} - \frac{\rho}{\lambda_f+\rho}\frac{\lambda_a}{\lambda_a+\rho}V(\boldsymbol{\zeta}_{ij})\middle|\zeta_{zij}, j \in \Omega_{it}^c\right]. \end{aligned} \tag{45}$$

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 (45) and (44), 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[32]

$$\Lambda_z(\boldsymbol{f})(\cdot) = \mathbb{E}\left[\frac{\pi_{\mathbb{J}}v_{ij} + \lambda_a \sum_{l=1}^2 x_z \int_{-\infty}^{\zeta_{lij}} f_l(s)\,\mathrm{d}O_l(s)}{\lambda_f + \rho + \lambda_a \sum_{l=1}^2 x_l O_l(\zeta_{zij})}\middle|\zeta_{zij} = \cdot, j \in \Omega_{it}^c\right]. \tag{46}$$

The map $\boldsymbol{\Lambda}$ 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 $\boldsymbol{\Lambda}$ 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 $\{\boldsymbol{\Lambda}^n(\boldsymbol{B})\}_{n=1}^\infty$ is a decreasing sequence which converges to the fixed point. Thus we can compute the equilibrium bidding strategies by iterating on (46).

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

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

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

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

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

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

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,

$$\ell_{\mathbb{K}z} = \frac{\varphi\delta\pi_{\mathbb{K}z}A(\epsilon-1)}{\rho+\delta}\frac{1}{m_z + m_{-z}\left(\frac{\pi_{\mathbb{K}-z}}{\pi_{\mathbb{K}z}}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}}} \tag{47}$$

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 (47) we also have the quality level

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

and attention share

$$x_z = \frac{m_z q_z^{\epsilon-1}}{m_z q_z^{\epsilon-1} + m_{-z}q_{-z}^{\epsilon-1}} = \frac{m_z\left(\frac{\pi_{\mathbb{K}z}}{\pi_{\mathbb{K}-z}}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}}}{m_z\left(\frac{\pi_{\mathbb{K}z}}{\pi_{\mathbb{K}-z}}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}} + m_{-z}} \tag{48}$$

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. $\square$

D.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 (43) forward to compute h.
3. Iterate (46) to compute per-unit income bid functions and average ad prices.
4. Check whether the guess of x_1 aligns with (48).
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.

E Supplemental Material for Section 8

Proof. 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 (21), let $\hat{\pi}_{\mathbb{K}} = \pi_{\mathbb{K}}/I$ be the average ad price per unit of income. From (17),

$$\hat{\pi}_{\mathbb{K}} = \frac{1}{\sigma M \mu_H} \int_0^{\infty} \frac{1 - NH^c(s)^{N-1} + (N-1)H^c(s)^N}{\rho + \lambda_f + \lambda_e(s)} \, \mathrm{d}s. \tag{49}$$

Then, rewriting (18), we have

$$\ell_{\mathbb{K}} = \frac{\varphi\delta\hat{\pi}_{\mathbb{K}} A(\epsilon - 1)}{K(\rho + \delta)} I. \tag{50}$$

Given investment,

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

Solving this linear system for $\ell_{\mathbb{K}}$ yields (21). 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 A^* and suppose that we have already reached steady state for H_t. The Hamiltonian for the planner's problem for investment is

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

where λ_t satisfies

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

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

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

Substituting ths into the previous equation, we have

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

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

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

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 (25). □

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

□

F 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.

G Extension: Network Effects

I extend the baseline model to allow for network effects.

G.1 Setup

I redefine the CES aggregate for platform consumption to be

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

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

G.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. □

H Extension: Heterogeneous Platform Productivity

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

H.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.

H.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

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

$$\dot{q}_t = \ell_t^{\varphi} - \delta q_t$$

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. □

I 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 8. 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

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

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

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

Canceling out some terms and rearranging yields

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

which is linear in the price p.

Solving for p yields the inverse demand curve:[33]

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

[33] One can see that the inverse demand curve is monotone since the objective is submodular in (p, 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:

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

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

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

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

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

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 $\tau \leq 1/2$.

J Extension: Firm and Platform Entry

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

J.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.

J.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 11 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 B to this setting.

K Extension: Reserve Prices

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

K.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.

K.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

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

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

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

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,

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

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

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

for $s \geq Y$. 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

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

Then using an accounting identity, we have

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

which rearranges to

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

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

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

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 1 it is almost a necessary condition for equilibrium existence.

$\square$