The Market for Attention

Daniel Chen*

dtchen@princeton.edu

Princeton University

Feb 2022, revised May 2024

Abstract

This paper develops a general equilibrium model of competing platforms that profit from targeted advertising. In the model, platforms offer consumers quality services at zero prices in exchange for attention. Platforms then monetize the attention by selling targeted ads to firms in the product market. The model shows how ad revenue, the quality of platforms' services, and the product market allocation are determined in equilibrium. I find that a multi-sided analysis is critical: intuitive comparative statics based on single-sided analyses can flip, short-run effects of policies may differ drastically from long-run effects, and policies must balance tradeoffs across the market sides. I illustrate these lessons in the context of data and interoperability policies, two of the leading regulatory tools.

In the past two decades, the various platforms on the Internet that profit from targeted advertising have become a large and important part of the economy.[1] Though these platforms offer distinct services, many say that they compete in a broader "market for attention" (Evans, 2020). Some key stylized features of this market are as follows.

1. Platforms provide distinct services (social media, streaming, podcasts etc.) to consumers often at a zero price.

*I am indebted to my advisors Yuliy Sannikov and Andy Skrzypacz and committee members Darrell Duffie and Bob Wilson for their suggestions and helpful comments. I am also thankful to Mohammad Akbarpour, Susan Athey, Kyle Bagwell, Anirudha Balasubramanian, Martino Banchio, Markus Brunnermeier, Christoph Carnehl, Cindy Chung, Sebastian Di Tella, Ravi Jagadeesan, Ramesh Johari, Chad Jones, Bruno Jullien, Nadia Kotova, David Kreps, Ernest Liu, Alessandro Lizzeri, Moritz Lenel, Erik Madsen, Suraj Malladi, Matthew Mitchell, Mike Ostrovsky, David Ritzwoller, Marzena Rostek, Karthik Sastry, Takuo Sugaya, Timur Sobolev, Chris Tonetti, Laura Veldkamp, Frank Yang, Ali Yurukoglu, Anthony Zhang, Weijie Zhong, and several seminar participants. This research is supported by the Bradley Graduate Fellowship through a grant to SIEPR.

[1] See https://www.iab.com/wp-content/uploads/2023/04/IAB_PwC_Internet_Advertising_Revenue_Report_2022.pdf and https://datareportal.com/reports/digital-2024-deep-dive-the-time-we-spend-on-social-media.

2. Platforms compete for attention by investing in the quality of their services.[2]

3. Platforms earn revenue by selling advertisements (ads) to firms in the product market.

4. Platforms sell ads via consumer-specific auctions that arise in real time as consumers engage with platforms' services.

5. Platforms provide firms with data on consumers when firms bid in ad auctions.

This paper develops a general equilibrium model with a platform sector that is consistent with these features. Because the model is in general equilibrium, it also includes an endogenous product market. We can therefore calculate the welfare impact of platforms from their roles in both features 1 and 3 in a single internally consistent model. I use the model to analyze the effects of policies and to compare the equilibrium to an economy under a beneficent planner. A distinctive property of my analysis is that it accounts for interactions among each of the sides of the market for attention, including the product market.[3]

The model is a reaction to three points that often come up in regulatory discussion. First, traditional competition policy relies heavily on markups as a gauge of market efficiency (OECD, 2022; Wu, 2018). Regulators are in need of competition analysis that applies to zero-price markets. Second, the market for attention is complex and multisided: the product market allocation, ad revenue, and the quality of platforms are jointly determined. It is thus difficult to assess the net effects of a policy without a formal model. Third, many leading policies relate to data and platform interoperability (OECD, 2022, 2020). Interoperability refers to the extent that platforms can communicate or work with each other. A lack of interoperability limits platform substitutability and inhibits competition (OECD, 2022).

The model I develop shows how outcomes on each of the market sides are determined in equilibrium and how they depend on data and platform substitutability. I find that data and interoperability policies typically must trade off allocative efficiency in the product market with the quality of platforms' services. Also, spillovers across the market sides can be strong and lead to counterintuitive effects. For example, in the short run, an interoperability policy can lead platforms to invest *less* in the quality of their services and in the long run, relaxing data privacy laws can lead to *lower* platforms' profits. Further, these effects can flip, depending on the time horizon, as shocks gradually propagate across the market sides: in the long run, an interoperability policy leads to *greater* investment by platforms and in the short run, relaxing data privacy laws typically leads to *higher* platforms' profits. Whether

[2]Platforms that charge zero prices typically compete in quality (OECD, 2018; Ambrus et al., 2016).

[3]With abuse of terminology, I will refer to the product market as one of the "market sides."

policies are, in the end, beneficial will often depend on the economy's ex ante efficiency. I derive a sufficient statistic to gauge the efficiency of equilibrium investment in the model.

To develop a tractable general equilibrium analysis, I must make some modeling concessions. In particular, I model platforms that are monopolistic competitors: each platform is the dominant player in its category of service and thus has market power, but is small relative to the entire platform sector. This allows me to abstract from complicated strategic interactions but I may omit potentially important effects for platforms that are large relative to the sector.[4] I also assume that the product and platform sectors both have a constant-elasticity-of-substitution (CES) structure. This implies that firms charge a fixed markup on products and have no incentives to personalize prices. Nonetheless, the model is able to shed light on some basic issues that can not be addressed by other models at this time, which generally treat one or more of the market sides as exogenous.

In the model, consumers are aware of only some firms in the product market. Platforms facilitate search and matching by displaying ads for firms' products to consumers. As in common practice, each platform sells ads, at the individual consumer level, via auctions that arise dynamically as consumers engage with its services. A platform displays ads to a consumer at a Poisson rate per unit of attention that it receives from the consumer. To be able to sell more ads, each platform competes to attract attention by investing in the quality of its services. Each platform is endowed with data on consumers' preferences for products that it supplies to firms when firms bid. Consumers derive higher flow utility from paying attention to higher quality platforms but dislike viewing ads. Each consumer has a taste for variety and splits a unit of attention among the platforms. To ensure that frictions persist in the long run, I assume that consumers gradually "forget" about firms whose ads they have seen, at independent exponential times.

In equilibrium, a firm matches with (i.e, shows its ad to) a given consumer more quickly when platforms display ads at a faster rate. Firms also match with consumers who value their products highly more quickly when data is more informative. Ceteris paribus, the longer it takes firms to match with consumers, the greater the profit that platforms extract in each ad auction. Thus, search frictions are a source of platform market power. Platforms therefore typically invest more and are of higher quality when the product market is less efficient. In equilibrium, a key state variable is the distribution of consumers' values for firms in their consideration sets. This distribution evolves gradually as consumers discover and forget firms and so the economy has a potentially lengthy transition to steady state.

[4] In 2023, Meta earned roughly 23 percent of global digital ad revenue. YouTube earned 5.5 percent. X, TikTok, Snapchat, Pinterest, Twitch and many others each earned far less. For more stats, see footnote 32.

I consider two policy experiments whose effects I trace out over time and across the market sides. In the first experiment, an interoperability mandate leads to an increase in platform substitutability.[5] The goal is to encourage platforms to invest in better quality services by giving consumers a meaningful choice of how to spend their attention.[6] In the short run, platforms reduce their ad rates because attention is now more sensitive and consumers dislike viewing ads. This turns out to cause ad revenues to decline and so platforms actually invest less. In the long run, these effects reverse. First, it takes longer for a firm to match with a consumer in the future if it loses an ad auction for that consumer. Second, firms face competition from fewer rivals in consumers' consideration sets. Both of these factors raise a firm's present value for an ad. Firms therefore bid more to the extent that ad revenue is higher than before the shock. This, in turn, incentivizes platforms to invest more. In the end, platforms are of higher quality but the product allocation is less efficient.

In the second experiment, a new policy relaxes existing privacy laws, giving platforms access to more informative data.[7] Initially, platforms' profits increase because winning firms are on average more highly valued by consumers and platforms are able to extract some of this surplus gain. However, this effect reverses in the long run. First, firms face stronger competition as consumers' consideration sets fill up with firms that are more highly valued on average than before. Second, it takes less time for firms to match with consumers who value their products highly in the future. Both factors reduce a firm's present value for an ad, ultimately leading ad revenues (and platforms' profits) to decrease. In the end, platforms invest less and are of lower quality though the product allocation is more efficient.

The counterintuitive result that giving all platforms access to better data may reduce their profits relies on the impact of data on competition in the product market. To offer suggestive evidence, I show that, because of this channel, an extended version of the model can rationalize two empirical trends many consider puzzling (Silk et al., 2021). Namely, (i) digital advertising as a share of ad revenue has grown dramatically in the past decade at the expense of traditional advertising; (ii) total ad revenue as a fraction of gross domestic product (GDP) has declined slightly in the past decade despite the rise of digital platforms with their unprecedented abilities to use consumer data for ad targeting.

I next compare the equilibrium to the economy of a beneficent planner who sets the rates that platforms invest and display ads. I find that the two key sources of inefficiency are businesses-stealing externalities and appropriability issues. Namely, platforms internalize

[5] E.g., the mandate could be that platforms now must let users post links to content on other platforms.

[6] "Interoperability is a key aspect of many digital platform services, as it gives users access to a wide range of choices" (OECD, 2021).

[7] Examples of privacy laws are General Data Protection Regulation and California Consumer Privacy Act.

the nuisance costs of ads or the benefits of their investments only to the extent that they can steal ad revenue from each other. This is so because they do not appropriate any surplus that they generate for consumers with their services because they charge zero prices. For the case when consumers have Cobb-Douglas utility over a CES product aggregate and a CES platform aggregate, I derive a simple sufficient statistic (that depends on only ad revenue, income, the product markup, platform substitutability, and the utility parameter) that can gauge the efficiency of equilibrium investment.

In the online appendices, I explore several extensions of the baseline model, including: (i) network effects (i.e., a consumer's utility from platform use depends on attention allocated by other consumers); (ii) firm and platform entry; (iii) ad auction reserve prices; (iv) heterogeneity in platform data; (v) heterogeneity in platform productivity.

1. Related Literature

This paper contributes to the literature on platforms and two-sided markets. Early seminal work by Rochet & Tirole (2003), Caillaud & Jullien (2003), Armstrong (2006), and others study platforms' pricing decisions in abstract settings with broad applications. In these models, the surplus generated by interactions of agents on the two sides of the market is typically exogenous or not microfounded. For a survey of this literature, see Jullien et al. (2021). Subsequent work has studied models more closely tailored to modern advertising platforms, often with a role for consumer data (e.g., Ambrus et al. 2016; Bergemann et al. 2019; Bergemann & Bonatti 2023; Prat & Valletti 2021; Galperti & Perego 2022; Jullien & Bouvard 2022). However, these papers do not simultaneously endogenize the product market and provision of content by platforms. Also, these papers typically consider either one or two platforms (whereas I consider the other extreme of monopolistic competition).

Another related paper is Kirpalani & Philippon (2021) which models a monopolistic platform that matches consumers with firms using data on consumers' preferences. Kirpalani & Philippon (2021) analyze issues related to consumers sharing data with the platform. Because they focus on online marketplaces like Amazon (which fall outside the scope of my analysis), they do not incorporate platform content provision.

This paper also relates to a long literature on traditional advertising. For a survey, see Bagwell (2007). Most papers do not model endogenous content provision to attract consumers. An important exception is Anderson & Coate (2005) upon which many subsequent models that do this are based. They study investment by platforms on the extensive margin in the form of entry costs. They likewise identify appropriability issues and business stealing as sources of inefficiency. In their model, firms extract all surplus from consumers who have binary types so there is no benefit to consumers from advertising. They study the case of either one or two platforms in a static setup with no role for data or ad targeting.

This paper also relates to the literature on ad auctions (e.g., Edelman et al. 2007; Athey & Ellison 2011; Varian 2007). These models usually do not include the product market. More informative data is shown to typically raise ad revenue (Board, 2009; Bergemann et al., 2021; Hummel & McAfee, 2016). I find that the opposite can be true in the long run due to competition in the product market. My model may also be of interest to empirical researchers using ad auction data to measure the effects of privacy policies (Alcobendas et al., 2021; Johnson et al., 2020; Beales & Eisenach, 2014; Marotta et al., 2019). A structural model may be useful for welfare and counterfactual analysis.

There is a growing body of work that studies the role of data and markets for data in the macroeconomy (Jones & Tonetti, 2020; Farboodi & Veldkamp, 2021). These papers focus on the nonrival property of data and the implications for growth. They do not explicitly model platforms according to the definition in Hagiu & Wright (2015). A contemporaneous paper Cavenaile et al. (2023) builds a macroeconomic model of targeted advertising with evolving product awareness but does not include platforms. Also related is Rachel (2021) which studies the macroeconomic effects of leisure-enhancing technological change in a model where consumer attention is monetized to finance zero-price products.

In terms of methodology, this paper combines elements from the Melitz (2003) model of international trade with firm heterogeneity, the Duffie et al. (2005) model of trade in over-the-counter markets, and the Wolinsky (1988) model of dynamic auctions. As in Melitz (2003), I model a CES product market to tractably incorporate firm heterogeneity and entry/exit dynamics. One can show that the heterogeneity in firms' production costs in Melitz (2003) is isomorphic to heterogeneity in consumers' values for firms' products.

As in Duffie et al. (2005), I rely on search frictions to explain the existence of intermediaries. Like the dealers in Duffie et al. (2005), platforms extract more rents when it takes longer for firms to find consumers. As match delay vanishes, the equilibrium of both of our models converges to that of a classical frictionless benchmark.

When there are search and matching frictions, prices can not be determined in a Walrasian way. Following Wolinsky (1988), I combine search and matching with auctions. In contrast to Wolinsky (1988), the auction model is tailored to fit the microstructure of bidding in online ad auctions. In Wolinsky (1988), once a buyer wins an auction hosted by a seller, both leave the market. In my model, firms advertise to the same consumers over and over again. Whereas Wolinsky (1988) analyzes steady state, I solve for the full dynamics.

2. The Baseline Model

Time $t \in [0, \infty)$. There is a unit measure of consumers $i \in \mathbb{I}$.[8] The economy has two sectors. The product sector contains a measure J of monopolistically competitive firms $j \in \mathbb{J}$. The platform sector contains a measure K of monopolistically competitive platforms $k \in \mathbb{K}$. Firms profit by selling products to consumers. Platforms profit by selling targeted ads for products which they display to consumers. To attract consumer attention, platforms provide valuable services (e.g., social media, streaming, podcasts) to consumers at zero prices. Each platform competes for attention by investing in the quality of its services.

Consumers' Preferences.—Consumer i has preferences

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

where $\rho > 0$, u is increasing, C_{it} is a product aggregate, and X_{it} is a platform aggregate.

Given consumption $\{c_{ijt}\}$ of individual products,

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

where $\sigma > 1$ is product substitutability and $\{v_{ij}\}$ are tastes or values for products. Each v_{ij} is an independent draw from the cumulative distribution function (cdf) F supported on $[0, \infty)$. I assume that F has a finite mean.

Given consumption $\{x_{ikt}\}$ of individual platforms,

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

where $\epsilon > 1$ is platform substitutability and $\{q_{kt}\}$ are the quality levels of platforms' services. Platform k's quality is multiplied by the factor $\nu(a_{kt})$ where a_{kt} is the rate that platform k displays ads and ν is positive and weakly decreasing. Thus, ads are nuisances that reduce effective platform quality.

Consumers' Problem.—Consumer i selects product and platform consumption at each time t. With respect to these choices, products and platforms differ in two key ways: 1. Consumer i is aware of all platforms but only some firms $\Omega_{it} \subset \mathbb{J}$. As I soon describe, the *consideration set* Ω_{it} evolves as she views ads on platforms; 2. Firms charge positive

[8] To be concrete, let $\mathbb{I}$, $\mathbb{J}$, and $\mathbb{K}$ be intervals in $\mathbb{R}$.

prices (set optimally) but platforms charge zero prices (set exogenously) and instead require attention to consume.

At each t, consumer i has I units of income to spend on products and a unit of attention to spend on platforms. To simplify the analysis, I assume she does not allocate attention to purposefully view ads—i.e., she does not internalize how platform use affects her consideration set in the future. (In equilibrium this is without loss of optimality as all platforms display ads at the same rate so how she splits attention among them has no effect on her consideration set.) She therefore maximizes flow utility which simplifies to

$$\max_{\{c_{ijt}\}} C_{it} \text{ s.t } \int_{\Omega_{it}} c_{ijt} p_{jt} \,\mathrm{d}j \leq I; \quad c_{ijt} = 0, \forall j \notin \Omega_{it}$$
$$\max_{\{x_{ikt}\}} X_{it} \text{ s.t } \int_{\mathbb{K}} x_{ikt} \,\mathrm{d}k \leq 1.$$

The choices of the two types of consumption separates in this way because platforms charge zero prices. As a result, u does not appear anywhere and there is no link between the marginal utilities for product consumption and platform consumption, a fundamental source of inefficiency in the model (discussed in Section 7). Further, note that all consumers choose attention in the same way, so in what follows, I omit the index i on x_{ikt}.

Size of Consideration Sets.—While using platform k, consumer i views ads for products at a Poisson rate $a_{kt} x_{ikt}$. Each time an ad is viewed, the relevant firm enters the consideration set. By the law of large numbers (LLN), firms flow into Ω_{it} at rate[9]

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

To ensure that frictions persist in the long run, I assume that once inside Ω_{it}, a firm is forgotten at an independent exponential time with rate λ_f. Thus, by the LLN, the mass M_t of firms in Ω_{it} evolves according to

$$\dot{M}_t = A_t - \lambda_f M_t \tag{1}$$

starting from an initial condition $M_0 < J$ that I assume is common to all consumers. Since (1) applies for all i, I do not index M_t by i. Also, note that (1) applies only if $M_t < J$. I impose conditions on parameters so that this is always so in Section 4.

Ad Auctions.— To determine which ads to show which consumers, platforms host auc-

[9] I assume a LLN throughout. I expect a LLN can be proven using methods in Duffie et al. (2020).

tions. Each time platform k has an opportunity to display an ad to consumer i it initiates a second-price auction with no reserve price.[10] Platform k then sends out requests to firms to submit bids. Due to latency, only $N > 1$ firms are able to respond within the instant. For simplicity, N is exogenous and deterministic. The N firms are drawn uniformly at random from outside the consumer's consideration set. I assume only firms outside Ω_{it} submit bids to obtain analytical solutions for equilibrium bidding strategies. However, the economic mechanisms behind the main results do not hinge on this assumption.[11]

Data.—To be able to study the implications of data policies, I assume that platforms send firms data on consumers along with their bid requests. For the baseline model, all platforms have the same data though I allow for heterogeneity in Section 6. Data are modeled as Blackwell experiments on consumers' values for products. It will suffice to specify only the distribution of firms' posterior expectations of consumers' values conditional on the data. I assume that firm j's expectation $\hat{v}_{ij}$ of consumer i's value v_{ij} is drawn from a continuous cdf G independently across i and j. I also assume that G is a mean-preserving contraction of the prior F, which by Blackwell (1953) is a necessary and sufficient condition for there to exist some underlying data that generates G. Firms' expectations are drawn at $t = 0$ and fixed thereafter.[12]

Composition of Consideration Sets.—In each auction for consumer i, each firm submits its bid *after* observing the data on the consumer. The firm that bids the most pays the second-highest bid and then displays its ad to the consumer at the end of the instant. Suppose that the winning firm in each auction is also the firm with the highest expectation of the consumer's value for its product, as will be the case in equilibrium. Let H_t denote the cdf of the expected values of firms in Ω_{it} and H_t^c the cdf for the complement Ω_{it}^c. Then, starting from a given initial H_0 common to all consumers,

$$(\dot{M_t H_t}) = A_t (H_t^c)^N - \lambda_f M_t H_t. \tag{2}$$

In (2), the cdf of expected values flowing in is $(H_t^c)^N$ since these firms are the winners in the auctions while the cdf flowing out is H_t since each firm is forgotten at the same rate.

[10] Revenue equivalence holds in the model so the results extend to any other standard auction format and in online Appendix J, I analyze a version of the model where platforms set reserve prices.

[11] Since only firms outside the consideration set have a positive value for the ad opportunity, this amounts to an assumption that firms know whether they are currently in consideration. In practice, firms have some knowledge of consumers' consideration sets by tracking consumer visits to their retail websites.

[12] That is, firms do not observe consumer i's past purchase history as otherwise they could infer her values for their products perfectly. For example, consumers might browse firms' retail website but make purchases in the brick and mortar store. Future work may explore a setup with time-varying values and learning.

By accounting, H_t and H_t^c must satisfy the identity

$$M_t H_t + (J - M_t) H_t^c = JG \tag{3}$$

which states that the frequency distribution of expected values in Ω_{it} added to that outside Ω_{it} must coincide with that of the whole population. As with M_t, I do not index H_t and H_t^c by i since (2) applies to all consumers. To ensure that M_0 and H_0 are *feasible*, I assume that $M_0 \, \mathrm{d}H_0 \leq J \, \mathrm{d}G$ since there can not be more firms with a given expectation in Ω_{it} than there are in the entire economy.

Firms.—Each firm maximizes the net-present-value (NPV) of its flow profits. Firms each produce at a constant marginal cost that I set as the economy's numeraire. A firm's static problem is to set a price for its product. In equilibrium, given M_t and H_t, consumer i's demand c_{ijt} for product j is a known function $c_t(v_{ij}, p_{jt})$ of her value v_{ij} and the price p_{jt}.[13] Firm j sets p_{jt} to maximize the expected flow profit from selling to a consumer:

$$\pi_{\mathbb{J}t}(\hat{v}_{ij}) = \max_{p_{jt}} \mathbb{E}\left[(p_{jt} - 1) c_t(v_{ij}, p_{jt}) | \hat{v}_{ij}\right].$$

Firm j's dynamic problem is to set bids in the ad auctions. The problem is dynamic because of the outside option to wait to bid in future auctions for a given consumer which limits a firm's willingness to pay in any given auction. To state the problem formally, let

$$\lambda_{at} = \frac{NA_t}{J - M_t} \tag{4}$$

denote the Poisson rate that a firm enters an auction for a given consumer i while outside Ω_{it}. This is the total rate NA_t that bid responses are collected divided by the measure $J - M_t$ of firms that are eligible to respond. Let τ_z denote the zth time of entry into an auction for consumer i.[14] In equilibrium, firm j takes as given the bidding strategies of its rivals which are of the following form. At time t, each firm l bids according to an increasing function B_t that maps firm l's expectation $\hat{v}_{il}$ to a bid $B_t(\hat{v}_{il})$ in an auction for consumer i at time t.

Taking B_t as given, firm j sets bids to maximize the NPV of flow profits (including the costs of advertising) from selling to a typical consumer i:

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

[13] The price p_{jt} may depend on $\hat{v}_{ij}$ but I later show it is *optimal* for firm j *not* to personalize prices.

[14] τ_z is the zth arrival of a Poisson process that ticks at rate $\lambda_{at} + \lambda_f \mathbb{1}_{\{j \in \Omega_{it}\}}$.

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

Platforms.—Each platform maximizes the NPV of its flow profits. Given B_t, the average ad price is $\mathbb{E}\left[B_t\left(v_t^{(2)}\right)\right]$ where $v_t^{(2)}$ denotes the second highest of N independent draws from the cdf H_t^c. At each t, platform k sets its ad rate to maximize its flow profits

$$\pi_{\mathbb{K}t}(q_{kt}) = \max_{a_{kt}} \mathbb{E}\left[B_t\left(v_t^{(2)}\right)\right] a_{kt} x_{kt}(a_{kt}, q_{kt})$$

taking as given the dependence of attention on its ad rate and quality. In later sections, I abuse notation and let $\pi_{\mathbb{K}t}$ also denote the average ad price $\mathbb{E}\left[B_t\left(v_t^{(2)}\right)\right]$.

Platform k's dynamic problem is to set a path for investment ℓ_{kt} to solve

$$\Pi_{\mathbb{K}} = \max_{\{\ell_{kt}\}} \int_0^\infty e^{-\rho t}\left(\pi_{\mathbb{K}t}(q_{kt}) - \ell_{kt}\right) \mathrm{d}t$$

subject to the law of motion of its quality

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

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. These assumptions are necessary for the existence of a long run stationary level of quality in equilibrium.

Equilibrium Concept.—An *equilibrium* for the economy with initial conditions M_0, H_0, and q_0 is a collection of processes for product demands $\{c_{ijt}\}$, platform demands $\{x_{kt}\}$, the mass of varieties in consideration sets $\{M_t\}$, the cdf of the expected values of consumers for those varieties $\{H_t\}$, prices $\{p_{jt}\}$, bidding functions $\{B_t\}$, ad rates $\{a_{kt}\}$, investment rates $\{\ell_{kt}\}$, and quality levels $\{q_{kt}\}$ such that consumers, firms, and platforms solve their respective problems and $\{M_t\}$, $\{H_t\}$, and $\{q_{kt}\}$ satisfy their respective laws of motion.[15]

To simplify the exposition, I have introduced the baseline model in partial equilibrium with the income I of consumers exogenously given. I consider the full general equilibrium model in Section 7 when I analyze welfare. The full model is just as tractable (see online Appendix C) and all of the main qualitative results of the baseline model extend to it.

[15] I have defined equilibrium with firms bidding symmetrically for ease of exposition. One can extend the definition to allow firms to bid asymmetrically, but, only symmetric bidding equilibria exist.

3. Discussion of the Model

The model is designed to study how outcomes on each of the market sides are determined in equilibrium and how they might evolve together in response to policies. Below, I briefly comment on several modeling decisions.

Monopolistic Competition.—Each platform has market power and offers a unique type of content, but is small relative to the whole sector. When substitutability ϵ is high, each platform may represent a small website on the Internet. When ϵ is low, each platform may represent the dominant player in its content category, effectively holding captive its share of consumer attention. In either case, however, there are no strategic interactions among platforms and each platform individually has no impact on consideration sets or firms' continuation payoffs. Thus, I may be abstracting from effects that are relevant for platforms that are large relative to the whole sector.[16] This sacrifice enables a tractable analysis that isolates the role of interactions among the market sides, the main focus of this paper.

Dynamics.—A dynamic model allows me to show how endogenous search and matching frictions in the product market affect platforms' profits and serve as a source of market power. An implication of these frictions (beyond explaining the existence of platforms) is the economy's gradual adjustment dynamics: I compare the short run and long run effects of policies which I find may differ dramatically. This is especially important as transition dynamics can be slow. Further, the long run outcomes do not depend on any ad hoc assumptions on initial conditions. Dynamics also has methodological advantages: there does not exist a symmetric bidding equilibrium if a firm can participate in multiple auctions for a given consumer at a given time. I avoid this issue by spreading competition out dynamically. Because the reasons for this are technical I discuss them in online Appendix B. There, I also discuss issues with other static setups, including setups with a Walrasian ad market.[17]

Zero Prices for Platforms' Services.—Zero prices are common, but not universal. In online Appendix E, I show zero prices emerge endogenously, for a range of parameters, in a model variant where platforms can charge nonnegative prices (see also Corrao et al. (2023)). Zero prices are often cited as a challenge for regulators who traditionally rely on markups to gauge market efficiency (Khan, 2017; OECD, 2022). Rather than through prices, platforms often compete through quality (Ambrus et al., 2016; OECD, 2018).

[16] In 2023, Meta (Facebook and Instagram) earned roughly 23 percent of global digital ad revenue. YouTube earned 5.5 percent. X, TikTok, Snapchat, Pinterest, Twitch, Spotify and many others each earned far less. Even when excluding search-engine advertising, the numbers are qualitatively the same: see footnote 32.

[17] There are issues with modeling data with a Walrasian ad market. For example, I would not be able to extend model to study data heterogeneity (see Section 6) as elaborated on in online Appendix D.

The Role of Advertising.—Advertising, in the model, is informative. This is a common assumption though other roles of advertising have been proposed in the literature. For example, one view is that advertising is persuasive and alters consumers' tastes. Though these and other effects are potentially important, they are beyond the scope of this paper.

Other Assumptions.—The analysis, at least for steady state, can be easily extended to include an outside option for attention and a consumption-savings tradeoff (under standard homotheticity assumptions on flow utility u). CES preferences for products is a more critical assumption for tractability because it ensures that prices are a fixed markup and there are no incentives to personalize prices. In online Appendices F-J, I study, in separate extensions, heterogeneity in platform data, heterogeneity in platform productivity, network effects, firm and platform entry, and reserve prices.

4. The Equilibrium

The objective of this section is to present the main properties of the equilibrium, provide intuition for them, and give some sense of how they are derived. These properties are summarized below in Theorem 1. In what follows, I walk through each part of Theorem 1 in succession, and in the process, sketch the proof of the theorem. The intuition developed here is critical to understand the results of later sections.

Theorem 1. *Suppose that A is the unique solution of $\max_a a\nu(a)^{\epsilon-1}$. If $A/\lambda_f < J$ and $\epsilon - 1 < 1/\varphi$, then there exists a unique equilibrium 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. *Consumers' demands for products and platforms are as in* (5) *and* (6) *respectively.*
2. *Firms' prices are as in* (8).
3. *Platforms' ad rates are as in* (9).
4. *The mass of firms in Ω_{it} and the cdf of expected values in Ω_{it} are characterized by* (10) *and* (11) *respectively given accounting identity* (3).
5. *Firms' expected flow profits from selling to consumer i and Poisson rates of entry into Ω_{it} are as in* (12) *and* (13) *respectively.*
6. *Firms bid according to* (14) *and the average ad price is* (15).
7. *Platforms' investment rates and quality levels solve the boundary-value problem* (16).

8. *Consumer, firm, and platform surplus are as in* (17), (18), *and* (19) *respectively.*

Moreover, the sufficient conditions are almost necessary: if either $A/\lambda_f \geq J$ *or* $\epsilon - 1 > 1/\varphi$, *then an equilibrium does not exist.*

Solving for the equilibrium (i.e., proving Theorem 1) may appear difficult because outcomes on the different market sides are mutually determined and agents must forecast the future paths of consideration sets which themselves depend on agents' actions. However, the model is set up so that each component speaks to the others in as minimal a way as possible so that the model is solvable nearly explicitly while remaining internally consistent. I sketch the proof of Theorem 1 below, step-by-step, using objects derived in previous parts to derive objects in subsequent parts.

Step 1: Consumers' Demands.—Consumer i's CES preferences imply a demand for product $j \in \Omega_{it}$ of

$$c_{ijt} = \frac{I v_{ij}}{\int_{\Omega_{it}} v_{iz} p_{zt}^{1-\sigma} \mathrm{d}z} p_{jt}^{-\sigma} \tag{5}$$

and a demand for platform $k \in \mathbb{K}$ of

$$x_{kt} = \frac{[\nu(a_{kt}) q_{kt}]^{\epsilon - 1}}{\int_{\mathbb{K}} [\nu(a_{zt}) q_{zt}]^{\epsilon - 1} \mathrm{d}z}. \tag{6}$$

Her demand for product j is increasing in her income I and her value v_{ij} for the product. It is decreasing in her values for the rival products in the consideration set. The elasticity of demand with respect to firm j's price is higher when σ is higher. Similarly, her demand for platform k is increasing in the effective quality $\nu(a_{kt}) q_{kt}$ of the platform and decreasing in the effective quality of the rival platforms in the economy. The elasticity of demand with respect to platform k's effective quality is higher when ϵ is higher.

Step 2: Firms' Prices.—Given the demand in (5), firm j's flow profit is

$$\frac{I v_{ij}}{\int_{\Omega_{it}} v_{iz} p_{zt}^{1-\sigma} \mathrm{d}z} p_{jt}^{-\sigma} (p_{jt} - 1). \tag{7}$$

from selling to consumer i. Since the consumer's value v_{ij} for product j and consideration set Ω_{it} only 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

$$p_{jt} = \frac{\sigma}{\sigma - 1}. \tag{8}$$

Thus, it is optimal for firm j not to personalize prices.

Step 3: Platforms' Ad Rates.— Given the demand in (6), platform k sets its ad rate to maximize flow profit:

$$A = \arg\max_{a_{kt}} \pi_{\mathbb{K}t} a_{kt} \frac{[\nu(a_{kt})q_{kt}]^{\epsilon-1}}{\int_{\mathbb{K}} [\nu(a_{zt})q_{zt}]^{\epsilon-1} \mathrm{d}z} = \arg\max_{a_{kt}} a_{kt} \nu(a_{kt})^{\epsilon-1} \tag{9}$$

where $\pi_{\mathbb{K}t}$ denotes the average ad price at time t. By assumption of Theorem 1, A is well-defined. Thus, all platforms set a common ad rate A that is constant over time. The ad rate depends on only the nuisance cost function ν and platform substitutability ϵ. When ϵ increases, A decreases since platforms compete more aggressively for attention, which is more sensitive to ad rates, by reducing their ad rates.

Step 4: Size and Composition of Consideration Sets.—Given the ad rate A, the law of motion (1), implies that the size M_t of consideration sets satisfies

$$M_t = \frac{A}{\lambda_f} - \left(\frac{A}{\lambda_f} - M_0\right) e^{-\lambda_f t}. \tag{10}$$

As $t \to \infty$, $M_t \to M = A/\lambda_f$, its steady state level. Intuitively, when the ad rate is higher, consideration sets are larger, and when the forget rate is higher, consideration sets are smaller. For (10) to be valid, M_t must be less than the total mass J of firms in the economy at all t. This is so if $A/\lambda_f < J$, a necessary condition for equilibrium existence.

Given (10), using the law of motion (2) and accounting identity (3), in online Appendix A I characterize the cdf H_t^c of expected values outside Ω_{it} via the equation[18]

$$\int_{H_0^c(\hat{v})}^{H_t^c(\hat{v})} \frac{1}{JG(\hat{v}) - Mu^N - (J-M)u} \mathrm{d}u = \frac{\ln\left[M - M_0 + (J-M)e^{\lambda_f t}\right]}{J-M}. \tag{11}$$

Given H_t^c, I derive H_t, the cdf of expected values in Ω_{it} from accounting identity (3).

Using (11), in online Appendix A I prove that H_t^c and therefore H_t eventually converge to their steady state levels H and H^c. I also prove intuitive comparative statics in online Appendix B. I show that the positive selection in ad auctions leads to a better matching in Ω_{it} relative to the population: $H \succeq G \succeq H^c$ all in first-order stochastic dominance. Moreover, when A increases, the total value $M_t \mu_{H_t} = \int_{\Omega_{it}} v_{ij} \, \mathrm{d}j$ in Ω_{it} increases at all t (where

[18] Equation (11) applies at each point $\hat{v} \in [0, \infty)$ such that $H_0^c(\hat{v})$ is not already at its steady-state level $H^c(\hat{v})$. The steady state level $H^c(\hat{v})$ is such that the denominator of the integrand in (11) is zero when $u = H^c(\hat{v})$.

here μ_{H_t} is the mean of the cdf H_t).[19] Similarly, when G increases in the mean-preserving spread order so that data is more informative, the total value $M_t \mu_{H_t}$ in Ω_{it} likewise increases for all t sufficiently large.[20] In Step 8, we will see that $M_t \mu_{H_t}$ is a sufficient statistic for the product aggregate C_{it} which therefore increases in response to these changes.

Step 5: Firms' Expected Flow Profits and Match Rates.—Given M_t, H_t, and H_t^c, the next step is to compute a firm's expected flow profit from selling to consumer i and the rate that it matches with the consumer (i.e. enters Ω_{it}). These are necessary precursors to solving a firm's dynamic bidding problem. In particular, a firm's match rate with the consumer determines the value of the firm's outside option to wait to bid in future auctions.

Using (7) and (8), firm j's expected flow profit from selling to consumer i is

$$\pi_{\mathbb{J}t} \hat{v}_{ij}. \tag{12}$$

where

$$\pi_{\mathbb{J}t} = \frac{I}{\sigma M_t \mu_{H_t}}.$$

Thus, the expected flow profit is linear in the the firm's expected value $\hat{v}_{ij}$ to the consumer and decreasing in the total value $M_t \mu_{H_t}$ of the rival firms in Ω_{it}. As a result, data G affects flow profits and thus bidding and ad revenues which is important for comparative statics as we will see in Section 5.

Next, I compute that the Poisson rate that firm j enters Ω_{it} while outside Ω_{it} is

$$\lambda_{et}(\hat{v}_{ij}) = \lambda_{at} H_t^c(\hat{v}_{ij})^{N-1} \tag{13}$$

where recall that λ_{at} in (4) is the Poisson rate that firm j enters an auction for consumer i and $H_t^c(\hat{v}_{ij})^{N-1}$ is the probability that firm j wins an auction. Thus, when H_t^c increases in first-order stochastic dominance, match rates are lower since a firm faces stiffer competition in the ad auctions. In online Appendix B I show that when ad rate A is lower, match rates decrease since λ_{at} decreases and H_t^c increases in first-order stochastic dominance.

Step 6: Firms' Bidding Strategies.—Given the expected flow profits and match rates, I solve firm j's bidding problem using Bellman's principle of optimality. Let $V_t^{\text{In}}(\hat{v})$ denote firm j's continuation value from selling to consumer i at all points in the future if it is in Ω_{it}

[19] The total expected value in Ω_{it} coincides with the total value in Ω_{it} in that $\int_{\Omega_{it}} \hat{v}_{ij} \, \mathrm{d}j = \int_{\Omega_{it}} v_{ij} \, \mathrm{d}j$ because expectations are unbiased and errors are washed out in the aggregate.

[20] I suspect that this is true for all t but have not proven it.

at time t and if $\hat{v}_{ij} = \hat{v}$. Let $V_t^{\mathrm{Out}}(\hat{v})$ be the continuation value if j is not in Ω_{it} at time t. In online Appendix A I prove that V_t^{In} and V_t^{Out} satisfy the Hamilton-Jacobi-Bellman (HJB) equations

$$\dot{V}_t^{\mathrm{In}}(\hat{v}) = \rho V_t^{\mathrm{In}}(\hat{v}) - \lambda_f \left[V_t^{\mathrm{Out}}(\hat{v}) - V_t^{\mathrm{In}}(\hat{v})\right] - \pi_{\mathbb{J}t}\hat{v}$$
$$\dot{V}_t^{\mathrm{Out}}(\hat{v}) = \rho V_t^{\mathrm{Out}}(\hat{v}) - \lambda_{et}(\hat{v}) \left(V_t^{\mathrm{In}}(\hat{v}) - V_t^{\mathrm{Out}}(\hat{v}) - \mathbb{E}\left[B_t^{(1)} \middle| B_t(\hat{v}) > B_t^{(1)}\right]\right).$$

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. That is, $B_t^{(1)} \overset{d}{=} B_t(\hat{v}^{(1)})$ where $\hat{v}^{(1)} \sim \left(H_t^C\right)^{N-1}$.

Because the auction is second price, the optimal bid is the gain in continuation value from winning the auction: $B_t = V_t^{\mathrm{In}} - V_t^{\mathrm{Out}}$. Using this fact, in online Appendix A I solve the HJB equations to derive that

$$B_t(\hat{v}) = \int_0^{\hat{v}} \int_t^{\infty} \pi_{\mathbb{J}s} e^{-\int_t^s [\rho + \lambda_f + \lambda_{ez}(y)]\,\mathrm{d}z}\,\mathrm{d}s\,\mathrm{d}y, \quad \hat{v} \in [0, \infty). \tag{14}$$

From there, I compute that the expected ad price is

$$\pi_{\mathbb{K}t} = \int_0^{\infty} \int_t^{\infty} \pi_{\mathbb{J}s} e^{-\int_t^s [\rho + \lambda_f + \lambda_{ez}(\hat{v})]\,\mathrm{d}z}\,\mathrm{d}s\,\left[1 - O_t(\hat{v})\right] \mathrm{d}\hat{v} \tag{15}$$

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 the auction.

Intuitively, both the bid function B_t and expected ad price $\pi_{\mathbb{K}t}$ are increasing in the coefficients $\{\pi_{\mathbb{J}s}, s \geq t\}$ on future flow profits. Also, B_t and $\pi_{\mathbb{K}t}$ are decreasing in future match rates $\{\lambda_{es}, s \geq t\}$. This is because when firms match with the consumer at higher rates, the values of their outside options are higher so they bid less aggressively and the average ad price is lower. In fact, a firm's optimal bid depends on the match rates of all firms with lower expected values: this is because when firms with lower expected values reduce their bids, the values of the outside options for firms with higher expected values are further raised causing them to further reduce their bids in response.

Though λ_{et} is endogenous, if we treat it as given and send it to infinity at each point in time pointwise, we see that the average ad price vanishes, consideration sets grow to include almost all firms in the economy, and in the limit we obtain a classical frictionless economy where there are, in effect, no platforms.

Step 7: Platforms' Investment Rates.—Given the average ad price in (15) and consumers' demand in (6), in online Appendix A I solve platform k's problem using Pontryagin's Maximum Principle. There, I prove that when $\epsilon - 1 < 1/\varphi$, investment ℓ_{kt} and quality q_{kt} solve the ordinary differential equation (ODE) system

$$\begin{aligned} \dot{\ell}_{kt} &= \frac{\rho+\delta}{1-\varphi}\ell_{kt} - \frac{\varphi}{1-\varphi}\frac{\pi_{\mathbb{K}t}A(\epsilon-1)}{Kq_{kt}}\ell_{kt}^{\varphi} \\ \dot{q}_{kt} &= \ell_{kt}^{\varphi} - \delta q_{kt} \end{aligned} \tag{16}$$

with boundary conditions

$$\begin{aligned} \lim_{t\to\infty} \ell_{kt} &= \frac{\varphi\delta\pi_{\mathbb{K}}A(\epsilon-1)}{K(\rho+\delta)} \\ q_{k0} &= q_0 \end{aligned}$$

where $\pi_{\mathbb{K}} = \lim_{t\to\infty} \pi_{\mathbb{K}t}$ denotes the steady state average ad price. Numerical solutions to (16) are easily computed via the shooting method.

While the ODE for investment is not very intuitive, the long run steady state level of investment $\lim_{t\to\infty} \ell_{kt}$ *is* intuitive. It is increasing in the steady state total ad revenue $\pi_{\mathbb{K}}A$ and the elasticity of attention with respect to platform quality $\epsilon - 1$. Together, these two objects determine the marginal benefit of investing which is the ad revenue that can be stolen from rival platforms by attracting more attention.

Step 8: Surpluses.—Given the equilibrium objects from previous steps, it is a matter of accounting to show that

$$\text{Total consumer surplus} = \int_0^\infty e^{-\rho t} u(C_{it}, X_{it})\, \mathrm{d}t \tag{17}$$

$$\text{where } C_{it} = I(M_t \mu_{H_t})^{\frac{1}{\sigma-1}}$$

$$X_{it} = K^{\frac{1}{\epsilon-1}} \nu(A) q_t$$

$$\text{Total firm surplus} = \int_0^\infty e^{-\rho t} \left(I - \pi_{\mathbb{K}t}A\right)\, \mathrm{d}t \tag{18}$$

$$\text{Total platform surplus} = \int_0^\infty e^{-\rho t} \left(\pi_{\mathbb{K}t}A - K\ell_t\right)\, \mathrm{d}t. \tag{19}$$

Above, in the expression for the platform aggregate X_{it}, q_t denotes the common quality level of all platforms in equilibrium. Also, note that the product aggregate C_{it} depends on only the total value $M_t \mu_{H_t}$ of the firms in the consideration set.

5. Comparative Statics

Using the equilibrium characterization, we can now investigate the potential effects of policies on the economy. In what follows, I analyze the effect of an interoperability policy that raises the substitutability among platforms and a data policy that leads to a mean-preserving spread of firms' expectations of consumers' values. Because of search frictions and because investment is gradual, the economy's transition dynamics in response to these policies are potentially significant and lengthy. I illustrate the joint dynamics of outcomes on each of the market sides with numerical examples. However, the comparative statics across steady states hold more generally and I present analytical results in Table 1.

Because the model is highly stylized, the results in this section are intended to be pedagogical and not a definitive policy analysis. However, the economic channels that I identify appear natural and are strongly indicative of the importance of a multisided analysis for policy. I find that: (i) cross-side spillovers can be strong and reverse seemingly intuitive comparative statics based on single-sided analyses; (ii) the effects of policies can flip over time as shocks propagate across the market sides; (iii) policies typically must trade off product consumption with platform consumption.

A Shock to Platform Substitutability.—At $t = 0$, the economy begins in steady state for the parameters listed below Figure 1. At $t = 1$ there is an unanticipated shock to ϵ which increases by 1 percent from its initial level of 1.33. This shock may be the result of a policy that mandates greater platform interoperability. For example, it may be that platforms must now allow users to post links to other platforms on their websites or any other technological change that makes it more convenient for users to shift attention across platforms. Presumably, the logic behind the policy is that it will promote competition by making it so that consumers have a meaningful choice as to how to spend their attention.[21]

Policies like these are often proposed (OECD, 2021). An example is the Digital Markets Act, which seeks to make large platforms more interoperable, both with each other, and with smaller platforms.[22] As discussed in Section 3, a limitation is that I consider only platforms that are strategically small. At the end of this section, I discuss in depth, why I expect

[21] "Interoperability is a key aspect of many digital platform services, as it gives users access to a wide range of choices" (OECD, 2021). Interoperability policies are technological changes that alter platforms' services. Thus, one must take a stance on how consumers' preferences for platforms change in response. In line with the spirit of these policies, I model the change in terms of the substitutability parameter of the CES technology.

[22] There has also been discussion about whether interoperability mandates should apply more broadly, regardless of a platform's size (e.g., see the section on reciprocity at `https://www.newamerica.org/oti/briefs/how-to-make-the-access-act-a-success/` or page 38 of OECD (2021) or the first paragraph of page 25 of `https://crsreports.congress.gov/product/pdf/R/R47662/3`).

the qualitative patterns in Figure 1 to persist, at least for some parameters, when platforms are atomic. However, one can also interpret platforms in the model as individual content creators who are small though they operate on large platforms. Content creators invest in their own content and often exercise some discretion over their ad rates.[23]

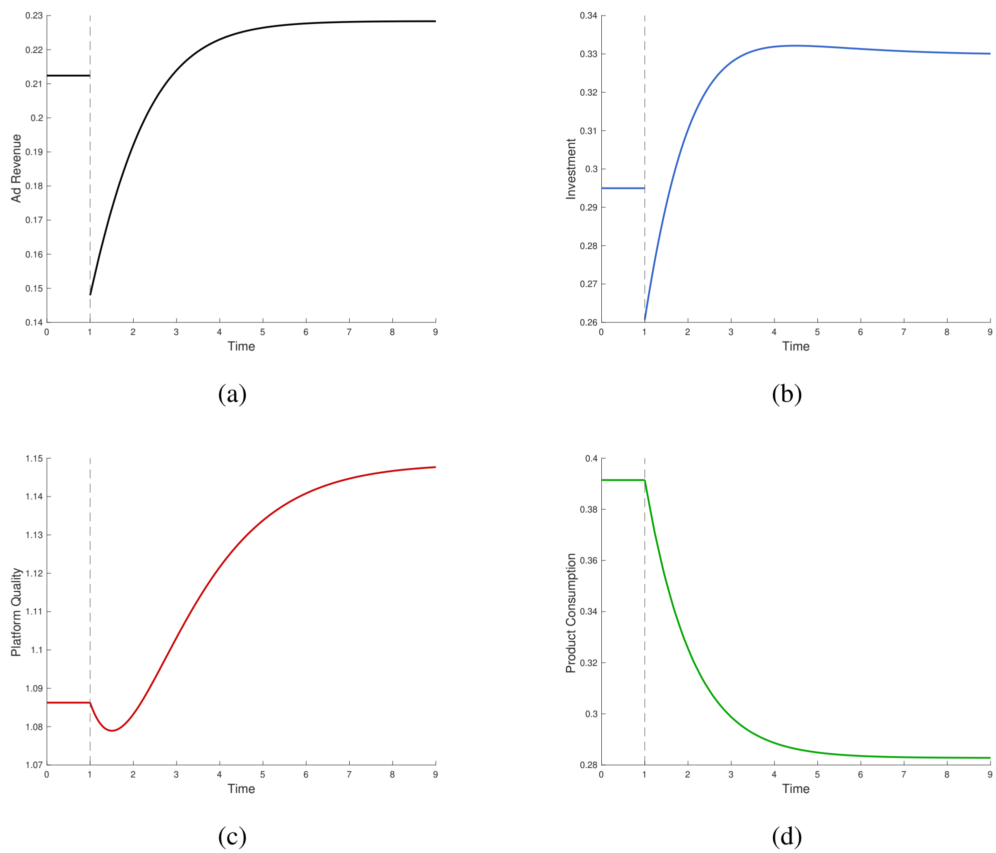


Figure 1: Transition dynamics of a shock to ϵ

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 $\lambda_f = 1$; $K = .1$; $J = 1$; $\rho = .1$; $\sigma = 3$; $N = 5$; $G = \mathrm{U}[0, 1]$; $I = 1$; $\varphi = .5$; $\delta = .5$; $\nu(a) = (1 - .8395a^{.01})^{63.1472}$.

Figure 1 plots the economy's transition dynamics for several outcomes of interest. As seen in Panel (a), ad revenue initially spikes down. This is because the ad rate A drops

[23] Content creators on Twitch and YouTube often bargain over ad rates when signing contracts (in the case of Twitch, see https://www.washingtonpost.com/video-games/2022/06/20/twitch-ad-incentive-money-payout-55-percent/). An issue with this analogy is that one would expect an interoperability policy to raise the substitutability of content relatively more between creators on different platforms whereas the experiment raises substitutability among all creators by the same amount.

(in this case from .2 to .1): platforms display ads at a lower rate because they can more easily steal attention from each other after the shock. Despite the drop in A, ad revenue eventually increases beyond its initial steady-state level for two reasons. First, consumers' consideration sets decrease in size causing firms' flow profits to increase. Second, the value of a firm's outside option when bidding decreases since there is greater delay in the time it takes a firm to enter a consumer's consideration set. For these reasons, firms bid more in the ad auctions, causing the average ad price $\pi_{\mathbb{K}t}$ to increase to the extent that even though A is lower, ad revenue $\pi_{\mathbb{K}t}A$ is higher in the long run.

As seen in Panel (b), investment follows a similar pattern (though it slightly overshoots before declining to its steady-state level). Perhaps surprisingly, investment *decreases* in the short run. However, after a not insignificant amount of time, platforms do indeed invest more. If we were to ignore the effect of the shock on ad revenue, we would predict that the new steady-state level of investment is 4 percent higher as opposed to the 10 percent higher that it is in the example.

As seen in Panel (c), platform quality decreases in the short run before increasing past its initial steady-state level. In the long run, since platform quality is higher and ads are displayed at a lower rate, the platform aggregate X_{it} is also higher. However, as seen in Panel (d), the product aggregate C_{it} decreases over time as consumers' consideration sets shrink because A is lower. The interoperability policy therefore must trade off product consumption with platform consumption in the long run.

A Shock to Data.—At $t = 0$, the economy begins in steady state for the parameters listed below Figure 3. At $t = 1$ there is an unanticipated shock to the cdf G which increases in the mean-preserving spread order from $\mathrm{U}[.2, .8]$ to $\mathrm{U}[0, 1]$. This shock may be the result of a new policy that scales back on privacy laws, allowing platforms to collect more information on consumers.[24] As a result, firms' expected values become more disperse as they Bayesian update on the new information. Namely, a firm's expectation $\hat{v} \sim \mathrm{U}[.2, .8]$ of a consumer's value for its product is shocked to a new level $\tilde{v} \sim \mathrm{U}[0, 1]$.

The particular joint distribution of $\hat{v}$ and $\tilde{v}$ is irrelevant for the new steady state. However, it is relevant for transition dynamics since it determines the cdf of expectations in consideration sets right after the shock. As a result, I specify the joint distribution illustrated in Figure 2. 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

[24] Examples of privacy laws include the General Data Protection Regulation and California Consumer Privacy Act which apply to all websites on the Internet.

distributed uniformly across the upper (lower) black region. The probability that $\hat{v}$ jumps

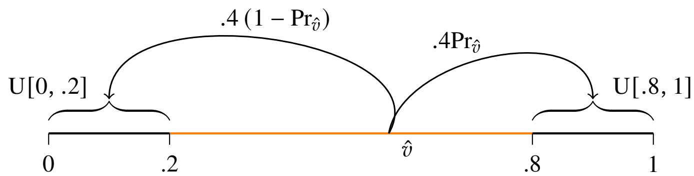


Figure 2: Joint distribution of pre and post shock expected values

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

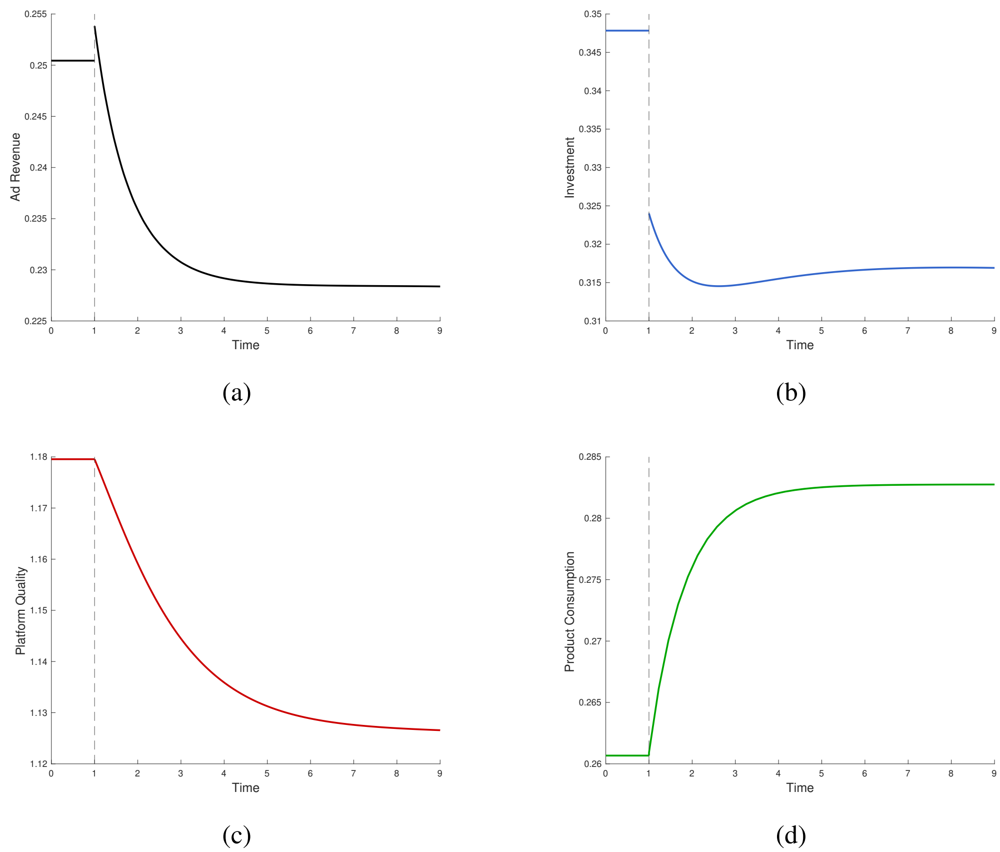


Figure 3: Transition dynamics of a shock to G

Notes: I plot the transition between steady states following an unanticipated change in G from U[.2, 8] to U[0, 1] at $t = 1$ described in Figure 2. The other parameters are $\rho = .1$; $\epsilon = 1.33$; $\lambda_f = 1$; $P = .1$; $F = 1$; $\sigma = 3$; $N = 5$; $I = 1$; $\varphi = .5$; $\delta = .5$; $\nu(a) = 1 - 7.5a$.

Figure 3 plots the economy's transition dynamics for several outcomes of interest. As seen in Panel (a), ad revenue initially spikes up. This is because more informative data leads to a better matching of firms and consumers and platforms are able to extract some of this surplus gain in the ad auctions (despite the fact that information rents also increase). However, over time, ad revenue *decreases* and eventually settles at a lower steady-state level. There are two forces that lead to this. First, each firm internalizes that it faces competition from firms in consideration sets that are now on average more highly valued relative to before the shock. Thus, any given firm's flow profits from selling to any given consumer are lower. Firms therefore bid less for ads. Second, more informative data leads firms to match more quickly on average with consumers who value their products highly raising the values of firms' outside options further causing them to bid less for those consumers. These two effects dominate, causing ad revenue to decrease in the long run. To see why these effects must dominate, note that firms' total flow revenue is fixed and equal to the income I of consumers. Thus, over time, the gain in surplus to firms from better matches must be competed away as consideration sets fill up with rival firms that are also better matches.

As seen in Panel (b), because the increase in ad revenue is short-lived, platform investment ℓ_{kt} after the shock is lower than before at each point in time. A failure to account for the effects of data on the product market may lead policymakers to predict an increase in ad revenue and thus greater platform investment in the long run, contrary to what happens in this example. As seen in Panel (c), the decrease in investment leads platform quality q_t (and thus the platform aggregate X_{it}) to decrease at all times. However, the opposite is true for the product aggregate C_{it} as seen in Panel (d) which increases as consumers match with firms they value more highly on average. Thus, the data policy also leads to a tradeoff between platform consumption and product consumption.

Analytical Results for Steady State.—The comparative statics across steady states in the above examples hold more generally. Table 1 summarizes results proven for all parameters.

Table 1: Comparative Statics Across Steady States

		C	X	$\pi_{\mathbb{K}} A$	ℓ
$\epsilon \uparrow$	$\Rightarrow$	$\downarrow$	$\uparrow$ if $K \leq 1$	$\uparrow$	$\uparrow$
$G(\hat{v}) = \mathbb{1}_{\{\hat{v} \geq \mu_F\}}$	$\Rightarrow$	minimal	maximal	maximal	maximal
$J \uparrow$	$\Rightarrow$	$\uparrow$	$\uparrow$	$\uparrow$	$\uparrow$
$K \uparrow$	$\Rightarrow$	$-$	$\uparrow$	$-$	$\downarrow$

Notes: When $G(\hat{v}) = \mathbb{1}_{\{\hat{v} \geq \mu_F\}}$, data is uninformative because all mass is concentrated on the mean μ_F of the prior F. The symbol $\uparrow$ denotes increase, $\downarrow$ denotes decrease, and $-$ denotes no change.

Formal statements and proofs are found in online Appendix B.

What if There Are A Finite Number of Platforms?—If there are a finite number of platforms, then each one must account for the strategic effects of its actions on other platforms as well as the effect of its ad rates on consideration sets and firms' bids. I do not conduct a formal analysis, but I expect these factors might mute certain effects but the qualitative patterns in Figures 1 and 3 should persist (though perhaps for a different set of parameters).

For example, a positive shock to platform substitutability ϵ, would still incentivize each platform to reduce its ad rates to steal business, though the effect might be lessened due to the anticipated retaliation by rival platforms. On the other hand, it may be amplified as now platforms recognize the positive effect of greater frictions in the product market on their ad revenues in the long run. Under some assumptions on the nuisance cost function ν, the decline in ad rate should be large enough for ad revenue to decrease in the short run. Similarly, the relationship between ad revenues and investment may be muted due to strategic effects but should still be positive. Thus, the short-run patterns of the shock will likely be qualitatively the same as in the numerical example for a range of parameters (perhaps requiring a higher rate of quality depreciation δ or discounting ρ). The long-run patterns should also continue to hold since the decline in ad rate eventually leads to an increase in ad revenue regardless of whether platforms are large.

Similarly, I expect a positive shock to data informativeness to typically have the same patterns as in the example. In the short run, what might change is that platforms may raise their ad rates A to capitalize on the higher average ad price. This would mute the short-term increase in the average ad price (which is decreasing in A as shown in online Appendix B) but not reverse it. Thus ad revenue would rise in the short run as in the example. In the long run, the ad rate would decline once the average ad price declines as consumers' consideration sets fill up with more favorable firms. This would mute the decline in average ad price, but clearly can not reverse it. Thus, in the long run, ad revenue will decline and so too will investment though perhaps by not as much as in the absence of strategic effects.

6. Empirical Trends

> *"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)

In Section 5, I showed that giving all platforms access to more informative data may be harmful to them in that it reduces their ad revenues (and profits) in the long run. This counterintuitive result relied on the impact of data on the product market. To offer suggestive evidence in support of this channel, I show that by accounting for this channel, we can qualitatively explain two empirical trends that many consider puzzling. In the process, we will see that individual platforms do benefit from better data, even though platforms are collectively made worse off if they *all* have access to better data.

Silk et al. (2021) document that

1. Ad revenue as a fraction of GDP in the United States has been relatively stable since the 1920s but has experienced a slight decline in the past decade.

2. As a share of ad revenue, digital advertising has grown dramatically in the past decade while traditional advertising has declined.

In light of the unprecedented use of data for ad targeting, one might have expected entry of digital platforms to lead to a large increase in ad revenue. On the contrary, though digital advertising has grown dramatically, total ad revenue as a fraction of GDP has been in slight decline over the past decade. I now show that the model, when extended to include traditional platforms, can qualitatively match both trends simultaneously.

Extension: Data Heterogeneity.—I extend the model to allow for two groups of platforms that differ in their data. Group 1 platforms are data poor and represent traditional media like newspapers or television. Group 2 platforms are data rich and represent online platforms. The general formulation of this extension is found in online Appendix G (and may interest readers seeking to explore the effects of policies that force platforms to share data or more generally, analyze the relationship among data, platform quality, and platform market share). Below, I describe only the setting for this section's numerical illustration.

I assume that 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. For simplicity, the measure of group 1 platforms is equal to that of group 2 platforms. Firm j sees the signal ζ_{lij} of consumer i's value v_{ij} when bidding on a group l platform. I assume that

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

and

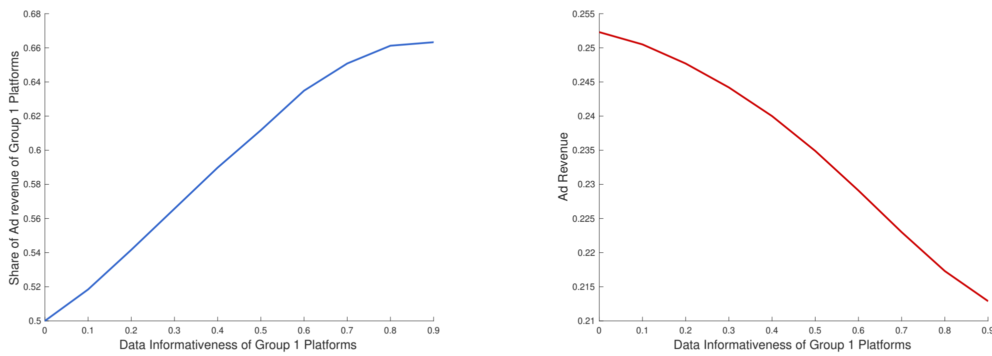
$$\zeta_{2ij} = Z_{ij} + u$$

where $0 \leq \Delta \leq 1$ and $u \sim N(0, \sigma_u^2)$ is independent of all other model primitives.[25]

[25]When platforms have heterogeneous data, it no longer suffices to specify only the cdfs of posterior ex-

Thus, group 1 platforms are data rich while group 2 platforms are data poor. All other aspects of the model are as in the baseline. In equilibrium, there are now two bidding functions, one for each platform group. Attention shares and average ad prices now also differ across the two groups.

Numerical Illustration.—The figure below plots the steady state share of ad revenue accruing to group 1 platforms as a function of group 1 platforms' data advantage for the parameters listed below the figure.



(a) (b)

Figure 4: Group 1 share of total ad revenue and total ad revenue

Notes: We plot the group 1 share of ad revenue and total ad revenue 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.

The quantity on the x-axis is $1 - \Delta$. The left-most point when $\Delta = 1$ can be thought of as the case when both platform groups comprise only traditional media while any point to the right can be thought of as arising after entry of digital platforms. Note that total ad revenue and the shares of ad revenue of the respective groups are invariant to the total measure of platforms in the economy. Thus, when comparing the point at $\Delta = 1$ to any other point, the comparison is consistent with any pattern of entry of data-rich and exit of data-poor platforms as long as the new steady state comprises an equal share of them both. Consistent with the first empirical trend, group 1 platforms have a higher share of the ad revenue and this share is increasing in their data advantage. In this example, the higher share of ad revenue results because group 1 platforms command both higher ad prices and

pectations. This is because a firm must conduct inference on its expected value to a consumer on the other platform group when bidding in a given platform's auction to assess the value of its outside option. To form this expectation, the firm needs to know the joint distribution of the underlying signals.

higher shares of attention. However, consistent with the second empirical trend, total ad revenue decreases in the data advantage of group 1 platforms.

7. Welfare Analysis

This section analyzes welfare in a general equilibrium extension of the baseline model of Section 2. I solve the problem of a social planner who sets the rates that platforms display ads and invest in their services to maximize welfare, taking into account the balance of objectives on the different market sides. For tractability, the analysis focuses on the planner's steady-state solution (defined below). I compare the planner's steady-state solution to equilibrium investment and ad rates in steady state to identify the sources of inefficiency. I then show that a tax or subsidy on platforms' ad revenues funded by or redistributed to consumers can restore platforms' steady-state investment to the efficient level. I also characterize the efficient tax or subsidy.

Extension: General Equilibrium.—In the baseline model, firm and platform surplus are expressed in different units from consumer surplus. To obtain a natural welfare measure, I first extend to general equilibrium.

I endogenize the income I of consumers in the baseline model. Each consumer supplies L units of labor inelastically at each point in time. Labor is the only productive resource in the economy and is used by platforms for investment and by firms for production. The wage is the numeraire in the economy and is set to 1. Thus, ℓ_{kt} now denotes the labor hired by platform k for investment at time t. Also, assuming each unit of output requires a unit of labor, the marginal cost of production is 1 just as in the baseline model. Let

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

denote the total quantity of labor hired by firm j at time t. The labor market clears at t if

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

I assume that each consumer owns an equal share of all firms and platforms in the economy. This implies that the relevant measure of social welfare is then consumer surplus. The income of a consumer at t is

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

which is the total revenue earned by firms. This is because platforms, who charge zero prices, merely extract revenues from firms and all costs are labor costs. Note also that this condition is implied by product market clearing at t.

Except for the changes above, I retain all other aspects of the baseline model. An equilibrium for initial conditions M_0, H_0, and q_0 is defined as before, except with the additions of a process for income $\{I_t\}$ and the condition that the labor market and product market clear at all times.

In online Appendix C, I prove an analog of Theorem 4 characterizing the full general equilibrium. Proposition 1 presents only equilibrium ad rates and investment rates in steady state.

Proposition 1. *Suppose that A is the unique solution of $\max_a \nu(a)^{\epsilon-1}$. If $A/\lambda_f < F$ and $\epsilon - 1 < 1/\varphi$, then there exists a unique equilibrium. The equilibrium converges to a steady state where*

1. *Each platform k displays ads at rate $a_{kt} = A$ at each time t.*
2. *Each platform k invests at rate $\ell_{kt} = \ell_{\mathbb{K}}$ where*

$$\ell_{\mathbb{K}} = \frac{\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathbb{K}}A(\epsilon-1)}{\rho+\delta+\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathbb{K}}A(\epsilon-1)}\frac{L}{K} \tag{20}$$

and $\hat{\pi}_{\mathbb{K}} = \pi_{\mathbb{K}}/I$ denotes the average ad price price per unit of income and satisfies equation (34) in online Appendix C.

Moreover, the sufficient conditions are almost necessary: if either $A/\lambda_f \geq J$ or $\epsilon-1 > 1/\varphi$, then an equilibrium does not exist.

Planner's Problem.—The planner sets platforms' investment rates and ad 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 NPV of their flow profits. I assume that the planner is constrained to treat platforms symmetrically. I could also allow the planner to set firms' prices but those will be efficient anyway since each firm will choose the same price and thus produce the same amount. Further, the allocation of ad opportunities will be efficient as well since the firms with the highest expected values win in each auction.[26]

Formally, the planner solves

$$\max_{\{\ell_{\mathbb{K}t}\},\{A_t\}} \int_0^\infty e^{-\rho t} u(C_t, X_t)\,\mathrm{d}t \tag{21}$$

[26] This assumes that the planner is subject to the same data that platforms have.

subject to

$$\begin{aligned}
C_t &= (L - K\ell_{\mathbb{K}t})(M_t \mu_{H_t})^{\frac{1}{\sigma-1}},\\
X_t &= K^{\frac{1}{\rho-1}} \nu(A_t) q_{\mathbb{K}t},\\
\dot{M}_t &= A_t - \lambda_f M_t,\\
(M_t\dot{H}_t) &= A_t (H_t^c)^N - \lambda_f H_t,\\
M_t H_t + (J - M_t) H_t^c &= G,\\
\dot{q}_{\mathbb{K}t} &= \ell_{\mathbb{K}t}^{\varphi} - \delta q_{\mathbb{K}t},
\end{aligned}$$

given initial conditions M_0, H_0, and q_0.

To obtain a sharp characterization of the planner's solution, I assume that utility is Cobb-Douglas for the rest of this section:

$$u(C_t, X_t) = C_t^{1-\tau} X_t^{\tau} \tag{22}$$

where $0 < \tau < 1$ is the weight on platform consumption.

I also restrict attention to the *steady-state solution* of the planner's problem defined by a constant ad rate A^*, investment rate $\ell_{\mathbb{K}}^*$, and initial conditions M^*, H^*, q^* such that the planner solves (21) by setting $A_t = A^*$ and $\ell_{\mathbb{K}t} = \ell_{\mathbb{K}}^*$ at each t and moreover, $M_t = M^*$ and $H_t = H^*$ at each t.[27]

In online Appendix C, I prove the following Theorem 2, which characterizes investment in the steady-state solution for an arbitrary discount rate ρ and the ad rate in the steady-state solution in the limit as $\rho \to 0$.[28]

Theorem 2. *Let u be as in (22). Then, any steady-state solution to (21) has investment*

$$\ell_{\mathbb{K}}^* = \frac{\varphi\delta\frac{\tau}{1-\tau}}{\rho + \delta + \varphi\delta\frac{\tau}{1-\tau}} \frac{L}{K}. \tag{23}$$

Suppose that a steady-state solution to (21) exists for all ρ in a neighborhood of zero. Let $A^*(\rho)$ denote the ad rate in the steady-state solution when the discount rate is ρ whenever

[27] I conjecture that under some technical conditions, any solution to the planner's problem eventually converges to the steady state solution but have not investigated this formally.

[28] I expect that the solution for the ad rate for arbitrary ρ can be obtained using a version of Pointryagin's Maximum Principle extended to an infinite-dimensional state space (since the planner must keep track of the cdf H_t).

the steady-state solution exists. Then

$$\lim_{\rho \to 0} A^*(\rho) = \arg\max_a \, [a \mu_H(a)]^{\frac{1-\tau}{\sigma-1}} \, \nu(a)^\tau \tag{24}$$

whenever the right-hand side is well-defined where $\mu_H(a)$ 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 rates but not their ad rates, then the unique steady-state solution for the planner's choice of investment remains as in (23).*

I now compare the planner's choice of investment and ad rate to their equilibrium counterparts, beginning first with investment.

Investment.—By comparing equilibrium investment (20) in steady state and the planner's steady-state choice of investment (23) we can prove the following proposition.

Proposition 2. *Holding τ fixed, the deviation $\ell_{\mathbb{K}} - \ell^*_{\mathbb{K}}$ between steady-state equilibrium investment (20) and the planner's steady-state solution for investment (23) is increasing in*

$$\frac{\sigma}{\sigma - 1} \hat{\pi}_{\mathbb{K}} A(\epsilon - 1) - \frac{\tau}{1 - \tau}. \tag{25}$$

When (25) is zero, equilibrium investment in steady state coincides with the planner's steady-state investment. When (25) is positive (negative), equilibrium investment in steady state is too high (too low) relative to the planner's steady-state investment. Each of these outcomes is possible as seen by changing τ (since $\hat{\pi}_{\mathbb{K}}$ does not depend on τ).

The statistic (25) gives a simple way of gauging the efficiency of equilibrium investment in the model and depends on relatively few objects: (i) the markup $\sigma/(\sigma - 1)$; (ii) the ad revenue $\hat{\pi}_{\mathbb{K}} A$ per unit of income; (iii) the elasticity $\epsilon - 1$ of attention with respect to platform quality; and (iv) the weight τ in utility on platform consumption.[29]

Each of these objects represents a source of inefficiency. The markup $\sigma/(\sigma - 1)$ appears due to the monopoly power of product firms. This monopoly power leads them to demand less labor for production resulting in more labor allocated to investment. The average ad revenue per unit of income $\hat{\pi}_{\mathbb{K}} A$ is closely related to the business stealing incentives of product firms. By entering the consideration set of a consumer, a firm steals business from the rival firms in the set. If it can steal more, it will bid more aggressively in the ad auction, leading $\hat{\pi}_{\mathbb{K}} A$ to increase. These business-stealing incentives by firms also generate

[29]One potential method for estimating ϵ and τ jointly would be to run an experiment in which participants are charged prices for using platforms.

business-stealing incentives for platforms who invest solely to steal ad revenue from each other. How much ad revenue platforms can steal from each other in turn depends on the elasticity $\epsilon - 1$ of attention with respect to platform quality.

As a result, equilibrium investment is determined by factors that are entirely distortionary: monopoly power and business-stealing externalities. Neither of these are directly related to the surplus generated for consumers through investment which is what determines the planner's investment incentives. This is why $\tau/(1-\tau)$ appears in (25). Platforms fail to internalize the surplus that they generate for consumers with their investments because they do not appropriate any of that surplus since platforms charge a price of zero.

These distortions can be corrected by simply placing a proportional tax/subsidy on platforms' ad revenues that is financed by/redistributed to consumers. The tax/subsidy is set so that the ad revenue per income after adjusting for the tax/subsidy is such that (25) is zero.

Proposition 3. *A proportional tax/subsidy on platforms' ad revenues equal to*

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

that is funded by/redistributed to consumers can restore steady-state equilibrium investment to the efficient level $\ell^*_{\mathbb{K}}$ defined in (23).

Another useful property of (25) is that it can be used to sign the effects of changes to data and interoperability on the efficiency of investment. That is, using the results in Table 1 (which can be shown to extend to the general equilibrium setting if ad revenue $\pi_{\mathbb{K}}A$ is replaced with ad revenue per unit of income $\hat{\pi}_{\mathbb{K}}A$), we see that an increase in ϵ increases $\hat{\pi}_{\mathbb{K}}A$, bringing investment closer to its efficient level if it is initially too low or pushing it farther from its efficient level if it is initially too high. Typically, improving the informativeness of data reduces $\hat{\pi}_{\mathbb{K}}A$ leading to the opposite effects.

Ad Rate.—In equilibrium, the 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 depending on parameters. For example, suppose that ν is continuous and $\nu(\overline{a}) = 0$ after some level

$\overline{a} > 0$.[30] As ϵ tends to 1, the equilibrium ad rate converges to $\overline{a}$ whereas the planner's ad rate is unaffected. Thus, the ad rate can be higher in equilibrium than under the planner. This is because platforms internalize the nuisance costs from ads only to the extent that the nuisance costs affect how much attention platforms can steal from their rivals. Therefore, when ϵ is small, platforms do not internalize nuisance costs enough. The key reason for this is because platforms charge zero prices.

On the other hand, when τ tends to zero, the planner's choice of ad rate converges to $\overline{a}$ whereas the equilibrium ad rate is unaffected. Thus the ad rate can be lower in equilibrium than under the planner. The intuition is that when τ is near zero, platform consumption is not important for welfare, while product consumption is very important. Thus the planner seeks to expose consumers to as many products as possible. However, in equilibrium, platforms do not internalize the impact of exposing consumers to ads on product consumption. Thus the ad rate can be lower in equilibrium than under the planner.

8. Summary of Extensions

I summarize five extensions of the baseline model of Section 2.[31]

Network Effects.—The first extension, found in online Appendix F, introduces network effects. I assume that a consumer's flow utility from platform use depends on how other consumers allocate their attention. That is, the effective quality of platform k is modified to $\eta(x_{kt})\nu(a_{kt})q_{kt}$ where η is increasing in $x_{kt} = \int_{\mathbb{I}} x_{ikt}\,\mathrm{d}i$. Analogous to (6) of the baseline model, the equilibrium attention received by platform k is

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

Note that x_{kt} appears on both sides above. To obtain explicit solutions for x_{kt} we assume $\eta(x) = x^{\zeta}$ where $\zeta > 0$ controls the strength of the network effects.

For each subset $E_t \subset \mathbb{K}$ of positive measure, there is a solution that sets

$$x_{kt} = \frac{[\nu(a_{kt})q_{kt}]^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}}}{\int_{E_t}[\nu(a_{zt})q_{zt}]^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}}\mathrm{d}z}$$

[30] In the baseline model, I assumed $\nu > 0$. This is just for ease of exposition for the technical reason that demand (6) is not well-defined if almost all platforms set an ad rate a such that $\nu(a) = 0$. A version of Theorem 1 still applies even if $\nu(a)$ vanishes at some point $\overline{a} > 0$.

[31] Each of these extensions can also be analyzed in general equilibrium.

if $k \in E_t$ and otherwise sets $x_{kt} = 0$. Thus, network effects lead to equilibrium multiplicity. Under the refinement that $E_t = \mathbb{K}$ at each t, there is a unique equilibrium, characterized just as in the baseline model, except with a higher elasticity of platform demand equal to $(\epsilon - 1)/[1 - \zeta(\epsilon - 1)] > \epsilon - 1$. Thus, network effects lead platforms to display ads at a lower rate and invest more in the long run. Under the refinement, $\mathbb{K}$ represents the set of platforms that remain active in the economy at each t with stable market shares.

Heterogeneous Platform Data.—The second extension, found in online Appendix G, gives a more general formulation of the model discussed in Section 6. I provide an algorithm to compute steady-state equilibrium for the model with two groups of platforms that may differ in the informativeness of their data. This extension can be used to study policies that mandate that platforms share their data or more generally, the relationship among data, platforms' market shares, and quality levels.

Heterogeneous Platform Productivity.—The third extension, found in online Appendix H, allows platforms to be heterogeneous with respect to the productivities of their investments. I characterize full equilibrium dynamics using analogous methods to those of the baseline model. This extension can be used to generate an essentially arbitrary non-atomic distribution of platform size by varying the distribution of productivities, which improves the model's potential for a quantitative analysis.

Firm and Platform Entry.—The fourth extension, found in online Appendix I, introduces firm and platform entry. To enter, a firm or platform must pay an up front cost. I characterize the unique steady-state equilibrium and find that an increase in data informativeness often leads to entry of firms and exit of platforms while an increase in platform substitutability leads to exit of both firms and platforms.

Reserve Prices.—The fifth extension, found in online Appendix J, introduces reserve prices that are set optimally by platforms. I find that in candidate steady-state equilibria reserve prices are generally positive. This is so even with a continuum of platforms because of search frictions in the ad market. In any steady-state equilibrium, ad revenues are maximal when data is uninformative just as in the baseline model.

9. Conclusion

This paper has presented a general equilibrium model of the market for attention. The model is among the first to endogenize outcomes on each of the market sides, including the allocation in the product market, ad revenue, and the quality of platforms' services.

Because these outcomes are jointly determined and are relevant for welfare, such a model is desirable. I show how these outcomes relate to each other, are determined in equilibrium, and how they evolve together in response to changes to data and platform substitutability.

I apply the model to investigate the potential effects of interoperability and data policies. I show that (i) seemingly intuitive comparative statics based on single-sided analyses can flip. For example, in the long run, allowing platforms to have access to more informative data may be *harmful* to them and in the short run, interoperability policies may appear to reduce platform competition in that platforms invest *less* in their services. (ii) The short run effects of policies may look very different from the long run effects as shocks gradually propagate across the market sides. Extrapolating short run trends to predict long run outcomes may thus be misleading. In the short run, more informative data may appear to benefit platforms and in the long run, interoperability policies do raise platforms' investment rates. (iii) There are tradeoffs across the market sides: platform consumption is often traded off with product consumption. To assess whether a policy is on the whole beneficial or harmful, both types of consumption should be taken into account. The welfare analysis in this paper provides some sense of the economic conditions when we may expect policies to bring the economy closer to first best.

The model is stylized and abstracts from some potentially important effects. Because I model monopolistically competitive platforms, the analysis may omit forces that are relevant for platforms that are a large relative to the sector, who may consider the strategic effects of their actions on other platforms or on the sizes of consumers' consideration sets and firms' continuation values.[32] However, this does not mean, that the platforms in the model do not have market power. Each platform is a monopolist in its category of service and when substitutability is low, effectively holds its share of consumer attention captive. Further, platforms also have market power in the ad market due to search frictions and compete only via the outside options of firms to match with consumers through other platforms.

The model also abstracts from the potential effects of personalized pricing. Because

[32] In 2023, Meta (Facebook and Instagram) earned close to 23 percent of global digital ad revenue which was 602 billion according to https://www.statista.com/statistics/237974/online-advertising-spending-worldwide/. YouTube earned 5.5 percent. X, TikTok, Snapchat, Pinterest, Twitch, and Spotify and many others each earned far less. Even if we remove search-engine advertising (which perhaps fits my model assumptions somewhat less well) from consideration, the numbers remain qualitatively the same. Specifically, according to https://www.oberlo.com/statistics/google-ad-revenue, Google's search-engine ad revenue was 175 billion in 2023. Since Google search has 92 percent of the global search market, an estimate of global search-engine ad revenue in 2023 is 190 billion. Thus global digital ad revenue excluding search-engine ad revenue was roughly 412 billion in 2023. Of this, Meta (Facebook and Instagram) earned close to 33 percent. YouTube earned 7.7 percent. X, TikTok, Snapchat, Pinterest, Twitch, and Spotify and many others earned far less.

I assume CES preferences for products, it is optimal for the firms in the model to charge a fixed markup and not to personalize prices. This greatly simplifies the analysis but is a special property of CES preferences. For an analysis of the effects of personalized pricing, see Rhodes & Zhou (2022) and the literature discussed there.

Because I have abstracted from these and other potentially important effects, the policy analysis in this paper is only suggestive. The primary purpose of the analysis is to shed light on some economic channels resulting from interactions among the market sides that can potentially be strong and lead to counterintuitive effects. We can not analyze these channels in other models, which generally treat one or more sides of the market as exogenous. Yet, these channels appear natural and worthy of serious consideration by policymakers. Future work may seek to incorporate additional effects into this paper's setup in order to make progress toward a serious quantitative exploration of policies.

References

Alcobendas, M., Kobayashi, S., & Shum, M. (2021). The impact of privacy measures on online advertising markets. *Available at SSRN 3782889*.

Ambrus, A., Calvano, E., & 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. & Coate, S. (2005). Market provision of broadcasting: A welfare analysis. *The review of Economic studies*, 72(4), 947–972.

Armstrong, M. (2006). Competition in two-sided markets. *The RAND journal of economics*, 37(3), 668–691.

Athey, S. & Ellison, G. (2011). Position auctions with consumer search. *The Quarterly Journal of Economics*, 126(3), 1213–1270.

Bagwell, K. (2007). The economic analysis of advertising. *Handbook of industrial organization*, 3, 1701–1844.

Beales, H. & Eisenach, J. A. (2014). An empirical analysis of the value of information sharing in the market for online content. *Available at SSRN 2421405*.

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

Bergemann, D., Bonatti, A., & Gan, T. (2019). The economics of social data.

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

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

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

Caillaud, B. & Jullien, B. (2003). Chicken & egg: Competition among intermediation service providers. *RAND journal of Economics*, (pp. 309–328).

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

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

Duffie, D., Gârleanu, N., & Pedersen, L. H. (2005). Over-the-counter markets. *Econometrica*, 73(6), 1815–1847.

Duffie, D., Qiao, L., & Sun, Y. (2020). Continuous-time random matching. *Graduate School of Business, Stanford University*.

Edelman, B., Ostrovsky, M., & Schwarz, M. (2007). Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 97(1), 242–259.

Engelking, R. (1977). General topology.

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

Farboodi, M. & Veldkamp, L. (2021). *A Growth Model of the Data Economy*. Technical report, National Bureau of Economic Research.

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

Hagiu, A. & Wright, J. (2015). Multi-sided platforms. *International Journal of Industrial Organization*, 43, 162–174.

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

Johnson, G. A., Shriver, S. K., & Du, S. (2020). Consumer privacy choice in online advertising: Who opts out and at what cost to industry? *Marketing Science*, 39(1), 33–51.

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

Jullien, B. & Bouvard, M. (2022). Fair cost sharing: big tech vs telcos.

Jullien, B., Pavan, A., & Rysman, M. (2021). Two-sided markets, pricing, and network effects. In *Handbook of Industrial Organization*, volume 4 (pp. 485–592). Elsevier.

Khan, L. (2017). Amazon's antitrust paradox. *Yale law journal*, 126.

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

Marotta, V., Abhishek, V., & Acquisti, A. (2019). Online tracking and publishers' revenues: An empirical analysis. In *Workshop on the Economics of Information Security*.

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 (2020). Consumer data rights and competition - background note.

OECD (2021). Data portability, interoperability, and digital platform competition. *OECD Competition Committee Discussion Paper*.

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

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

Rachel, L. (2021). Leisure-enhancing technological change. *Job market paper, London School of Economics*.

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

Rochet, J.-C. & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4), 990–1029.

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.

Varian, H. R. (2007). Position auctions. *International Journal of Industrial Organization*, 25(6), 1163–1178.

Wolinsky, A. (1988). Dynamic markets with competitive bidding. *The Review of Economic Studies*, 55(1), 71–84.

Wu, T. (2018). Blind spot: The attention economy and the law. *Antitrust LJ*, 82, 771.

Online Appendix

A. Proofs for Section 4

This appendix completes the proof of Theorem 1. I start with some useful lemmas.

Lemma 1. *In any equilibrium, platforms invest at the same rates if $\epsilon - 1 < 1/\varphi$. An equilibrium does not exist if $\epsilon - 1 > 1/\varphi$.*

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. Thus, all platforms must follow the same strategy in an equilibrium.

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*

$$\begin{aligned}\dot{\ell}_{kt} &= \frac{\rho+\delta}{1-\varphi}\ell_{kt} - \frac{\varphi}{1-\varphi}\frac{\pi_{\mathbb{K}t}A(\epsilon-1)}{Kq_{kt}}\ell_{kt}^{\varphi} \\ \dot{q}_{kt} &= \ell_{kt}^{\varphi} - \delta q_{kt}\end{aligned} \tag{26}$$

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 will be continuous in initial conditions. Observe that the ODE for ℓ_{kt} has an absorbing point at 0 and by Gronwall's inequality

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

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

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

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

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

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

Let $f_t = \ln(\ell_{kt})$. Then

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

with

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

Now I make the observation 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. By my initial observiation, it is easy to see 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 (26) 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 (26). If $q_{kt} \to q_{SS}$ then it must be that $\ell_{kt} \to \ell_{SS}$. To see why, fix $\alpha > 0$. There exists $\zeta > 0$ sufficiently small such that

$$\frac{\rho+\delta}{1-\varphi} - \frac{\varphi}{1-\varphi}\frac{(\pi_{\mathbb{K}}+\zeta)A(\epsilon-1)}{K(q_{SS}-\zeta)}(\ell_{SS}+\alpha)^{\varphi-1} > 0. \tag{27}$$

Note that if α and ζ were 0, the above would be 0 since (ℓ_{SS}, q_{SS}) are by definition steady state. Since $\pi_{\mathbb{K}t} \to \pi_{\mathbb{K}}$ and $q_{kt} \to q_{SS}$, for all t sufficiently large, we have

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

Then if ever $\ell_{kt} > \ell_{SS} + \alpha$ for some t sufficiently large, then $\dot{\ell}_{ks}/\ell_{ks}$ will be bounded below by (27) 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 $\ell_{kt} \to \ell_{SS}$.

Thus, going forward we suppose that $q_{kt} \nrightarrow q_{SS}$. Then there exists $\alpha > 0$ such that for any T, there exists $t > T$ such that $|q_{kt} - q_{SS}| > \alpha$.

Suppose that $q_{kt} \geq q_{SS} + \alpha$ for some t sufficiently large so that $|\pi_{\mathbb{K}t} - \pi_{\mathbb{K}}| < \zeta$. We may without loss assume that $\ell_{kt} \geq \ell_{SS}$ because after a long enough time, there must be t such that $\ell_{kt} \geq \ell_{SS}$ if $q_{kt} \geq q_{SS} + \alpha$ since otherwise q_{kt} could not have been reached. Therefore $\dot{\ell}_{ks} > 0$ for all $s \geq t$ and in fact if ζ was chosen sufficiently small,

$$\frac{\dot{\ell}_{ks}}{\ell_{ks}} > \frac{\rho + \delta}{1 - \varphi} - \frac{\varphi}{1 - \varphi} \frac{(\pi_{\mathbb{K}t} + \zeta) A(\epsilon - 1)}{K(q_{SS} + \alpha)} \ell_{SS}^{\varphi - 1} > 0.$$

Thus ℓ_{ks} must diverge as $s \to \infty$.

Analogous logic shows that if we are in the other case when $\ell_{kt} \leq \ell_{SS}$ and $q_{kt} \leq q_{SS} - \alpha$ for some t sufficiently large then $\ell_{kt} \to 0$. $\square$

Lemma 4. *There exists a unique solution to the boundary-value problem in Part 7 of Theorem 1.*

Proof. I first argue that there can be at most one solution to the boundary value problem. 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.

I now prove existence of a solution to the boundary value problem. To do this, consider a version of the problem 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 (26) 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 0 as we would like. Thus, by continuity (established in Lemma 2), there must be some initial condition for which $\ell_{kT} = \ell_{SS}$ which I denote by $\ell_{k0}(T)$. 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. Consider ℓ_{k0}^* defined as the infimum of the set of initial conditions of the diverging solutions. 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. Suppose for contradiction that ℓ^*_{kt} diverges and let T^* denote the last time that $\ell^*_{kt} = \ell_{SS}$. Then ℓ^*_{kt} is the solution to the boundary value problem with boundary at T^*. Next consider $T_n > T^*$ where n is 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}.

Now suppose that ℓ^*_{kt} eventually converges to 0. By inspecting (26) 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 the ℓ_{kt} must converge to 0. 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 solutions must diverge by the definition of ℓ^*_{k0}. This contradicts the continuity of the solutions of the initial value problem in the initial condition established in Lemma 2.

The case when there are an infinite number of solutions that converge to 0 is analogous. Since there must be either an infinite number of solutions that diverge or converge to 0, the proof is complete.

□

Proof of Theorem 1. I start by taking equilibrium existence as given in which case Parts 1-3 of the theorem follow from standard arguments.

To complete the proof of Part 4, one can directly verify that (10) solves the ODE (1) whenever $A/\lambda_f < J$. To derive (11), I conjecture and later verify (in the proof of Part 6 below) that bidding strategies are necessarily monotone so that (2) applies:

$$(M_t \dot{H}_t) = A(H^c_t)^N - \lambda_f M_t H_t.$$

Using the accounting identity $M_t H_t + (J - M_t)H^c_t = JG$, we have

$$\begin{aligned}(J - M_t)\dot{H}^c_t &= \lambda_f [JG - (J - M_t)H^c_t] - A(H^c_t)^N + (A - \lambda_f M_t)H^c_t \\ &= \lambda_f [JG - M(H^c_t)^N - (J - M)H^c_t].\end{aligned}$$

The steady state H^c is such that the right-hand side above is zero. Since M on the right-hand side is the steady state level and is just a constant, the ODE is separable and we can solve it analytically for H^c_t. I find that at each point $\hat{v} \in [0, \infty)$ where $H^c_0(\hat{v})$ is not already at its steady state level $H^c(\hat{v})$, $H_t^c(\hat{v})$ satisfies (11). By inspecting (11), we see that H_t^c and therefore H_t must eventually converge to their steady state levels H^c and H.

Part 5 follows straightforwardly from Parts 1 and 2.

To prove Part 6, let $V_t^{\text{In}}(\hat{v})$ denote firm j's continuation value from selling to consumer i at all points in the future if it is in Ω_{it} at time t when $\hat{v}_{ij} = \hat{v}$. Let $V_t^{\text{Out}}(\hat{v})$ be defined analogously. V_t^{In} and V_t^{Out} must satisfy the Hamilton-Jacobi Bellman 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

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

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 distributed according 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

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

The solution to this ODE is

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

One can show that V_t^{In} and V_t^{Out} must be Lipschitz and thus, since B_t' can never diverge in an equilibrium, the term in parenthesis must vanish as $t \to \infty$. 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.

To prove Part 7, given the average ad prices and demand from (6), 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

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

By Lemma 1, all equilibria are neccessarily symmetric when $\epsilon < 2$. Thus, 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

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

where recall that

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

starting from $q_{k0} = q_0$.

By Lemmas 2 and 3 the unique solution of this ODE system that is consistent with an equilibrium must solve (16). Note that 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 optimizes by investing according to $\{\ell_{kt}\}$ that solves (16) provided its rivals do the same. By Lemma 4, a unique solution to (16) exists and thus there exists a unique equilibrium that converges to steady state.

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 (16), 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 seen from the proof of Lemma 1. I omit the details of these steps for brevity. □

B. Proofs for Section 5 and Additional Comparative Statics

Since I prove many comparative statics for steady state only, it is useful to first write down the steady state equilibrium properties in Lemma 5.

Lemma 5. *Suppose that $\epsilon - 1 < 1/\varphi$ and that A is the unique solution of $\max_a a\nu(a)^{\epsilon-1}$. In a steady state equilibrium, the following hold:*

1. *The measure of varieties M in a consideration set is A/λ_f.*
2. *The cdf H^c of the expected values of a consumer for the firms outside of her consideration set solves*
$$M(H^c)^N + (J - M)H^c = JG. \tag{28}$$
3. *The cdf H of the expected values of a consumer for the firms in her consideration set satisfies $H = (H^c)^N$.*
4. *Each firm sets price $p = \sigma/(\sigma - 1)$.*
5. *Firm j's flow profit from selling to consumer i is $\pi_{\mathbb{J}}\hat{v}_{ij}$ where*
$$\pi_{\mathbb{J}} = \frac{I}{\sigma M \mu_H}.$$
6. *Firm j bids*
$$B(\hat{v}_{ij}) = \pi_{\mathbb{J}} \int_0^{\hat{v}_{ij}} \frac{1}{\rho + \lambda_f + \lambda_e(s)}\,\mathrm{d}s \tag{29}$$
in an auction for consumer i where $\lambda_e = \lambda_a(H^c)^{N-1}$.
7. *The Poisson rate that firm $j \in \Omega^c_{it}$ enters Ω_{it} is $\lambda_e(\hat{v}_{ij})$.*

8. *The average ad price is*

$$\pi_{\mathbb{K}} = \pi_{\mathbb{J}} \int_0^\infty \frac{1 - NH^c(s)^{N-1} + (N-1)H^c(s)^N}{\rho + \lambda_f + \lambda_e(s)} \, \mathrm{d}s. \tag{30}$$

9. *Each platform invests at rate*

$$\ell_{\mathbb{K}} = \frac{\varphi\delta\pi_{\mathbb{K}}A(\epsilon - 1)}{K(\rho + \delta)}. \tag{31}$$

Proof. Lemma 5 follows almost immediately from the derivation in Section 4. The condition $\epsilon - 1 < 1/\varphi$ is almost a necessary condition as seen in Lemma 1. I only use it here because it guarantees that each platform employs the same investment strategy. □

I will now prove a several comparative statics, including those in Table 1.

Lemma 6. *An increase in G in the mean-preserving spread order leads to an increase in the steady state cumulative value $M\mu_H$ 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 2 of Lemma 5, 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

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

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

$$\int_0^s \nu(y)\,\mathrm{d}y\left[\frac{J - M}{J} + \frac{M}{J}\hat{H}^c(s)^{N-1}\right] - \frac{M}{J}\int_0^s \int_0^y \nu(l)\,\mathrm{d}y\,\mathrm{d}\hat{H}^c(y)^{N-1} \geq 0. \tag{32}$$

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

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

If $l^* > 0$, then (32) is violated at l^* which is a contradiction. Then it must be that $l^* = 0$. But by inspecting (32), we see that $\int_0^s \nu(y)\,\mathrm{d}y$ must be increasing in s when it fist departs from 0 as otherwise (32) 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.

□

In the baseline model I assumed that G is nonatomic. This was simply to avoid ties in the auctions. I will show that steady-state ad revenue is maximal when data is uninformative in a limiting sense formalized below in Lemma 7. To do this, let us make explicit the dependency of the equilibirum ad price in steady state on the cdf G by denoting it by $\pi_{\mathbb{K}}(G)$.

Lemma 7. *Let $\{G_n\}$ be a sequence of continuous cdfs converging pointwise to the Heaviside function centered at μ_F. That is $\lim_{n\to\infty} G_n(\hat{v}) = \mathbb{1}_{\hat{v}\geq\mu_F}$ 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 (as seen from the equation (28)).

Using (30) in Lemma 5 we have

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

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

Though the ad rate A is endogenous, it is useful to derive comparative statics in which we treat it as exogenous and vary it directly, holding all other parameters in the model fixed as we do below in Lemmas 8 and 9.

Lemma 8. *An increase in A leads to a decrease in H_t and H_t^c in first-order stochastic dominance and an increase in $M_t\mu_{H_t}$ and C_{it} at all t.*

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

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

When A goes up, both M and M_t increase as seen from (10) and so $\dot{H}_t^c$ is higher holding fixed the value of H_t^c at time t. Using standard comparison arguments for differential equations, it is easy to see that this implies 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

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

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

Lemma 9. *An increase in A leads to an decrease in steady state ad revenue $\pi_{\mathbb{K}}A$.*

Proof. We have from Lemma 5

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

where $\lambda_e(s) = \lambda_a H^c(s)^{N-1}$ for each $s \in [0, \infty)$. By Lemma 8, an increase in A leads to a decrease in H^c in first-order stochastic dominance. Since λ_e increases pointwise, to show that $\pi_{\mathbb{K}}A$ decreases as A increases, it suffices to observe 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 shown simply by computing the derivative. I omit this step. □

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

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

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

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

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

Lemma 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 $\pi_{\mathbb{K}} A$. □

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 (9) 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 (31) and thus platform quality. Since $X = K^{\frac{1}{\epsilon-1}} \nu(A) q_t$ and q_t increased while A decreased, if $K \leq 1$ then X must increase. □

Lemma 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 (14), the equilibrium ad revenue $\pi_{\mathbb{K}t} A$ does not depend on K. In steady state, using (31)

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

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. □

C. Proofs for Section 7

The following Theorem 3 characterizes the general equilibrium of the model introduced at the start of Section 7.

Theorem 3. *Theorem 1 applies except with with I_t replacing I in the equations listed in the theorem. Moreover, in equilibrium, $I_t = \frac{\sigma}{\sigma-1}(L - K\ell_{\mathbb{K}t})$ at each $t \in [0, \infty)$. Therefore, under the sufficient conditions for equilibrium existence and uniqueness given in Theorem 1, in equilibrium, investment and quality solve the ODE system*

$$\begin{aligned} \dot{\ell}_{kt} &= \frac{\rho+\delta}{1-\varphi}\ell_{kt} - \frac{\varphi}{1-\varphi}\frac{\frac{\sigma}{\sigma-1}(L-K\ell_{\mathbb{K}t})\hat{\pi}_{\mathbb{K}t}A(\epsilon-1)}{Kq_{kt}}\ell_{kt}^{\varphi} \\ \dot{q}_{kt} &= \ell_{kt}^{\varphi} - \delta q_{kt} \end{aligned} \tag{33}$$

with boundary conditions

$$\begin{aligned} \lim_{t\to\infty}\ell_{kt} &= \frac{\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathbb{K}}A(\epsilon-1)}{r+\delta+\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathbb{K}}A(\epsilon-1)}\frac{L}{K} \\ q_{k0} &= q_0. \end{aligned}$$

Above,

$$\hat{\pi}_{\mathbb{K}t} = \int_0^{\infty}\int_t^{\infty}\frac{1}{\sigma M_t\mu_{H_t}}e^{-\int_t^s[\rho+\lambda_f+\lambda_{ez}(\hat{v})]\,\mathrm{d}z}\,\mathrm{d}s\;[1-O_t(\hat{v})]\,\mathrm{d}\hat{v},$$

is the average ad price per unit of income at time t where $O_t = (H_t^c)^N + N(H_t^c)^{N-1}(1-H_t^c)$ and $\pi_{\mathbb{K}} = \lim_{t\to\infty}\pi_{\mathbb{K}t}$ is the steady-state average ad price,

$$\hat{\pi}_{\mathbb{K}} = \frac{1}{\sigma M\mu_H}\int_0^{\infty}\frac{1-NH^c(s)^{N-1}+(N-1)H^c(s)^N}{\rho+\lambda_f+\lambda_e(s)}\,\mathrm{d}s. \tag{34}$$

Proof. That $I_t = \frac{\sigma}{\sigma-1}(L-K\ell_{\mathbb{K}t})$ follows from product market clearing. All other parts of Theorem 3 are proven using analogous methods to those used to prove Theorem 1. □

Proof of Proposition 1. The Proposition follows immediately from Theorem 3.

□

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

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

$$+ \lambda_t \left(\ell_{\mathbb{K}t}^{\varphi} - \delta q_t\right)$$

where λ_t satisfies

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

Maximizing the Hamiltonian with respect to the control yields

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

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

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

Substituting ths into the previous equation, we have

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

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

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

Rearranging gives

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

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

In the limit as $\rho \to 0$, the steady state ad rate chosen by the planner must maximize the flow utility of consumers which amounts to (24). □

Proofs of Propositions 2 and 3. These propositions follow straghtforwardly from Proposition 1 and Theorem 2.

□

D. 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 2. 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.

E. 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 7. 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

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

Canceling out some terms and rearranging yields

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

which is linear in the price p.

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

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

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

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

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

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

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

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

$$\frac{\mathrm{d}[p(x_1)x_1]}{\mathrm{d}x_1} + \pi_{\mathbb{K}} A = p'(x_1)x_1 + p(x_1) + \pi_{\mathbb{K}} A > 0 \tag{35}$$

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 (35) 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

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

whenever $x \le \epsilon$. But

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

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

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

and the upper bound holds uniformly over all $\tau \le \frac{1}{2}$. 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 $\tau \le 1/2$.

F. Extension: Network Effects

I extend the baseline model to allow for network effects.

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 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 (5) and her demands for platforms are as in (38).*
2. *Firm j sets prices as in (8).*
3. *Platform k displays ads at rate $\hat{A}$.*
4. *The size of consideration sets is given by (10) and the cdfs of the expected values of firms inside and outside of them are characterized by (11) and (3) with $\hat{A}$ in place of A.*
5. *Firm j's expected flow profits from sales are as in (12) and the rates at which firm j matches with consumers are as in (13).*
6. *Firm j bids according to (14).*
7. *Platform k's quality and investment solve the boundary-value problem (16) 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 8, the attention received by platform k solves

$$x_{kt}^{\zeta} = \frac{x_{kt}^{\zeta(\epsilon-1)}[\nu(a_{kt})q_{kt}]^{\epsilon-1}}{Y} \tag{36}$$

where

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

Solving (36) for x_{kt} yields two possibilities:

$$x_{kt} = \frac{[\nu(a_{kt})q_{kt}]^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}}}{Y^{\frac{1}{1-\zeta(\epsilon-1)}}} \tag{37}$$

or $x_{kt} = 0$. Under the equilibrium refinement, all platforms must receive positive attention share (37). Integrating both sides of (37) 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 (37) gives

$$x_{kt} = \frac{[\nu(a_{kt})q_{kt}]^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}}}{\int_{\mathbb{K}} [\nu(a_{kt})q_{kt}]^{\frac{\epsilon-1}{1-\zeta(\epsilon-1)}} \, \mathrm{d}k}. \tag{38}$$

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. □

G. Extension: Data Heterogeneity

I extend the baseline model to allow for two groups of platforms, each of positive measure, who may have different data.

Setup

There are two groups of platforms $z \in \{1, 2\}$. Let m_z denote the measure of platforms in group z. As before, let v_{ij} denote consumer i's value for firm j's product. Firm j receives signal ζ_{zij} when bidding on a platform in group z. I assume that $(v_{ij}, \zeta_{1ij}, \zeta_{2ij}) \sim Q$ defined on $[0, \infty) \times \mathbb{R}^2$ independently across i and j and that v_{ij} has a finite mean. Let G denote the joint cdf of ζ_{1ij} and ζ_{2ij} derived from Q. I assume that G has a continuous density g. I retain the other aspects of the baseline model.

Solving for the Steady State Equilibrium

I now sketch the procedure to solve for the steady state equilibrium of the model. The main properties of the equilibrium are summarized below in Theorem 5.

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 (5), (6), (9), and (10). However, we now must keep track of the joint distribution of the two signals inside and outside of consideration sets.

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. Let us assume for now 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

$$\begin{aligned}
&\left[M_t \dot{h_t}(\boldsymbol{\zeta})\right] \\
&\qquad = A\left[x_{1t} N H_t^c(\zeta_1, \infty)^{N-1} h_t^c(\boldsymbol{\zeta}) + x_{2t} N H_t^c(\infty, \zeta_2)^{N-1} h_t^c(\boldsymbol{\zeta}) - h_t(\boldsymbol{\zeta})\right]. \quad (39)
\end{aligned}$$

In (39), 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 (39) forward to convergence at each point $\boldsymbol{\zeta}$ in a fine grid on a region that contains almost all of G's mass.

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

More precisely, V satisfies the recursive equation

$$\begin{aligned} V(\boldsymbol{\zeta}_{ij}) = \sum_{z=1}^{2} x_z O_z(\zeta_{zij}) & \left[\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} \mathbb{E}[v_{ij} | \boldsymbol{\zeta}_{ij}] + \frac{\lambda_f}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \right] \\ & - x_l \left[\int_{-\infty}^{\zeta_{zij}} B_z(s) \, \mathrm{d}O_z(s) + \left[1 - O_z(\zeta_{zij})\right] \frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \right] \end{aligned} \tag{40}$$

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,

$$\begin{aligned} B_z(\zeta_{lij}) &= \mathbb{E} \left[\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} v_{ij} + \frac{\lambda_f}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \middle| \zeta_{zij}, j \in \Omega_{it}^c \right] \\ &\quad - \mathbb{E} \left[\frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \middle| \zeta_{zij}, j \in \Omega_{it}^c \right] \\ &= \mathbb{E} \left[\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} v_{ij} - \frac{\rho}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \middle| \zeta_{zij}, j \in \Omega_{it}^c \right]. \end{aligned} \tag{41}$$

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 (41) and (40), 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[34]

$$\Lambda_z(\boldsymbol{f})(\cdot) = \mathbb{E} \left[\frac{\pi_{\mathbb{J}} v_{ij} + \lambda_a \sum_{l=1}^{2} x_z \int_{-\infty}^{\zeta_{lij}} f_l(s) \, \mathrm{d}O_l(s)}{\lambda_f + \rho + \lambda_a \sum_{l=1}^{2} x_l O_l(\zeta_{zij})} \middle| \zeta_{zij} = \cdot, j \in \Omega_{it}^c \right]. \tag{42}$$

It is easy to show that $\boldsymbol{\Lambda}$ 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

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

purposes, this will always be the case. Even without this bounded support assumption, we see that $\mathbf{\Lambda}$ is increasing. Thus, starting from an initial $\boldsymbol{B}$ such that $B_z > \Lambda_z(\boldsymbol{B})$ for each $z \in \{1, 2\}$ it follows that $\{\mathbf{\Lambda}^n(\boldsymbol{B})\}_{n=1}^{\infty}$ is a decreasing sequence which converges to the fixed point. Thus we can compute the equilibrium bidding strategies by iterating on (42).

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

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

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

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,

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

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.

$$\ell_{\mathbb{K}z}=\frac{\varphi\delta\pi_{\mathbb{K}z}A\left(\epsilon-1\right)}{\rho+\delta}\frac{1}{m_z+m_{-z}\left(\frac{\pi_{\mathbb{K}-z}}{\pi_{\mathbb{K}z}}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}}}\tag{43}$$

From (43) we also have the quality level

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

and attention share

$$x_z=\frac{m_z q_z^{\epsilon-1}}{m_z q_z^{\epsilon-1}+m_{-z}q_{-z}^{\epsilon-1}}=\frac{m_z\left(\frac{\pi_{\mathbb{K}z}}{\pi_{\mathbb{K}-z}}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}}}{m_z\left(\frac{\pi_{\mathbb{K}z}}{\pi_{\mathbb{K}-z}}\right)^{\frac{\varphi(\epsilon-1)}{1-\varphi(\epsilon-1)}}+m_{-z}}\tag{44}$$

of group z platforms.

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.

Summary of Computational Procedure

1. Guess a value of x_1.
2. Iterate (39) forward to compute h.
3. Iterate (42) to compute per-unit income bid functions and average ad prices.
4. Check whether the guess of x_1 aligns with (44).
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.

Steady State Equilibrium Characterization

The following Theorem 5 summarizes the steady state equilibrium properties.

Theorem 5. *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 (5) and her demands for platforms are as in (6).*
2. *Firm j sets prices as in (8).*
3. *Platform k displays ads at rate A as in (9).*
4. *The size of consideration sets is $M = A/\lambda_f$.*

5. *Firm j's expected flow profits from sales are as in* (12).

6. *Bidding functions $\boldsymbol{B} = (B_1, B_2)$ for the two groups are the fixed point of the operator $\boldsymbol{\Lambda}$ defined by* (42) *which is a contraction map whenever values are bounded.*

7. *The attention shares received by the two groups are given by* (44).

8. *The rates of investment by the two groups are given by* (43).

Above, the condition $\epsilon - 1 < 1/\varphi$ is almost a necessary condition for existence following an anologous argument to Lemma 1. I only use it here because it guarantees that each platform in a group employs the same investment strategy, a fact we took for granted in our sketch of the steady state equilibrium derivation.

In order for a steady state equilibrium with increasing bidding strategies to exist, the fixed point of $\boldsymbol{\Lambda}$ must be increasing. The following condition, is sufficient for this to be so.

Condition 1. *The following hold:*[35]

1. *The conditional distribution of ζ_{-zij} given ζ_{zij} under the steady state solution for H^c to* (39) *increases in first-order stochastic dominance when ζ_{zij} increases for each $z \in \{1, 2\}$.*

2. *The conditional expectation $\mathbb{E}\left[v_{ij} | \zeta_{1ij}, \zeta_{2ij}\right]$ is nondecreasing in both ζ_{1ij} and ζ_{2ij} and increasing in at least one of ζ_{1ij} or ζ_{2ij}.*

This is a stochastic monotonicity condition which essentially states that having a higher signal on either platform is good news to a firm about the consumer's value for its product.

To prove that Condition 1 is sufficient for monotone bidding strategies I will use the following generalization of Arzelà-Ascoli, which follows from Theorems 3.4.20 and 8.2.10 of Engelking (1977).

Theorem 6. *$\mathcal{A} \subset C(\mathbb{R}_+)$ is relatively compact in the topology of uniform convergence on compact subsets of $\mathbb{R}_+$ if and only if $\mathcal{A}$ is equicontinuous at each $x \in \mathbb{R}_+$ and $\{f(x) | f \in \mathcal{A}\} \subset \mathbb{R}$ is bounded for each $x \in \mathbb{R}_+$.*

Lemma 14. *Under Condition 1 the fixed point of $\boldsymbol{\Lambda}$ defined in* (42) *is a pair of increasing functions.*

[35] With abuse of notation, we use the same symbols for both the signals and their realizations.

Proof. Let B_1, B_2 be arbitrary bidding functions. Let τ_k be the time that firm j is invited to its kth auction for consumer i. Let l_k be the platform group that hosts the kth auction. If all firms other than j bid according to B_1, B_2, then firm j's solves

$$V(\boldsymbol{\zeta}_{ij}) = \max_{(b_k)_{k=1}^{\infty}} \mathbb{E}\left[\int_0^{\infty} \pi_{\mathbb{J}} v_{ij} \mathbb{1}_{j \in \Omega_{it}}\, ds - \sum_{k=1}^{\infty} e^{-\rho \tau_k} B_{l_k}^{(1)} \mathbb{1}_{b_k > B_{l_k}^{(1)}} \middle| \boldsymbol{\zeta}_{ij}\right] \tag{45}$$

such that b_k is $\sigma(\zeta_{l_k ij}, \tau_k, l_k)$-measurable. Above, $B_{l_k}^{(1)}$ denotes the highest bid of the $N-1$ other bidders in the kth auction. I argue that if it is also optimal for firm j to bid according to B_1, B_2, then B_1, B_2 must be increasing and therefore must satisfy (41), (40) and thus must be the fixed point of the contraction (42). This guarantees that the fixed point of (42) is necessarily increasing provided there exists such a pair of bidding functions B_1, B_2.

I will prove that

$$\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} \mathbb{E}\left[v_{ij} | \boldsymbol{\zeta}_{ij}\right] - \frac{\rho}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\boldsymbol{\zeta}_{ij}) \tag{46}$$

is nondecreasing in both of its arguments and increasing in one of them. This will ensure that bidding strategies which satisfy (41) are increasing using Condition 1. Without loss of generality suppose that $\mathbb{E}[v_{ij} | \zeta_{1ij}, \zeta_{2ij}]$ is increasing in its first argument.

Suppose for contradiction that there exists $\overline{\zeta}_1$ and $\underline{\zeta}_2$ with $\overline{\zeta}_1 > \underline{\zeta}_1$ such that

$$\begin{aligned}\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} \mathbb{E}[v_{ij} | \zeta_{1ij} = \overline{\zeta}_1, \zeta_{2ij}] - \frac{\rho}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\overline{\zeta}_1, \zeta_{2ij}) \\ < \frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} \mathbb{E}[v_{ij} | \zeta_{1ij} = \underline{\zeta}_1, \zeta_{2ij}] - \frac{\rho}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\underline{\zeta}_1, \zeta_{2ij}).\end{aligned}$$

Rearranging yields

$$\begin{aligned}\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} \mathbb{E}[v_{ij} | \zeta_{1ij} = \overline{\zeta}_1, \zeta_{2ij}] - \frac{\pi_{\mathbb{J}}}{\lambda_f + \rho} \mathbb{E}[v_{ij} | \zeta_{1ij} = \underline{\zeta}_1, \zeta_{2ij}] \\ < \frac{\rho}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\overline{\zeta}_1, \zeta_{2ij}) - \frac{\rho}{\lambda_f + \rho} \frac{\lambda_a}{\lambda_a + \rho} V(\underline{\zeta}_1, \zeta_{2ij}).\end{aligned} \tag{47}$$

I will show that the above inequality can not hold.

As seen in (45), the value function can be decomposed into two parts: one which arises from flow profits from sales (the expectation of the first sum in (45)) and one which arises from costs of advertising (the expectation of the second sum (45)). Let us write the value function of a firm j with signals $\overline{\zeta}_1, \zeta_{-lij}$ to reflect this:

$$V(\overline{\zeta}_{lij}, \zeta_{-lij}) = \overline{\Pi}_{\text{sales}} - \overline{C}_{\text{ad cost}}.$$

Now suppose a firm j with signal $\underline{\zeta}_{lij}$ deviates and bids as though its signal was $\overline{\zeta}_{lij}$. Then its payoff is

$$\frac{\mathbb{E}\left[v_{ij}\middle|\zeta_{lij} = \underline{\zeta}_l, \zeta_{-lij}\right]}{\mathbb{E}\left[v_{ij}\middle|\zeta_{lij} = \overline{\zeta}_l, \zeta_{-lij}\right]}\overline{\Pi}_{\text{sales}} - \overline{C}_{\text{ad cost}} \leq V(\underline{\zeta}_{lij}, \zeta_{-lij}).$$

Substituting into (47), we therefore have,

$$\begin{aligned}\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho}\mathbb{E}[v_{ij}|\zeta_{1ij} = \overline{\zeta}_1, \zeta_{2ij}] - \frac{\pi_{\mathbb{J}}}{\lambda_f + \rho}\mathbb{E}[v_{ij}|\zeta_{1ij} = \underline{\zeta}_1, \zeta_{2ij}] \\ < \frac{\rho}{\lambda_f + \rho}\left(1 - \frac{\mathbb{E}\left[v_{ij}\middle|\zeta_{lij} = \underline{\zeta}_l, \zeta_{-lij}\right]}{\mathbb{E}\left[v_{ij}\middle|\zeta_{lij} = \overline{\zeta}_l, \zeta_{-lij}\right]}\right)\overline{\Pi}_{\text{sales}}.\end{aligned}$$

Since there are times when the firm is not in the consideration set,

$$\overline{\Pi}_{\text{sales}} \leq \frac{\pi_{\mathbb{J}}}{\rho}\mathbb{E}\left[v_{ij}\middle|\zeta_{lij} = \overline{\zeta}_l, \zeta_{-lij}\right].$$

Substituting into the RHS above, we obtain a contradiction. The proof that (46) is also nondecreasing in its second argument is analogous. Now I have completed the first step of the proof. The second step is to prove that there in fact exists a pair of bidding functions B_1, B_2 such that each firm optimally bids according to them if its rivals do (that is, firm j solves (45)).

Let $\text{Lip}(\mathbb{R}_+)$ denote the set of $\frac{\pi_{\mathbb{J}}}{\lambda_f+\rho}$-Lipschitz functions f such that $f(y) \leq \frac{\pi_{\mathbb{J}}}{\lambda_f+\rho}y$ at each $y \in \mathbb{R}_+$. By Theorem 6 and Tychonoff's Theorem, $\text{Lip}(\mathbb{R}_+)^2$ is compact in the product topology. To ensure bidding strategies B_1, B_2 live in $\text{Lip}(\mathbb{R}_+)^2$ we redefine signals so that, with abuse of notation

$$\zeta_{lij} = \mathbb{E}[v_{ij}|\zeta_{lij}, j \in \Omega_{it}^c].$$

Thus $\zeta_{lij} \in \mathbb{R}_+$. Condition 1 ensures that there is a one to one mapping between old signals and new signals. Moreover, now

$$B_l(\zeta_{lij}) = \frac{\pi_{\mathbb{J}}}{\lambda_f + \rho}\zeta_{lij} - \mathbb{E}\left[\frac{\rho}{\lambda_f + \rho}\frac{\lambda_a}{\lambda_a + \rho}V(\boldsymbol{\zeta}_{ij})\middle|\zeta_{lij}, j \in \Omega_{it}^c\right] \tag{48}$$

which is $\frac{\pi_{\mathbb{J}}}{\lambda_f+\rho}$-Lipschitz since V is nondecreasing in both signals.

Let us take as given some input bidding functions $\boldsymbol{B}_{input} \in \text{Lip}(\mathbb{R}_+)^2$. Let V be the value function bidding optimally given this: set V equal to the RHS of (45) when B_1, B_2 are given by $\boldsymbol{B}_{input}$. Next, define $\boldsymbol{B}_{output}$ using (48). This map from $\boldsymbol{B}_{input}$ to $\boldsymbol{B}_{output}$ is continuous. Moreover it maps from $\text{Lip}(\mathbb{R}_+)^2$ into itself. $\text{Lip}(\mathbb{R}_+)^2$ is closed, convex, and compact. Thus by Schauder's fixed point theorem there exists a fixed point. Any such fixed point is in increasing bidding stragies and therefore, must be the fixed point of (42).

□

H. Extension: Heterogeneity in Platform Productivity

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

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 α_k, $k \in \mathbb{K}$. I retain all other aspects of the baseline model.

Equilibrium Characterization

Theorem 7. *Suppose that A is the unique solution of $\max_a a\nu(a)^{\epsilon-1}$ and $\epsilon < 1/\varphi$. 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. 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

$$(\rho+\delta)\frac{\alpha_k}{\varphi}\ell_{kt}^{1-\varphi} - \alpha_k\frac{1-\varphi}{\varphi}\ell_{kt}^{-\varphi}\dot{\ell}_{kt} = \pi_{\mathbb{K}t}A(\epsilon-1)\frac{q_{kt}^{\epsilon-2}}{Q_t} \tag{49}$$

where

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

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

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

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 (49)

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

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, it is easy to use the same method as in the proof of Theorem 1 to verify the conjecture and complete the rest of the proof. □

I. Extension: Firm and Platform Entry

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

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 2 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.

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 8. *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 properties as in Lemma 5 for a given J and K which satisfy*

$$K = \frac{\pi_{\mathbb{K}} A}{e_{\mathbb{K}}} \left(1 - \frac{\varphi\delta(\epsilon - 1)}{\rho + (1-\alpha)\delta}\right), \tag{50}$$

and

$$J = \frac{\frac{I}{\sigma} - \pi_{\mathbb{K}} A}{e_{\mathbb{J}}}. \tag{51}$$

Proof. Equations 50 and (51) are simply the zero profit conditions. Note that in (51), $\pi_{\mathbb{K}}$ depends on J. Thus, to prove uniqueness I must prove there is a unique solution for J in (51). 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 Proposition 1. □

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

J. Extension: Reserve Prices

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

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 2.

Steady State Equilibrium Characterization

Theorem 9. *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 (5) and her demands for platforms are as in (6).*
2. *Firm j sets prices as in (8).*
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 (12).*
6. *Firm j sets reserve price $R = \frac{\pi_{\mathbb{J}}}{\lambda_f+\rho}Y$ where Y solves*

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

   where

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

   and

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

7. *Firm j bids according to*

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

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

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

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

for $s \geq Y$.

9. *Each platform k invests at rate* (31).

Proof. It is clear that consumers' demands and firms' flow profits and prices will be the same as in the baseline model of Section 2.

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

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

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

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

whereas before $M = A/\lambda_f$.

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

$$B(\hat{v}_{ij}) = \frac{\pi_{\mathbb{J}}}{\lambda_f + \rho}\hat{v}_{ij} + \frac{\lambda_f}{\lambda_f + \rho}\frac{\lambda_a}{\lambda_a + \rho}V(\hat{v}_{ij}) - \frac{\lambda_a}{\lambda_a + \rho}V(\hat{v}_{ij}) \tag{52}$$

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

$$\begin{aligned} V(\hat{v}_{ij}) = \left[1 - H(\hat{v}_{ij})^{N-1}\right] \frac{\lambda_a}{\lambda_a + \rho} V(\hat{v}_{ij}) + \\ H^c(\hat{v}_{ij})^{N-1}\Bigg(\frac{\pi_{\mathbb{J}}}{\lambda_f + \rho}\hat{v}_{ij} + \frac{\lambda_f}{\lambda_f + \rho}\frac{\lambda_a}{\lambda_a + \rho}V(\hat{v}_{ij}) \\ - \mathbb{E}\left[\max\{B(\hat{v}^{(1)}), R\} | \hat{v}_{ij} > \hat{v}^{(1)}\right]\Bigg) \end{aligned} \tag{53}$$

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 (53) 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 (52)

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

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,

$$\frac{NH^c(s)^{N-1}h^c(s)}{1 - H^c(Y)^N} = h(s). \tag{54}$$

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

$$H^c(Y) = \frac{K}{J-M}G(Y). \tag{55}$$

Substituting into the above equation, we find that

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

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

It will also be useful to derive $h^c(Y)$ which we can do by using the equation (54) which implues

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

Then using an accounting identity, we have

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

which rearranges to

$$h^c(Y) = \frac{\frac{K}{M}g(Y)\left[1 - H^c(Y)^N\right]}{NH^c(Y)^{N-1} + \frac{J-M}{M}\left[1 - H^c(Y)^N\right]}. \tag{56}$$

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:

$$\int_{\hat{Y}}^{\infty} B(s)[N(N-1)H^c(s)^{N-2}(1-H^c(s))h^c(s)]\,\mathrm{d}s$$
$$+\hat{Y}\frac{\pi_{\mathbb{J}}}{\lambda_f+\rho}N\left[1-H^c(\hat{Y})\right]H^c(\hat{Y})^{N-1}. \quad (57)$$

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

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

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 (55) and (56) respectively.

From here, given the average ad price $\pi_{\mathbb{K}}$ which coincides with (57) 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 paltforms follow the same investment strategy. As seen in Lemma 1 it is “almost" a necessary condition for equilibrium existence.

□