The Market for Attention

Daniel Chen*

link to latest version

2 Feb 2022; Revised 21 Nov 2022

Abstract

This paper develops a dynamic general equilibrium model of the market for attention. Digital platforms compete for the attention of consumers by investing in the quality of their services, which they provide for free. Platforms then sell the attention, in the form of advertisements, to firms in the product market via auctions that use consumer data for targeting. We characterize outcomes in the product market, ad revenue, and platform investment in the unique stationary equilibrium. When data is more informative for all platforms, typically product consumption improves but ad revenues and investment decline. When platforms are more interoperable, investment rises but product consumption worsens. Compared with the first best, investment can be either too high or too low. The model predicts variation in ad prices, bid pacing, and delay in the matching of a firm to a consumer and relates these to platform market power. It also predicts that platforms that are data-rich relative to their rivals typically have higher market shares, ad prices, and investment.

Keywords: digital platforms, advertising, data, auctions, regulation

*dtchen@stanford.edu; Graduate School of Business, Stanford University
I thank my advisors Yuliy Sannikov and Andy Skrzypacz and committee members Darrell Duffie and Bob Wilson for their guidance and support. I thank Mohammad Akbarpour, Susan Athey, Kyle Bagwell, Anirudha Balasubramanian, Martino Banchio, Christoph Carnehl, Cindy Chung, Sebastian Di Tella, Ravi Jagadeesan, Ramesh Johari, Chad Jones, Nadia Kotova, David Kreps, Ruilin Li, Erik Madsen, Suraj Malladi, Matthew Mitchell, Hala Moussawi, Mike Ostrovsky, David Ritzwoller, Chris Tonetti, Joseph Stiglitz, Takuo Sugaya, Timur Sobolev, David Yang, Frank Yang, Ali Yurukoglu, Anthony Zhang, Charles Zhang, and Weijie Zhong for insightful comments and suggestions. This research is supported by the Bradley Graduate Fellowship through a grant to the Stanford Institute for Economic Policy Research.

1. Introduction

This paper is motivated by the ongoing debate over how to regulate digital platforms that profit primarily from targeted advertising. Many argue that, while different platforms offer distinct services, they nonetheless compete in a *market for attention* (Evans, 2020; Wu, 2018; Newman, 2019).

Some key stylized features of this market are as follows.

1. Various platforms provide distinct services to consumers for free.
2. They compete for attention by investing in the quality of their services.
3. They earn revenue by selling ads to firms in the product market.
4. They sell ads via auctions that are consumer-specific and arise in real time as consumers engage with their services.
5. Ads are targeted using individual-level consumer data.

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 a fully endogenous product market. We can therefore calculate the welfare impact of platforms from their roles in *both* features 1 and 3 within a single internally consistent model.

The model is designed to address three points that often come up in regulatory discussion. The first point is that existing competition policy relies heavily on markups to gauge market efficiency (Khan, 2017). For a long time, platforms, with their free services, occupied a blind spot in policy (Wu, 2018). Regulators are in need of guiding principles that apply to platforms. The second point is that the market is complex and multisided: outcomes in the product market, ad revenue, and platform investment are jointly determined (Evans, 2019). It is therefore difficult to assess the net effects of policies without a formal model. The third point is that two of the leading policy tools concern regulating the data used for ad targeting and enforcing platform interoperability (OECD, 2021). Interoperability refers to the extent to which users can share information or content across platforms. When there is greater interoperability, platforms are more substitutable and it is more convenient for users to shift their attention among them.

The model in this paper allows for a formal investigation into these issues. We characterize outcomes in the product market, ad revenue, and platform investment in the unique stationary equilibrium. We find that when data is more informative for all platforms, consumers discover products that they value more highly and so product consumption improves. However, ad revenue typically *declines* and thus, so does investment by platforms in the quality of their services. On the other hand, a rise in interoperability leads platforms to invest more. However, it also comes at the cost of worse product consumption. Thus, the net welfare impacts of policies that regulate data or interoperability are generally ambiguous. Whether these policies are beneficial or harmful may crucially depend on whether investment is inefficiently high or low to begin with. We show that either can be true in equilibrium and derive a simple condition that can be used to discern which is the case.

A tractable general equilibrium analysis requires some degree of modeling compromise. We model platforms that are *monopolistically competitive*. Though they have market power, they are strategically small. The reality is somewhere between duopoly and monopolistic competition. As of 2022, Meta and Google together have roughly a 50 percent share of the digital advertising market.[1] However, the remaining 50 percent is split across a wide array of other platforms such as Twitter, TikTok, LinkedIn, Pinterest, and others.[2] Whether or not platforms are strategically large, the basic economic mechanisms that we uncover will be there.

We consider an economy with a single productive resource: labor. Consumers supply a fixed quantity of labor at each point in time. Firms use it to produce and platforms use it to invest in their services. Each firm supplies a product of a distinct variety. As in Dixit and Stiglitz (1977), a consumer derives utility from individual products through a constant elasticity of substitution (CES) aggregate. Each consumer has private tastes or values for the various products. These values appear as share parameters in the CES aggregate. At any given time, a consumer is aware of only a subset of products. This "consideration set" evolves as the consumer discovers firms through ads displayed by the platforms and gradually forgets about the firms that are currently in the set.

At each time, each consumer splits a unit of attention across the various platforms. As with products, platform use enters utility through a CES aggregate. A platform's quality level is also its share parameter. To retain its share of attention, in equilibrium, a platform must invest in its quality level which otherwise depreciates over time (as, for instance, the platform's content grows stale). The attention is valuable because it determines the rate that the platform will display ads. Each time there is an opportunity for a platform to display an ad to a consumer, it invites a finite number of randomly chosen firms to bid in an auction. The platform then supplies each invited firm with data (a Blackwell experiment) that is informative of the consumer's value for the firm's product. Firms then set their bids. The winning firm displays its ad and enters the consumer's consideration set.

To close the model, we assume that consumers own the firms and platforms. An unambiguous measure of social welfare is therefore consumer surplus. We analyze stationary equilibria in which the distribution of consumers' values for the firms in their consideration sets remains constant over time, as do platforms' investment rates. For the baseline model, we assume that all platforms have the same data, though in Section 7, we allow data to differ across them.

We show that when data is uninformative, the aggregate product consumption of a consumer is *minimal*. However, ad revenue and investment by platforms are *maximal*. Thus, an increase in data informativeness, in the Blackwell order, typically leads to higher product consumption but causes ad revenue and investment to decrease. If

[1] See https://www.insiderintelligence.com/content/meta-google-s-hold-on-digital-advertising-loosens-tiktok-others-gain-share.

[2] Platforms in the model can also be interpreted as individual content creators which are small though they operate on large platforms. All of the analysis of the baseline model will apply.

platforms could collude to not use data, they would be better off. The intuition is as follows. In a standard auction setting, where bidders' values are taken as given, more informative data raises the expected value of the highest bidder. Though information rents rise, this effect typically dominates, and auction revenue increases. In our setting, there are two additional effects. First, bidders' values for displaying an ad are endogenous and depend on competition in the product market. When data is more informative, a firm anticipates that any given consumer is already aware of firms that it values highly. Displaying an ad is thus less valuable and so the firm bids less. Second, there is an option for firms to wait to bid in future auctions for a given consumer. When data is more informative, the option value is typically higher because ad prices are more variable. Firms exploit this variation by strategically reducing their bids to win only at favorable prices. In practice, this is known as bid pacing.[3] These latter two effects typically dominate and so ad revenue decreases which in turn leads platforms to invest less.

In contrast, we find that an increase in interoperability leads to a rise in investment but a decline in product consumption. In our analysis, we interpret an increase in interoperability as a technological change that results in platforms that are more substitutable. When platforms are more substitutable, consumer attention is more sensitive to platforms' quality levels. As a result, platforms invest more. Because less labor is allocated to the product sector, product consumption decreases. These effects are magnified when platforms also set the rate, per unit of attention, that they display ads to consumers.[4] Then, an increase in interoperability also causes attention to be more sensitive to platforms' ad frequencies. Since consumers dislike viewing ads, platforms respond by setting lower ad frequencies. Consumers are thus exposed to fewer products and product consumption worsens. The reduction in ad frequencies also leads ad revenues to rise, further raising investment.

Overall, these results suggest that regulations on data or interoperability must trade off between product consumption and investment in platform quality. Both of these factors determine welfare. Thus, whether a policy is, in the end, beneficial may crucially depend on whether investment is initially higher or lower than is efficient.

To address this, we compare the equilibrium to a first best benchmark in which a social planner sets the level of investment by platforms to maximize welfare. We assume that the planner relies on the same matching technology as do platforms (consideration sets evolve as in equilibrium). We find that, compared with the planner's choice, equilibrium investment may be either too high or too low. Investment may be too low because platforms do not appropriate any of the surplus that they generate for consumers through investment since their services are free. Investment may be too high because of business-stealing externalities: platforms invest only to steal ad revenues from each other. Moreover, the total ad revenue that is up for grabs is itself determined by the incentives of firms to steal profits from each other by advertising.

For the special case when consumers have Cobb-Douglas utility over aggregate product and platform consumption, we derive a simple condition that can be used

[3]See https://www.facebook.com/business/help/1754368491258883?id=561906377587030.

[4]All results described in the Introduction hold whether or not ad frequency is endogenous.

to distinguish whether investment is too high or too low. The condition depends on only product markups, ad revenue, platform substitutability, and the weight in the utility function on platform consumption relative to product consumption. Using this condition, together with the comparative statics described earlier, we prove results that show when the net welfare impacts of changes to data informativeness or interoperability can be signed analytically. For example, if investment is too high, then banning data use will reduce welfare since it both leads to worse product consumption and raises investment even farther from its efficient level. These results conclude the anaysis of the baseline model.

In the last part of the paper, we extend the model to allow platforms to have different data. This extension is especially relevant since policies may affect platforms that are data-rich differently than those that are less data-rich. The extension also allows us to explore the effects of regulations, such as forcing platforms to share data, that are only meaningful when data are heterogeneous. We find that platforms that have more informative data typically have higher ad prices, market shares, and investment. Mandating that platforms share data often raises the ad revenues and market shares of data-poor platforms at the expense of data-rich ones. While product consumption improves, total investment declines to the extent that overall platform consumption worsens. This is so even though consumers prefer a more even distribution of platform quality and despite the fact that there are decreasing returns to investment.

In the appendices, we consider several additional extensions of the baseline model that may be of first-order relevance: we endogenize the ad frequency set by platforms and allow for network effects, reserve prices, and firm and platform entry. Many of the main results continue to hold in each of these extensions, as briefly described in Section 8.

The rest of this paper proceeds as follows. Section 2 summarizes the related literature. Section 3 presents the baseline model. Section 4 characterizes the unique stationary equilibrium. Section 5 analyzes the effects of changes to data informativeness and platform interoperability. Section 6 compares the equilibrium to a first best benchmark. Section 7 extends the model to allow platforms to have different data. It then presents a numerical example to illustrate the effects of a policy that forces platforms to share data. Section 8 summarizes the results of further extensions. Section 9 concludes.

2. Related Literature

This paper is related to the literature on platforms and two-sided markets. To our knowledge, it supplies the first model that is consistent with the stylized features described at the start of this paper. It is also the first to feature monopolistic competition in a two-sided market.

The early seminal papers Rochet and Tirole (2003), Caillaud and Jullien (2003), Armstrong (2006) study platforms' pricing decisions. For a survey of subsequent work see Jullien, Pavan, and Rysman (2021). These models are abstract so as to apply to a wide variety of settings. As a result, they do not endogenize the surplus generated by the interactions between individuals on the two sides of the market. An important result of these papers is that, due to cross-side network effects, platforms may charge consumers on one side of the market a low price to encourage their participation. This is so that they can then profitably charge consumers on the other side of the market a high price. In contrast, the model in this paper is specialized to fit the market for attention. The surplus generated by the interactions between firms and consumers is determined endogenously in a product market. We assume, from the outset, that platforms give their services to consumers for free and that the prices that firms pay are set in auctions. We focus instead on platforms' decisions to invest in the quality of their services.

This paper is more closely related to the seminal paper Anderson and Coate (2005) which is among the first to study traditional advertising platforms with an endogenous product market. Many of the papers that do this use variants of their model. In their model, there may be up to two platforms in the market. To enter, a platform must pay a fixed cost. If both platforms enter, consumers single-home, taking into account their idiosyncratic preferences for the content provided by the platforms. There are a finite number of firms in the product market. Consumers have binary values for each product: each value is either 0 or a positive number ω. Their utilities are quasilinear in transfers and additively separable across products. There is therefore no competition among firms, who each set a price of ω, and thus extract all surplus from any sales. In contrast, we allow consumers to multi-home in order to study competition among platforms in the ad market. The consumers in our model benefit greatly from advertising since prices are not fully extractive. We study investment on the intensive margin in the form of gradual quality improvements as opposed to on the extensive margin in the form entry (though we later allow for entry as well). In both of our models, investment can be too high or too low due to business-stealing externalities and the inability of platforms to appropriate consumer surplus.

Because Anderson and Coate (2005) study traditional advertising, consumer data and ad targeting do not appear in their model. Ichihashi (2020) builds a model of competing data intermediaries. The intermediary compensates consumers for their data using transfers. It then sells the data to a single firm downstream which can then use the data to price discriminate against the consumer. Prat and Valletti (2021) model digital platforms that have data on users' preferences and sell targeted ads. Their main focus is on the anti-competitive effect of an incumbent firm that buys up ads to prevent an entrant from doing so. They show that platforms' market power can magnify this effect since platforms will sell fewer ads. In both Ichihashi (2021) and Prat and Valletti (2021), firms' profits from sales are exogenously given.[5] Motivated by targeted advertising, Bergemann, Bonatti, and Gan (2019) study a model where a single platform uses data to match consumers to products. The paper's focus is on the efficiency of data use when the data on one consumer is informative of other

[5] In Ichihashi (2021) the firm's profit and consumer utility is an exogenous function of data. In Prat and Valletti (2021) the incumbent firms' profit is an exogenous function of whether the consumer is aware of the entrant.

consumers. More informative data may be harmful in that model since firms may use it to price discriminate. In our model, this channel is not there. More informative data may be harmful because it can lead to lower platform investment.

This paper also relates to a literature that studies data intermediaries in a macroeconomic context. In Jones and Tonetti (2020) and Farboodi and Veldkamp (2021) data is a byproduct of production. It is also a nonrival good that can be bought or sold by firms to improve production. In Jones and Tonetti (2020) these transactions occur through a single intermediary. In Farboodi and Veldkamp (2021), they occur in a competitive market. But, some firms may emerge endogenously as intermediaries by offering their products at low prices to profit primarily from data sales. Neither of these papers study targeted advertising. Rather, in Jones and Tonetti (2020), data enters directly into firms' production functions. In Farboodi and Veldkamp (2021), data is used by firms to forecast an unknown state which affects the optimal choice of production technology. In our model we show that the data used for ad targeting leads to a better matching of firms to consumers and that this is equivalent to an increase in the aggregate productivity of firms.

This paper relates to the literature on ad auctions (Edelman, Ostrovsky, and Schwarz, 2007; Athey and Gans, 2010; Varian, 2007; Hummel and McAfee, 2016; Board, 2009; Bergemann, Heumann, Morris, Sorokin, and Winter, 2021). Most existing work studies the auctions in isolation without modeling a product market. Bidders' values are therefore assumed to be exogenously given. An important result in this literature is that revealing more information to bidders about their values often leads to an increase in revenue. We show that when the product market is accounted for, this result is reversed. The auction model we develop is also novel and introduces significant dynamic effects. Match delay, variation in ad prices, and bid pacing are empirically relevant features of real-world ad auctions. Future empirical work using ad auction data may benefit from the model's ability to capture these. There is a growing empirical literature that uses ad auction data to analyze various policies such as GDPR (Alcobendas, Kobayashi, and Shum, 2021; Johnson, Shriver, and Du, 2020; Beales and Eisenach, 2014; Marotta, Abhishek, and Acquisti, 2019).[6] A structural model is needed to apply the data to counterfactual and welfare analysis.

This paper relates to the literature on competing auctions (McAfee, 1993; Wolinsky, 1988; Iyer, Johari, and Sundararajan, 2014; Chen and Duffie, 2021). As far as we know, it is the first to model competing auctions with different data. Our model is most closely related to that of Wolinsky (1988). There, a dynamic option to wait also matters for bidding. In Wolinsky (1988), once a buyer wins an auction hosted by a given seller, both the buyer and seller exit the market which contrasts with the model in this paper. In practice, for real-time bidding, multiple firms may advertise to the same consumer again and again over time. In Wolinsky (1988), the continuation value upon losing an auction is the same for all buyers. In our model, it differs across firms depending on their expectations of the consumer's values for their products.

[6] General Data Protection Regulation is a regulation on data and privacy protection in the EU.

3. Baseline Model

Model Overview

The model builds on the monopolistic competition framework introduced by Dixit and Stiglitz (1977). A measure F of firms $j \in \mathcal{F}$ each produce a product of a distinct variety. A unit measure of consumers $i \in \mathcal{C}$ each have a taste for variety.[7] The new ingredient is that firms and consumers must search to find each other. They search through intermediaries that we call digital platforms or platforms for short.

There is a measure D of platforms $k \in \mathcal{D}$. At any given time, each consumer splits a unit of attention across platforms. While using platforms, they discover firms through ads displayed there. They also derive flow utility directly from the various services. Each platform offers a free service of a distinct variety. Services may differ in quality. Better quality platforms yield higher flow utility for consumers who choose to spend more attention on them. To compete for attention, platforms invest in quality.

Attention is sold as ads. An opportunity arises for a platform to display an ad to a consumer at random times with a rate proportional to the attention the consumer spends on the platform. At each such display time, the platform invites a finite number of randomly chosen firms to bid in a second-price auction. The winning firm displays its ad. To bid optimally, each firm forms an expectation of the consumer's idiosyncratic value for its product using data provided by the platform.

Consumers

Consumer i's flow utility at time t is an increasing function $u : \mathbb{R}^2_+ \to \mathbb{R}$ of CES aggregates over product and platform consumption:

$$(1) \qquad C_{it} = \left[\int_{\mathcal{F}} v_{ij}^{\frac{1}{\sigma}} c_{ijt}^{\frac{\sigma-1}{\sigma}}\, dj\right]^{\frac{\sigma}{\sigma-1}} \quad \text{and} \qquad (2) \qquad X_{it} = \left[\int_{\mathcal{D}} (q_{kt} x_{ikt})^{\frac{\rho-1}{\rho}}\, dk\right]^{\frac{\rho}{\rho-1}}.$$

In (1), $\sigma > 1$ is the level of product substitutability, c_{ijt} is the consumption of product j, and v_{ij} is the value for product j.[8] Each v_{ij} is drawn from a CDF P on $\mathbb{R}_+$ with a finite mean, independently across i and j. Values are realized at $t = 0$ and fixed thereafter.

In (2), $\rho > 1$ is the level of platform interoperability,[9] x_{ikt} is the attention spent on platform k, and q_{kt} is platform k's quality. In contrast with products, consumers all agree on a platform's quality, which may represent the quantity of its content, the efficacy of its recommendation algorithm, or the aesthetics of its user interface, etcetera.

We assume that, at any time t, consumer i is aware of all platforms in $\mathcal{D}$ but

[7] Firms, consumers, and platforms are each associated with a nonatomic finite-measure space.

[8] We refer to v_{ij} as value but it is *not* the willingness to pay in units of numeraire.

[9] Motivated by the discussion in the Introduction, we refer to platform substitutability as platform interoperability to emphasize its connection with policy.

is aware of only a subset of products $\Omega_{it} \subset \mathcal{F}$ which we call the *consideration set*. Positive amounts can be purchased from this set only: $c_{ijt} = 0$ if $j \notin \Omega_{it}$.

At time t, consumer i selects product and platform consumption to maximize flow utility. This amounts to solving two separate problems:

$$\text{(3)} \quad \max_{\{c_{ijt}\}_{j \in \Omega_{it}}} C_{it} \quad \text{s.t.} \int_{\Omega_{it}} p_{jt} c_{ijt}\, dj = I_t \qquad \text{(4)} \quad \max_{\{x_{ikt}\}_{k \in \mathcal{D}}} X_{it} \quad \text{s.t.} \int_{\mathcal{D}} x_{ikt}\, dk = 1$$

where p_{jt} is firm j's price and I_t is the income at time t. Note that (4) is the same for all consumers. We therefore omit the index i on x_{ikt} going forward.

Consideration Sets

Consumer i's consideration set Ω_{it} evolves as the consumer discovers firms via ads and forgets about those which are already in awareness. Consumer i views ads on platform k at the tick times of a Poisson process with intensity Ax_{kt} that is proportional to the attention spent on that platform. Tick times are independent across platforms conditional on the consumer's choice of attention allocation. The parameter $A > 0$ is the *ad frequency*, which, by the exact law of large numbers (ELLN) is the almost-sure rate that the consumer views ads when aggregated across all platforms.[10] We take A as given, but explain in Section 5 how it can be endogenized, at no cost to tractability, as the outcome of individual optimizations by the platforms.

If, at time t, consumer i views an ad for a firm j which is not currently in the consideration set ($j \notin \Omega_{it-}$), then that firm will enter the set at "the end of the instant" ($j \in \Omega_{it}$).[11] It will not remain there forever. Firms in the set are forgotten at independent exponential times with rate λ_f. This is to ensure that frictions persist in the long run. This property also delivers the real-world feature that the same firms often repeatedly advertise to the same consumers.

Ad Auctions and Consumer Data

Ads are sold via a process modeled after real-time bidding which is widely used in practice.[12] We assume that each ad opportunity is auctioned off by the platform where it arises. When firm j bids in an auction for a given consumer i, it is given data by the platform which is informative of the consumer's value for its product. We assume, for the baseline model, that the data is the same for all platforms, but later allow data to differ across platforms in Section 7. It will turn out that the firm's expectation $\hat{v}_{ij}$ of the value v_{ij} will be a sufficient statistic for it to bid optimally. Rather than explicitly model data and belief-updating, we assume that $\hat{v}_{ij}$ is drawn from a continuous CDF G on $\mathbb{R}_+$, independently across i and j. We also assume that

[10] Duffie, Qiao, and Sun (2020) derive an ELLN for continuous-time random matching.

[11] $\mathbb{1}_{\{j \in \Omega_{it}\}}$ is càdlàg.

[12] E.g. see www.facebook.com/business/help/430291176997542?id=561906377587030.

$P \succeq_{MPS} G$ since this is a necessary and sufficient condition for there to exist some data that generates G by Blackwell (1953).[13] In what follows, we refer to G as data informativeness. We say that data is more informative when G increases in $\succ_{MPS}$. Expectations are realized at $t = 0$ and fixed thereafter.

If, at time t, an opportunity arises for platform k to display an ad to consumer i, then the following events happen within the instant.

1. $N > 1$ firms are invited,[14] uniformly at random from Ω_{it}^c, to bid on the ad.[15]
2. The bidders observe their expectations of the consumer's values.
3. The bidders submit bids in a second-price auction with no reserve price.[16]
4. The highest bidder displays its ad to the consumer.

The parameter N is an exogenous integer. It is meant to reflect real-world latency constraints that limit the number of bidders that can feasibly participate in the auction.[17] We assume that only firms from outside the consideration set are invited (these are the only firms with positive values for the ad). This simplifies the analysis, but the mechanisms behind the main results do not hinge on this. In reality, platforms often place tracking pixels on firms' retail websites. These pixels tell them when a consumer has stopped visiting which is to some extent informative of whether the firms' products are still in consideration.[18]

The Evolution of Consideration Sets

We can now give a formal mathematical description of the evolution of consideration sets, depicted in Figure 1. First, some notation. Let $M_t = |\Omega_{it}|$ denote the size of the consideration set. Let H_t and H_t^c denote the (empirical) CDFs of the expectations $\hat{v}_{ij}$ of firms j in Ω_{it} and Ω_{it}^c respectively. M_t, H_t, and H_t^c are not indexed by i because they are the same for all consumers provided this is so at $t = 0$, which we assume. We take M_0, H_0 as given and assume that firms are distributed symmetrically across initial consideration sets $(\Omega_{i0})_{i \in \mathcal{C}}$.[19] We also assume that $M_0 \, dH_0 < F \, dG$ since the measure of firms with a given expectation in a consideration set must be less than that of the entire economy.[20]

[13] $\succ_{MPS}$ denotes the mean-preserving spread order. $P \succeq_{MPS} G$ if P and G have the same mean and $\int_0^x P(s)\, ds \geq \int_0^x G(s)\, ds$ for all $x \in \mathbb{R}_+$.

[14] When $N = 1$ the same analysis goes through except there is equilibrium multiplicity since any positive bid is optimal for a firm.

[15] The N firms are drawn from the uniform measure on $(\Omega_{it}^c)^N$.

[16] One can show that the results extend to all other standard auction formats in the sense of revenue equivalence. We allow for reserve prices in Appendix G.

[17] In Open Market RTB, latencies in bid response times play an important role. See https://cloud.google.com/architecture/infrastructure-options-for-rtb-bidders and https://medium.com/@datapath_io/how-network-latency-affects-the-rtb-process-for-adtech-6ecbf29d025

[18] See https://www.facebook.com/business/help/742478679120153?id=1205376682832142.

[19] Ω_{it}^c is the relative complement of Ω_{it} in $\mathcal{F}$.

[20] We say M_0, H_0 are *admissible* if this condition is satisfied.

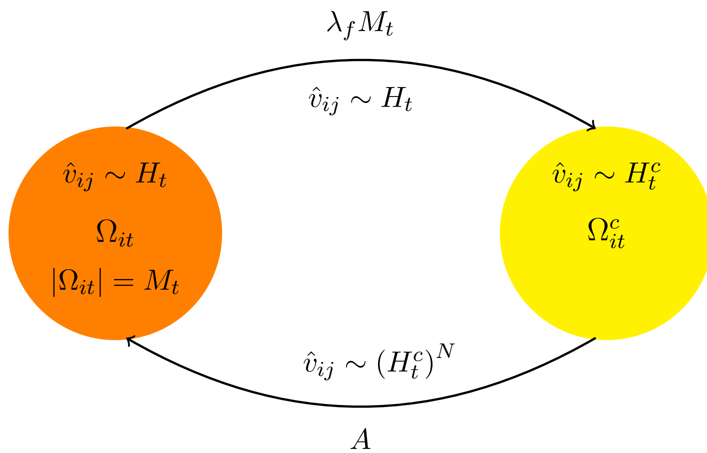

Figure 1: The evolution of consideration sets

Starting from their initial conditions, M_t and H_t evolve deterministically. By the ELLN,

$$dM_t = (A - \lambda_f M_t)\, dt \tag{5}$$

whenever $M_t < F$. To ensure that $M_t < F$ at all times, we assume the following condition for the rest of this paper.

Condition 1. *$A/\lambda_f < F$.*

In equilibrium, it will turn out that the firm with the highest expected value wins in each auction. Then, by the ELLN,

$$d\,(M_t H_t) = \Big(A\,(H_t^c)^N - \lambda_f M_t H_t\Big)\, dt. \tag{6}$$

In (6), the distribution of the inflow is that of the maximum of N independent draws from H_t^c. The distribution of the outflow is H_t since firms are forgotten uniformly at random. To find H_t^c in terms of H_t, we use the accounting identity

$$M_t H_t + (F - M_t)\, H_t^c = FG \tag{7}$$

which closes the system.

Firms

Each firm selects prices and bids to maximize the net present value of its flow profits. The pricing problem is static. At each time t, firm j sets a price p_{jt} to maximize its total flow profit taking as given M_t, H_t, and the prices set by its rivals $p_{lt}, l \neq j$. Given these, in equilibrium, the demand of any consumer i is a known function of the firm's price p_{jt} and value v_{ij}. To meet demand, firm j must produce output. Each unit of output requires a unit of labor. Labor is the only productive resource in the economy. It is used by firms for production and, as we later describe, by platforms for investment. There are L units of labor supplied inelastically by consumers at each time. We set the wage as the numeraire. The marginal cost of production is therefore 1. Given this, firm j solves

$$\max_{p_{jt}} \mathbb{E}[c_{ijt}(p_{jt}, v_{ij})(p_{jt} - 1)] \tag{8}$$

where i is arbitrary since firms are ex-ante symmetric.

While the pricing problem is static, the bidding problem is dynamic: to determine how much to bid in an auction for consumer i, firm j must internalize the value of its outside option to bid in future auctions for that consumer. To write down the firm's problem, let

$$\pi_{\mathcal{F}t}(\hat{v}_{ij}) = \mathbb{E}\left[c_{ijt}(p_{jt}, v_{ij})(p_{jt} - 1)|\hat{v}_{ij}, j \in \Omega_{it}\right]$$

be the expected flow profit from selling to consumer i at time t conditional on the data supplied by the platforms. Also, let

$$\lambda_{at} = \frac{NA}{F - M_t} \tag{9}$$

denote the Poisson rate that firm j enters auctions for consumer i while outside Ω_{it}. In (9), λ_{at} equals the rate that auction invitations are sent out, normalized by the measure of eligible firms. Let τ_z denote the time of zth entry into an auction for consumer i by firm j. In equilibrium, firm j takes as given the bidding strategies of its rivals which are of the following form. Each firm l bids according to an increasing function $B_t : \mathbb{R}_+ \to \mathbb{R}_+$ at time t.[21] B_t maps firm l's expectation $\hat{v}_{lij}$ to a bid $B_t(\hat{v}_{lij})$ in an auction for consumer i at time t. Given this, firm j solves

$$\Pi_{\mathcal{F}} = \max_{\{b_z\}_{z\in\mathbb{N}}} \mathbb{E}\left[\int_0^\infty e^{-rs}\pi_{\mathcal{F}s}(\hat{v}_{ij})\mathbb{1}_{\{j\in\Omega_{is}\}}\,ds - \sum_{z=1}^{\infty} e^{-r\tau_z} B_{\tau_z}\left(\hat{v}_z^{(1)}\right)\mathbb{1}_{\left\{b_z > B_{\tau_z}\left(\hat{v}_z^{(1)}\right)\right\}}\right] \tag{10}$$

where $\hat{v}_z^{(1)}$ is the highest expectation of the $N-1$ other bidders in the zth auction: $\hat{v}_z^{(1)} \sim (H^c_{\tau_z})^{N-1}$ conditional on τ_z. In (10), the bid in the zth auction b_z is measurable with respect to the expectation $\hat{v}_{ij}$ and auction time τ_z. As in (8) i is arbitrary since consumers are ex-ante symmetric.

Platforms

Each platform k selects an investment strategy to maximize the net present value of its flow profits taking as given the paths of average ad prices $\pi_{\mathcal{D}t}$,[22] consumers'

[21] It is without loss to assume a bidding function which is common to all firms and increasing: see Lemma 8 in Appendix A.

[22] The average ad price $\pi_{\mathcal{D}t} = \mathbb{E}[B_t(\hat{v}^{(2)})]$ where $\hat{v}^{(2)}$ is the second highest of N independent draws from H^c_t.

demands for the platform x_{kt}, and the quality levels of its rivals q_{lt}, $l \neq k$. Formally, by the ELLN, platform k solves

$$\Pi_{\mathcal{D}} = \max_{\{L_{kt}\}_{t\geq 0}} \int_0^\infty e^{-rt} \left(\pi_{\mathcal{D}t} A x_{kt} - L_{kt}\right) dt \tag{11}$$

subject to

$$dq_{kt} = (L_{kt}^{\varphi} - \delta q_{kt})\, dt,\ \ q_{k0} = q_0.$$

In (11), we have suppressed the dependency of demand x_{kt} on platforms' quality levels $q_{lt}, l \in \mathcal{D}$. The investment technology $0 < \varphi < 1$ has decreasing returns. The depreciation rate $\delta > 0$ represents the platform's content becoming stale or less relevant over time in the absence of investment.

Income and Welfare

To close the model, we assume that consumers own the firms and platforms. Income is equal to the sum of labor compensation and firm and platform flow profits. Since any transfers from advertising net out, this gives

$$I_t = L + \underbrace{M \int_0^\infty \pi_{\mathcal{F}t}(\hat{v})\, dH_t(\hat{v})}_{\text{total sales profits}} - \underbrace{\int_{\mathcal{D}} L_{kt}\, dk}_{\text{total investment}} .$$

Social welfare is defined to be consumer surplus

$$\int_0^\infty e^{-rt} u(C_{it}, X_{it})\, dt.$$

We could instead define welfare to be a weighted sum of consumer surplus and firm and platform profits while taking income as given and not modeling the labor market. But, surplus and profits are measured in different units: utility is not quasilinear in transfers.[23] Not only this, but welfare analysis would be more difficult. Endogenizing income as we do leads to the cleanest and most interpretable results.

Equilibrium

An *equilibrium* given initial conditions M_0, H_0, q_0 is a collection of processes $\{c_{ijt}\}$, $\{x_{kt}\}$, $\{M_t\}$, $\{H_t\}$, $\{p_{jt}\}$, $\{B_t\}$, $\{L_{kt}\}$, $\{q_{kt}\}$ and $\{I_t\}$ such that at each $t \geq 0$:

1. Each consumer i selects product and platform demands c_{ijt} and x_{ikt} optimally.
2. The measure of firms in Ω_{it} is M_t and the CDF of expected values in Ω_{it} is H_t.
3. Each firm j optimally sets price p_{jt} and bids according to B_t.
4. Each platform k optimally invests at rate L_{kt}.

[23] A version of the model that is quasilinear in transfers is also less tractable because it leads to a fixed point in the total amount spent on products that can not be solved explicitly.

5. Each platform k has quality q_{kt}.

6. Income I_t is the sum of labor and firms' and platforms' flow profits.

This completes the model setup. In the model, product market outcomes, ad revenue, and investment by platforms are all endogenously determined in equilibrium. Such a model is well-suited to addressing the regulatory issues described in the Introduction. For this, we must first obtain a tractable equilibrium characterization.

4. Unique Stationary Equilibrium

The rest of this paper analyzes stationary equilibria. A *stationary equilibrium* is an equilibrium with initial conditions M_0, H_0, q_0 such that:[24]

1. M_t and H_t are in steady state: $M_t = M_0$ and $H_t = H_0$ at each $t \geq 0$.

2. Each platform k's quality is in steady state: $q_{kt} = q_0$ at each $t \geq 0$.

We now sketch the procedure to solve for a stationary equilibrium. In the process, we develop some results that apply to all equilibria. Namely, Lemmas 1–3 (see also Lemmas 7 and 8 in Appendix A). All proofs are in Appendix A.

Lemma 1 reports the solutions to consumer i's problems (3) and (4).

Lemma 1. *In any equilibrium, the following hold:*

1. *Consumer i's demand for product $j \in \Omega_{it}$ is*

$$c_{ijt} = \frac{I_t v_{ij}}{\int_{\Omega_{it}} v_{il} p_{lt}^{1-\sigma} dl} p_{jt}^{-\sigma}. \tag{12}$$

2. *Consumer i's demand for platform $k \in \mathcal{D}$ is*

$$x_{kt} = \frac{q_{kt}^{\rho-1}}{\int_{\mathcal{D}} q_{lt}^{\rho-1} dl}. \tag{13}$$

The demand for product j is increasing in the consumer's value for that product but decreasing in the consumer's values for the other products in the consideration set Ω_{it}. Similarly, the demand for platform k is increasing in the quality of that platform but decreasing in the quality levels of all other platforms. When platforms are more interoperable, demand is more sensitive to quality.

Lemma 2. *In any equilibrium, the following hold:*

[24] In the setup we assume initial conditions that are symmetric across consumers and platforms. All results generalize when initial conditions are allowed to be asymmetric. However, one can show that all stationary equilibria necessarily have symmetric initial conditions.

1. *Firm j's flow profit from selling to consumer i is*

$$\frac{I_t v_{ij}}{\int_{\Omega_{it}} v_{il} p_{lt}^{1-\sigma} dl} p_{jt}^{-\sigma} (p_{jt} - 1) \tag{14}$$

   if $j \in \Omega_{it}$.

2. *Firm j's profit-maximizing price is*

$$p_{jt} = \frac{\sigma}{\sigma - 1}. \tag{15}$$

The same price in (15) maximizes the flow profit in (14) of *any* consumer i. It does not depend on the firm's beliefs about the consumer's values or consideration set or even the prices set by the other firms as one might have thought. These terms only scale the flow profit by a factor that is the same for any price that the firm sets. As a result, the optimal price is a function of only product substitutability. This is a key property that allows a tractable general equilibrium analysis with data modeled nonparametrically. It also implies that even if we allow firms to personalize prices,[25] they would choose not to. In reality, personalized pricing is not a prevalent practice (though perhaps for other unmodeled reasons). We shut this practice down to focus on other issues.[26]

Though a firm's choice of price will not depend on what it knows about consumers' consideration sets, its bidding strategy will. Recall that M_t is the measure of firms in consideration sets and H_t (H_t^c) is the CDF of expectations inside (outside) consideration sets. In Appendix A we use (5)–(7) to derive the paths of M_t, H_t, and H_t^c which we then use to prove Lemma 3.

Lemma 3. *The unique steady state solution M, H, H^c to (5)–(7) satisfies*

$$M = \frac{A}{\lambda_f}, \tag{16}$$

$$H = (H^c)^N, \tag{17}$$

and

$$M (H^c)^N + (F - M) H^c = FG. \tag{18}$$

Moreover, starting from any admissible M_0, H_0, the unique solution to (5)–(7) converges to the steady state as $t \to \infty$, in that $M_t, H_t, H_t^c \to M, H, H^c$.

Lemma 3 asserts the stability of the steady state. It also implies that H can be computed by solving pointwise for the unique nonnegative root of (18). Figure 2 graphs H, G, H^c and their densities for the case when G is uniform on $[0, 1]$. As seen

[25] That is, we allow p_{jt} in (8) to be measurable with respect to $\hat{v}_{ij}$.

[26] This yields a clean benchmark of analysis. Bergemann, Brooks, and Morris (2015) show that the welfare effects of personalized pricing may depend sensitively on the information structure.

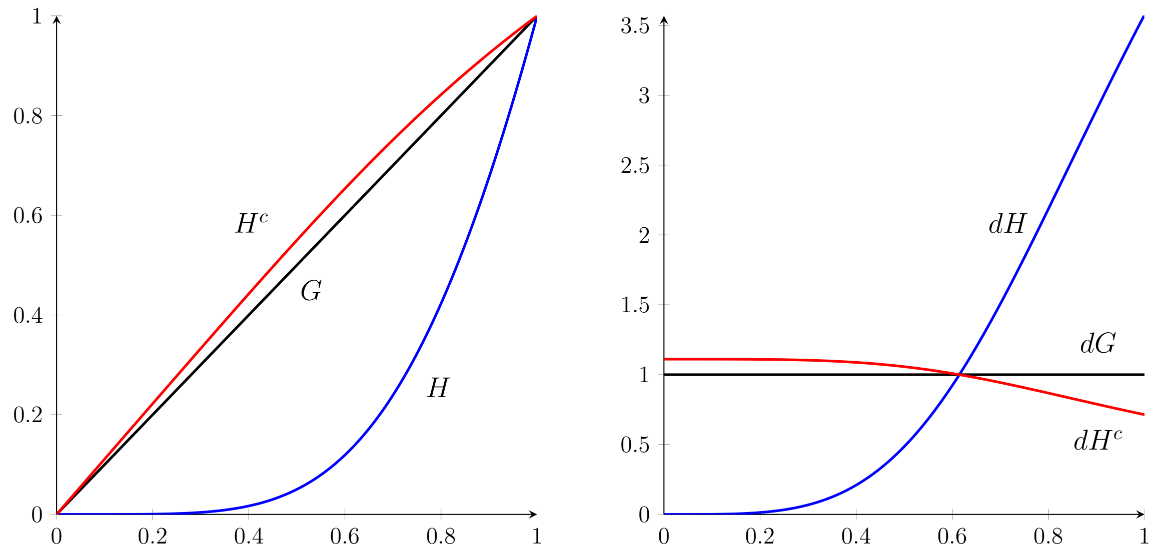

Figure 2: The steady state distributions and densities when $N = 5$, $M/F = 1/10$, and $G = U[0, 1]$.

in the figure, the positive selection in the auctions leads to $H \succ_{FOSD} G \succ_{FOSD} H^c$.[27]

We are now ready to derive firms' bidding strategies. To bid optimally in an auction for consumer i, firm j must assess its expected flow profit from selling to consumer i. Suppose that M_t and H_t are in steady state. Let μ_H denote the mean of the CDF H which by the ELLN is the average value of firms in consideration sets. Then, using Lemma 2, the expected flow profit is

$$\mathbb{E}\left[\frac{I}{\sigma \int_{\Omega_{it}} v_{il} dl} v_{ij} \middle| \hat{v}_{ij}\right] = \pi_{\mathcal{F}} \hat{v}_{ij} \tag{19}$$

where

$$\pi_{\mathcal{F}} = \frac{I}{\sigma M \mu_H}$$

and I denotes the stationary equilibrium level of income. The expected flow profit is linear in the expectation $\hat{v}_{ij}$ of the consumer's value with a coefficient $\pi_{\mathcal{F}}$ that is decreasing in the cumulative value $M\mu_H$ of the firms in the consideration set. Changes to data informativeness G or ad frequency A will affect flow profits through the cumulative value. Firm j will internalize this when bidding in the ad auctions. This feature is missing from most existing models which study the ad auctions in isolation of the product market. We show in Section 5 that more informative data or higher ad frequency typically lead to a decline in ad revenue once this is accounted for but the opposite is true when it is not.

In a stationary equilibrium, each firm bids according to a function B that does not vary over time.[28] Let $V(\hat{v}_{ij})$ denote firm j's continuation value for consumer i at

[27] $\succ_{FOSD}$ denotes the usual stochastic order.

[28] This can be shown using Lemma 8 in Appendix A where we characterize the *entire path* of bid functions B_t in any equilibrium taking as given only the path of income I_t.

the time of auction entry. Because the auction is second-price, firm j's optimal bid is its gain in continuation value from winning. Thus, it must be that

$$B(\hat{v}_{ij}) = \underbrace{\frac{\pi_{\mathcal{F}}}{\lambda_f + r}\hat{v}_{ij} + \frac{\lambda_f}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r}V(\hat{v}_{ij})}_{\text{continuation value if win}} - \underbrace{\frac{\lambda_a}{\lambda_a + r}V(\hat{v}_{ij})}_{\text{continuation value if lose}}. \tag{20}$$

The first term is the net present value of sales to the consumer while in the consideration set which the firm exits at rate λ_f. It then enters an auction for consumer i again at rate $\lambda_a = NA/(F - M)$ which corresponds to the second term (see (9)). If the firm loses the auction, it is instantly eligible to enter another one which corresponds to the third term.

To derive a second condition relating B and V we apply Bellman's principle of optimality which yields

$$V(\hat{v}_{ij}) = \underbrace{\frac{\lambda_a}{\lambda_a + r}V(\hat{v}_{ij})}_{\text{continuation value if lose}} + \underbrace{H^c(\hat{v}_{ij})^{N-1}}_{\text{win probability}}\left(B(\hat{v}_{ij}) - \underbrace{\mathbb{E}\left[B(\hat{v}^{(1)})|\hat{v}_{ij} > \hat{v}^{(1)}\right]}_{\text{expected payment}}\right) \tag{21}$$

where $\hat{v}^{(1)} \sim (H^c)^{N-1}$. In (21), the probability that firm j wins the auction is the probability that the other $N-1$ bidders have lower expected values. The unique solution to (20) and (21) for the the bidding function B is reported below.

Lemma 4. *In a stationary equilibrium, the following hold:*

1. *Firm j bids*

$$B(\hat{v}_{ij}) = \pi_{\mathcal{F}} \int_0^{\hat{v}_{ij}} \frac{1}{r + \lambda_f + \lambda_e(s)}\, ds \tag{22}$$

   in an auction for consumer i where $\lambda_e = \lambda_a(H^c)^{N-1}$.

2. *The Poisson rate that firm $j \in \Omega_{it}^c$ enters Ω_{it} is $\lambda_e(\hat{v}_{ij})$.*

3. *The average ad price is*

$$\pi_{\mathcal{D}} = \pi_{\mathcal{F}} \int_0^{\infty} \frac{1 - NH^c(s)^{N-1} + (N-1)H^c(s)^N}{r + \lambda_f + \lambda_e(s)}\, ds. \tag{23}$$

4. *The net present value of a firm's flow profits is*

$$\Pi_{\mathcal{F}} = \frac{I/\sigma - \pi_{\mathcal{D}}A}{rF}. \tag{24}$$

The bid is increasing in the coefficient of firms' flow profits $\pi_{\mathcal{F}}$ and decreasing in the rates that firms match with consumers λ_e. The latter is because the value of a firm's outside option is higher when λ_e is higher. If λ_e was 0 everywhere, the bid would equal the firm's value for the ad disregarding the outside option. Though platforms have a monopoly over their ad opportunities, they are in effective competition with each other through the outside option. Thus, $1/\lambda_e$ is a measure of platforms' market power in the ad market.

In the model, it is strictly optimal for a firm to reduce its bid due to the outside option only because there is variation in auction competition (and thus ad prices). This is for the following reason. Suppose that a firm bids its value and wins at a price close to its bid, thus earning little surplus. It could instead reduce its bid so that it would have lost the auction. It then might have been able to enter an auction in the near future where by chance competition was less stiff and thus win at a price significantly below its bid. For some level of reduction, this deviation would be profitable. The strategic bid reduction in anticipation of variation in auction competition is known as "bid pacing" in online ad auctions where it plays a prominent role. Variation in auction competition and bid pacing are empirically significant[29] and thus may be relevant to future empirical work.[30]

Given the bidding function, the expected auction revenue is computed by the formula $\pi_{\mathcal{D}} = \mathbb{E}[B(\hat{v}^{(2)})]$ where $\hat{v}^{(2)}$ is the second highest of N independent draws from H^c. This gives (23). As the ad frequency approaches its upper limit $A \to \lambda_f F$, the match rate diverges $\lambda_e \to \infty$ leading to a frictionless product market in the limit and vanishing auction revenue $\pi_{\mathcal{D}} \to 0$.

Given (23), firms' profits $\Pi_{\mathcal{F}}$ follow from accounting: firms extract the income of consumers but must pay labor costs as well as advertising costs. The total flow advertising cost is $\pi_{\mathcal{D}} A$ when aggregated across all firms.

Firms' advertising costs are platforms' ad revenues. Each platform k takes total ad revenue $\pi_{\mathcal{D}} A$ and platform demand x_{kt} as given when selecting its investment strategy to solve (11). Recall from (13) that interoperability ρ determines the elasticity of demand x_{kt} with respect to the platform's quality q_{kt}. Intuitively, if ρ is too high, a stationary equilibrium can not exist because each platform would have an incentive to deviate by investing more aggressively to steal more of the market. To ensure the existence of a stationary equilibrium, we assume the following condition for the rest of this paper.

Condition 2. $\rho \leq 2$

However, even when Condition 2 is violated, we will identify the unique *candidate* stationary equilibrium: all of our results will apply so long as equilibrium exists.

In Appendix A we derive platforms' investment rates as a function of ad revenue and primitives via the Maximum Principle.

Lemma 5. *In a stationary equilibrium, the following hold:*

[29] See https://www.facebook.com/business/help/2024547657774300 which states that "the difference between the competitive bid on this day and the average competitive bid for the previous 3 days...is considered significant if it is over 20%." See also https://www.facebook.com/business/help/1754368491258883?id=561906377587030 which states that "Pacing helps us deliver your ads in a way that accounts for that variation so you can meet your cost goals..."

[30] See Section 2 for a discussion of the recent empirical literature that studies ad auction level data. The structural model of this paper may be useful for counterfactual and welfare analysis.

1. *Each platforms invests at the constant rate*

$$L_{\mathcal{D}} = \frac{\varphi\delta\pi_{\mathcal{D}}A(\rho-1)}{D\,(r+\delta)}. \tag{25}$$

2. *The net present value of a platform's flow profits is*

$$\Pi_{\mathcal{D}} = \frac{\pi_{\mathcal{D}}A - DL_{\mathcal{D}}}{rD}. \tag{26}$$

As seen in (25), investment is increasing in ad revenue $\pi_{\mathcal{D}}A$ and interoperability ρ which together determine the benefit of investment to a platform: the amount of revenue the platform can steal from its rivals by improving its quality.

All that is left is to compute income I. Let $\hat{\pi}_{\mathcal{D}} = \pi_{\mathcal{D}}/I$ be the average ad price per unit of income. From (23),

$$\hat{\pi}_{\mathcal{D}} = \frac{1}{\sigma M\mu_H}\int_0^\infty \frac{1 - NH^c(s)^{N-1} + (N-1)H^c(s)^N}{r+\lambda_f+\lambda_e(s)}\,ds. \tag{27}$$

Then, rewriting (25), we have

$$L_{\mathcal{D}} = \frac{\varphi\delta\hat{\pi}_{\mathcal{D}}A(\rho-1)}{D\,(r+\delta)}I. \tag{28}$$

Investment is increasing in income because ad revenue is higher when income is higher.

Given investment, income must equal

$$I = \frac{\sigma}{\sigma-1}(L - DL_{\mathcal{D}}). \tag{29}$$

In (29), the right-hand side is the total revenue of firms: recall that $\frac{\sigma}{\sigma-1}$ is the price of output and $L - DL_{\mathcal{D}}$ is total production by labor market clearing. Income must equal this because all costs are labor costs and platforms extract all of their revenues from firms. Income is decreasing in investment since investment diverts resources away from production.

The solution of the linear system (28), (29) determines the unique stationary equilibrium levels of income and investment as represented graphically in Figure 3.

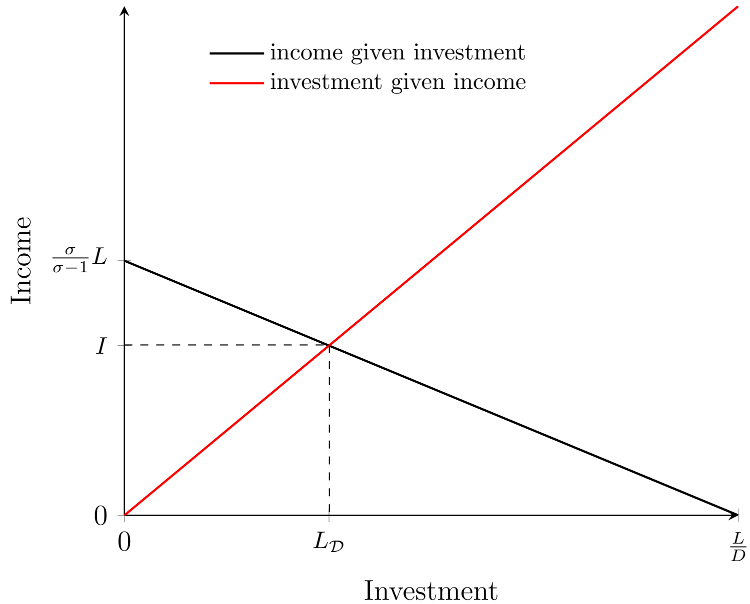

Figure 3: Determination of the stationary equilibrium levels of investment and income.

This completes the derivation of the stationary equilibrium. Theorem 1 summarizes its properties. Each subsequent part of the theorem describes objects which are defined explicitly using only objects from earlier parts.

Theorem 1. *In the unique stationary equilibrium:*

1. *M, the measure of firms in Ω_{it}, and H, H^c, the CDFs of expected values in Ω_{it} and Ω_{it}^c respectively are as in Lemma 3.*

2. *Platform investment is*

$$L_{\mathcal{D}} = \frac{\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}}A(\rho-1)}{r+\delta+\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}}A(\rho-1)}\frac{L}{D}. \tag{30}$$

   where $\hat{\pi}_{\mathcal{D}}$ is the average ad price per unit of income defined by (27).

3. *Platform quality is $q = L_{\mathcal{D}}^{\varphi}/\delta$.*

4. *Income I is given by (29).*

5. *Product demands $\{c_{ijt}\}$, platform demands $\{x_{ikt}\}$, prices $\{p_{jt}\}$, and the bidding function B are as in Lemmas 1, 2, and 4.*

6. *Welfare is $u(C, X)/r$ where $C = (L - DL_{\mathcal{D}})(M\mu_H)^{\frac{1}{\sigma-1}}$ and $X = D^{\frac{1}{\rho-1}}q$ are aggregate product and platform consumption respectively.*

As seen in Part 6, welfare depends on only investment $L_{\mathcal{D}}$ and the cumulative value $M\mu_H$. Given investment, it does not depend on any other moments of the CDF H. Higher investment $L_{\mathcal{D}}$ leads to higher aggregate platform consumption X but lower aggregate product consumption C since less labor is used for production. When the cumulative value $M\mu_H$ is higher, each unit of production yields higher aggregate product consumption.

5. Comparative Statics

Theorem 1 is a nearly explicit characterization of the stationary equilibrium. It shows how various parameters such as data informativeness or interoperability determine equilibrium outcomes. With this in hand, we are ready to begin thinking about some important questions. What will be the long run effects of policies such as GDPR or CCPA which degrade data informativeness by banning data tracking without user consent? What about policies such as the DMA or ACCESS Act which enforce platform interoperability?[31] Will these policies necessarily improve welfare? We will answer these questions, within the model, next and in the process, uncover some compelling economic mechanisms. Our answers will come in the form of comparative statics with respect to data informativeness G and interoperability ρ. We will also derive comparative statics with respect to ad frequency A. These latter comparative statics are intrinsically interesting. However, we also use them, at the end of the section, to argue that the main comparative statics with respect to G and ρ persist when A is endogenous (in a manner to be described later) and in some cases, are strengthened.

Data Informativeness

When data is more informative, firms are more likely to win auctions for consumers who value their products highly. As a result, the cumulative value $M\mu_H$ in consideration sets increases by Lemma 10 in Appendix B. It is difficult to prove other global monotone comparative statics with respect to data informativeness. However, it turns out that ad revenue is maximal when data is uninformative in that G places all of its mass on the mean of the prior P.[32] This in turn implies that investment is maximal and aggregate product consumption is minimal when data is uninformative.

These results may seem counterintuitive. Existing results in the literature show that more informative data typically leads to higher ad revenue (Board, 2009; Bergemann, Bonatti, and Gan, 2022). However, these papers take bidders' values for the auctioned item as given. We instead assume that bidders' values for ads are the endogenous outcome of competition in the product market.

For intuition, consider the equilibrium bid of firm j reproduced below:

$$B(\hat{v}_{ij}) = \pi_{\mathcal{F}} \int_0^{\hat{v}_{ij}} \frac{1}{r + \lambda_f + \lambda_e(s)}\, ds.$$

If we ignore the effect of the outside option by setting $\lambda_e = 0$ then we have

$$\pi_{\mathcal{F}} \frac{\hat{v}_{ij}}{r + \lambda_f} = \frac{I}{\sigma M \mu_H} \frac{\hat{v}_{ij}}{r + \lambda_f}. \tag{31}$$

[31] GDPR: General Data Privacy Regulation; CCPA: California Consumer Privacy Act; DMA: Digital Markets Act; ACCESS Act: Augmenting Compatibility and Competition by Enabling Service Switching Act. The ACCESS Act is in the legislative process but the others are laws.

[32] Strictly speaking, this does not fall within the baseline model because we assumed G was continuous to avoid ties in the auctions. But it is clear how the analysis extends if we assume ties are broken uniformly at random. Proposition 1 can also be understood in the sense of a limit.

This is firm j's value for the ad if there were no future auctions. In the settings studied in the previously referenced literature, more informative data leads to higher revenue because it raises the available surplus which the auctioneer can extract (some fraction of). Namely, due to convexity, the expectation of the maximum of bidders' values is higher. But, in our setting, the expectation of the maximum of bidders' values for the ad (31) is

$$\frac{I}{\sigma M(r+\lambda_f)}$$

for any data informativeness G. This is because the winning bidder has an expected value which is on average no higher than those of the firms already in the consideration set. When data is more informative, the expectation of the maximum $\hat{v}_{ij}$ increases, but so does competition in the product market which causes firms' flow profits to decline since μ_H rises. Thus there is no increase in bidder surplus. Rather, an increase in data informativeness typically leads to higher information rents and also variation in auction competition which exacerbates bid pacing (recall that this second effect appears in the equilibrium bid through λ_e). As a result, ad revenue typically declines.

When data is uninformative, there is no information rent and there is no variation in auction competition. Thus, ad revenue is maximal.

Proposition 1. *If data G is uninformative, then among all feasible data:*

1. *Aggregate product consumption C is minimal.*
2. *Ad revenue $\pi_{\mathcal{D}}A$ is maximal.*
3. *Investment $L_{\mathcal{D}}$ is maximal.*

A common concern for regulators is that banning data tracking without user consent will lead to a reduction in ad revenue and thus investment by platforms. Proposition 1 shows that this is not necessarily the case and that the opposite may often happen.

Figure 4 plots ad revenue as a function of data informativeness for a numerically computed example when G is uniform on $[.5-\epsilon, .5+\epsilon]$ as ϵ ranges from 0 to .5. Ad revenues decline monotonically over the entire interval: platforms are better off if data is less informative. If they could collude, they would prefer to provide no data. Yet, it is often said that data is integral to the profitability of platforms. Indeed, in the model, more informative data *is* typically valuable to the *individual* platform. We can compute its value.

To illustrate, continue to assume that G is uniform on $[.5-\epsilon, .5+\epsilon]$. Now take an individual platform k and garble its data such that firm j's expectation given platform k's data is now $\tilde{v}_{ij} \sim \tilde{G}$ where $\tilde{G}$ is uniform on $[.5-\tilde{\epsilon}, .5+\tilde{\epsilon}]$ with $\tilde{\epsilon} < \epsilon$. The explicit construction of $\tilde{v}_{ij}$ from v_{ij} is in Appendix B. Since platform k is infinitessimal, the bidding strategy B on all other platforms is unchanged. For this example, one can show that the optimal bid on platform k is $\mathbb{E}[B(\hat{v}_{ij})|\tilde{v}_{ij}, j \in \Omega_{it}^c]$. This is the expected gain in continuation value from winning the auction conditional on the signal $\tilde{v}_{ij}$ and the knowledge that the firm is outside the consideration set.

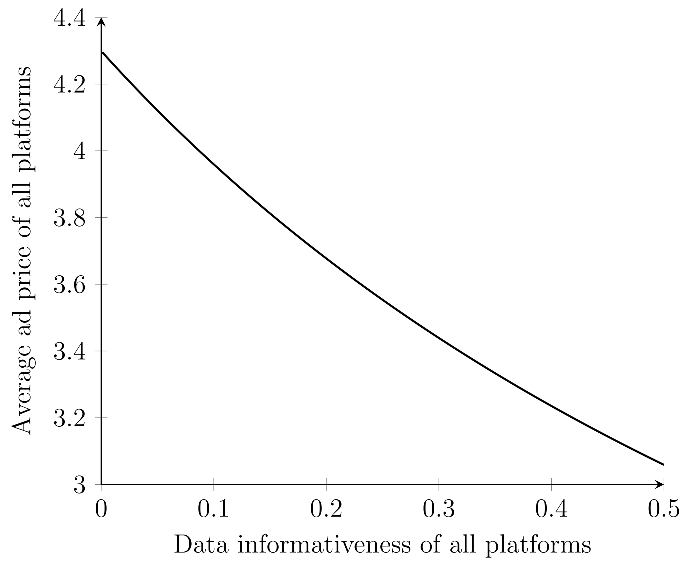

Figure 4: The average ad price $\pi_{\mathcal{D}}$ is plotted as against data informativeness $G = \text{Uniform}[.5-\epsilon, .5+\epsilon]$ as ϵ ranges from 0 to .5 for parameter values $F = .5$, $A = .1$, $\lambda_f = 1$, $N = 5$, $\sigma = 3$, $\rho = 1.5$, $r = .1$, $\varphi = .5$, $\delta = .1$, $L = 1$, where D can be any positive value.

Figure 5 plots platform k's average ad price as a function of its data informativeness as $\tilde{\epsilon}$ ranges from 0 to .5 when ϵ is .5.

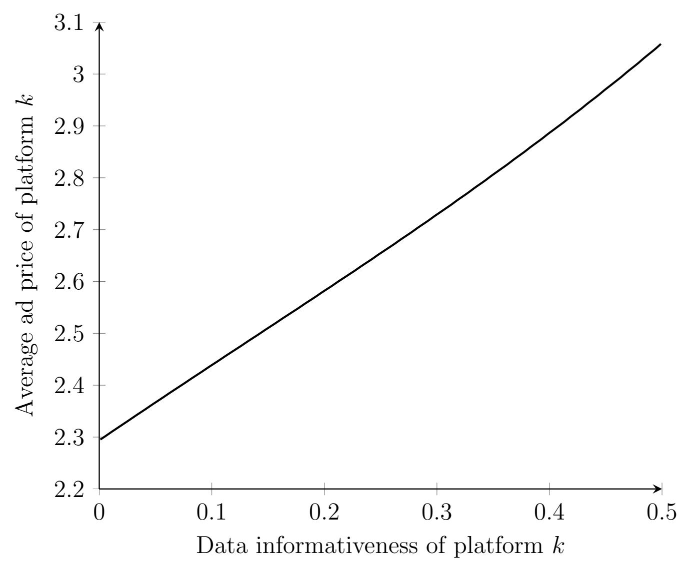

Figure 5: Platform k's average ad price plotted against its data informativeness $\tilde{G} = \text{Uniform}[.5 - \tilde{\epsilon}, .5 + \tilde{\epsilon}]$ as $\tilde{\epsilon}$ ranges from 0 to .5 when the other platforms have data informativeness $G = \text{Uniform}[0, 1]$, all other parameters are as in Figure 4, and for the garbling described in Appendix B.

As data informativeness improves, platform k's average ad price increases over the entire interval.

Figure 6 plots platform k's average ad price as a function of the data informativeness of its rivals as ϵ ranges from .25 to .5 when $\tilde{\epsilon}$ is .25.

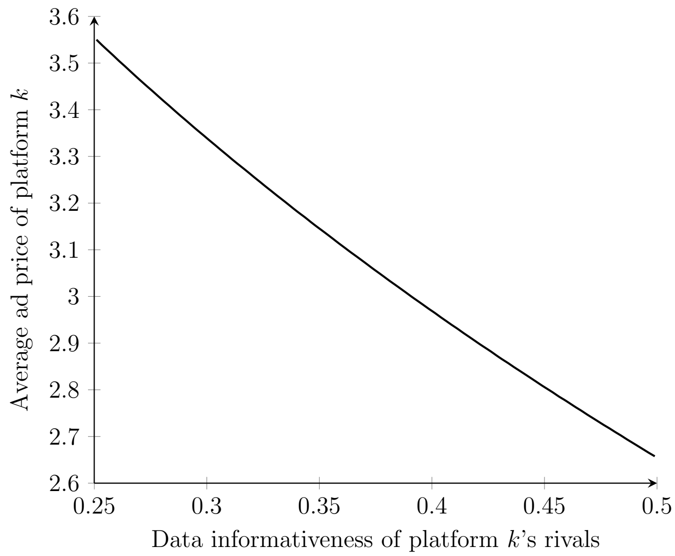

Figure 6: Platform k's average ad price is plotted against the data informativeness of the other platforms $G = \text{Uniform}[.5-\epsilon, .5+\epsilon]$ as ϵ ranges from .25 to .5 when platform k's data informativeness is $G = \text{Uniform}[.25, .75]$, all other parameters are as in Figure 4, and for the garbling described in Appendix B.

As data informativeness improves, the average ad price decreases over the entire interval. This effect occurs because of firms' outside options to wait to bid in an auction with more informative data in the future. Thus, data informativeness has *competitive* value.

These results suggest that the model may be able to match, at least qualitatively, certain empirical trends. In the US, the total advertising expenditure as a fraction of GDP has been relatively stable over time at roughly 1.5 percent with some sources saying that it has been declining (Silk, Berndt, et al., 2021). This is despite the introduction of digital platforms with data and ad targeting capabilities that traditional media do not have (or, in light of Proposition 1, perhaps *because* of it). This suggests that the rapid growth in digital advertising is because digital platforms have stolen business away from traditional media perhaps because of these capabilities as hinted at by Figure 6. In Section 7 we extend the model to allow for two groups of platforms, each of positive measure, which differ in terms of their data. If we interpret traditional media as platforms with less informative data, this extended model can reproduce these trends for a wide variety of parameters.

Interoperability

When interoperability is higher, attention is more sensitive to quality and so platforms can more easily steal it from each other. This leads them to invest more aggressively. This implies that less labor is allocated to production which leads to a decline in aggregate product consumption. Less production also implies that firms' revenues are lower and so they bid less in the ad auctions leading to a decline in ad revenue. Proposition 2 summarizes.

Proposition 2. *If interoperability ρ increases then:*

1. *Aggregate product consumption C decreases.*
2. *Ad revenue $\pi_{\mathcal{D}} A$ decreases.*
3. *Investment $L_{\mathcal{D}}$ increases.*

Thus, consistent with intuition in regulatory discussion, mandating interoperability can promote competition among platforms and lead to more investment (OECD, 2021). However, it may also have the unintended consequence of leading to worse aggregate product consumption. In the baseline model, the mechanism for this operates through the labor market. When we endogenize ad frequency at the end of this section, we will see that there is a second important channel: when platforms are more interoperable, they set lower ad frequencies which further reduces aggregate product consumption. To show this we must first present some comparative statics with respect to ad frequency.

Ad Frequency

When ad frequency is higher, consumers discover more firms and so the size of their consideration sets increases $M = A/\lambda_f$. However, the average value μ_H of the firms in the set decreases. In fact, both H and H^c decrease in $\succ_{FOSD}$ as shown in Lemma 10 of Appendix B. The intuition is that firms which remain outside the set are more likely to have participated in more auctions and lost and thus have lower expected values. As a result, the firms inside the set are also more likely to have lower expected values since they won auctions against competition which is on average less stiff. Though μ_H declines, we show in Appendix B that overall, the growth in size dominates in that the cumulative value $M\mu_H$ rises. This serves to increase aggregate product consumption which recall is $C = (L - DL_{\mathcal{D}})(M\mu_H)^{\frac{1}{\sigma-1}}$. Further, it turns out that ad revenue also declines leading to a decline in investment and a rise in production which again serves to increase aggregate product consumption.

To understand the decline in ad revenue, recall that

$$\pi_{\mathcal{D}} A = \pi_{\mathcal{F}} \mathbb{E}\left[\int_0^{v^{(2)}} \frac{1}{r + \lambda_f + \lambda_e(s)} ds\right] A$$

where $v^{(2)}$ is the second highest of N independent draws from H^c, $\pi_{\mathcal{F}} = I/(\sigma M \mu_H)$, and $\lambda_e = \lambda_a (H^c)^{N-1}$. Though a larger quantity of ads are sold, firms' flow profits decline since the size of consideration sets increases. If the average value μ_H remained constant, these effects would cancel each other out but it decreases which serves to increase ad revenue. However, H^c decreases in $\succ_{FOSD}$ which causes the distribution of $v^{(2)}$ to decrease in $\succ_{FOSD}$ and match rates λ_e to increase. This serves to decrease ad revenue. We show in Appendix B that these latter effects dominate.

Proposition 3. *If ad frequency A increases then:*

1. *Aggregate product consumption C increases.*

2. *Ad revenue $\pi_{\mathcal{D}} A$ decreases.*

3. *Investment $L_{\mathcal{D}}$ decreases.*

In reality, one might wonder if changes to data informativeness or interoperability will also affect platforms' optimal choices of ad frequency, and thus whether the comparative statics in Propositions 1 or 2 are robust to this.

What if Ad Frequency is Endogenous?

Suppose that the ad frequency A is endogenous as follows. Suppose that consumers dislike viewing ads so that the effective quality level of platform k is now $\nu(A_{kt})q_{kt}$ where $\nu : \mathbb{R}_+ \to (0, 1]$ is a decreasing function of A_{kt}, the ad frequency set by platform k. Then the attention that platform k will receive is

$$x_{kt} = \frac{[\nu(A_{kt})q_{kt}]^{\rho-1}}{\int_{\mathcal{D}} [\nu(A_{lt})q_{lt}]^{\rho-1} \, dl}.$$

If each platform k selects A_{kt} to maximize its flow profits, then

$$A = \arg\max_{A_{kt} \leq \overline{A}} \pi_{\mathcal{D}} A x_{kt} = \arg\max_{A_{kt} \leq \overline{A}} A_{kt} \nu(A_{kt})^{\rho-1}$$

where $\overline{A}$ is some exogenous maximal ad frequency. Thus, A is a function of just ρ and ν. Moreover, it is weakly decreasing in ρ. Using this result, Proposition 3 implies that Proposition 1 and comparative statics 1 and 3 of Proposition 2 continue to hold when A is endogenous (but it is unclear whether comparative static 2 does). Thus, regulators should take into account that while enforcing interoperability, may raise investment, it may also lead to lower ad frequency and thus worse aggregate product consumption.

Discussion

We wrap up this section with three important comments. First, the comparative statics with respect to ad frequency and data informativeness depend on the endogeneity of the product market. Second, they rely on the change in steady state consideration sets which takes place over time. As a result, they should be interpreted as comparisons of long run outcomes. Intuition suggests that these results might flip in the short run. Third, Propositions 1–3 show that parameter changes typically cause aggregate product consumption and investment to move in opposing directions. In general, parameter changes will have nonmonotone effects on welfare. To understand the impact of policies on welfare as a whole requires further analysis. In the next section, we investigate how the equilibrium compares to a first best benchmark. As a byproduct of the analysis, we will derive monotone comparative statics with respect to welfare under some conditions.

6. Stationary Equilibrium Versus First Best

This section compares the stationary equilibrium to a first best benchmark. The analysis is done for the special case when consumers' utilities are Cobb-Douglas. It turns out that investment by platforms may be either too high or too low relative to first best. We derive a simple condition that can be used to assess which is the case. We briefly discuss how the results extend to other utility functions later in the section.

Preliminaries

Assume that the flow utility of consumer i is

$$u(C_{it}, X_{it}) = C_{it}^{1-\tau} X_{it}^{\tau}$$

where $\tau \in [0, 1]$. We can give τ the following interpretation. Suppose that consumer i must devote attention in order to consume the products that are purchased.[33] Let $\tilde{\tau}$ denote the fraction of attention devoted to platform use with the residual $1 - \tilde{\tau}$ devoted to product consumption.[34] Suppose consumption of products and platforms occurs at a constant rate per unit of attention. If consumer i selects $\tilde{\tau}$ to maximize flow utility

$$\left[(1 - \tilde{\tau})C_{it}\right]^{1-\tau} \left[\tilde{\tau} X_{it}\right]^{\tau},$$

then we have $\tilde{\tau} = \tau$. The resulting flow utility is just a constant scaling of u. Thus, we can interpret τ as the fraction of attention spent on platforms. This extension with endogenous attention reduces to the baseline model except with $A\tau$ in place of A (since now only τ units of attention are spent viewing ads). For the rest of this section, we will go by this interpretation. Otherwise, one might suspect that the result that investment can be either too high or too low hinges on the assumption that consumers can not shift their attention away from platforms to do other things.

First Best Benchmark

The social planner has only one decision to make which is how much labor to allocate to investment versus production. The planner can not affect the ad frequency, data, or any other aspect of the matching technology.[35] This is without loss since first best investment will turn out to not depend on any of these properties.

[33] For example, suppose that one of the products is a guitar, purchased or rented at the "start of time t". To enjoy the guitar, consumer i must devote attention to playing it.

[34] Attention should be thought of as coming from leisure since we assume labor is supplied inelastically. In reality, not all products require leisure attention for consumption. For instance, wearing a pair of shoes. In Appendix C we extend the model to include two consumption sectors: one which requires leisure attention and one which does not. This comes at no cost to tractability and may be useful for future empirical work.

[35] It turns out that an equivalent interpretation is that the planner can only set the demand by platforms for labor with all other aspects of the economy determined as in equilibrium.

Since the planner seeks to maximize welfare, it is optimal to invest the same amount in each platform and to produce the same amount of each product. Using this fact, the planner's problem is

$$\max_{\{L_{\mathcal{D}t}\}_{t\geq 0}} \int_0^\infty e^{-rt} u(C_t, X_t)\, dt \tag{32}$$

such that

$$C_t = (L - DL_{\mathcal{D}t})(M_t \mu_{H_t})^{\frac{1}{\sigma-1}},$$
$$M_t, H_t \text{ solve (5)–(7)},$$
$$X_t = D^{\frac{1}{\rho-1}} q_t,$$
$$dq_t = (L_{\mathcal{D}t}^{\varphi} - \delta q_t)\, dt$$

given initial quality q_0.

We will only analyze the *steady state solution to the planner's problem* which is a constant rate of investment $L_{\mathcal{D}}^{fb}$ that solves (32) when M_0, H_0, and q_0 are at their steady state levels.

Lemma 6. *The steady state investment $L_{\mathcal{D}}^{fb}$ chosen by the planner is*

$$L_{\mathcal{D}}^{fb} = \frac{\varphi\delta\frac{\tau}{1-\tau}}{r + \delta + \varphi\delta\frac{\tau}{1-\tau}} \frac{L}{D}. \tag{33}$$

We derive Lemma 6 via Maximum Principle in Appendix C. There, we also show, under some mild technical conditions, that (33) is the solution to a version of the planner's problem when u is arbitrary provided the planner takes the cumulative value $M_t \mu_{H_t}$ as given. This is because the consumer equates the marginal utility for product consumption with that of platform consumption when selecting attention. The social planner therefore faces essentially the same trade-offs for any utility function. However, when u is arbitrary, it is with loss to take $M_t \mu_{H_t}$ as given. The planner can affect it with investment since the consumer's choice of attention $\tilde{\tau}$ will generally depend on platforms' quality levels and $\tilde{\tau}$ affects the rate at which the consumer views ads. We conjecture that if the planner were to internalize this, (33) would be a lower bound on first best investment.[36]

Comparison of Stationary Equilibrium and First Best

Comparing equilibrium investment (30) to first best (33) yields Theorem 2.

Theorem 2. *If, in a stationary equilibrium,*

$$\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}} A\tau(\rho-1) - \frac{\tau}{1-\tau} \begin{cases} < 0, & \textit{then investment is too low.} \\ = 0, & \textit{then investment is efficient.} \\ > 0, & \textit{then investment is too high.} \end{cases} \tag{34}$$

Given τ, the deviation $L_{\mathcal{D}} - L_{\mathcal{D}}^{fb}$ is increasing in $\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}} A\tau(\rho-1) - \frac{\tau}{1-\tau}$.

[36]This conjecture assumes u is submodular so that the total attention spent on platforms is increasing in platform quality.

Equilibrium investment is generally inefficient because platforms' investment incentives are based on how much ad revenue they can steal from rivals and not on the surplus that they generate for consumers. They do not appropriate any of this surplus since their services are free. By Theorem 3, it suffices to use only the expression in (34) to assess whether policies that affect data informativeness, interoperability, or ad frequency will bring investment closer to or farther from first best.

The first term in (34) is $\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}}A\tau(\rho-1)$ which comes from equilibrium investment. $\frac{\sigma}{\sigma-1}$ is the markup. When it is higher, firms earn higher revenues and so income is higher which raises ad revenue. $\hat{\pi}_{\mathcal{D}}A\tau$ is the ad revenue per unit of income. It is related to the strength of firms' business-stealing incentives. Interestingly, platforms appropriate ad revenues from firms even though they do not generate any surplus for them: advertising only shifts demand among firms. Firms individually value advertising because it allows them to steal business from their rivals. These business-stealing incentives are actually lower when τ is higher. This is because the higher ad frequency $A\tau$ leads firms outside consideration sets to have lower expected values but increases the cumulative value of firms inside them. Thus the fraction of flow profits that a firm can steal when enters a consideration set decreases. In fact, recall from Proposition 3 that ad revenue is decreasing in ad frequency $A\tau$ (as is the ad revenue per unit of income $\hat{\pi}_{\mathcal{D}}A\tau$). Thus, when τ is higher and consumers value platform quality more, equilibrium investment is actually *lower* which contrasts sharply with first best. The parameter $\rho - 1$ appears because it is the elasticity of attention with respect to quality. When it is higher, platforms can steal attention from each other more easily. To sum up, equilibrium investment depends on the extent of firms' market power and the business-stealing externalities among firms and among platforms.

These are irrelevant to the planner who seeks to maximize welfare. The second term in (34) is $\frac{\tau}{1-\tau}$ which comes from first best investment. $\frac{\tau}{1-\tau}$ determines the ratio, for consumers, of the marginal benefit of investment resulting from higher quality platforms to the marginal cost resulting from lower product consumption.

One might wonder if equilibrium investment is always too high or always too low.

Corollary 2.1. *Equilibrium investment may be either too high or too low relative to first best depending on model parameters.*

Corollary 2.1 can be proven by inspecting condition (34) in Theorem 2. As τ tends to 1, $\frac{\tau}{1-\tau}$ tends to ∞ while $\hat{\pi}_{\mathcal{D}}A\tau$ converges to a finite number. Thus, investment is too low. On the other hand, when τ tends to 0, $\frac{\tau}{1-\tau}$ tends to 0 while $\hat{\pi}_{\mathcal{D}}A\tau$ increases. Thus, investment is too high.

In either case, a tax or subsidy on auction revenues can bring equilibrium investment to first best.

Corollary 2.2. *A proportional tax or subsidy on auction revenues that is financed by or redistributed to consumers and is equal to*

$$\frac{1}{1-\tau}\frac{\sigma-1}{\sigma}\frac{1}{\hat{\pi}_{\mathcal{D}}A}\frac{1}{\rho-1}$$

achieves first best.

In the absence of a tax or subsidy, there are other policies that can bring investment closer to first best.

Corollary 2.3. *The following comparative statics hold in stationary equilibrium:*

1. *An increase in interoperability ρ causes investment to be farther from (closer to) first best when it is too high (too low).*

2. *An increase in ad frequency A causes investment to be closer to (farther from) first best when it is too high (too low).*

3. *An increase in data informativeness G can cause investment to be closer to (farther from) first best when it is too high (too low).*

We prove Corollary 2.3 by combining Theorem 2 with the earlier Propositions 1–3. Since changes to ρ, A, or G will also affect either the cumulative value $M\mu_H$ or the gains to variety $D^{\frac{1}{\rho-1}}$, the above comparative statics are not equivalent to comparative statics on welfare. However, there are cases when a parameter change both leads investment to be closer to first best and increases match value or gains to variety. In these cases we can infer the impact on welfare.

Corollary 2.4. *Suppose that welfare under the social planner exceeds that of stationary equilibrium. Then the following comparative statics hold in stationary equlibrium:*

1. *If investment is too high, banning all data so that G is uninformative reduces welfare.*

2. *If investment is too high, an increase in ad frequency A raises welfare.*

3. *If investment is too low and $D < 1$, an increase in interoperability ρ raises welfare.*

Corollary 2.4 is almost immediate since welfare is a single-peaked function of investment $L_{\mathcal{D}}$. Note that it is necessary to assume that welfare under the social planner is higher than in stationary equilibrium. This is because the social planner's economy starts with different initial conditions than those of stationary equilibrium. In terms of primitives, welfare under the planner is always higher provided the discount rate r is sufficiently low. Corollary 2.4 concludes the analysis of the baseline model. It provides some conditions when the welfare impacts of various policy-relevant parameter changes can be signed analytically. These results account for the welfare impact of platforms from both matching firms with consumers and providing quality services. The rest of this paper explores several important extensions of the baseline model.

7. Platforms with Different Data

So far, we have assumed that all platforms have the same data, except for possibly a measure zero set. This section extends the model to allow for two groups of platforms, each of positive measure, which have distinct data. This extension is especially relevant because policies that affect data informativeness will impact platforms which are inherently data-rich differently from those which are not. For instance, in 2024, Google will voluntarily remove third-party cookies on its Chrome web browser, an event referred to as "Chrome-ageddon" by many in the digital advertising industry. Some argue that this is in Google's own interests because it is less reliant on cookies for its ad business than other platforms such as Facebook. We will give a formal basis for these concerns.

The main prediction of this section is that platforms with more informative data will have larger market shares, higher ad prices, and also invest more. This is not very surprising. The contribution is to show that it can be captured by a formal model. Doing so opens the door to some interesting questions: How much of a *competitive* advantage does a platform with better data have? What will happen if platforms are forced to share their data? Because of the model's generality, though it may be possible, with difficulty, to obtain analytical results, it seems sensible to explore these questions numerically. As a result, the qualitative results of this section are illustrated by a numerical example.

Setup

There are now two groups of platforms $l \in \{1, 2\}$. The measure of platforms in group l is m_l. As before, v_{ij} denotes consumer i's value for firm j's product. Firm j receives a signal ζ_{lij} which may be informative of v_{ij} when it bids on a platform in group l. We assume that $(v_{ij}, \zeta_{1ij}, \zeta_{2ij}) \sim P$ on $\mathbb{R}_+ \times \mathbb{R}^2$ independently across i and j and that v_{ij} has a finite mean. Let G denote the joint CDF of the signals ζ_{1ij}, ζ_{2ij}, derived from P. We assume that G has a continuous density g.

To ensure that stationary equilibrium bidding strategies are monotone in signal realizations, we assume the following stochastic monotonicity condition.

Condition 3. *P is such that*[37]

1. *the conditional distribution of ζ_{-lij} given ζ_{lij} increases in $\succ_{FOSD}$ when ζ_{lij} increases for each $l \in \{1, 2\}$.*
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}.*

Thus, a high realization of either signal is always good news to a firm. The rest of the setup is as in the baseline model of Section 3. We continue to assume Conditions 1 and 2 throughout.

Computing a Stationary Equilibrium

We now give an informal sketch of the procedure to compute a stationary equilibrium. Formal statements of results and omitted proofs are in Appendix D. We start by observing that consumers' demands as well as firms' prices and flow profits

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

remain as in Lemmas 1 and 2 of the baseline model. The evolution of the size of consideration sets M_t also remains as in (5) with a steady state level $M = A/\lambda_f$.

Let h_t denote the PDF of signals in consideration sets. Also, let H_t^c denote the CDF and h_t^c the PDF of signals outside of consideration sets. Since the bidder with the highest group l signal wins in each group l auction, it follows that

(35)
$$d\big[M_t h_t(\boldsymbol{\zeta})\big] \;=\; A\left[x_{1t} N H_t^c(\zeta_1, \infty)^{N-1} h_t^c(\boldsymbol{\zeta}) + x_{2t} N H_t^c(\infty, \zeta_2)^{N-1} h_t^c(\boldsymbol{\zeta}) - h_t(\boldsymbol{\zeta})\right] dt$$

at each $\boldsymbol{\zeta} = (\zeta_1, \zeta_2) \in \mathbb{R}^2$, by the ELLN. Above, with abuse of notation, x_{lt} denotes the total attention share of group l platforms. The first two terms in brackets sum to equal the distribution of the inflowing signals. A fraction x_{lt} of this inflow is from the winners in the group l auctions. The last term in brackets is the distribution of the outflowing firms which are forgotten uniformly at random. To derive the steady state h, first fix an initial guess of x_1, the stationary equilibrium level of x_{1t}. Then set $x_{1t} = x_1$, $x_{2t} = 1 - x_1$, $M_t = M$ and use the accounting identing $M_t h_t + (F - M_t) h_t^c = Fg$ to iterate (35) forward to convergence at each point $\boldsymbol{\zeta}$ in a fine grid on a region that contains almost all of G's mass.

With M and h in hand, we next compute the bidding strategies. First, some notation. As in the baseline model, let

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

be the average value of firms in consideration sets and let

$$\pi_{\mathcal{F}} = \frac{I}{\sigma M \mu_H}$$

be the coefficient on firms' flow profits. Note that we have not yet computed income I. Finally, let

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

and

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

be the probabilities that firm j wins an auction if it takes place on a platform in group 1 and in group 2, respectively.

In a stationary equilibrium, bidding strategies are determined by a pair of functions $\boldsymbol{B} = (B_1, B_2)$. $B_l : \mathbb{R} \Rightarrow \mathbb{R}_+$ maps firm j's group l signal ζ_{lij} to its bid $B_l(\zeta_{lij})$ in a group l 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 for consumer i at the time of auction entry if it knows $\boldsymbol{\zeta}_{ij}$ but does not know which platform hosts the auction. Then, analogous to (20),

(36)
$$B_l(\zeta_{lij}) = \mathbb{E}\left[\frac{\pi_{\mathcal{F}}}{\lambda_f + r} v_{ij} - \frac{r}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r} V(\boldsymbol{\zeta}_{ij}) \Big| \zeta_{lij}, j \in \Omega_{it}^c\right]$$

which is the gain in continuation value from winning the group l auction. As before, $\lambda_a = NA/(F-M)$ is the rate of auction entry. The expectation is conditional on only the group l signal and the fact that the firm is outside the consideration set since this is all that the firm knows when it bids.

Given $\boldsymbol{B}$, V must satisfy,

$$V(\boldsymbol{\zeta}_{ij}) = \sum_{l=1}^{2} x_l O_l(\zeta_{lij}) \left[\frac{\pi_{\mathcal{F}}}{\lambda_f + r}\mathbb{E}[v_{ij}|\boldsymbol{\zeta}_{ij}] + \frac{\lambda_f}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r}V^*(\boldsymbol{\zeta}_{ij})\right] - x_l\left[\int_{-\infty}^{\zeta_{lij}} B_l(s)\,dO_l(s) + \left(1 - O_l(\zeta_{lij})\right)\frac{\lambda_a}{\lambda_a + r}V(\boldsymbol{\zeta}_{ij})\right] \tag{37}$$

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 l auction. The last term is the continuation value in the event that firm j loses the auction, weighted by the probability that this happens.

Using (36) and (37), we can show that $\boldsymbol{B}$ is the fixed point of an operator $\boldsymbol{\Lambda} := (\Lambda_1, \Lambda_2)$. $\Lambda_l : C^+(\mathbb{R})^2 \Rightarrow C^+(\mathbb{R})$ takes in a pair of functions $\boldsymbol{f} = (f_1, f_2)$ and outputs

$$\Lambda_l(\boldsymbol{f})(\cdot) = \mathbb{E}\left[\frac{\pi_{\mathcal{F}} v_{ij} + \lambda_a \sum_{z=1}^{2} x_z \int_{-\infty}^{\zeta_{zij}} f_z(s)\,dO_z(s)}{\lambda_f + r + \lambda_a \sum_{z=1}^{2} x_z O_z(\zeta_{zij})}\middle|\zeta_{lij} = \cdot, j \in \Omega_{it}^c\right]. \tag{38}$$ [38]

In Appendix D, we show that $\boldsymbol{\Lambda}$ is a contraction with modulus $\lambda_a/(\lambda_a + \lambda_f + r)$. If we knew income I (which appears in $\pi_{\mathcal{F}}$), we could iterate $\boldsymbol{\Lambda}$ starting from some initial guess to derive $\boldsymbol{B}$. By Lemma 13 in Appendix D, the resulting bidding functions would be increasing as needed for (35) to apply. Though we do not know income, the model is homothetic in it. We therefore proceed by first computing the bid functions *per unit of income* by iterating $\boldsymbol{\Lambda}$ with $I = 1$. From these, we obtain the platforms' average ad prices per-unit of income $\hat{\pi}_{\mathcal{D}l}$, $l \in \{1,2\}$. Later, we compute the actual income I in terms of $\hat{\pi}_{\mathcal{D}l}$, $l \in \{1,2\}$ and from there, the average ad prices $\pi_{\mathcal{D}l} = I\hat{\pi}_{\mathcal{D}l}$, $l \in \{1,2\}$.

Given $\pi_{\mathcal{D}l}$, each platform in group l solves the analog of (11) (with $\pi_{\mathcal{D}l}$ in place of $\pi_{\mathcal{D}}$) by investing at a constant rate $L_{\mathcal{D}l}$. In Appendix D, we derive that

$$L_{\mathcal{D}l} = \frac{\varphi\delta\pi_{\mathcal{D}l}A(\rho - 1)}{r + \delta}\frac{1}{m_l + m_{-l}\left(\frac{\pi_{\mathcal{D}-l}}{\pi_{\mathcal{D}l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}} \tag{39}$$

using the Maximum Principle. Together with the condition that income equals the total of firms' revenues $I = \frac{\sigma}{\sigma-1}(L - \sum_{l=1}^{2} m_l L_{\mathcal{D}l})$, (39) implies

$$I = L\left[1 + \frac{\varphi\delta A(\rho - 1)}{r + \delta}\frac{m_l\hat{\pi}_{\mathcal{D}l} + m_{-l}\left(\frac{\hat{\pi}_{\mathcal{D}-l}}{\hat{\pi}_{\mathcal{D}l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}\hat{\pi}_{\mathcal{D}-l}}{m_l + m_{-l}\left(\frac{\hat{\pi}_{\mathcal{D}-l}}{\hat{\pi}_{\mathcal{D}l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}}\right]^{-1}. \tag{40}$$

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

From (39) we also have the quality level

$$q_l = \frac{L_{\mathcal{D}l}^{\varphi}}{\delta}$$

and attention share

$$x_l = \frac{m_l q_l^{\rho-1}}{m_l q_l^{\rho-1} + m_{-l} q_{-l}^{\rho-1}} = \frac{m_l \left(\frac{\hat{\pi}_{\mathcal{D}l}}{\hat{\pi}_{\mathcal{D}-l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}}{m_l \left(\frac{\hat{\pi}_{\mathcal{D}l}}{\hat{\pi}_{\mathcal{D}-l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}} + m_{-l}} \tag{41}$$

of group l platforms. (41) gives x_1 as a function of $\hat{\pi}_{\mathcal{D}1}/\hat{\pi}_{\mathcal{D}2}$. Recall that the latter is itself a function of the initial guess of x_1. To finish the equilibrium computation, we adjust the initial guess until (41) is satisfied. While we have not proven the uniqueness of an x_1 such that this holds, it is simple to check numerically. If platforms in group 1 have more informative data, then $\hat{\pi}_{\mathcal{D}1}/\hat{\pi}_{\mathcal{D}2}$ is typically decreasing in the initial guess of x_1 and so the stationary equilibrium is unique as in the numerical example we consider later.

From here, it is immediate to compute welfare $u(C, X)/r$ using

$$C = \left(L - \sum_{l=1}^{2} m_l L_{\mathcal{D}l}\right) (M\mu_H)^{\frac{1}{\sigma-1}}$$

and

$$X = \left(\sum_{l=1}^{2} m_l q_l^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}.$$

In summary, the computational procedure is as follows:

1. Guess a value of x_1.
2. Iterate (35) forward to compute h.
3. Iterate (38) to compute per-unit income bid functions and average ad prices.
4. Check whether the guess of x_1 aligns with (41).
5. If yes, done. If not, repeat with a revised guess.

All other equilibrium quantities 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* stationary equilibria and in particular, check uniqueness.

Remark. In Appendix D, we analyze an especially tractable special case where group 2 platforms have uninformative data and may therefore better represent traditional media than digital media.[39] As in the baseline model, H can be computed

[39]Strictly speaking, G does not have a density in this case but essentially the same analysis goes through. The only difference is that firms may tie in an auction with positive probability. We assume uniform tie breaking.

pointwise by solving for the unique nonnegative root of a polynomial. We are also able to compute B_1 and B_2 as explicit functions of H and x_1.

Numerical Example

In this example, values are log-normal: each $v_{ij} = e^{Z_{ij}}$ where $Z_{ij} \sim N(0, .5)$. The signal in group 1 is perfectly informative: $\zeta_{1ij} = Z_{ij}$. The signal in group 2 is noisy: $\zeta_{2ij} = Z_{ij} + \epsilon_{ij}$ where $\epsilon_{ij} \sim N(0, 2)$. This implicitly defines the distribution P. All other parameters are as in Table 1.

Table 1: Parameter Values

Description	Parameter	Value
measure of group 1 platforms	m_1	.05
measure of group 2 platforms	m_2	.05
ad frequency	A	.01
forget rate	λ_f	1
measure of firms	F	.1
number of bidders	N	20
investment technology	φ	.75
quality depreciation rate	δ	.1
product substitutability	σ	3
platform interoperability	ρ	1.6
labor supply	L	1

We now describe the numerically computed stationary equilibrium for these parameters. Figure 7 plots the relationships that determine group 1's attention share x_1 and the ratio of the ad prices $\pi_{\mathcal{D}1}/\pi_{\mathcal{D}2}$ in equilibrium.

The red curve plots (41) which expresses x_1 explicitly as a function of $\pi_{\mathcal{D}1}/\pi_{\mathcal{D}2}$. The black curve plots the computed value of $\pi_{\mathcal{D}1}/\pi_{\mathcal{D}2}$ as a function of the initial guess of x_1. Since these curves intersect only once, the stationary equilibrium is unique.

For the black curve, $\pi_{\mathcal{D}1}/\pi_{\mathcal{D}2}$ declines from around 1.51 to around 1.43 as x_1 increases from 0 to 1. This is for the most part because the distribution of group 1 signals outside of consideration sets worsens because of the greater positive selection into consideration sets. As a result, bidders in group 1 auctions are more likely to have lower group 1 signals. On the flip side, since x_2 is lower, there is less positive selection of group 2 signals into consideration sets. As a result, bidders in group 2 are more likely to have *higher* signals. Overall, $\pi_{\mathcal{D}1}/\pi_{\mathcal{D}2}$ decreases.

The red curve is increasing because group 1's investment is higher relative to group 2's investment when $\pi_{\mathcal{D}1}/\pi_{\mathcal{D}2}$ is higher. From the intersection of the two curves we see that, in stationary equilibrium, group 1 platforms have around 57.5 percent of the attention and an average ad price that is around 1.44 times as high as that of group 2 platforms.

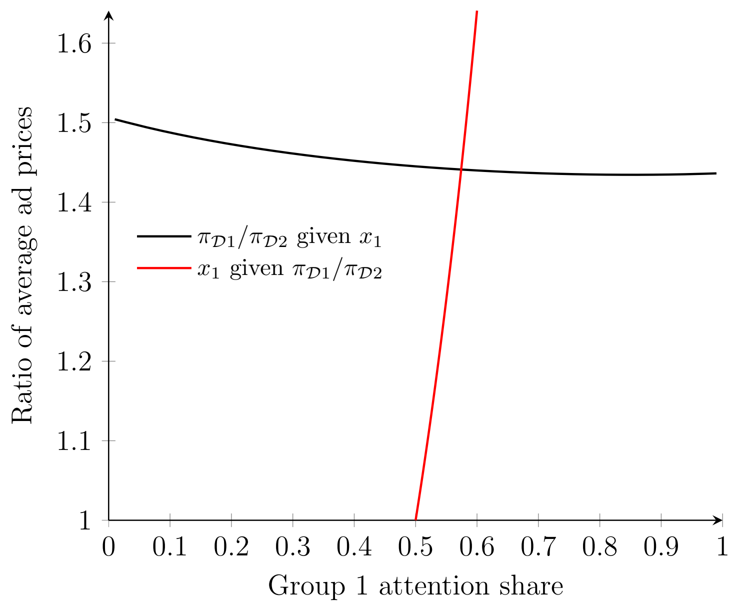

Figure 7: Determination of the unique stationary equilibrium

By inspecting (41), we see that the red curve increases at a slower rate when the investment technology φ or interoperability ρ is higher. The black line, however, does not depend on either φ or ρ. This yields the following comparative static: an increase in φ or ρ leads to to a decrease in x_1 and an increase in $\pi_{\mathcal{D}1}/\pi_{\mathcal{D}2}$. Thus, having better data is more advantageous when it comes to ad prices but less advantageous when it comes to attention share when interoperability is higher or the investment technology is better.

Below, Figure 8 plots the bidding functions of the two groups. We plot them for the interval -2 to 2. This is around 2.8 standard deviations above and below the mean for group 1 signals and around 1.3 standard deviations above and below the mean for group 2 signals.[40] As one might expect, for low signal realizations, the group 1 bids are below the group 2 bids while for the high realizations the opposite is true. This is because group 1 signals are more informative and so bids should be more sensitive to group 1 signals.

Table 2 reports the remaining equilibrium properties of interest in column 3. Column 4 reports their counterparts in the case when all platforms have fully informative data (which fits into the baseline model setting). Going from column 3 to column 4 captures the effects of a policy that forces platforms to share data. Alternatively, if we start in the regime when all platforms have fully informative data, going from column 4 to column 3 captures the effects of a policy that degrades the data of only some platforms.

When platforms are forced to share data, product consumption improves from around .157 to around .168.[41] This is both because the average value in consideration

[40] We numerically computed the bidding functions for signal realizations out to 5 standard deviations from the mean.

[41] The expectation of the value v_{ij} under the prior is approximately 1.284.

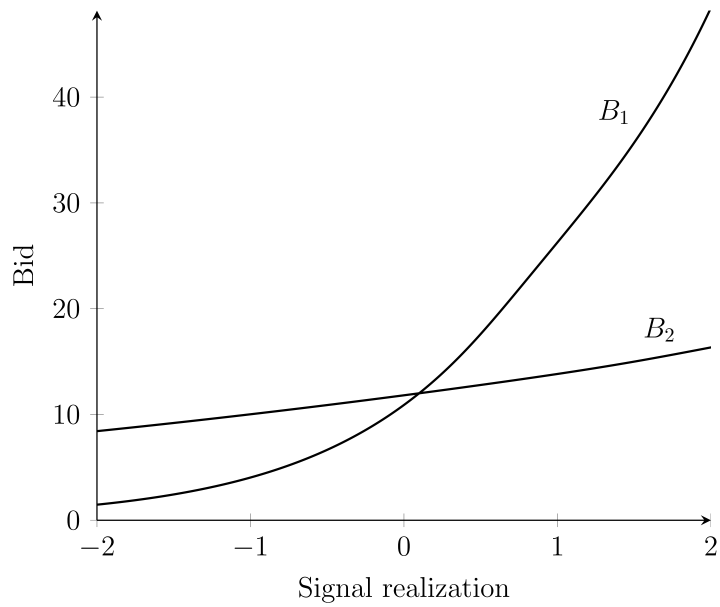

Figure 8: Bidding functions

Table 2: Separate Data Versus Shared Data

Description	Variable	Separate Data	Shared Data
agg. product consumption	C	.157	.168
agg. platform consumption	X	.0115	.0112
avg. value in consideration sets	μ_H	2.687	3.084
group 1 investment	$L_{\mathcal{D}1}$	.600	.422
group 2 investment	$L_{\mathcal{D}2}$	.310	.422
group 1 avg. ad price	$\pi_{\mathcal{D}1}$	23.129	18.783
group 2 avg. ad price	$\pi_{\mathcal{D}2}$	16.050	18.783

sets increases and because total investment declines. Though group 1 investment declines, group 2 investment increases. Overall, platform consumption decreases from around .0115 to .0112, a small amount. This is so even though consumers prefer a more even distribution of platform quality and the investment technology has decreasing returns. Lastly, we see that the group 1 ad price declines while that of group 2 increases. Thus, data informativeness has *competitive* value, which the model can quantify, as we had highlighted already in Section 5.

8. Discussion of Model Extensions

This section summarizes the results of some further extensions of the baseline model of Section 3. The model is able to tractably accommodate several additional actors that may be of first-order relevance. As described below, many of the main results continue to hold in each of these extensions.

Ad Frequency

In the first extension, described already at the end of Section 5, each platform sets its own ad frequency in light of nuisance costs that consumers incur from viewing ads. In the stationary equilibrium, each platform sets the same ad frequency. In Appendix E, we show that the ad frequency may be either too high or too low relative to what is socially efficient as in Anderson and Coate (2005). Inefficiency may arise for two reasons. First, platforms do not internalize the beneficial impact of ads on consumers' consideration sets. This is a force that may lead ad frequency to be too low. Second, they only internalize consumers' nuisance costs to the extent that they can steal business from each other by reducing ad frequency. This is a force that may lead ad frequency to be too low when interoperability is low or too high when interoperability is high. All of the results of the baseline model continue to hold in this setting with the exception of Part 2 of Proposition 2 which states that an increase in interoperability leads to a decline in ad revenues. The effect of interoperability on ad revenues is now ambiguous.

Network Effects

In a second extension, found in Appendix F, a consumer's utility from using a platform may depend on the amount of attention spent by the other consumers on that platform. The effective quality of platform k is now $\eta(x_{kt})q_{kt}$ where $x_{kt} = \int_{\mathcal{C}} x_{ikt} di$ and $\eta : [0,1] \Rightarrow \mathbb{R}_+$ is increasing. Then, analogous to (13), the attention that platform k recieves must satisfy

$$x_{kt} = \frac{[\eta(x_{kt})q_{kt}]^{\rho-1}}{\int_{\mathcal{D}} [\eta(x_{lt})q_{lt}]^{\rho-1}\, dl}.$$

Suppose that $\eta(x) = x^{\zeta}$ where $\zeta > 0$ parameterizes the strength of the network effects. Under this assumption, we can solve explicitly for $x_{kt}, k \in \mathcal{D}$. There is a continuum of solutions. For each $\mathcal{E}_t \subset \mathcal{D}$ of positive measure, there is a solution that sets

$$x_{kt} = \frac{q_{kt}^{\frac{\rho-1}{1-\zeta(\rho-1)}}}{\int_{\mathcal{E}_t} q_{lt}^{\frac{\rho-1}{1-\zeta(\rho-1)}}\, dl}$$

if $k \in \mathcal{E}_t$ and otherwise sets $x_{kt} = 0$. This expression is the same as in (13) of the baseline model except for a change in the exponent. Compared with (13), attention is more elastic when there are network effects. Thus platforms' business-stealing incentives are greater.

Under the refinement that all platforms in $\mathcal{D}$ receive a positive share of attention at all points in time (that is, $\mathcal{E}_t = \mathcal{D}$ at each t), there is a unique stationary equilibrium provided that ζ is not too high. Here, $\mathcal{D}$ should be interpreted as the set of platforms that remain active with stable market shares in the long run. $\mathcal{D}$ is taken as a primitive—the model has no predictions for which set of platforms will prevail. If ζ is sufficiently small, one can apply an analogous analysis to that of Section 6 to solve the planner's problem and derive an alternative condition to the one in Theorem 2.

Reserve Prices

In a third extension, found in Appendix G, we allow platforms to set reserve prices. In stationary equilibrium, though there is a continuum of platforms, reserve prices are generally positive because ad opportunities are only partially substitutable since they occur at different points in time. Reserve prices can be inefficient since they reduce the effective rate that consumers see ads leading to worse product consumption. The comparative statics in Section 5 with respect to data informativeness continue to hold as does Theorem 2 in Section 6 which compares equilibrium investment to first best. The comparative statics with respect to interoperability also hold as long as there is a unique equilibrium.

Entry

In a final extension, found in Appendix H, we endogenize entry of platforms and firms. Entry incurs fixed labor costs. In the unique stationary equilibrium, the measures of platforms and firms are such that they earn zero profits. A decrease in the entry cost for firms can drive entry of both firms and platforms. An improvement in data informativeness typically leads to entry of firms and exit of platforms. While the welfare analysis of Section 6 does not apply, all of the comparative statics of Section 5 continue to hold except for Parts 1 and 2 of Proposition 2 which states that an increase in interoperability leads to a decline in product consumption and ad revenue. However, if the ad frequency is also endogenous as in the first extension, then Part 1 is restored while Part 2 is reversed: an increase in interoperability leads to a decline in product consumption but an increase in ad revenue.

9. Conclusions

The goal of this paper is to contribute to our understanding of the complex market for attention. We approach this goal by tackling the complexity head-on in a general equilibrium model that can capture the different sides of the market at once. We are therefore able to analyze the welfare impact of platforms from both the matching of firms to consumers in the product market and the provision of quality services. The rationale for this is that a model that studies one or the other in isolation can not address some of the *fundamental* tradeoffs that regulators must contend with. Let us review some of our findings. We have shown that an increase in data informativeness or a decline in platform interoperability typically leads to an improvement in product consumption but at the expense of lower investment and thus worse platform consumption. Equilibrium investment may be either too high or too low depending on the extent of firms' market power and business-stealing externalities among firms and among platforms. We have shown how empirically relevant phenomena such as variation in ad prices, bid pacing, and delay in the matching of a firm to a consumer arise in equilibrium and relate to platforms' market power. These phenomena figure into real-world bidding algorithms in ad auctions and may be relevant to future empirical work. We have also shown that platforms with more informative data typically have higher market shares, ad prices, and investment. Forcing platforms to share their data can improve product consumption but may come at the expense of worse platform consumption.

A. Proofs for Section 4

Proof of Lemma 1. Part 1 of Lemma 1 is derived by formulating the Lagrangian and taking a first order condition as in Dixit and Stiglitz (1977). We omit the details. Part 2 is a special case of Part 1. □

Proof of Lemma 2. Part 1 follows from Part 1 of Lemma 1. Given this, Part 2 is immediate. □

Proof of Lemma 3. The solution to (5) has

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

and thus $M_t \to M$ as $t \to \infty$. Using (7), we rewrite (6) in terms of only H_t^c, M_t:

$$(F - M_t)\, dH_t^c = \lambda_f \left[M_t (H_t^c)^N + (F - M_t) H_t^c - FG\right] dt.$$

If $H_0^c \neq H^c$, then

$$\int_{H_0^c}^{H_t^c} \frac{du}{\lambda_f \left[FG - M_u u^N - (F - M_u) u\right]} = \int_0^t \frac{du}{F - M_u}.$$

The LHS is strictly monotone in H_t^c which must lie between H_0^c and H^c. It follows that there is a unique H_t^c satisfying the above equation. Since the RHS is increasing in t, H_t^c is either increasing or decreasing in t and thus converges as $t \to \infty$. Since $M_t \to M < F$, the RHS grows without bound. The only way for equality to hold is if $H_t^c \to H$. □

Note that in the proof of Lemma 3 we have characterized the paths of M_t, H_t, and H_t^c analytically. For this characterization to apply to equilibrium, the bidding function B_t must be increasing at all times t. This will follow from Lemmas 7 and 8 which we now prove. We will in fact analytically compute the entire path of bidding functions taking as given the path of income.

Lemma 7. *In an equilibrium, the expected flow profit of a firm $j \in \Omega_{it}$ from selling to consumer i at time t is*

$$\pi_{\mathcal{F}t} \hat{v}_{ij}$$

where

$$\pi_{\mathcal{F}t} = \frac{I_t}{\sigma M_t \mu_{H_t}}.$$

Proof. By Lemma 2, the flow profit is

$$\frac{I_t v_{ij}}{\sigma \int_{\Omega_{it}} v_{ij}\, dj}.$$

By the ELLN, the integral in the denominator is equal to $M_t \mu_{H_t}$ since firms' expectations are unbiased. □

Lemma 8. *In an equilibrium, the bidding function B_t is increasing at all times t. Moreover, B_t satisfies*

$$B_t'(\hat{v}) = \int_t^\infty \pi_{\mathcal{F}s} e^{-\int_t^s [r+\lambda_f+\lambda_a H_z^c(\hat{v})^{N-1}]\, dz}\, ds, \quad \hat{v} \in \mathbb{R}_+$$

where the coefficients $(\pi_{\mathcal{F}s})_{s=0}^\infty$ are as in Lemma 7.

Proof. Let τ_a, τ_f, and $\hat{\tau}_a$ denote independent exponentially distributed random variables with rates λ_a, λ_f, and λ_a respectively. Suppose that a firm enters an auction for a consumer at time t whose value for its product is in expectation $\hat{v}$. Let $V_t^W(\hat{v})$ denote the firm's continuation value conditional on winning the auction (gross of the expected payment). Let $V_t^L(\hat{v})$ denote the firm's continuation value conditional on losing the auction. Let $V_t(\hat{v})$ denote the firm's continuation value at the time of auction entry but prior to knowing whether it has won or lost the auction.

It follows that

$$\begin{aligned}(42)\quad V_t^W(\hat{v}) = \mathbb{E}\left[\int_t^{t+\tau_a} e^{-rs}\pi_{\mathcal{F}s}\hat{v}\, ds + e^{-r(t+\tau_a)} V_{t+\tau_a}^W \middle| \tau_f > \tau_a\right] \mathbb{P}\{\tau_f > \tau_a\} \\ + \mathbb{E}\left[\int_t^{t+\tau_f} e^{-rs}\pi_{\mathcal{F}s}\hat{v}\, ds + e^{-r(t+\tau_f+\hat{\tau}_a)} V_{t+\tau_f+\hat{\tau}_a} \middle| \tau_f < \tau_a\right] \mathbb{P}\{\tau_f < \tau_a\}.\end{aligned}$$

Similarly,

$$\begin{aligned}(43)\quad V_t^L(\hat{v}) = \mathbb{E}\left[e^{-r(t+\tau_a)} V_{t+\tau_a}(\hat{v}) \middle| \tau_f > \tau_a\right] \mathbb{P}\{\tau_f > \tau_a\} \\ + \mathbb{E}\left[e^{-r(t+\tau_a)} V_{t+\tau_a}(\hat{v}) \middle| \tau_f < \tau_a\right] \mathbb{P}\{\tau_f < \tau_a\}.\end{aligned}$$

It will become apparent later as to why we decompose the value functions in this way.

Moving on, let W_t denote the distribution of the highest among the $N-1$ other bids in an auction at time t. That is, $W_t(B_t(\hat{v}))$ is the probability that a firm with expectation $\hat{v}$ wins an auction at time t. To ease notation, let $\tilde{W}_t(\hat{v}) = W_t(B_t(\hat{v}))$. Also, let $B_t^{(1)} \sim W_t$. Then we can rewrite (43) as follows

(44)

$$\begin{aligned}V_t^L(\hat{v}) = &\ \mathbb{E}\left[e^{-r(t+\tau_a)} \tilde{W}_{t+\tau_a}(\hat{v}) V_{t+\tau_a}^W(\hat{v}) | \tau_f > \tau_a\right] \mathbb{P}\{\tau_f > \tau_a\} \\ &- \mathbb{E}\left[e^{-r(t+\tau_a)} \tilde{W}_{t+\tau_a}(\hat{v}) \mathbb{E}\left[B_{t+\tau_a}^{(1)} \middle| \tau_a, B_{t+\tau_a}(\hat{v}) > B_{t+\tau_a}^{(1)}\right] \middle| \tau_f > \tau_a\right] \mathbb{P}\{\tau_f > \tau_a\} \\ &+ \mathbb{E}\left[e^{-r(t+\tau_a)} \left(1 - \tilde{W}_{t+\tau_a}(\hat{v})\right) V_{t+\tau_a}^L(\hat{v}) | \tau_f > \tau_a\right] \mathbb{P}\{\tau_f > \tau_a\} \\ &+ \mathbb{E}\left[e^{-r(t+\tau_f+\hat{\tau}_a)} V_{t+\tau_f+\hat{\tau}_a}(\hat{v}) \middle| \tau_f < \tau_a\right] \mathbb{P}\{\tau_f < \tau_a\}.\end{aligned}$$

The second term on the RHS is the expected payment in the auction. Together, the first three terms on the RHS in (44) correspond to the first term in (43). The last term on the RHS in (44) corresponds to the second term in (43) by the memorylessness of the exponential distribution.

Since firms bid optimally in equilibrium, $B_t = V_t^W - V_t^L$ at all t. Subtracting (44) from (42) yields

(45)

$$\begin{aligned}
B_t(\hat{v}) = \mathbb{E}&\left[\int_t^{t+\tau_a} e^{-rs}\pi_{\mathcal{F}s}\hat{v}\,ds\middle|\tau_f > \tau_a\right]\mathbb{P}\{\tau_f > \tau_a\} \\
&+ \mathbb{E}\left[e^{-r(t+\tau_a)}\tilde{W}_{t+\tau_a}(\hat{v})\mathbb{E}\left[B^{(1)}_{t+\tau_a}\middle|\tau_a, B_{t+\tau_a}(\hat{v}) > B^{(1)}_{t+\tau_a}\right]\middle|\tau_f > \tau_a\right]\mathbb{P}\{\tau_f > \tau_a\} \\
&+ \mathbb{E}\left[e^{-r(t+\tau_a)}\left(1 - \tilde{W}_{t+\tau_a}(\hat{v})\right)B_{t+\tau_a}(\hat{v})\middle|\tau_f > \tau_a\right]\mathbb{P}\{\tau_f > \tau_a\} \\
&+ \mathbb{E}\left[\int_t^{t+\tau_f} e^{-rs}\pi_{\mathcal{F}s}\hat{v}\,ds\middle|\tau_f < \tau_a\right]\mathbb{P}\{\tau_f < \tau_a\}.
\end{aligned}$$

Writing out the expectations, differentiating both sides with respect to $\hat{v}$, and combining some like terms yields

$$\begin{aligned}
(46)\quad e^{-(r+\lambda_a+\lambda_f)t}B_t'(\hat{v}) = (\lambda_a+\lambda_f)&\int_t^\infty \int_t^y e^{-rs}\pi_{\mathcal{F}s}\,ds\,e^{-(\lambda_a+\lambda_f)y}dy \\
&+ \lambda_a\int_t^\infty B_y'(\hat{v})\big(1-\tilde{W}_y(\hat{v})\big)e^{-(r+\lambda_a+\lambda_f)y}\,dy.
\end{aligned}$$

Note that, after differentiating, only the derivatives of the bidding functions remain (though recall that $\tilde{W}_t$ implicitly depends on B_t). Using this convenient property, we will simplify the derivation of bidding strategies down to a single differential equation.

First, we simplify the double integral in (46) using integration by parts:

$$(\lambda_a+\lambda_f)\int_t^\infty \int_t^y e^{-rs}\pi_{\mathcal{F}s}\,ds\,e^{-(\lambda_a+\lambda_f)y}\,dy = \int_t^\infty e^{-rs}\pi_{\mathcal{F}s}e^{-(r+\lambda_a+\lambda_f)s}\,ds.$$

Then

$$\begin{aligned}
e^{-(r+\lambda_a+\lambda_f)t}B_t'(\hat{v}) = &\int_t^\infty \pi_{\mathcal{F}s}e^{-(r+\lambda_a+\lambda_f)s}\,ds \\
&+ \lambda_a\int_t^\infty B_y'(\hat{v})(1-\tilde{W}_y(\hat{v}))e^{-(r+\lambda_a+\lambda_f)y}\,dy.
\end{aligned}$$

Let $f_t(\hat{v}) = e^{-(r+\lambda_a+\lambda_f)t}B_t'(\hat{v})$. Differentiating both sides of the above equation with respect to time yields

$$\dot{f}_t(\hat{v}) + \lambda_a f_t(\hat{v})\big(1-\tilde{W}_t(\hat{v})\big) = -\pi_{\mathcal{F}t}e^{-(r+\lambda_a+\lambda_f)t}$$

This is a first order linear differential equation and thus has an explicit solution:

$$f_t(\hat{v}) = e^{-\int_0^t \lambda_a(1-\tilde{W}_s(\hat{v}))\,ds}\left(f_0(\hat{v}) - \int_0^t \pi_{\mathcal{F}s}e^{-(r+\lambda_f)s-\int_0^s \lambda_a\tilde{W}_z(\hat{v})\,dz}\,ds\right).$$

Equivalently,

$$B_t'(\hat{v}) = e^{\int_0^t (r+\lambda_f+\lambda_a\tilde{W}_s(\hat{v}))\,ds}\left(B_0'(\hat{v}) - \int_0^t \pi_{\mathcal{F}s}e^{-(r+\lambda_f)s-\int_0^s \lambda_a\tilde{W}_z(\hat{v})\,dz}\,ds\right).$$

As $t \to \infty$ the integral in the parenthesis converges to a positive constant. If this constant is not $B_0'(\hat{v})$ then $B_t'(\hat{v})$ diverges. However, a bidding strategy with this property is never optimal. We omit the formal proof of this claim but it is trivial to see that the continuation value functions are uniformly Lipschitz in $\hat{v}$ and thus so are bidding functions. Using this fact, we arrive at

$$B_t'(\hat{v}) = \int_t^\infty \pi_{\mathcal{F}s} e^{-\int_t^z [r+\lambda_f+\lambda_a \tilde{W}_z(\hat{v})]\, dz}\, ds.$$

Since $B_t'(\hat{v})$ is always positive, it follows that $\tilde{W}_t(\hat{v}) = H_t^c(\hat{v})^{N-1}$ at all t. This completes the proof. □

Proof of Lemma 4. Part 1 is a special case of Lemma 8. To prove Part 2, recall that the Poisson rate that a firm enters an auction for consumer i is λ_a and the probability that it wins an auction is $H^c(\hat{v}_{ij})^{N-1}$. Part 3 follows from Part 1 and integration by parts. Part 4 follows from accounting. □

Proof of Lemma 5. To prove Part 1, suppose that $L_{\mathcal{D}}$ is a stationary equilibrium level of investment. The (present-value) Hamiltonian associated with platform k's problem (11) is

$$\mathcal{H}(q_{kt}, \lambda_t, L_{kt}) = \pi_{\mathcal{D}} A \frac{q_{kt}^{\rho-1}}{Dq^{\rho-1}} - L_{kt} + \lambda_t \left(L_{kt}^{\varphi} - \delta q_{kt}\right)$$

where the costate variable λ_t satisfies

$$r\lambda_t - \dot{\lambda}_t = \pi_{\mathcal{D}} A \left(\rho - 1\right) \frac{q_{kt}^{\rho-2}}{Dq^{\rho-1}} - \lambda_t \delta.$$

By the Maximum Principle, $L_{kt} = L_{\mathcal{D}}$ must maximize the Hamiltonian along the equilibrium trajectory. This yields:

$$\lambda_t \varphi L_{\mathcal{D}}^{\varphi-1} = 1.$$

Then λ_t is a constant and by the costate evolution satisfies

$$\lambda_t = \frac{\pi_{\mathcal{D}} A \left(\rho - 1\right)}{D(r+\delta)} \frac{1}{q}.$$

Substituting gives

$$\frac{\pi_{\mathcal{D}} A \left(\rho - 1\right)}{qD(r+\delta)} \varphi L_{\mathcal{D}}^{\varphi-1} = 1.$$

Since $q = L_{\mathcal{D}}^{\varphi}/\delta$, we have

$$L_{\mathcal{D}} = \frac{\varphi \delta \pi_{\mathcal{D}} A(\rho - 1)}{D\left(r+\delta\right)}.$$

This is a necessary condition for $L_{\mathcal{D}}$ to be a stationary equilibrium level of investment. □

To finish the proof of Theorem 1, we use the following lemma.

Lemma 9. *Let $L_{\mathcal{D}}$ and q be as in Lemma 5. Platform k solves its problem (11) when $q_0 = q$, x_{kt} is as in Lemma 1, and $q_{lt} = q$ for all $l \neq k$ and all t by setting $L_{kt} = L_{\mathcal{D}}$ at each t.*

Proof. Let

$$\mathcal{H}^*(q_{kt}, \lambda) = \max_{L_{kt}} \mathcal{H}(q_{kt}, \lambda, L_{kt}) = \pi_{\mathcal{D}} A \frac{q_{kt}^{\rho-1}}{Dq^{\rho-1}} - (\lambda\varphi)^{\frac{1}{1-\varphi}} + \lambda \left((\lambda\varphi)^{\frac{\varphi}{1-\varphi}} - \delta q_{kt} \right)$$

where $\mathcal{H}$ is the Hamiltonian defined in the proof of Lemma 5. Consider a deviation to an arbitrary strategy $(\hat{L}_{kt})$ with associated quality path $(\hat{q}_{kt})$. Recall from Condition 2 that $\rho \leq 2$. Since $\mathcal{H}^*$ is concave in the state,

$$\int_0^s e^{-rt} \mathcal{H}^*(\hat{q}_{kt}, \lambda) dt \leq \int_0^s e^{-rt} \mathcal{H}^*(q, \lambda) dt + \int_0^s e^{-rt} \mathcal{H}_1^*(q, \lambda) \left(\hat{q}_{kt} - q\right) dt$$

which implies

$$\int_0^s e^{-rt} \mathcal{H}^*(\hat{q}_{kt}, \lambda) dt \leq \int_0^s e^{-rt} \mathcal{H}^*(q, \lambda) dt + \int_0^s e^{-rt} r\lambda \left(q - \hat{q}_{kt}\right) dt.$$

By integration by parts, the second integral on the RHS is

$$\lambda \left[e^{-rs} \left(q - \hat{q}_{ks}\right) + \int_0^s e^{-rt} \dot{\hat{q}}_{kt} dt \right].$$

We therefore have

$$\int_0^s e^{-rt} \mathcal{H}^*(\hat{q}_{kt}, \lambda) dt \leq \int_0^s e^{-rt} \mathcal{H}^*(q, \lambda) dt + \lambda \int_0^s e^{-rt} \dot{\hat{q}}_{kt} dt + \lambda e^{-rt} q.$$

Rearranging,

$$\int_0^s e^{-rt} \mathcal{H}^*(\hat{q}_{kt}, \lambda) dt + \lambda \int_0^s e^{-rt} \dot{\hat{q}}_{kt} dt \leq \int_0^s e^{-rt} \mathcal{H}^*(q, \lambda) dt + \lambda e^{-rt} q.$$

Using the definition of $\mathcal{H}^*$, we have

$$\int_0^s e^{-rt} \left[\mathcal{H}(\hat{q}_{kt}, \lambda, \hat{L}_{kt}) + \lambda \dot{\hat{q}}_{kt} \right] dt \leq \int_0^s e^{-rt} \mathcal{H}^*(q, \lambda) + \lambda e^{-rt} q.$$

This implies

$$\int_0^s e^{-rt} \left[\pi_{\mathcal{D}} A \frac{\hat{q}_{kt}^{\rho-1}}{Dq^{\rho-1}} - \hat{L}_{kt} \right] dt \leq \int_0^s e^{-rt} \left[\frac{\pi_{\mathcal{D}} A}{D} - L_{\mathcal{D}} \right] + \lambda e^{-rt} q.$$

By taking limits as $s \to \infty$ we see that the deviation is not profitable. This proves Part 1. Part 2 follows from accounting. ☐

Proof Theorem 1. Parts 1-5 follow from Lemmas 1-5 and 7-9 or were proven in the main text with one caveat: we have not done a formal verification that the bidding strategy in Lemma 4, which solves the Bellman, does in fact solve the firm's problem. This proof is standard, since the value function clearly satisfies a transversality condition since flow profits are bounded. We therefore omit it. Part 6 follows by substituting the equilibrium demands and platform quality levels (using Parts 1-5 of Theorem 1) into equations (1) and (2) for the CES aggregates. ☐

B. Proofs for Section 5

The following Lemma 10 is used to prove Propositions 1 and 3.

Lemma 10. *The following hold:*

1. *An increase in data informativeness G leads the cumulative value $M\mu_H$ of firms in Ω_{it} to increase and the CDF H^c to decrease in $\succ_{SOSD}$.*

2. *An increase in ad frequency A leads the CDFs H and H^c to decrease in $\succ_{FOSD}$ but the cumulative value $M\mu_H$ of firms in Ω_{it} to increase.*

Proof. To prove Part 1, suppose that G increases in $\succ_{MPS}$ to $\hat{G}$. Let $\hat{H}^c$ denote the steady state distribution under $\hat{G}$. Define $\gamma : \mathbb{R}_+ \to [-1, 1]$ and $\nu : \mathbb{R}_+ \to [-1, 1]$ such that

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

for all $y \in \mathbb{R}_+$. Then by Lemma 3, it follows that

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

and

$$FG(y) = MH^c(y)^N + (F - M)H^c(y).$$

Substracting the bottom equation from the top equation gives

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

This holds for each for each $y \in \mathbb{R}_+$. Integrating both sides from 0 to $s \in \mathbb{R}_+$ we derive

$$\int_0^s \nu(y)\, dy\left(\frac{F - M}{F} + \frac{M}{F}\hat{H}^c(s)^{N-1}\right) - \frac{M}{F}\int_0^s \int_0^y \nu(l)\, dl\, d\hat{H}^c(y)^{N-1} \geq 0. \tag{47}$$

Above, we have used integration by parts and the fact that $\hat{G} \succ_{MPS} G$ implies that $\int_0^s \gamma(y)\, dy \geq 0$ for each $s \in \mathbb{R}_+$. We now argue that $\int_0^s \nu(y)\, dy \geq 0$ for all $s \in \mathbb{R}_+$ with strict inequality at some point $s \in \mathbb{R}_+$. This implies both that $H^c \succ_{SOSD} \hat{H}^c$ and $\mu_{\hat{H}} > \mu_H$. Suppose for contradiction that there exists a point $s \in \mathbb{R}_+$ such that $\int_0^s \nu(y)dy < 0$. Let

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

If $l^* > 0$, then (47) is violated at l^* which is a contradiction. Then it must be that $l^* = 0$. But by inspecting (47), we see that $\int_0^s \nu(y)\, dy$ must be increasing in s when it fist departs from 0 as otherwise (47) is violated for s close to the point of departure. Thus $l^* \neq 0$, a contradiction. It follows that $\int_0^s \nu(y)\, dy \leq 0$ for each $s \in \mathbb{R}_+$. Strict inequality must occur at a some point since $\hat{G} \succ_{MPS} G$.

To prove Part 2, recall from Lemma 3 that an increase in A leads to an increase in M. By inspecting the equation for H^c in Lemma 3, it follows that H^c must decrease in $\succ_{FOSD}$ (which implies that H also decreases in $\succ_{FOSD}$). This implies that $(F-M)\mu_{H^c}$ must decrease. Since $M\mu_H + (F-M)\mu_{H^c} = F\mu_G$, then $M\mu_H$ must increase.

□

Proof of Proposition 2. By Part 2 of Theorem 1, an increase in interoperability leads to an increase in investment $L_{\mathcal{D}}$ which proves Part 3 of Proposition 2. By Part 4 of Theorem 1 the rise in investment $L_{\mathcal{D}}$ leads income I to decrease which in turn leads to a decrease in ad revenue $\pi_{\mathcal{D}}A$. By Part 6 of Theorem 1 the rise in investment $L_{\mathcal{D}}$ also leads aggregate product consumption C to decline. □

Proof of Proposition 3. We first prove Part 2. Recall that by definition $\pi_{\mathcal{D}}A = I\hat{\pi}_{\mathcal{D}}A$. We will prove that an increase in A leads to a decrease in $\hat{\pi}_{\mathcal{D}}A$. Later, we argue that the general equilibrium feedback through income I can not overturn the sign of the direct effect on $\pi_{\mathcal{D}}A$ from a decrease in $\hat{\pi}_{\mathcal{D}}A$.

By Propositions 3 and 4,

$$\hat{\pi}_{\mathcal{D}}A = \frac{\lambda_f}{\sigma \int_0^\infty 1 - H^c(s)^N \, ds} \int_0^\infty \frac{1 - NH^c(s)^{N-1} + (N-1)H^c(s)^N}{r + \lambda_f + \lambda_e(s)} \, ds$$

where $\lambda_e(s) = \lambda_a H^c(s)^{N-1}$ for each $s \in \mathbb{R}_+$. By Part 2 of Lemma 10, an increase in A leads to a decrease in H^c in $\succ_{FOSD}$. Since λ_e increases pointwise, to show that $\hat{\pi}_{\mathcal{D}}A$ decreases as A increases, it suffices to prove 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 $H^c(s)$. This can be verified by taking a derivative. We omit this step. To see that the general equilibrium feedback through income can not offset the decrease in $\hat{\pi}_{\mathcal{D}}A$ and that ad revenue $\pi_{\mathcal{D}}A$ declines we observe the following. If $\pi_{\mathcal{D}}A$ were to increase, then $L_{\mathcal{D}}$ would increase by Proposition 5. But this contradicts Part 2 of Theorem 1 where we see that a decrease in $\hat{\pi}_{\mathcal{D}}A$ leads to a decrease in $L_{\mathcal{D}}$.

Since ad revenue declines, Part 3 follows from Part 1 of Lemma 5. Part 1 follows from Part 3, Lemma 10, and Part 6 of Theorem 1.

□

Proof of Proposition 1. By inspecting (27) and based on the discussion in the main text, $\hat{\pi}_{\mathcal{D}}$ is bounded above by

$$\frac{1}{\sigma M(r + \lambda_f)}.$$

This upper bound is obtained when data is uninformative since each firm bids this amount in each auction. Ad revenue $\pi_{\mathcal{D}}A$ is maximal across all data when $\hat{\pi}_{\mathcal{D}}$ is maximal across all data as general equilibrium feedback effects through income are never strong enough to overturn the direct effect on ad revenues as described in the proof of Proposition 3. We therefore have Part 2 of Theorem 1. Part 3 follows from Part 2 of Theorem 1. Part 1 then follows from Lemma 10 and Part 6 of 1. □

Construction of Figures 5 and 6. We explain the garbling used to construct Figures 5 and 6 and depicted graphically below in Figure 9.

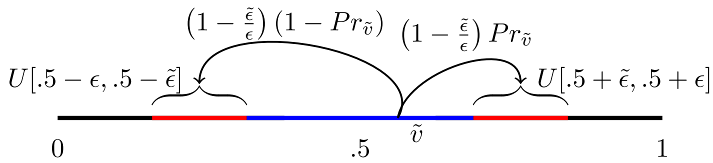

Figure 9: Construction of the conditional distribution over $\hat{v}$ given $\tilde{v}$.

We construct the less informative signal with realizations denoted by $\tilde{v}_{ij}$ by garbling $\hat{v}_{ij}$ as follows. Fix $\tilde{v}$ in the blue region. If $\hat{v}_{ij} = \tilde{v}$ then set $\tilde{v}_{ij} = \tilde{v}$. If $\hat{v}_{ij}$ is in the red region above the blue region, then $\mathbb{P}\{\tilde{v}_{ij} \in d\tilde{v}|\hat{v}_{ij}\} = Pr_{\tilde{v}}\frac{1}{\epsilon}$. If $\hat{v}$ is in the red region below the blue region, then $\mathbb{P}\{\tilde{v}_{ij} \in d\tilde{v}|\hat{v}_{ij}\} = (1-Pr_{\tilde{v}})\frac{1}{\epsilon}$ where $Pr_{\tilde{v}}$ is chosen so that

$$\tilde{v} = Pr_{\tilde{v}} \underbrace{\frac{.5+\epsilon+.5+\tilde{\epsilon}}{2}}_{\text{mean of upper red region}} + (1-Pr_{\tilde{v}}) \underbrace{\frac{.5-\epsilon+.5-\tilde{\epsilon}}{2}}_{\text{mean of lower red region}}. \tag{48}$$

Then $\mathbb{E}[v_{ij}|\tilde{v}_{ij} = \tilde{v}] = \tilde{v}$. Moreover $\tilde{v}_{ij} \sim U[.5-\tilde{\epsilon}, .5+\tilde{\epsilon}]$. From this, we compute the joint distribution of $\hat{v}, \tilde{v}$ conditional on $j \in \Omega_{it}^c$. We then use that to numerically compute the optimal bid on the platform with the garbled data using the formula $\mathbb{E}[B(\hat{v}_{ij})|\tilde{v}_{ij}, j \in \Omega_{it}^c]$.

We now compute the joint distribution of $\hat{v}, \tilde{v}$ conditional on $j \in \Omega_{it}^c$. Based on the construction described above, each point $\hat{v}_{ij}$ in $[.5-\epsilon, .5+\epsilon]$ stays with probability $\tilde{\epsilon}/\epsilon$. Otherwise, it jumps. Conditional on jumping, the probability it jumps up is $Pr_{\tilde{v}}$ which solves (48):

$$Pr_{\tilde{v}} = \frac{2\tilde{v}-1+\epsilon+\tilde{\epsilon}}{2\tilde{\epsilon}+2\epsilon}.$$

Conditional on jumping, the probability that it jumps down is $1 - Pr_{\tilde{v}}$. Let $\chi(\hat{v}|\tilde{v})$ denote $\mathbb{P}\{\hat{v}_{ij} \in d\hat{v}|\tilde{v}_{ij} = \tilde{v}\}$. Then if $\hat{v}$ is in the upper red region $[.5+\tilde{\epsilon}, .5+\epsilon]$,

$$\chi(\hat{v}|\tilde{v}) = \left(1-\frac{\tilde{\epsilon}}{\epsilon}\right)\frac{2\tilde{v}-1+\epsilon+\tilde{\epsilon}}{2\tilde{\epsilon}+2\epsilon}\frac{1}{\epsilon-\tilde{\epsilon}} = \frac{2\tilde{v}-1+\epsilon+\tilde{\epsilon}}{(2\tilde{\epsilon}+2\epsilon)\epsilon}.$$

If $\hat{v}$ is in the lower red region $[.5-\epsilon, .5-\tilde{\epsilon}]$,

$$\chi(\hat{v}|\tilde{v}) = \left(1-\frac{\tilde{\epsilon}}{\epsilon}\right)\left(1-\frac{2\tilde{v}-1+\epsilon+\tilde{\epsilon}}{2\tilde{\epsilon}+2\epsilon}\right)\frac{1}{\epsilon-\tilde{\epsilon}} = \left(1-\frac{2\tilde{v}-1+\epsilon+\tilde{\epsilon}}{2\tilde{\epsilon}+2\epsilon}\right)\frac{1}{\epsilon}.$$

The residual mass $\tilde{\epsilon}/\epsilon$ is concentrated on the event $\hat{v} = \tilde{v}$. With abuse of notation let $\chi(\tilde{v}|\hat{v})$ denote $\mathbb{P}\{\tilde{v}_{ij} \in d\tilde{v}|\hat{v}_{ij} = \hat{v}\}$. Then, if $\hat{v}$ is in one of the red regions $[.5-\epsilon, .5-\tilde{\epsilon}] \cup [.5+\tilde{\epsilon}, .5+\epsilon]$,

$$\chi(\tilde{v}|\hat{v}) = \chi(\hat{v}|\tilde{v})\frac{\epsilon}{\tilde{\epsilon}}$$

for each $\tilde{v}$. Otherwise, if $\hat{v}$ is in the blue region $[.5-\tilde{\epsilon}, .5+\tilde{\epsilon}]$ all of the mass is placed on $\tilde{v} = \hat{v}$. Then

$$\tilde{H}^c(\tilde{v})' = \int_{[.5-\epsilon,.5-\tilde{\epsilon}]\cup[.5+\tilde{\epsilon},.5+\epsilon]} \chi(\tilde{v}|\hat{v}) dH^c(\hat{v}) + H^c(\tilde{v})'$$

where $\tilde{H}^c(\tilde{v})$ is the distribution over $\tilde{v}$ outside the consideration set. Thus, if $\hat{v}$ is in one of the red regions, it is

$$\mathbb{P}\{\hat{v}_{ij} \in d\hat{v} | j \in \Omega^c_{it}, \tilde{v}_{ij} = \tilde{v}\} = \frac{\chi(\tilde{v}|\hat{v}) H^c(\hat{v})'}{\tilde{H}^c(\tilde{v})'}.$$

If $\hat{v} = \tilde{v}$, then

$$\mathbb{P}\{\hat{v}_{ij} \in d\hat{v} | j \in \Omega^c_{it}, \tilde{v}_{ij} = \tilde{v}\} = \frac{H^c(\tilde{v})'}{\tilde{H}^c(\tilde{v})'}.$$

No density is placed on any other point in the blue region. □

Additional Comparative Statics

We present additional comparative statics results with respect to the measure of firms F, the measure of platforms, D, and product substitutability σ.

Proposition 4. *An increase in the measure of firms F leads to*

1. *an increase in ad revenue $\pi_{\mathcal{D}} A$.*

2. *an increase in investment $L_{\mathcal{D}}$.*

Proof. An increase in F leads H and H^c to increase in $\succ_{FOSD}$ by an analogous argument used to show Part 2 of Lemma 10. This in turn leads $\pi_{\mathcal{D}} A$ to increase by analogous argument to that of Part 2 of Proposition 3. That then leads $L_{\mathcal{D}}$ to increase. □

Proposition 5. *If the measure of platforms D increases then:*

1. *Aggregate platform consumption X increases if $1/(\rho-1) > \varphi$ and decreases if $1/(\rho-1) < \varphi$.*

2. *Aggregate product consumption C does not change.*

3. *Ad revenue $\pi_{\mathcal{D}} A$ does not change.*

Proof. By Part 2 of Theorem 1, the total investment by platforms is invariant to the measure of platforms D. This immediately implies Parts 2 and 3 of Proposition 5. To prove Part 1, by Theorem 1 it follows that

$$X = D^{\frac{1}{\rho-1}} \frac{L_{\mathcal{D}}^{\varphi}}{\delta} = \frac{1}{\delta} D^{\frac{1}{\rho-1}-\varphi} \left(\frac{\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}} A(\rho-1)}{r+\delta+\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}} A(\rho-1)} L \right)^{\varphi}.$$

Part 1 follows by inspection. □

Proposition 6. *If product substitutability σ increases then:*

1. *Ad revenue $\pi_{\mathcal{D}} A$ decreases.*
2. *Investment $L_{\mathcal{D}}$ decreases.*

Proof. To prove Part 1, we see that $\hat{\pi}_{\mathcal{D}}$ in (27) is decreasing in σ. As argued in the proof of Part 2 of Proposition 3, the general equilibrium effect through a change in income I can not offset this direct effect of a decrease in $\hat{\pi}_{\mathcal{D}}$ so that $\pi_{\mathcal{D}}$ decreases. Part 2 follows from Part 1. $\square$

C. Proofs for Section 6

We prove results for a more general formulation of the model which we next describe. Lemma 6 and Theorem 2 will follow as special cases of Lemma 11 and Theorem 3 proven below.

Setup

We extend the baseline model as follows. There are now two sectors of the product market. One sector is a leisure sector which requires attention as well as income to consume. All attention comes from leisure and is freely substitutable between platform use and leisure products. The second sector consists of all other products which do not require attention to consume. Both sectors have a CES structure with the same substitutability parameter σ. The total measure of products in the market is F as in the baseline model. We continue to denote the set of all firms by $\mathcal{F}$ and index individual firms by j. We assume that a fraction β of products are leisure products. As before, $v_{ij} \sim P$ on $\mathbb{R}_+$ and is drawn independently across i and j. Similarly, $\hat{v}_{ij} \sim G$ where $P \succ_{MPS} G$ independently across i and j.

We assume that a consumer's flow utility is

$$\tilde{u}(C_{Nt}, \tau C_{Lt}, (1-\tau)X_t) = \left[C_{Nt}^{1-\beta}\left(\left(1-\tau\right)C_{Lt}\right)^{\beta}\right]^{1-\gamma}\left[\tau X_t\right]^{\gamma} \tag{49}$$

where τ is the fraction of attention devoted to leisure products, C_{Nt} is the CES aggregate over nonleisure products, C_{Lt} is the CES aggregate over leisure products, and X_t is the CES aggregate over platform use. Note that β is not only the fraction of products which are leisure products but also a parameter of the utility function. We will explain the logic behind this assumption shortly.

We assume that the consumer chooses attention τ myopically to maximize flow utility: the consumer does not allocate attention to platforms to purposefully see ads. Because the utility is Cobb-Douglas, it is immediate that the consumer spends a fraction $\tau = \gamma/[\gamma + (1-\gamma)\beta]$ of attention on platforms. Moreover, the consumer spends a fraction β of income on leisure products. With this, let us define

$$C_t = C_{Nt}^{1-\beta} C_{Lt}^{\beta}.$$

Then, up to a constant scaling factor, the consumer's flow utility is

$$u(C_t, X_t) = C_t^{1-\gamma} X_t^{\gamma}. \tag{50}$$

Given the utility function in (49), since β is the fraction of income spent on leisure products, if we endogenize the choice by firms concerning which type of product to produce at ex-ante, then a fraction β would choose to produce leisure products. Rather than explicitly microfounding this we simply assume it from the outset.

Since it is efficient for 1. any labor allocated to investment to be split evenly among the platforms and 2. production labor to be split the same way as in equilibrium (a fraction β of production labor is used for leisure products) the social planner's problem is (32).

Analysis

It is easy to verify that the stationary equilibrium is characterized by the same equations as in the baseline model. Namely C, X, M, H, and $\pi_{\mathcal{D}}$ are as in Theorem 1 except with $A\tau$ in place of A, since now only τ units of attention are spent on platforms. In equilibrium, the fraction of leisure products in the consideration set is equal to the fraction in the population: β. Consequently, a firm's flow profit from selling to a consumer depends on the match value but not on the product type. To see this the flow profit accruing to firm j from selling a leisure product to consumer i is

$$\beta I \frac{\hat{v}_{ij}}{\sigma \beta M \mu_H} = I \frac{\hat{v}_{ij}}{\sigma M \mu_H}.$$

Thus β fraction of income cancels out with the β fraction of firms in the consideration set. Analogous logic applies to the flow profit of nonleisure products. Moreover, H in Proposition 3 is the (empirical) CDF of firms' expectations for both product types within the consideration set of any consumer. To see this, note that by the same logic as in the baseline model, the distribution of expectations in the consideration set corresponding to leisure products must satisfy

$$\beta M (H^c)^N + (\beta F - \beta M) H^c = \beta F G.$$

Since β cancels out from both sides we have the same condition as in the baseline model. The same also holds for non-leisure products. Thus, the same bidding function applies for both products in the equilibrium of the extended model. Moreover, all of the comparative statics results of the previous section continue to hold.

The following Lemma 11 reports the steady state solution to the planner's problem.

Lemma 11. *The steady state level of investment by any given platform under the social planner is*

$$L_{\mathcal{D}}^{fb} = \frac{\varphi \delta \frac{1-\tau}{\tau} \beta}{r + \delta + \varphi \delta \frac{1-\tau}{\tau} \beta} \frac{L}{D}. \tag{51}$$

Proof. Let $\hat{C} = (M\mu_H)^{\frac{1}{\sigma-1}}$. The (present-value) Hamiltonian for the social planner's problem is

$$\mathcal{H}(q_t, \lambda_t, L_{\mathcal{D}t}) = u\left((1-\tau)^{\beta}\frac{\sigma}{\sigma-1}\left(L - DL_{\mathcal{D}t}\right)\hat{C}, \tau D^{\frac{1}{\rho-1}}q_t\right) + \lambda_t\left(L_{\mathcal{D}t}^{\varphi} - \delta q_t\right)$$

where λ_t is the costate variable which evolves according to

$$r\lambda_t - \dot{\lambda}_t = -D^{\frac{1}{\rho-1}}(1-\tau)u_2 - \delta\lambda_t$$

Note that we are implicitly assuming that M and H are in steady state since we are only interested in the steady state solution. By the Maximum Principle, investment must maximize the Hamiltonian along the optimal trajectory. Since τ is an interior solution for the optimal choice of attention it follows that

$$u_1([1-\tau]^{\beta}C, \tau X)\beta\tau^{\beta-1}C = u_2([1-\tau]^{\beta}C, \tau X)X. \tag{52}$$

Using (52), the first-order condition at the steady state q is

$$\frac{\sigma}{\sigma-1}\hat{C}D[1-\tau]^{\beta}\frac{1}{\varphi}(\delta q)^{\frac{1-\varphi}{\varphi}}u_1 = \lambda_t q^{\alpha}. \tag{53}$$

Above, we have used the fact that $\delta q = L_{\mathcal{D}}^{\varphi}$ in steady state. This condition implies that λ_t is a constant λ in steady state. By the costate evolution equation,

$$\lambda = \frac{D^{\frac{1}{\rho-1}}\tau u_2}{r+\delta}.$$

Substituting into the first order condition gives

$$\frac{\sigma}{\sigma-1}\hat{C}D[1-\tau]^{\beta}\frac{1}{\varphi}(\delta q)^{\frac{1-\varphi}{\varphi}}\frac{u_1}{u_2} = \frac{D^{\frac{1}{\rho-1}}\tau}{r+\delta}q^{\alpha}. \tag{54}$$

From (52), we have

$$\frac{u_1}{u_2} = \frac{X}{\beta[1-\tau]^{\beta-1}C}.$$

Substituting into (54), we have

$$\frac{\sigma}{\sigma-1}D\frac{1}{\beta}\frac{1-\tau}{\tau}\frac{1}{\frac{\sigma}{\sigma-1}\left(L - DL_{\mathcal{D}}\right)}L_{\mathcal{D}}\frac{1}{\varphi\delta} = \frac{1}{r+\delta}$$

which rearranges to

$$L_{\mathcal{D}} = \frac{\varphi\delta\frac{1-\tau}{\tau}\beta}{D(r+\delta)}.$$

Using this relationship, we find that

$$L_{\mathcal{D}} = \frac{\varphi\delta\frac{1-\tau}{\tau}\beta}{r+\delta+\varphi\frac{1-\tau}{\tau}\beta}\frac{L}{D}.$$

This is a necessary condition for $L_{\mathcal{D}}$ to be a steady state solution to the planner's problem. Since the Hamiltonian is concave in the state and control, an analogous verification to that of Lemma 9 can be used to prove sufficiency. $\square$

Note that in the proof of Lemma 11, we did not use any properties of the Cobb-Douglass utility function beyond the interiority of the optimal choice of attention and concavity. Namely (53) would continue to hold by the envelope theorem even though τ will depend on $L_{\mathcal{D}t}$ for other utility functions. However, in the formulation of the planner's problem, we assumed that the planner takes as given the law of motion of M_t and H_t. Though this is without loss for Cobb-Douglas utility, it is not generally so.

Theorem 3. *If, in a stationary equilibrium,*

$$\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}}A\tau(\rho-1)-\frac{\tau}{1-\tau}\beta\begin{cases}<0, & \textit{then investment is too low.}\\=0, & \textit{then investment is efficient.}\\>0, & \textit{then investment is too high.}\end{cases}$$

Given τ and β, the deviation $L_{\mathcal{D}}-L_{\mathcal{D}}^{fb}$ is increasing in $\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}}A\tau(\rho-1)-\frac{\tau}{1-\tau}\beta$.

Proof. Theorem 3 follows from Lemma 11 and Part 2 of Theorem 1. □

D. Proofs for Section 7

We prove that $\boldsymbol{\Lambda}$ is a contraction with respect to the sup norm when values are bounded above by some constant $\overline{v}$. That is, all of the mass of the prior P is contained in $[0,\overline{v}]$. This result can be generalized, using a different norm, to allow for cases when values are unbounded. However, for practical purposes of computation, the following Lemma 12 suffices and requires less effort to prove. Note also that $\boldsymbol{\Lambda}$ is increasing. Thus, even without the bounded support assumption, starting from an initial $\boldsymbol{B}$ such that $B_l>\Lambda_l(B)$ for each $l\in\{1,2\}$ it follows that $\{\boldsymbol{\Lambda}_n(B)\}_{n=1}^{\infty}$ is a decreasing sequence which converges to the fixed point.

Lemma 12. *Suppose that each value v_{ij} is supported on $[0,\overline{v}]$. The operator $\boldsymbol{\Lambda}$ defined in (38) is a contraction map with modulus $\lambda_a/(\lambda_a+\lambda_f+r)$ with respect to the sup norm when restricted to the domain consisting of functions in $C^+(\mathbb{R})^2$ which are bounded above by $\frac{\pi_{\mathcal{F}}}{\lambda_f+r}\overline{v}$.*

Proof. Let $||\cdot||$ denote the sup norm. We have

$$||\boldsymbol{\Lambda}(\boldsymbol{f})-\boldsymbol{\Lambda}(\hat{\boldsymbol{f}})||=\max_l\left|\left|\mathbb{E}\left[\frac{\lambda_a\sum_{z=1}^2 x_z\int_{-\infty}^{\zeta_{zij}}\left(f_z(s)-\hat{f}_z(s)\right)dO_z(s)}{\lambda_f+r+\lambda_a\sum_{z=1}^2 x_zO_z(\zeta_{zij})}\middle|\zeta_{lij}=\cdot,j\in\Omega_{it}^c\right]\right|\right|$$

$$\leq\left|\left|\mathbb{E}\left[\frac{||\boldsymbol{f}-\hat{\boldsymbol{f}}||\lambda_a\sum_{z=1}^2 x_zO_z(\zeta_{z_{ij}})}{\lambda_f+r+\lambda_a\sum_{z=1}^2 x_zO_z(\zeta_{zij})}\middle|\zeta_{lij}=\cdot,j\in\Omega_{it}^c\right]\right|\right|$$
$$\leq\frac{\lambda_a}{r+\lambda_a+\lambda_f}||\boldsymbol{f}-\hat{\boldsymbol{f}}||.$$

□

Note that in Lemma 12 restricting the domain to functions in $C^+(\mathbb{R})^2$ which are bounded above by $\frac{\pi_{\mathcal{F}}}{\lambda_f+r}\overline{v}$ does not entail any loss of generality since it is never optimal for a firm to bid more than its perceived value for being in the consideration set.

We next prove that the fixed point is in increasing bidding functions. For this, we will use the following generalization of the Arzelà-Ascoli Theorem, which follows from Theorems 3.4.20 and 8.2.10 of Engelking (1977).

Theorem 4. *$\mathcal{A} \subset C(\mathbb{R}_+)$ 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 13. *The fixed point of the operator $\mathbf{\Lambda}$ defined in (38) is a pair of increasing functions.*

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

$$V(\boldsymbol{\zeta}_{ij}) = \max_{(b_k)_{k=1}^{\infty}} \mathbb{E}\left[\int_0^{\infty} \pi_{\mathcal{F}} v_{ij} \mathbb{1}_{\{j \in \Omega_{it}\}}\, ds - \sum_{k=1}^{\infty} e^{-r\tau_k} B_{l_k}^{(1)} \mathbb{1}_{\{b_k > B_{l_k}^{(1)}\}} \middle| \boldsymbol{\zeta}_{ij}\right] \tag{55}$$

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. We 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 (36), (37) and thus must be the fixed point of the contraction (38). This guarantees that the fixed point of (38) is necessarily increasing provided there exists such a pair of bidding functions B_1, B_2.

We will prove that

$$\frac{\pi_{\mathcal{F}}}{\lambda_f + r}\mathbb{E}\left[v_{ij}|\boldsymbol{\zeta}_{ij}\right] - \frac{r}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r}V(\boldsymbol{\zeta}_{ij}) \tag{56}$$

is nondecreasing in both of its arguments and increasing in one of them. This will ensure that bidding strategies which satisfy (36) are increasing using Condition 3. 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

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

Rearranging yields

$$\begin{aligned}\frac{\pi_{\mathcal{F}}}{\lambda_f + r}&\mathbb{E}[v_{ij}|\zeta_{1ij} = \overline{\zeta}_1, \zeta_{2ij}] - \frac{\pi_{\mathcal{F}}}{\lambda_f + r}\mathbb{E}[v_{ij}|\zeta_{1ij} = \underline{\zeta}_1, \zeta_{2ij}] \\ &< \frac{r}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r}V(\overline{\zeta}_1, \zeta_{2ij}) - \frac{r}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r}V(\underline{\zeta}_1, \zeta_{2ij}).\end{aligned} \tag{57}$$

We will show that the above inequality can not hold.

As seen in (55), the value function can be decomposed into two parts: one which arises from flow profits from sales (the expectation of the first sum in (55)) and one which arises from costs of advertising (the expectation of the second sum (55)). Let us write the value function of a firm j with signals $\overline{\zeta}_1$, ζ_{-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 (57), we therefore have,

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

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

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

Substituting into the RHS above, we obtain a contradiction. The proof that (56) is also nondecreasing in its second argument is analogous. Now we 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 (55)).

Let $\text{Lip}(\mathbb{R}_+)$ denote the set of $\frac{\pi_{\mathcal{F}}}{\lambda_f+r}$-Lipschitz functions f such that $f(y) \leq \frac{\pi_{\mathcal{F}}}{\lambda_f+r}y$ at each $y \in \mathbb{R}_+$. By Theorem 4 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 3 ensures that there is a one to one mapping between old signals and new signals. Moreover, now

$$B_l(\zeta_{lij}) = \frac{\pi_{\mathcal{F}}}{\lambda_f + r}\zeta_{lij} - \mathbb{E}\left[\frac{r}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r}V(\boldsymbol{\zeta}_{ij})\middle|\zeta_{lij}, j \in \Omega_{it}^c\right] \tag{58}$$

which is $\frac{\pi_{\mathcal{F}}}{\lambda_f+r}$-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 (55) when B_1, B_2 are given by $\boldsymbol{B}_{input}$. Next, define $\boldsymbol{B}_{output}$ using (58). 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 (38).

□

Lemma 14. *In any stationary equilibrium, the following hold:*

1. *A platform in group l invests at constant rate*

$$L_{\mathcal{D}l} = \frac{\varphi\delta\pi_{\mathcal{D}l}A\left(\rho-1\right)}{r+\delta}\frac{1}{m_l + m_{-l}\left(\frac{\pi_{\mathcal{D}-l}}{\pi_{\mathcal{D}l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}}$$

2. *A platform in group l has constant quality level*

$$q_l = \frac{L_{\mathcal{D}l}^{\varphi}}{\delta}.$$

3. *The total share of attention held by platforms in group l is*

$$x_l = \frac{m_l q_l^{\rho-1}}{m_l q_l^{\rho-1} + m_{-l} q_{-l}^{\rho-1}}.$$

Proof. In a stationary equilibrium, a platform in group $l \in \{1,2\}$ invests a constant level $L_{\mathcal{D}l}$ to mantain quality level $q_l = L_{\mathcal{D}l}^{\varphi}/\delta$.

Let platform k belong to group l. Let m_l denote the measure of platforms in group l. The (present-value) Hamiltonian for platform k's optimization problem is

$$\mathcal{H}(q_{kt}, \lambda_t, L_{kt}) = \pi_{\mathcal{D}l}A\frac{q_{kt}^{\rho-1}}{m_l q_l^{\rho-1} + m_{-l} q_{-l}^{\rho-1}} - L_{kt} + \lambda_t\left(L_{kt}^{\varphi} - \delta q_{kt}\right)$$

where λ_t, the costate variable, evolves according to

$$r\lambda_t - \dot{\lambda}_t = \pi_{\mathcal{D}l}A\left(\rho-1\right)\frac{q_{kt}^{\rho-2}}{m_l q_l^{\rho-1} + m_{-l} q_{-l}^{\rho-1}} - \lambda_t\delta.$$

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

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

Under the conjectured stationary strategy then

$$\lambda_t\varphi L_{\mathcal{D}l}^{\varphi-1} = 1.$$

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

$$\lambda = \frac{\pi_{\mathcal{D}l} A\,(\rho-1)}{r+\delta} \frac{q_l^{\rho-2}}{m_l q_l^{\rho-1} + m_{-l} q_{-l}^{\rho-1}}.$$

Substituting, we have

$$\frac{\pi_{\mathcal{D}l} A\,(\rho-1)}{r+\delta} \varphi L_{\mathcal{D}l}^{\varphi-1} = m_l q_l + m_{-l} \left(\frac{q_{-l}}{q_l}\right)^{\rho-2} q_{-l}.$$

This implies that

$$\frac{\pi_{\mathcal{D}l} A\,(\rho-1)}{r+\delta} \varphi L_{\mathcal{D}l}^{\varphi-1} = m_l \frac{L_{\mathcal{D}l}^{\varphi}}{\delta} + m_{-l} \left(\frac{L_{\mathcal{D}-l}}{L_{\mathcal{D}l}}\right)^{\varphi(\rho-2)} \frac{L_{\mathcal{D}-l}^{\varphi}}{\delta}.$$

Dividing both sides by $L_{\mathcal{D}l}^{\varphi}/\delta$ we arrive at

$$\frac{\delta \pi_{\mathcal{D}l} A\,(\rho-1)}{r+\delta} \varphi L_{\mathcal{D}l}^{-1} = m_l + m_{-l} \left(\frac{L_{\mathcal{D}-l}}{L_{\mathcal{D}l}}\right)^{\varphi(\rho-1)}.$$

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

$$\frac{\delta \pi_{\mathcal{D}-l} A\,(\rho-1)}{r+\delta} \varphi L_{\mathcal{D}-l}^{-1} = m_{-l} + m_l \left(\frac{L_{\mathcal{D}l}}{L_{\mathcal{D}-l}}\right)^{\varphi(\rho-1)}.$$

Let $y := L_{\mathcal{D}l}/L_{\mathcal{D}-l}$. Using the above two equations, we derive

$$\frac{\pi_{\mathcal{D}l}}{\pi_{\mathcal{D}-l}} \frac{1}{y} = \frac{m_l + m_{-l} y^{-\varphi(\rho-1)}}{m_{-l} + m_l y^{\varphi(\rho-1)}}.$$

Equivalently,

$$y = \left(\frac{\pi_{\mathcal{D}l}}{\pi_{\mathcal{D}-l}}\right)^{\frac{1}{1-\varphi(\rho-1)}}.$$

Thus,

$$L_{\mathcal{D}l} = \frac{\varphi \delta \pi_{\mathcal{D}l} A\,(\rho-1)}{r+\delta} \frac{1}{m_l + m_{-l} \left(\frac{\pi_{\mathcal{D}-l}}{\pi_{\mathcal{D}l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}}.$$

Note that Condition 2 implies that the Hamiltonian is jointly concave in the state and control so that $L_{\mathcal{D}l}$ does indeed solve a platform in group l's problem by an analogous verification argument to that of Lemma 9. $\square$

The following theorem summarizes the analysis.

Theorem 5. *In any stationary equilibrium, the following hold:*

1. *Demands and prices are as in Lemmas 1 and 2.*

2. *Income is*

$$I = L\left(1 + \frac{\varphi\delta A\left(\rho-1\right)}{r+\delta}\frac{m_l\hat{\pi}_{\mathcal{D}l} + m_{-l}\left(\frac{\hat{\pi}_{\mathcal{D}-l}}{\hat{\pi}_{\mathcal{D}l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}\hat{\pi}_{\mathcal{D}-l}}{m_l + m_{-l}\left(\frac{\hat{\pi}_{\mathcal{D}-l}}{\hat{\pi}_{\mathcal{D}l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}}\right)^{-1}. \tag{59}$$

3. *The (empirical) PDF over signals within consideration sets solves*

$$h(\boldsymbol{\zeta}) = x_1 N H^c(\zeta_1,\infty)^{N-1}h^c(\boldsymbol{\zeta}) + x_2 N H^c(\infty,\zeta_2)^{N-1}h^c(\boldsymbol{\zeta})$$

and the accounting identity

$$M_t h_t + (F - M_t)h_t^c = Fg.$$

4. *Platforms' choices of investment are as in Lemma 14.*

5. *The bidding function used by a firm in an auction on a platform in group $l \in \{1,2\}$ is the unique fixed point of the contraction map (38).*

6. *The total attention held by platforms in group $l \in \{1,2\}$ satisfies*

$$x_l = \frac{m_l q_l^{\rho-1}}{m_l q_l^{\rho-1} + m_{-l}q_{-l}^{\rho-1}} = \frac{m_l\left(\frac{\hat{\pi}_{\mathcal{D}l}}{\hat{\pi}_{\mathcal{D}-l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}}{m_l\left(\frac{\hat{\pi}_{\mathcal{D}l}}{\hat{\pi}_{\mathcal{D}-l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}} + m_{-l}}$$

where $\hat{\pi}_{\mathcal{D}l}$, $l \in \{1,2\}$ are the average ad prices per unit of income derived from the bidding functions in Part 5.

Proof. It suffices to prove only Parts 2 and 6 since all other parts follow from earlier results. To prove Part 2, recall that market clearing implies income is equal to the revenue of product firms so that

$$I = \frac{\sigma}{\sigma-1}\left(L - m_l L_{\mathcal{D}l} - m_{-l}L_{\mathcal{D}-l}\right)$$

where $\sigma/(\sigma-1)$ is the markup. This implies

$$L_{\mathcal{D}l} = \frac{\varphi\delta I\hat{\pi}_{\mathcal{D}l}A\left(\rho-1\right)}{r+\delta}\frac{1}{m_l + m_{-l}\left(\frac{\hat{\pi}_{\mathcal{D}-l}}{\hat{\pi}_{\mathcal{D}l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}}.$$

must hold for $l = 1,2$. We derive the equations in Part 2 by solving this linear system of equations. We observe that the system implies

$$I = L - \frac{\varphi\delta I\hat{\pi}_{\mathcal{D}l}A\left(\rho-1\right)}{r+\delta}\frac{m_l}{m_l + m_{-l}\left(\frac{\hat{\pi}_{\mathcal{D}-l}}{\hat{\pi}_{\mathcal{D}l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}} \\ - \frac{\varphi\delta I\hat{\pi}_{\mathcal{D}-l}A\left(\rho-1\right)}{r+\delta}\frac{m\_l}{m\_l + m_l\left(\frac{\hat{\pi}_{\mathcal{D}l}}{\hat{\pi}_{\mathcal{D}-l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}}$$

which is equivalent to

$$I = L - \frac{\varphi\delta IA\left(\rho-1\right)}{r+\delta}\frac{m_l\hat{\pi}_{\mathcal{D}l} + m_{-l}\left(\frac{\hat{\pi}_{\mathcal{D}-l}}{\hat{\pi}_{\mathcal{D}l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}\hat{\pi}_{\mathcal{D}-l}}{m_l + m_{-l}\left(\frac{\hat{\pi}_{\mathcal{D}-l}}{\hat{\pi}_{\mathcal{D}l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}}$$

which implies (40). To prove Part 6 we observe that

$$x_l = \frac{m_l q_l^{\rho-1}}{m_l q_l^{\rho-1} + m_{-l}q_{-l}^{\rho-1}} = \frac{m_l\left(\frac{q_l}{q_{-l}}\right)^{\rho-1}}{m_l\left(\frac{q_l}{q_{-l}}\right)^{\rho-1} + m_{-l}} = \frac{m_l\left(\frac{\hat{\pi}_{\mathcal{D}l}}{\hat{\pi}_{\mathcal{D}-l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}}}{m_l\left(\frac{\hat{\pi}_{\mathcal{D}l}}{\hat{\pi}_{\mathcal{D}-l}}\right)^{\frac{\varphi(\rho-1)}{1-\varphi(\rho-1)}} + m_{-l}}.$$

$\square$

Special Case

Here we study an especially tractable special case of the model. Platforms in group 1 have data while platforms in group 2 have uninformative data. Let $\hat{v}_{ij}$ denote the posterior expectation of v_{ij} given the data on a platform in group 1 which was assume is distributed according to G supported on $\mathbb{R}_+$ such that $P \succ_{MPS} G$. Each v_{ij} is drawn from G independently across consumers i and firms j.

Proposition 7. *In a stationary equilibrium, the CDF H of expectations conditional on group 1's signal inside the consideration set Ω_{it} solves*

$$MH(\hat{v}) + (F-M)H^c(\hat{v}) = FG(\hat{v})$$

where

$$x_1 H^c(\hat{v})^N + \left(\frac{F}{M} - x_1\right)H^c(\hat{v}) = \frac{F}{M}G(\hat{v})$$

for each $\hat{v} \in \mathbb{R}_+$ and

$$M = \frac{A}{\lambda_f}.$$

Proof. The distribution must match inflows and outflows:

$$x_1 H^c(\hat{v})^N + x_2 H^c(\hat{v}) = H(\hat{v}).$$

By the ELLN, we have the accounting identity

$$MH(\hat{v}) + (F-M)H^c(\hat{v}) = FG(\hat{v}).$$

Together with the previous equation we therefore have

$$x_1 H^c(\hat{v})^N + x_2 H^c(\hat{v}) = \frac{F}{M}G(\hat{v}) - \frac{F-M}{M}H^c(\hat{v}).$$

Simplifying gives

$$x_1 H^c(\hat{v})^N + \left(\frac{F}{M} - x_1\right)H^c(\hat{v}) = \frac{F}{M}G(\hat{v}).$$

$\square$

Thus, computing the stationary distribution is as tractable as in the baseline model under this extension. From here, we can derive the equilibrium bidding functions. On the platform without data, the bid is the same for all bidders and equal to a constant B_2.

Proposition 8. *In a stationary equilibrium, B_1 and B_2 are characterized explicitly given x_1, x_2, H, and H^c by the equations:*

$$B_1(\hat{v}_{ij}) = \frac{\pi_{\mathcal{F}}}{\lambda_f + r}\hat{v}_{ij} - \frac{r}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r}\left(x_2 \frac{\frac{1}{N}\left(\frac{\pi_{\mathcal{F}}}{\lambda_f+r}\hat{v}_{ij} - B_2\right) + \left(x_1 - \frac{r}{\lambda_f+r}\frac{x_1^2\lambda_a}{\lambda_a+r}\right)V_1(\hat{v}_{ij})}{1 + \frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\right).$$

and

$$B_2 = \frac{\pi_{\mathcal{F}}\mu_{H^c}}{\lambda_f + r} - \frac{r}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r}\frac{x_1 + 2x_2x_1\frac{\lambda_a}{\lambda_a+r}}{1 + x_2\frac{\lambda_a}{\lambda_a+r}}\mathbb{E}[V_1(\hat{v}_{ij}), j \in \Omega_{it}^c].$$

Above,

$$V_1(\hat{v}_{ij}) = \int_0^{\hat{v}_{ij}} \psi(s)ds + a + \frac{b}{1-b}\left(a + \mathbb{E}\left[\int_0^{\hat{v}_j}\psi(s)ds\right]\right)$$

where

$$\psi(\hat{v}_{ij}) = \frac{\frac{\pi_{\mathcal{F}}}{\lambda_f+r}\left(H^c(\hat{v}_{ij})^{N-1} + \frac{1}{N}\frac{\frac{x_2\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\left(1 - H^c(\hat{v}_{ij})^{N-1}\frac{r}{\lambda_f+r}\right)\right)}{1 + \frac{x_1\lambda_a}{\lambda_a+r} + H^c(\hat{v}_{ij})^{N-1}\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r} + x_2x_1\left(\frac{\lambda_a}{\lambda_a+r}\right)^2\frac{\frac{r}{\lambda_f+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\left(1 - H^c(\hat{v}_{ij})^{N-1}\frac{r}{\lambda_f+r}\right)},$$

$$a = \frac{-\frac{1}{N}\frac{\frac{x_2\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\frac{\pi_{\mathcal{F}}\mu_{H^c}}{\lambda_f+r}}{1 + \frac{x_1\lambda_a}{\lambda_a+r} + \frac{\frac{x_2\lambda_a}{\lambda_a+r}\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}},$$

and

$$b = \frac{\frac{1}{N}\frac{\frac{x_2\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}\frac{x_1+2x_2x_1\frac{\lambda_a}{\lambda_a+r}}{1+x_2\frac{\lambda_a}{\lambda_a+r}}}{1 + \frac{x_1\lambda_a}{\lambda_a+r} + \frac{\frac{x_2\lambda_a}{\lambda_a+r}\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}}.$$

Proof. To ease notation, let $O(\hat{v}) = H^c(\hat{v})^{N-1}$ for each $\hat{v} \in [0, \overline{v}]$. This is the probability of a firm j winning an ad auction on a platform in group 1 if $\hat{v}_{ij} = \hat{v}$. The stationary equilibrium bid by firm j on platform 1 must satisfy

$$B_1(\hat{v}_{ij}) = \frac{\pi_{\mathcal{F}}}{\lambda_f + r}\hat{v}_{ij} - \frac{r}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r}\left(x_1V_1(\hat{v}_{ij}) + x_2V_{12}(\hat{v}_{ij})\right). \tag{60}$$

where

$$\pi_{\mathcal{F}} := \frac{I}{\sigma M \mu_H}.$$

Above, V_1 denotes the expected NPV of flow profits from the consumer's purchases of firm j's product at the time when firm j is about to submit a bid in an auction on Platform 1. V_{12} is the expected NPV of flow profits when firm j is about to submit a bid in an auction on platform 2, but where the expectation is taken given the data on platform 1. Above, V_1 satisfies the equation

$$\begin{aligned}(61)\quad & V_1(\hat{v}_{ij})\left(1+\frac{\lambda_a}{(\lambda_a+r)}x_1\right) \\ &\qquad = O(\hat{v}_{ij})\left(B_1(\hat{v}_{ij})-\mathbb{E}[B_1(\hat{v}^{(1)})|\hat{v}_{ij}\geq\hat{v}^{(1)}]\right)+\frac{\lambda_a}{\lambda_a+r}x_2V_{12}(\hat{v}_{ij}).\end{aligned}$$

Above $\hat{v}^{(1)}$ is distributed according to the maximum of $N-1$ random draws from outside the consideration set. Next, let V_2 denote the expected NPV of flow profits when the expectation is taken without conditioning on data from platform 1, but conditioning on firm j knowing that it is not in the consideration set Ω_{it}. Then

$$V_2\left(1+x_2\frac{\lambda_a}{(\lambda_a+r)}\right)=\frac{\lambda_a}{\lambda_a+r}x_1\mathbb{E}[V_1(\hat{v}_{ij}),j\in\Omega_{it}^c]$$

and so rearranging,

$$(62)\qquad V_2=\frac{x_1\frac{\lambda_a}{\lambda_a+r}}{1+x_2\frac{\lambda_a}{\lambda_a+r}}\mathbb{E}[V_1(\hat{v}_{ij}),j\in\Omega_{it}^c].$$

Thus the optimal bid on platform 2 is

$$(63)\qquad B_2=\pi_{\mathcal{F}}\mu_{H^c}-\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}\left(x_2V_2+x_1\mathbb{E}[V_1(\hat{v}_{ij}),j\in\Omega_{it}^c]\right)$$

where μ_{H^c} is the mean of the distribution H^c. Similarly, V_{12} solves

$$V_{12}(\hat{v}_{ij})=\frac{1}{N}\left(\frac{\pi_{\mathcal{F}}}{\lambda_f+r}\hat{v}_{ij}-B_2\right)-\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}\left(x_2V_{12}(\hat{v}_{ij})+x_1V_1(\hat{v}_{ij})\right).$$

Rearranging, we have

$$(64)\quad V_{12}(\hat{v}_{ij})\left(1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}\right)=\frac{1}{N}\left(\frac{\pi_{\mathcal{F}}}{\lambda_f+r}\hat{v}_{ij}-B_2\right)-\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}V_1(\hat{v}_{ij}).$$

Substituting (60) into (61), I derive

$$\begin{aligned}&V_1(\hat{v}_{ij})\left(1+\frac{x_1\lambda_a}{\lambda_a+r}\right)= \\ &O(\hat{v}_{ij})\frac{\pi_{\mathcal{F}}}{\lambda_f+r}\hat{v}_{ij}-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}\left(x_1V_1(\hat{v}_{ij})+x_2V_{12}(\hat{v}_{ij})\right)+\frac{\lambda_a}{\lambda_a+r}x_2V_{12}(\hat{v}_{ij}) \\ &\qquad -\int_0^{\hat{v}_{ij}}\left(\frac{\pi_{\mathcal{F}}s}{\lambda_f+r}-\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}\left(x_1V_1(\hat{v}_{ij})+x_2V_{12}(\hat{v}_{ij})\right)\right)O'(s)ds\end{aligned}$$

Combining like terms gives,

$$V_1(\hat{v}_{ij})\left(1+\frac{x_1\lambda_a}{\lambda_a+r}\right)=O(\hat{v}_{ij})\frac{\pi_{\mathcal{F}}\hat{v}_{ij}}{\lambda_f+r}-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}x_1V_1(\hat{v}_{ij})$$
$$-\int_0^{\hat{v}_{ij}}\left(\frac{\pi_{\mathcal{F}}s}{\lambda_f+r}-\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}\left(x_1V_1(s)+x_2V_{12}(s)\right)\right)O'(s)ds$$
$$+\frac{\lambda_a}{\lambda_a+r}x_2V_{12}(\hat{v}_{ij})\left(1-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\right).$$

Taking a derivative of both sides we derive,

$$V_1(\hat{v}_{ij})'\left(1+x_1\frac{\lambda_a}{\lambda_a+r}\right)=O(\hat{v}_{ij})'\frac{\pi_{\mathcal{F}}\hat{v}_{ij}}{\lambda_f+r}+O(\hat{v}_{ij})\frac{\pi_{\mathcal{F}}}{\lambda_f+r}$$
$$-O'(\hat{v}_{ij})\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}x_1V_1(\hat{v}_{ij})-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}x_1V_1(\hat{v}_{ij})'$$
$$-\left(\frac{\pi_{\mathcal{F}}\hat{v}_{ij}}{\lambda_f+r}-\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}\left(x_1V_1(\hat{v}_{ij})+x_2V_{12}(\hat{v}_{ij})\right)\right)O'(\hat{v}_{ij})$$
$$-\frac{\lambda_a}{\lambda_a+r}x_2V_{12}(\hat{v}_{ij})O'(\hat{v}_{ij})\frac{r}{\lambda_f+r}+\frac{\lambda_a}{\lambda_a+r}x_2V_{12}(\hat{v}_{ij})'\left(1-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\right).$$

Simplifying,

$$V_1(\hat{v}_{ij})'\left(1+\frac{x_1\lambda_a}{\lambda_a+r}\right)= \tag{65}$$
$$O(\hat{v}_{ij})\frac{\pi_{\mathcal{F}}}{\lambda_f+r}-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}x_1V_1(\hat{v}_{ij})'$$
$$+\frac{\lambda_a}{\lambda_a+r}x_2V_{12}(\hat{v}_{ij})'\left(1-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\right).$$

Using (64), we find that

$$V_{12}(\hat{v}_{ij})'\left(1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}\right)=\frac{1}{N}\frac{\pi_{\mathcal{F}}}{\lambda_f+r}-\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}V_1(\hat{v}_{ij})'.$$

Rearranging

$$V_{12}(\hat{v}_{ij})'=\frac{\frac{1}{N}\frac{1}{\sigma M\mu_H\left(\lambda_f+r\right)}-\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}V_1(\hat{v}_{ij})'}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}.$$

Substituting into (65) gives

$$V_1(\hat{v}_{ij})'\left(1+\frac{x_1\lambda_a}{\lambda_a+r}\right)=O(\hat{v}_{ij})\frac{\pi_{\mathcal{F}}}{\lambda_f+r}-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}x_1V_1(\hat{v}_{ij})'$$
$$+x_2\frac{\lambda_a}{\lambda_a+r}\frac{\frac{1}{N}\frac{1}{\sigma M\mu_H\left(\lambda_f+r\right)}-\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}V_1(\hat{v}_{ij})'}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\left(1-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\right).$$

We next rearrange the equation so that $V_1(\hat{v}_{ij})'$ appears on the left hand side and all remaining terms appear on the RHS. We find that the coefficient on $V_1(\hat{v}_{ij})'$ is

$$1+\frac{x_1\lambda_a}{\lambda_a+r}+O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}x_1 \\ +x_2x_1\frac{\lambda_a}{\lambda_a+r}\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}\frac{1}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\left(1-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\right).$$

Simplifying we have

$$1+\frac{x_1\lambda_a}{\lambda_a+r}+O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}+x_2x_1\left(\frac{\lambda_a}{\lambda_a+r}\right)^2\frac{\frac{r}{\lambda_f+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\left(1-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\right).$$

The term on the RHS is

$$\frac{\pi_{\mathcal{F}}}{\lambda_f+r}\left(O(\hat{v}_{ij})+\frac{1}{N}\frac{\frac{x_2\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\left(1-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\right)\right).$$

Define $\psi(\hat{v}_{ij})$ by

$$\psi(\hat{v}_{ij})=\frac{\frac{\pi_{\mathcal{F}}}{\lambda_f+r}\left(O(\hat{v}_{ij})+\frac{1}{N}\frac{\frac{x_2\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\left(1-O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\right)\right)}{1+\frac{x_1\lambda_a}{\lambda_a+r}+O(\hat{v}_{ij})\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}+x_2x_1\left(\frac{\lambda_a}{\lambda_a+r}\right)^2\frac{\frac{r}{\lambda_f+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\left(1-O(\hat{v}_j)\frac{r}{\lambda_f+r}\right)}.$$

Then we have

$$V_1(\hat{v}_{ij})=\int_0^{\hat{v}_{ij}}\psi(s)\,ds+constant.$$

What remains is to derive *constant*. By (61) and (64), we see that

$$V_1(0)\left(1+\frac{x_1\lambda_a}{\lambda_a+r}\right)=\frac{x_2\lambda_a}{\lambda_a+r}V_{12}(0)$$

and

$$V_{12}(0)\left(1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}\right)=-\frac{1}{N}B_2-\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}V_1(0).$$

Rearranging this second equation gives

$$V_{12}(0)=\frac{-\frac{1}{N}B_2-\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}V_1(0)}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}.$$

Substituting into the first equation gives

$$V_1(0)\left(1+\frac{x_1\lambda_a}{\lambda_a+r}\right)=\frac{x_2\lambda_a}{\lambda_a+r}\frac{-\frac{1}{N}B_2-\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}V_1(0)}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}.$$

Therefore,

$$V_1(0)\left(1+\frac{x_1\lambda_a}{\lambda_a+r}+\frac{\frac{x_2\lambda_a}{\lambda_a+r}\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\right)=-\frac{1}{N}\frac{\frac{x_2\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}B_2.$$

Next, using (63) with (62) we find that

$$B_2=\frac{\pi_{\mathcal{F}}\mu_{H^c}}{\lambda_f+r}-\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}\frac{x_1+2x_2x_1\frac{\lambda_a}{\lambda_a+r}}{1+x_2\frac{\lambda_a}{\lambda_a+r}}\mathbb{E}[V_1(\hat{v}_{ij}),j\in\Omega_{it}^c].$$

Therefore,

$$V_1(0)\left(1+\frac{x_1\lambda_a}{\lambda_a+r}+\frac{\frac{x_2\lambda_a}{\lambda_a+r}\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\right)=$$
$$-\frac{1}{N}\frac{\frac{x_2\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\left(\frac{\pi_{\mathcal{F}}\mu_{H^c}}{\lambda_f+r}-\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}\frac{x_1+2x_2x_1\frac{\lambda_a}{\lambda_a+r}}{1+x_2\frac{\lambda_a}{\lambda_a+r}}\mathbb{E}[V_1(\hat{v}_{ij}),j\in\Omega_{it}^c]\right)$$

Define $V_1(0)=a+b\mathbb{E}[V_1(\hat{v}_{ij})]$ where the coefficients are defined by the above equation. That is,

$$a=\frac{-\frac{1}{N}\frac{\frac{x_2\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\frac{\pi_{\mathcal{F}}\mu_{H^c}}{\lambda_f+r}}{1+\frac{x_1\lambda_a}{\lambda_a+r}+\frac{\frac{x_2\lambda_a}{\lambda_a+r}\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}}$$

and

$$b=\frac{\frac{1}{N}\frac{\frac{x_2\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}\frac{r}{\lambda_f+r}\frac{\lambda_a}{\lambda_a+r}\frac{x_1+2x_2x_1\frac{\lambda_a}{\lambda_a+r}}{1+x_2\frac{\lambda_a}{\lambda_a+r}}}{1+\frac{x_1\lambda_a}{\lambda_a+r}+\frac{\frac{x_2\lambda_a}{\lambda_a+r}\frac{r}{\lambda_f+r}\frac{x_1\lambda_a}{\lambda_a+r}}{1+\frac{r}{\lambda_f+r}\frac{x_2\lambda_a}{\lambda_a+r}}}.$$

Then

$$\mathbb{E}[V_1(\hat{v}_{ij})]=\frac{1}{1-b}\left(a+\mathbb{E}\left[\int_0^{\hat{v}_j}\psi(s)ds\right]\right).$$

To conclude,

$$V_1(\hat{v}_{ij})=\int_0^{\hat{v}_{ij}}\psi(s)ds+a+\frac{b}{1-b}\left(a+\mathbb{E}\left[\int_0^{\hat{v}_j}\psi(s)ds\right]\right).$$

□

E. Extension: Ad Frequency

This Appendix extends the baseline model to allow platforms to select their ad frequencies.

Setup

We assume now that the CES aggregate over platform consumption is instead

$$X_{it} = \left[\int_{i \in \mathcal{D}} \left(q_{kt} \nu(A_{kt}) x_{ikt} \right)^{\frac{\rho-1}{\rho}} dk \right]^{\frac{\rho}{\rho-1}}$$

where $\nu : \mathbb{R}_+ \to [0,1]$ is decreasing. Each platform selects its ad frequency to maximize its flow profits at each point in time:

$$A = \arg\max_{A_{kt} \leq \overline{A}} \pi_{\mathcal{D}} A x_{kt}$$

where $\overline{A} > 0$ is some exogenous maximal ad frequency. All other aspects of the model are as in the baseline model.

Analysis

Theorem 6. *In the unique stationary equilibrium:*

1. *Consumer i's demand for platform k is*

$$x_{kt} = \frac{[\nu(A_{kt}) q_{kt}]^{\rho-1}}{\int_{\mathcal{D}} [\nu(A_{lt}) q_{lt}]^{\rho-1} \, dl}.$$

2. *Each platform sets ad frequency*

$$A = \arg\max_{A_{kt} \leq \overline{A}} A_{kt} \nu(A_{kt})^{\rho-1}.$$

3. *Welfare is $u(C, X)/r$ where C is as in Theorem 1 but $X = \nu(A) D^{\frac{1}{\rho-1}} q$ where q is as in Theorem 1.*

All other equilibrium properties are as in Theorem 1 of the baseline model.

Proof. We have already shown this in the main text in Section 5. □

Proposition 9. *Ad frequency A can be either too high or too low relative to what is socially optimal depending on parameters.*

Proof. Suppose that the discount rate is near zero. Then the social planner essentially sets the ad frequency on each platform to maximize consumers' flow utilities. Suppose that the flow utility is Cobb-Douglas with parameter τ. The planner chooses A to maximize

$$(A \mu_H)^{\frac{1-\tau}{\sigma-1}} \nu(A)^{\tau}.$$

Recall that μ_H is decreasing in A by Lemma 10. Suppose that σ and ρ both equal $3/2$. Then, if $\tau = 1/2$, equilibrium advertising is too high. On the other hand, for τ near 0, equilibrium advertising is too low as then the planner's choice of ad frequency diverges. □

Proposition 10. *Proposition 1 and Parts 1 and 3 of Proposition 2 hold in this setting with endogenous ad frequency.*

Proof. Note that

$$A = \arg\max_{A_{kt} \leq \overline{A}} \{\ln(A_{kt}) + (\rho - 1)\ln(\nu(A_{kt}))\}.$$

Taking a cross partial it is clear that the objective is submodular in ρ and A_{kt}. Therefore A is nondecreasing in ρ. The result then follows from Propositions 1–3. □

F. Extension: Network Effects

This Appendix extends the baseline model to allow for network effects.

Setup

We redefine the CES aggregate as

$$X_{it} = \left(\int_{i \in \mathcal{D}} \left(q_{kt} \eta(x_{kt}) x_{ikt} \right)^{\frac{\rho-1}{\rho}} dk \right)^{\frac{\rho}{\rho-1}}$$

where $\eta(x) = x^{\zeta}$ where $\zeta > 0$. All other aspects of the model are as in the baseline model.

Analysis

We assume the following throughout Appendix F.

Condition 4. $\frac{\rho-1}{1-\zeta(\rho-1)} \leq 2$.

Theorem 7. *In the unique stationary equilibrium with the property that each platform $k \in \mathcal{D}$ has $x_{kt} > 0$ at each t:*

1. *Consumer i's demand for platform k is*

$$x_{kt} = \frac{q_{kt}^{\frac{\rho-1}{1-\zeta(\rho-1)}}}{\int_{\mathcal{D}} q_{kt}^{\frac{\rho-1}{1-\zeta(\rho-1)}} dk}.$$

2. *All other equilibrium properties except for q and welfare are as in Theorem 1 of the baseline model except with $\frac{\rho-1}{1-\zeta(\rho-1)}$ in place of ρ.*

3. *Welfare is $u(C, X)/r$ where C is as in Theorem 1 but $X = \nu(A) D^{\frac{1}{\rho-1}} q^{\frac{1}{1-\zeta(\rho-1)}}$ where q is as in Theorem 1.*

Proof. In Section 8, we showed that

$$x_{kt}^{\zeta} = \frac{x_{kt}^{\zeta(\rho-1)} q_{kt}^{\rho-1}}{Y}$$

where

$$Y = \int_{\mathcal{D}} x_{kt}^{\zeta(\rho-1)} q_{kt}^{\rho-1}\, dk.$$

This implies

$$x_{kt} = \frac{q_{kt}^{\frac{\rho-1}{1-\zeta(\rho-1)}}}{Y^{\frac{1}{1-\zeta(\rho-1)}}}$$

or $x_{kt} = 0$. Since each platform receives $x_{kt} > 0$, integrating both sides over $k \in \mathcal{D}$ yields

$$Y^{\frac{1}{1-\zeta(\rho-1)}} = \int_{\mathcal{D}} q_{kt}^{\frac{\rho-1}{1-\zeta(\rho-1)}}\, dk.$$

This implies that

$$x_{kt} = \frac{q_{kt}^{\frac{\rho-1}{1-\zeta(\rho-1)}}}{\int_{\mathcal{D}} q_{kt}^{\frac{\rho-1}{1-\zeta(\rho-1)}}\, dk}.$$

The remaining parts of Theorem 1 are straightforward. □

G. Extension: Reserve Prices

This Appendix extends the baseline model with endogenous ad frequency as described in Section 8 by allowing each platform 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.

Analysis

The proposition below provides a characterization of the candidate equilibrium reserve price R and stationary distribution H. When there is a reserve of price, the effective rate at which individuals see ads per unit of attention is no longer simply A since sometimes there may not be a winner in an ad auction. Let Y denote the cutoff expectation such that any firm with expectation below Y in an auction will not bid above the reserve price.

Proposition 11. *In a stationary equilibrium with reserve prices the following hold:*

1. *The measure of firms below the cutoff Y solves the quadratic equation*

$$H^c(Y)\left[F - \frac{A}{\lambda_f}(1 - H^c(Y))\right] = FG(Y).$$

2. *The measure of firms in Ω_{it} is*

$$M = \frac{A}{\lambda_f}(1 - H^c(Y)).$$

3. *The cutoff Y solves*

$$Y = \left(1 - \frac{F}{F-M}G(Y)\right) \frac{\left(\frac{F}{F-M}\right)^{N-1} G(Y)^{N-1} + \frac{F-M}{M}\left(1 - \left(\frac{F}{F-M}\right)^{N} G(Y)^N\right)}{\frac{F}{M}g(Y)(1 - \left(\frac{F}{F-M}\right)^{N} G(Y)^N)}$$

which, given Parts 1 and 2, is an equation where Y is the only unknown.

4. *The stationary distribution H^c solves*

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

for $s \geq Y$. For $s \leq Y$,

$$H^c(s) = \frac{F}{F-M}G(s).$$

5. *Given H and income I, the coefficient of flow profits $\pi_{\mathcal{F}}$ is as in the baseline model.*

6. *The reserve price is*

$$R = \frac{\pi_{\mathcal{F}} Y}{\lambda_f + r}.$$

7. *The auction entry rate is*

$$\lambda_a = \frac{NA(1 - H^c(Y))}{F - M}.$$

8. *The bidding function is*

$$B(\hat{v}_{ij}) = \int_R^{\hat{v}_{ij}} \pi_{\mathcal{F}} \frac{1}{\lambda_f + r + \lambda_a H^c(s)^{N-1}} ds + R$$

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

All other equilibrium properties are as in Theorem 1 of the baseline model except with $A(1 - H^c(Y))$ in place of A.

Proof. Firm j's optimal bid must satisfy

$$B(\hat{v}_{ij}) = \frac{\pi_{\mathcal{F}}}{\lambda_f + r}\hat{v}_{ij} + \frac{\lambda_f}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r}V(\hat{v}_{ij}) - \frac{\lambda_a}{\lambda_a + r}V(\hat{v}_{ij}).$$

Let $Y = R\frac{\lambda_f + r}{\pi_{\mathcal{F}}}$. Above,

$$V(\hat{v}_{ij}) = \left(1 - H(\hat{v}_{ij})^{N-1}\right)\frac{\lambda_a}{\lambda_a + r}V(\hat{v}_{ij}) + \\ H^c(\hat{v}_{ij})^{N-1}\left(\frac{\pi_{\mathcal{F}}}{\lambda_f + r}\hat{v}_{ij} + \frac{\lambda_f}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r}V(\hat{v}_{ij}) - \mathbb{E}\left[\max\{B(\hat{v}^{(1)}), R\}|\hat{v}_{ij} > \hat{v}^{(1)}\right]\right)$$

when $\hat{v}_{ij} \geq Y$. Here, $\hat{v}^{(1)}$ is distributed according to the highest of $N-1$ draws from the distribution outside the consideration set. Letting $O(\hat{v}_{ij}) = H^c(\hat{v}_{ij})^{N-1}$, we have

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

Then

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

for $\hat{v}_{ij} \geq Y$. Then

$$B(\hat{v}_{ij}) = \frac{\pi_{\mathcal{F}}}{\lambda_f + r}\hat{v}_{ij} - \frac{r}{\lambda_f + r}\frac{\lambda_a}{r}\left(O(\hat{v}_{ij})B(\hat{v}_{ij}) - RO(R) - \int_Y^{\hat{v}_{ij}} B(s)O'(s)ds\right).$$

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

$$B(\hat{v}_{ij}) = \frac{\pi_{\mathcal{F}}}{\lambda_f + r}\int_Y^{\hat{v}_{ij}} \frac{1}{1 + \frac{\lambda_a}{\lambda_f + r}H^c(s)^{N-1}}ds + R$$

for $\hat{v}_{ij} \geq Y$. There is a multiplicity of equilibrium bids for $\hat{v}_{ij} < Y$.

Next, we derive the stationary distribution over expected values H. Matching inflows with outflows gives,

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

for all $s \geq Y$. Then 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 $MH + (F-M)H^c = FG$. Then

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

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

$$H^c(Y) = \frac{F}{F-M}G(Y).$$

Substituting into the above equation, we find that

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

for $s \geq Y$. This is a polynomial equation in $H^c(s)$. Next, we derive the first order condition for optimality of the cutoff Y. First, we derive $h^c(Y)$ by observing that

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

which implies

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

which rearranges to

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

Now we derive the first order condition. Consider the equation for profit:

$$\int_Y^{\overline{v}} B(s)[N(N-1)H^c(s)^{N-2}(1-H^c(s))h^c(s)]ds + Y\frac{1-\sigma}{M\mu_H(\lambda_f + r)}N(1-H^c(Y))H^c(Y)^{N-1}.$$

Note that implicitly μ_H depends on the reserve price chosen by the other platforms. Optimizing the above expression with respect to Y is effectively equivalent to optimizing with respect to the reserve price. Taking a first order condition we obtain,

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

Simplifying, 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 itself is a function of Y. We write the equation for Y in terms of primitives:

$$Y = \left(1 - \frac{F}{F-M}G(Y)\right) \frac{\left(\frac{F}{F-M}\right)^{N-1} G(Y)^{N-1} + \frac{F-M}{M}\left(1 - \left(\frac{F}{F-M}\right)^N G(Y)^N\right)}{\frac{F}{M}g(Y)(1 - \left(\frac{F}{F-M}\right)^N G(Y)^N)}.$$

□

Corollary 7.1. *In the limit as F tends to infinity, the candidate stationary equilibrium cutoff converges to the solution to*

$$Y = \frac{1 - G(Y)}{g(Y)}.$$

We can show the following comparative statics.

Proposition 12. *Parts 2 and 3 of Proposition 1 and Theorem 2 continue to hold in this setting. If ad frequency is endogenous as in Appendix E, then Part 1 of Proposition 1 also holds.*

We omit the proof of the above claim which follows using analogous arguments to those of the baseline model. We have not proven equilibrium uniquenes. However, the comparative statics in Proposition 1 will continue to hold across all stationary equilibria. Namely, there is a unique stationary equilibrium when data is uninformative. Ad revenue and investment will be maximal, across all data and all equilibria when G is uninformative. Part 1 of Proposition 1 will hold when ad frequency is endogenous since each platform will adjust is ad frequency to offset the effects of the reserve price. The following analog of Proposition 2 will also hold as long as stationary equilibrium is unique.

Proposition 13. *As long as stationary equilibrium exists and is unique, if interoperability ρ increases then:*

1. *Aggregate product consumption C decreases.*
2. *Ad revenue $\pi_{\mathcal{D}}A$ decreases.*
3. *Investment $L_{\mathcal{D}}$ increases.*

Proof. The reserve price and cutoff do not depend on ρ. Thus the same argument as in Proposition 2 of the baseline model applies. □

We have not proven existence of stationary equilibrium but one can derive sufficient conditions. We omit a formal argument. As long as F is sufficiently large relative to A and G is strictly regular (in that the derivative of its virtual value is bounded below by a positive constant), one can show existence.

H. Extension: Entry

This Appendix extends the baseline model to allow for entry of firms and platforms.

Setup

Suppose that entry by a firm requires $e_{\mathcal{F}}$ units of labor and entry by a platform requires $e_{\mathcal{D}}$ units of labor. We retain all other aspects of the baseline model except that now the measure of firms F and the measure of platforms D are such that both firms and platforms earn zero profits from entry.

Analysis

The following Theorem 8 characterizes the stationary equilibrium with entry.

Theorem 8. *In the unique stationary equilibrium, the following hold:*

1. *The measure of firms in Ω_{it} is M and the CDFs of values inside Ω_{it} and Ω_{it}^c are H and H^c respectively where M, H, and H^c are as in Lemma 3.*

2. *Income is $I = L$.*

3. *Demands, prices, bidding, ad revenue, and platforms' investments and quality levels are as in Lemmas 1-5.*

4. *The measure of platforms satisfies*

$$D = \frac{\pi_{\mathcal{D}} A}{e_{\mathcal{D}}} \left(1 - \frac{\varphi\delta(\rho - 1)}{r + (1 - \alpha)\delta}\right).$$

5. *The measure of firms satisfies*

$$F = \frac{\frac{L}{\sigma} - \pi_{\mathcal{D}} A}{e_{\mathcal{F}}}.$$

6. *Welfare is $u(C, X)/r$ where $C = \frac{\sigma-1}{\sigma} L (M\mu_H)^{\frac{1}{\sigma-1}}$ and $X = D^{\frac{1}{\rho-1}} q$.*

Proof. Note that $I = L$ since firms' and platforms' profits are driven to zero by entry. All other equilibrium properties besides D and F are determined as before in the baseline model. This proves Parts 1-3. Parts 4 and 5 follow from zero profit conditions. The stationary equilibrium is unique because there is a unique solution for F to the equation in Part 5. This is because $\pi_{\mathcal{D}} A$ is an increasing function of F as shown in Proposition 4. Part 6 follows straightforwardly using earlier parts of Theorem 8. □

Using Theorem 8 together with previous results we can derive the following comparative statics. The proofs are straightforward so we omit them.

Proposition 14. *If interoperability ρ increases then:*

1. *There is no change in aggregate product consumption C.*
2. *Investment $L_{\mathcal{D}}$ increases.*
3. *There is no change in ad revenue $\pi_{\mathcal{D}}A$.*

Proposition 15. *If ad frequency A increases then:*

1. *Aggregate product consumption C increases.*
2. *Ad revenue $\pi_{\mathcal{D}}A$ decreases.*
3. *Investment $L_{\mathcal{D}}$ decreases.*

Proposition 16. *When data G is uninformative:*

1. *Aggregate product consumption C is minimal.*
2. *Ad revenue $\pi_{\mathcal{D}}A$ is maximal.*
3. *Investment is $L_{\mathcal{D}}$ maximal.*

Proposition 17. *If the entry cost of firms increases $\epsilon_{\mathcal{F}}$ then:*

1. *The measure of firms F decreases.*
2. *The measure of platforms D decreases.*

References

ALCOBENDAS, M., S. KOBAYASHI, AND M. SHUM (2021): "The Impact of Privacy Measures on Online Advertising Markets," *Available at SSRN 3782889*.

ANDERSON, S. P., AND S. COATE (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., AND J. S. GANS (2010): "The Impact of Targeting Technology on Advertising Markets and Media Competition," *American Economic Review*, 100(2), 608–13.

BEALES, H., AND J. A. EISENACH (2014): "An empirical analysis of the value of information sharing in the market for online content," *Available at SSRN 2421405*.

BERGEMANN, D., A. BONATTI, AND T. GAN (2019): "The Economics of Social Data," .

——— (2022): "The economics of social data," *The RAND Journal of Economics*.

BERGEMANN, D., B. BROOKS, AND S. MORRIS (2015): "The limits of price discrimination," *American Economic Review*, 105(3), 921–57.

BERGEMANN, D., T. HEUMANN, S. MORRIS, C. SOROKIN, AND E. WINTER (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., AND B. JULLIEN (2003): "Chicken & egg: Competition among intermediation service providers," *RAND journal of Economics*, pp. 309–328.

CHEN, D., AND D. DUFFIE (2021): "Market Fragmentation," *American Economic Review*, 111(7), 2247–74.

DIXIT, A. K., AND J. E. STIGLITZ (1977): "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67(3), 297–308.

DUFFIE, D., L. QIAO, AND Y. SUN (2020): "Continuous-Time Random Matching," *Graduate School of Business, Stanford University*.

EDELMAN, B., M. OSTROVSKY, AND M. SCHWARZ (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. (2019): “Attention platforms, the value of content, and public policy,” *Review of Industrial Organization*, 54(4), 775–792.

——— (2020): “The economics of attention markets,” *Available at SSRN 3044858.*

Farboodi, M., and L. Veldkamp (2021): “A Growth Model of the Data Economy,” Discussion paper, National Bureau of Economic Research.

Hummel, P., and R. P. McAfee (2016): “When Does Improved Targeting Increase Revenue?,” *ACM Transactions on Economics and Computation (TEAC)*, 5(1), 1–29.

Ichihashi, S. (2020): “Online Privacy and Information Disclosure by Consumers,” *American Economic Review*, 110(2), 569–95.

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

Iyer, K., R. Johari, and M. Sundararajan (2014): “Mean Field Equilibria of Dynamic Auctions with Learning,” *Management Science*, 60(12), 2949–2970.

Johnson, G. A., S. K. Shriver, and S. Du (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., and C. Tonetti (2020): “Nonrivalry and the Economics of Data,” *American Economic Review*, 110(9), 2819–58.

Jullien, B., A. Pavan, and M. Rysman (2021): “Two-sided markets, pricing, and network effects,” in *Handbook of Industrial Organization*, vol. 4, pp. 485–592. Elsevier.

Khan, L. (2017): “Amazon’s antitrust paradox,” *Yale law journal*, 126.

Marotta, V., V. Abhishek, and A. Acquisti (2019): “Online tracking and publishers’ revenues: An empirical analysis,” in *Workshop on the Economics of Information Security.*

McAfee, R. P. (1993): “Mechanism Design by Competing Sellers,” *Econometrica*, pp. 1281–1312.

Newman, J. M. (2019): “Antitrust in Attention Markets: Objections and Responses,” *Santa Clara L. Rev.*, 59, 743.

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

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

Rochet, J.-C., and J. Tirole (2003): "Platform Competition in Two-Sided Markets," *Journal of the European Economic Association*, 1(4), 990–1029.

Silk, A. J., E. R. Berndt, 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.