# The Market for Attention

Daniel T. Chen*

2 Feb 2022, revised 28 April 2022

PRELIMINARY AND INCOMPLETE

Click here for the latest version. This draft is updated frequently.

**Abstract**

This paper builds a dynamic general equilibrium model of the market for attention. In the model, digital platforms compete for consumer attention by investing in the quality of their services which they provide for free. They sell the attention, in the form of advertisements, to firms in the product market who use consumer data to target. We characterize search frictions in the product market, ad revenues, platforms' quality levels, and welfare in the unique stationary equilibrium. Banning data tracking, capping ad frequency, and enforcing platform interoperability may often lead to better platforms at the expense of worse product consumption. Platform investment and ad frequency may be too high or low depending on parameters. The sources of inefficiency concern appropriability, business stealing, and market power.

Keywords: digital platforms, advertising, consumer data, auctions, competition, search frictions, welfare

*dtchen@stanford.edu; Graduate School of Business, Stanford University

I thank Yuliy Sannikov, Andy Skrzypacz, Darrell Duffie, Bob Wilson, Mohammad Akbarpour, Susan Athey, Kyle Bagwell, Anirudha Balasubramanian, Martino Banchio, Sebastian Bauer, Sebastian Di Tella, Ramesh Johari, Chad Jones, Nadia Kotova, David Kreps, Ruilin Li, Lina Lukyantseva, Suraj Malladi, Hala Moussawi, Mike Ostrovsky, David Ritzwoller, Chris Tonetti, Takuo Sugaya, David Yang, Timur Sobolev, Frank Yang, Anthony Zhang, Charles Zhang, and Weijie Zhong for insightful comments and suggestions.

# 1 Introduction

There has been considerable debate over how to regulate digital platforms that profit primarily from targeted advertising. Some argue that while many platforms offer distinct services, they should still be considered competitors within a broader "market for attention" (Evans, 2020; Wu, 2018; Prat and Valletti, 2021; Newman, 2019). In this market, digital platforms compete for consumer attention by investing in the quality of their services which they provide for free. They then sell the attention, in the form of advertisements, to firms in the product market who use consumer data for targeting.

Are consumers well compensated for their attention? Do platforms extract too much profit from product firms? What are the welfare effects of policies that ban data tracking, cap ad frequency, or enforce platform interoperability?[1]

Answering these questions is key to advancing effective policy. However, most existing competition analysis is based on more traditional markets which are single-sided and where consumers are charged prices. As a result, regulators have primarily relied on markups to gauge market efficiency. But, this has no bite in the market for attention where platforms offer consumers their services for free. Moreover, the market for attention is complex and multi-sided: outcomes in the product market, ad revenues, and platform quality levels are jointly determined in equilibrium. It is therefore difficult to assess the effects of policies which, by affecting any one side of the market, affect them all. A formal model may help overcome some of these roadblocks.

This paper builds a general equilibrium model of the market for attention. The model shows how platforms provide quality services to consumers, mitigate search frictions in the product market, and extract ad revenues from product firms. We characterize platforms' quality levels, search frictions, and ad revenues in the unique stationary equilibrium. We show how these depend on data informativeness, platform interoperability, ad frequency

[1]Broadly speaking, platform interoperability refers to how easily consumers can switch from using one platform to another. It is often suggested as a tool for competition policy. See, for example Scott Morton, Crawford, Crémer, Dinielli, Fletcher, Heidhues, Schnitzer, and Seim (2021).

(exogenous in the baseline model but endogenized in an extension), and product firm market power. We compare the equilibrium to a first best benchmark. We identify the sources of inefficiency.

We now summarize the baseline model before discussing the main results. There is a monopolistically competitive product market as in Dixit and Stiglitz (1977). Consumers have CES utility over product consumption but have different values (share parameters) for the different products. A consumer's values are his private information. The model is standard except that each consumer is aware of only a subset of the available products in the market. A consumer's consideration set evolves over time as he discovers firms by viewing ads and gradually forgets about the ones he is currently aware of.

Consumers view ads while using platforms. Platforms are monopolistically competitive.[2] Consumers have CES utility over platform consumption. Each consumer splits a unit of attention across the various platforms at each point in time. (Later, when we analyze welfare, we endogenize the total share of attention spent on platforms.) In equilibrium, better quality platforms receive a higher share of attention. (For the baseline model, we ignore network effects, but we later introduce them in an extension.)

Platforms sell ads to product firms via a process that is modeled after Real-Time Bidding (RTB) which is widely used in practice.[3] Namely, an opportunity arises to display an ad to a given consumer on a given platform at a Poisson rate equal to that platform's attention share multiplied by a parameter which we refer to as the ad frequency. Each time an opportunity arises, the platform invites a random finite number of firms to bid in an ad auction. The highest bidding firm wins, displays its ad to the consumer, and then enters the consumer's consideration set. The identity of the platform where an ad is shown is inconsequential: all platforms are identical gateways into the consumer's consideration set.

To determine their optimal bids, firms rely on consumer data to form expectations of consumers' values and the profits from entering their consideration sets. When bidding in

[2]One may also interpret the platforms in the model as individual content creators.

[3]See https://www.facebook.com/business/help/430291176997542?id=561906377587030.

an auction for a given consumer, a firm internalizes the competition that it will face from the other firms already in consideration by the consumer. It also internalizes the outside option of bidding in future ad auctions. By bidding more aggressively, a firm can enter the consideration set of a consumer more quickly. We are therefore able to show how endogenous search frictions affect platforms' ad revenues.

Each platform seeks to steal attention away from its rivals by investing in the quality of its services in order to run more ad auctions and earn more ad revenue. We assume that investment requires labor which can also be used for production. Thus, society faces a tradeoff between better quality platforms and higher consumption of products. How easily platforms can steal attention from each other is determined by platform substitutability which we also refer to as platform interoperability.

We now summarize the model's main results. Our first set of results concern comparative statics on final product and platform consumption with respect to changes in the informativeness of consumer data, ad frequency, platform interoperability, and product firm market power. We find that a reduction in data informativeness (ex. from a ban on cookie tracking) or binding cap on ad frequency worsens search frictions but often leads to higher investment and thus better quality platforms. This is because when search frictions are worse, the value to a firm of displaying an ad to a consumer is higher and so firms bid more aggressively leading to higher ad revenues and more incentives for platforms to invest to steal attention. The end result is that typically, final product consumption decreases while final platform consumption increases.

We also find that enforcing greater platform interoperability and allowing higher product firm market power also leads to lower final product consumption and higher final platform consumption. When platform interoperability is higher, platforms can more easily steal attention from their rivals so that in equilibrium there is more investment in platform quality and less production. When firms have higher market power, they earn more profits from selling to consumers and thus the value of displaying ads is higher leading to higher platform ad revenue and thus greater investment in platform quality and less production. While greater interoperability can promote investment, once it reaches a certain level, stationary equilibrium fails to exist.

The main sources of inefficiency in the model are the familiar culprits of appropriability, business stealing, and market power. However, due to the multi-sided nature of the market, these inefficiencies interact in perhaps unfamiliar ways, which we now describe. Firstly, platforms are unable to appropriate any of the surplus which they generate for consumers. This includes both the surplus from exposing them to a wider variety of products and from the services which they provide. This is a direct consequence of offering their services for free and is a force tending towards underinvestment.

However, the surplus which platforms collectively appropriate from firms may *exceed* the total surplus that they generate for firms. Upon first thought this may appear surprising. The reason behind it is that platforms generate no surplus for product firms: advertising merely shifts business from one firm to another. A firm's incentive to purchase ads is driven entirely by business stealing. Platforms are able to profit from this business stealing motive. This is a force tending towards overinvestment.

On its own, ad revenues are not enough to incentivize platforms to invest. Attention must be elastic (as determined by platform interoperability) so that platforms can steal attention from each other. This externality also appears in Anderson and Coate (2005) where it is shown that business stealing may lead to excessive entry by television programs. This second business stealing externality is another force leading to excessive investment.

Finally, not only do platforms fail to internalize the social value of their investments, they also fail to internalize the social costs. That is, there are distortions in the price of labor. One reason for this is that the market power of product firms leads them to have lower demand for labor than then the socially efficient level. This is a force leading to excessive investment. A second reason is that because platforms fail to internalize the social value of investments, their demand for labor is also distorted causing additional distortions in the price of labor.

The rest of the paper briefly discusses several important extensions. In the first extention, we include a nuisance cost to consumers from encountering ads and endogenize the rate at which ads are shown. We find that ad frequency may be excessive or insufficient. In the second extension, we incorporate network effects. We show, under an equilibrium refinement, that network effects are equivalent to an increase in platform substitutability and so most of the baseline model's analyses extend almost immediately. In the third extension we allow platforms to set reserve prices which are shown to be positive and can be inefficient. In the last extension we extend the baseline model to allow platforms to have different data. Under some conditions, we show that bidding strategies differ across platforms and are determined by the fixed point of a contraction mapping and can therefore be computed by an iterative procedure.

The rest of the paper proceeds as follows. Section 2 reviews the related literature. Section 3 introduces the baseline model. Section 4 characterizes the unique stationary equilibrium. Section 5 presents various comparative statics results. Section 6 compares the unique stationary equilibrium to a first best benchmark and identifies the potential sources of inefficiency. Section 7 summarizes the results of various extensions. Section 8 concludes. Appendices contain omitted proofs and additional results.

## 2 Related Literature

This paper contributes to a growing literature studying various aspects of digital platforms. Following Rochet and Tirole (2003), one line of work studies the optimal prices set by platforms on both sides of a two-sided market. These papers show that platforms set a low price on one side of the market in support of a high price on the other side due to network effects. In contrast, we assume that platforms charge consumers zero prices from the outset and that the prices platforms charge firms are determined in ad auctions. We also focus on platforms' investments in the quality levels of their services. Relative to Rochet and Tirole (2003) and its subsequent literature, we tailor our analysis more closely to the market for attention.

This paper is also related to Bergemann, Bonatti, and Gan (2019) and Acemoglu, Makhdoumi, Malekian, and Ozdaglar (2019). As with these papers, we view platforms as intermediaries who facilitate the matching of consumers to products using data. But beyond this similarity, our papers focus on different issues. Bergemann, Bonatti, and Gan (2019) and Acemoglu, Makhdoumi, Malekian, and Ozdaglar (2019) focus on the efficient use of data when there are externalities where data on one consumer is informative of other consumers. They also assume that consumers have either an intrinsic benefit from privacy or there is personalized pricing. In contrast, we shut down personalized pricing and instead focus on whether investment by platforms is efficient. This also sets this paper apart from Ali, Lewis, and Vasserman (2020) and Ichihashi (2020) who study the welfare consequences of personalized pricing.

Another line of literature studies the macroeconomic implications of digital platforms (data intermediaries) for growth. In Jones and Tonetti (2020) and Farboodi and Veldkamp (2021), data is a nonrival good which can be used to improve production. In our paper, data serves a different purpose which is to help firms target consumers with high values for their products. In Jones and Tonetti (2020) and Farboodi and Veldkamp (2021) a loss of privacy may be harmful because consumers have an intrinsic value for privacy. In contrast, in our model, a loss of privacy may be harmful because of its effect on ad revenues and therefore platforms' incentives to invest in quality services.

This paper contributes to the literature on online advertising (Edelman, Ostrovsky, and Schwarz, 2007; Athey and Ellison, 2011; Athey and Gans, 2010; Varian, 2007; Hummel and McAfee, 2016; Board, 2009). Most of these papers are squarely auction-focused. They do not consider the endogenous competition by advertisers in the product market. An important exception is Athey and Ellison (2011). In Athey and Ellison (2011) there is a single auction and platform. In contrast we study competition by platforms which each host ad auctions. To our knowledge, our dynamic model introduces novel effects such as the effect of the outside option to bid in future ad auctions on equilibrium bids. We also show how search frictions in the product market affect bidding in the ad auctions and how bidding in turn affects the long run level of search frictions.

This paper also relates to the literature on traditional advertising (see Bagwell, 2007 for a survey). While many papers in this literature do incorporate the product market into their analyses, most do not account for the services which platforms provide. Thus, most papers subtract advertising costs from social welfare (see for example Grossman and Helpman, 1991). In contrast, we emphasize the impact of platforms' services on social welfare. An important exception and perhaps the most closely related paper is Anderson and Coate (2005). As in our paper, Anderson and Coate (2005) find that advertising and entry of new television programs can be excessive or insufficient due to issues with appropriability and business stealing externalities.

However, in Anderson and Coate (2005), consumers have binary valuations and therefore reap no surplus from matching with firms who are monopolists and face no competition in the product market. Anderson and Coate (2005) focuses on the market for television programming. In contrast, our model is built to match certain features of the microstructure for the sale of digital ads. For example, one feature we incorporate is the individual level ad-targeting which uses consumer data.

Methodologically, this paper contributes to the literature on competing auction designs in dynamic settings (Chen and Duffie, 2021; McAfee, 1993; Wolinsky, 1988; Iyer, Johari, and Sundararajan, 2014). Of these, the closest is Wolinsky (1988) where a dynamic option to wait matters for bidding. However, once a buyer is matched to an auction for a seller and wins, both the buyer and seller exit the market. The model does not match the microstructure of Real-Time Bidding (RTB) in dynamic ad auctions as ours does where multiple firms may advertise to the same consumer repeatedly over time. In our model, the dynamic option to wait is heterogeneous across firms and depends on a firm's expectation of the consumer's value for its product. In Wolinsky (1988), the option value of waiting is the same for all buyers regardless of their values for the items being sold.

Methodologically, this paper is related to Duffie, Gârleanu, and Pedersen (2005) which develops a search-based asset pricing theory in Over-the-Counter markets. Like the dealers in that model, digital platforms mitigate search frictions. Firms and consumers are brought together by a random search and matching process similar to the investors and dealers in Duffie, Gârleanu, and Pedersen (2005). Like Duffie, Gârleanu, and Pedersen (2005), we characterize long run search frictions and intermediaries' profits. Because the memory of a consumer in our model has some persistence, an impression is an asset for which we develop a search-based dynamic asset pricing theory. In contrast to Duffie, Gârleanu, and Pedersen (2005), we model the matching process after RTB. Moreover, in addition to setting intermediation intensity (ad frequency in our model), intermediaries provide valuable services which require investment.

# 3 Baseline Model

This section presents the setup of the baseline model. In the model, platforms compete for consumer attention by investing in the quality of their services. As consumers engage with the platforms, they view ads for different product firms. These ads are sold by platforms to firms in dynamic ad auctions. Thus, platforms provide utility directly to consumers with their services, ease search frictions in product markets by displaying ads, and extract ad revenues. In what follows, time is continuous, starts at $t = 0$, and goes on forever. All agents discount time at rate $r > 0$.

## 3.1 Demand

There is a unit continuum of consumers $i \in \mathcal{C}$. Each consumer $i$ derives utility from both products and platform services. There is a measure $F > 0$ of products $j \in \mathcal{F}$ and a measure $D > 0$ of platforms $k \in \mathcal{D}$. Consumer $i$'s flow utility at time $t$ is equal to $u(C_{it}, X_{it})$ where $C_{it}$ and $X_{it}$ are 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 consumer $i$'s consumption of product $j$, and $v_{ij}$ is consumer $i$'s value for product $j$.[4] Each value $v_{ij}$ is drawn from the cdf $P$, supported on $[0, \overline{v}]$, independently across $i \in \mathcal{C}$ and $j \in \mathcal{F}$. Values are realized at $t = 0$ and fixed forever after.

In (2), $\rho > 1$ is the level of platform substitutability, $x_{ikt}$ is the share of attention spent on platform $k$, and $q_{kt}$ is platform $k$'s quality level which all consumers agree on. A platform's quality level may capture how enticing its content is, the efficacy of its content recommendation alogrithm, or how enjoyable it is to use the user interface.

(1) and (2) imply that consumer $i$ enjoys using a variety of products and platforms. However, while consumer $i$ is aware of all platforms in $\mathcal{D}$, he is aware of only a subset $\Omega_{it}$ of products in $\mathcal{F}$. $\Omega_{it}$ is the consumer's time-$t$ *consideration set* and he may only consume products from this set at time $t$. At each time, consumer $i$ has income $I$ which he spends on products and a unit of attention which he spends on platforms.

We assume that the flow utility $u : \mathbb{R}^2_+ \to \mathbb{R}$ is increasing. Then, to maximize flow utility, consumer $i$ solves two separate problems:

[4]The exponent on $v_{ij}$ in (1) is a convenient normalization that eases notation.

(3) $$\max_{\{c_{ijt}\}_{j\in\Omega_{it}}} C_{it}$$

$$\text{s.t} \int_{\Omega_{it}} p_{jt} c_{ijt} dj = I.$$

(4) $$\max_{\{x_{ikt}\}_{k\in\mathcal{D}}} X_{it}$$

$$\text{s.t} \int_{\mathcal{D}} x_{ikt} dk = 1.$$

Following Dixit and Stiglitz (1977) we derive consumer $i$'s demands.

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

$$c_{ijt} = \frac{I v_{ij}}{\int_{\Omega_{it}} v_{il} p_{lt}^{1-\sigma} dl} p_{jt}^{-\sigma}. \tag{5}$$

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

$$x_{kt} = \frac{q_{kt}^{\rho-1}}{\int_{\mathcal{D}} q_{lt}^{\rho-1} dl}. \tag{6}$$

In (6) we omit the index $i$ since all consumers spend the same share of attention on any given platform. Notice that in (5), the consumer's values and the prices set by the other firms appear in a term that scales demand. However, demand is still CES with parameter $\sigma$.

## 3.2 Production and Pricing

Each product $j \in \mathcal{F}$ is produced by a distinct firm which we also refer to as $j$. Labor is the only factor of production. The production technology is linear: output of product $j$ is equal to the quantity of labor hired by firm $j$. Each consumer supplies $L$ units of labor inelastically at each point in time. We let the wage be the numeraire.

At each time $t$, firm $j$ selects a price $p_{jt}$ to maximize its flow profit accruing from all consumers $i$ such that $j \in \Omega_{it}$.

Proposition 2. *If, at time $t$, consumers' demands are as in (5), then*

*1. if firm $j$ is in $\Omega_{it}$, its flow profit from selling to consumer $i$ is*

$$\frac{I v_{ij}}{\int_{\Omega_{it}} v_{il} p_{lt}^{1-\sigma} dl} p_{jt}^{-\sigma}(p_{jt} - 1). \tag{7}$$

*2. firm $j$'s profit-maximizing price is*

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

*Even if firm $j$ could personalize prices, it is optimal for firm $j$ not to do so.*

Upon first thought, a firm's optimal price should depend on its beliefs over consumers' values, consideration sets, and prices set by other firms. However, Proposition 2 states that it is determined only by product substitutability $\sigma$. The price in (8) maximizes the flow profit in (7) corresponding to any consumer $i$. Thus, it is optimal for firm $j$ *not* to personalize prices. The simplicity of a firm's pricing problem under CES utility is key for tractability in this model. Changes to ad frequency or data informativeness (described later) will have no effects on a firm's optimal price. We therefore shut down any welfare effects operating through personalized pricing in order to focus on other issues.[5]

## 3.3 Consumer Data

Each firm is endowed with consumer data (signals of consumers' valuations) which it uses to form conditional expectations of consumers' values for their products. We do not model data and belief-updating explicitly. Instead, we assume that firm $j$'s expectation $\hat{v}_{ij}$ of consumer $i$'s value $v_{ij}$ is drawn from the continuous cdf $G$, supported on $[0, \overline{v}]$, independently across $i \in \mathcal{C}$ and $j \in \mathcal{F}$. By Blackwell (1953), $P \succ_{MPS} G$ is a necessary and sufficient condition for there to exist some signal structure which generates $G$. We

[5] The results of Bergemann, Brooks, and Morris (2015) show that the welfare effects of personalized pricing may depend sensitively on the information structure. Shutting down these effects provides a clean benchmark.

are therefore able to incorporate consumer data tractably and nonparametrically. Firms' expectations are realized at $t = 0$ and fixed forever after.

### 3.4 Search and Matching: Consideration Sets and Ad Auctions

A consumer's consideration set evolves over time as the consumer discovers firms by viewing ads and forgets about the firms it is currently aware of. The consumer views ads at a constant rate $A$ while using the platforms. Each ad corresponds to a firm which the consumer is not yet aware of. The firm enters the consumer's consideration set in the same instant the ad is shown. Once inside, the firm does not remain there forever. The consumer forgets about it at a random exponential time occurring at rate $\lambda_f$.

We assume that the Poisson rate at which the consumer sees ads on a given platform $k$ is proportional to its attention share $Ax_{kt}$.[6] The identity of the platform where the ad is shown does not matter: all platforms are identical gateways into the consumer's consideration set. Each time there is an opportunity to show consumer $i$ an ad, platform $k$ invites an exogenous finite number $N$ of firms from outside the consideration set to bid in a second-price auction. The firms are chosen uniformly at random. The auction is held at the "start of the same instant" the ad is shown. There is no reserve price.[7]

This model is based on the Real-Time Bidding process widely used in practice.[8] We assume that the platform invites only firms from outside the consideration set of the consumer for analytical tractability. In practice, platforms have some awareness of a consumer's consideration set which they may use to tailor auction invitations. For example, firms often allow platforms to place tracking cookies on their retail websites to determine when a consumer has stopped visiting.[9] We assume that the platform invites only a finite number

[6]By ELLN the total rate at which ads are seen is almost surely $A$ as stated earlier.

[7]We extend the model to incorporate reserve prices later in the Appendix. We also show that revenue-equivalence applies in this setting so that all results extend to other standard auction formats including the first-price auction.

[8]For example, see www.facebook.com/business/help/430291176997542?id=561906377587030.

[9]This is why ad retargeting, which is widespread and has a high return on ad spend, possible.

of firms due to the technological constraint of having to run an auction and display an ad within the instant.[10]

Initially, consumer $i$'s consideration set $\Omega_{i0}$ contains a positive measure $M_0 < F$ of products. The initial empirical cdf over expectations $\hat{v}_{ij}$ within $\Omega_{i0}$ is $H_0$ which satisfies $M_0 dH_0 < F dG$.[11] Firms are symmetrically distributed across initial consideration sets $(\Omega_{i0})_{i\in\mathcal{C}}$ in a way that is consistent with $M_0$ and $H_0$.

Assuming the exact law of large numbers (ELLN), the size of consumer $i$'s consideration set $M_t = |\Omega_{it}|$ evolves according to

$$dM_t = \left(A - \lambda_f M_t\right) dt. \tag{9}$$

In contrast, firms' bidding strategies determine the law of motion of the cdf $H_t$ over expectations within consideration sets. Assuming ELLN, if the bidder with the highest expectation wins in any auction, then $H_t$ evolves according to

$$d(M_t H_t) = \left(A\left(H_t^c\right)^N - \lambda_f M_t H_t\right) dt. \tag{10}$$

Above, $H_t^c$ denotes the distribution over expectations outside of the consideration set. It satisfies the accounting identity

$$M_t H_t + (F - M_t)\, H_t^c = FG. \tag{11}$$

Using equations (9) (10), and (11), we derive the steady state size and composition of consideration sets reported below in Proposition 3. To ensure that a positive measure of

[10]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. In Private Market RTB (which accounts for the majority of RTB ad revenues), only a few firms are invited to a given private auction. See https://www.emarketer.com/content/private-marketplace-ad-spending-to-surpass-open-exchange-in-2020.

[11]There can not be more firms of a given type in a consideration set than in the economy.

firms remain outside of consideration, we assume the following.

Condition 1. *It holds that* $A/\lambda_f < F$.

Proposition 3. *Suppose, in each auction for consumer* $i$*, that the firm* $j$ *with the highest expectation* $\hat{v}_{ij}$ *wins. Then under Condition 1 there is a unique size* $M$ *and cdf* $H$ *such that if* $H_0 = H$ *and* $M_0 = M$*, then* $M_t = M$ *and* $H_t = H$ *for all* $t > 0$*. Moreover,*

$$M = \frac{A}{\lambda_f} \tag{12}$$

*and*

$$H = (H^c)^N \tag{13}$$

*where*

$$M\,(H^c)^N + (F - M)\,H^c = FG. \tag{14}$$

Equation (14) states that the stationary distribution $H$ is determined by a simple polynomial equation which has a unique solution and can be solved pointwise. From it, we see that the positive selection of firms in ad auctions leads to $H \succ_{fosd} G \succ_{fosd} H^c$. The distribution $H$ determines the efficiency of the matching of firms and consumers in steady state. The distribution $H^c$ determines the demand and competition for the intermediation services provided by platforms in steady state.

We derive Proposition 3 using only the conjectured property that the firm with the highest expectation wins in any auction. We verify this next when we show that firms' bids are monotone in their expectations in steady state.

### 3.5 Real-Time Bidding

Each firm $j$ selects a bidding strategy to maximize the NPV of flow profits from product sales net of advertising costs. Flow profits accrue from the ever-changing set of consumers which are aware of firm $j$.

Firm $j$'s optimal bid in an auction for consumer $i$ depends on

1. it's expectation of the flow profit from selling to the consumer now and in the future.

2. the rate at which it enters ad auctions for consumer $i$.

3. the bidding strategies of his opponents.

Suppose that firm $j$ believes that the distribution and measure of varieties in consideration sets are fixed at a steady state level $H$ and $M$.

Let $\mu_H$ denote the mean of $H$. Then by (7), if all firms set prices optimally, firm $j$'s expected flow profit from selling to consumer $i$ at any time $t$ is

$$\mathbb{E}\left[\frac{I}{\sigma \int_{\Omega_{it}} v_{il}\, dl} v_{ij} \middle| \hat{v}_{ij}\right] = \pi_{\mathcal{F}} \hat{v}_{ij} \tag{15}$$

where

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

Above, (15) follows from ELLN. Notice that the flow profit is decreasing in the cumulative match value $M\mu_H$. Firm $j$ internalizes the match value of its rivals when bidding.

Given $M$, we assume that each firm perceives itself to be invited to an auction for consumer $i$ at an exponential rate $\lambda_a = NA/(F - M)$ when outside $\Omega_{it}$. This assumption can be given a formal foundation by considering the continuous-time limit of a sequence of discrete-time models.[12] Consumer $i$ internalizes the outside option of waiting for future ad

[12] We sketch the argument here. Suppose that at each $t = n\Delta$ a positive measure $A\Delta$ of auctions are held for a given consumer $i$ which implies that $NA\Delta$ measure of firms are invited to an auction at $t = n\Delta$. Since these are chosen uniformly at random from a measure $F - M$ of firms, each firm perceives itself to enter an auction at time $t = n\Delta$ with probability $NA\Delta/(F - M)$ if it has not yet entered an auction at a prior

auctions when bidding.

The bidding strategies of firm $j$'s rivals are determined in the stationary equilibrium of the bidding subgame. Namely, a stationary equilibrium of the bidding subgame is a function $B : \mathbb{R}_+ \to \mathbb{R}$ such that each firm $j$ optimizes by submitting a bid equal to $B(\hat{v}_{ij})$ in an ad auction for consumer $i$ at any time $t$ if the other firms act analogously.

To derive it, suppose that firm $j$ believes that all other firms bid according to a monotone bidding function $B : [0, \overline{v}] \to \mathbb{R}$. In a second-price auction, it is optimal for firm $j$ to bid its continuation value conditional on winning the auction net of its continuation value conditional on losing. In the classic static setting, the value conditional on losing is 0. But in our setting, it is positive because of the option to wait for future auctions.

Let $V$ denote the NPV of all future flow profits from selling to consumer $i$ net of the costs of repeatedly displaying ads to consumer $i$. Then it is optimal for firm $j$ to bid according to $B$ if

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

Moreover, $V$ satisfies the recursive relationship

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

This system of equations reduces to a simple ODE which can be solved explicitly.

Proposition 4. *There exists a unique stationary equilibrium of the bidding subgame. In the unique stationary equilibrium, the following hold:*

date. Thus the first time that a firm enters an ad auction for consumer $i$ is distributed geometrically with parameter $NA\Delta/(F - M)$. As $\Delta \to 0$, this random variable converges to an exponential random variable with parameter $NA/(F - M)$.

1. *When invited to an auction for consumer $i$, firm $j$ bids*

$$B(\hat{v}_{ij}) = \pi_{\mathcal{F}} \int_0^{\hat{v}_{ij}} \frac{1}{r + \lambda_f + \lambda_e(s)} ds \tag{16}$$

*where $\lambda_e(\hat{v}_{ij}) = \lambda_a H^c(\hat{v}_{ij})^{N-1}$ is the exponential rate at which firm $j$ enters $\Omega_{it}$ while in $\Omega^c_{it}$.*

2. *The expected revenue from an auction for any consumer at any time is*

$$\pi_{\mathcal{D}} = \pi_{\mathcal{F}} \int_0^{\infty} \frac{1 - NH^c(s)^{N-1} + (N-1)H^c(s)^N}{r + \lambda_f + \lambda_e(s)} ds. \tag{17}$$

3. *The NPV of any given firm's flow profits aggregated across all consumers is*

$$\Pi_{\mathcal{F}} = \frac{I/\sigma - \pi_{\mathcal{D}} A}{rF}. \tag{18}$$

Proposition 4 unifies three different measures of search frictions: 1. average match delay $1/\lambda_e$, 2. deviation of cumulative match value from its potential level $F\mu_G - M\mu_H$,[13] and 3. average search costs $\pi_{\mathcal{D}}$.[14]

In (16), $\lambda_e$ captures the effect of the outside option.[15] It depends on both auction entry rate $\lambda_a$ and the probability that firm $j$ wins an auction. Thus, even if $\lambda_a$ is high, if competition is high, $\lambda_e$ may be low and a large fraction of surplus is extracted from firm $j$. The integral in (16) is due to the cumulative effects of bid shading. When lower types shade their bids, higher types shade their bids even more because the outside option is more attractive. When $A$ converges to its upper limit, $\lambda_e$ diverges pointwise and ad auction profits

[13] $\mu_G$ denotes the mean of the distribution $G$.

[14] One can interpret the model as a search-based dynamic asset pricing theory as in Duffie, Gârleanu, and Pedersen (2005) where the asset is a spot in the consumer's memory. Bidders' tastes are shocked when they are forgotten. The main distinction is that intermediaries not only provide liquidity but also compete for attention by providing quality services to consumers. Total intermediation intensity is equal to the ad frequency $A$.

[15] In practice, as in the model, bidders "pace" their bids in anticipation of variation in impression prices. See `https://www.facebook.com/business/help/1754368491258883?id=561906377587030`. For information on variation in impression prices see `https://www.facebook.com/business/help/2024547657774300`.

decay to zero as the product market becomes frictionless.

The cumulative match value $M\mu_H$ affects bidding through the flow profit and appears in $\pi_{\mathcal{F}}$. When the cumulative match value is high, a firm expects to steal only a small fraction of the consumer's income from its rivals if it wins the ad auction.

As seen in (17), $\lambda_e$, $M\mu_H$ and $H^c$ determine $\pi_{\mathcal{D}}$, the expected auction revenue. Given income $I$, firms' profits are decreasing in platforms' ad revenues as seen in (18).

## 3.6 Investment in Quality

Platforms invest in quality to steal attention from their rivals and earn a larger share of ad revenue $\pi_{\mathcal{D}}$. Each platform $k$ solves

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

subject to

$$\dot{q}_{kt} = L_{kt}^{\varphi} - \delta q_{kt}, \quad q_{k0} = q$$

where $0 < \varphi < 1$ implies diminishing returns to investment. Quality depreciates at rate $\delta > 0$ as a platform's content grows stale or less relevant over time.

A stationary equilibrium of the investment subgame is a constant investment level $L_{\mathcal{D}}$ such that each platform $k$ optimizes by setting $L_{kt} = L_{\mathcal{D}}$ at each $t \geq 0$ if

1. $q = L_D^{\varphi}/\delta$,

2. consumer demand is as in (6),

3. and all other platforms do the same.[16]

Property 1 together with the quality accumulation equation implies $q_{kt} = q$ for all $t \geq 0$.

[16]We have defined stationary equilibrium such that the investment level is the same for all platforms. There do not exist stationary equilibria (with the definition extended appropriately) where investment differs across platforms.

We solve for the stationary equilibrium of the subgame using Pontryagin's maximmum principle.[17] This gives necessary conditions which only need to be satisfied along the equilibrium path. We use these to derive a unique candidate $L_{\mathcal{D}}$. Details are in the Appendix. The candidate $L_{\mathcal{D}}$ does in fact correspond to a stationary equilibrium of the subgame if and only if the following condition is satisfied.

Condition 2. *It holds that*

$$\rho \leq 2.$$

Intuitively, if attention is sufficiently elastic there can not exist a stationary equilibrium. Under this condition, Proposition 5 below characterizes the stationary equilibrium of the investment subgame.

Proposition 5. *There exists a unique stationary equilibrium of the investment subgame if and only if Condition 2 holds. In this equilibrium, each platform $k$ invests $L_{kt} = L_{\mathcal{D}}$ where*

$$L_{\mathcal{D}} = \frac{\varphi\delta\pi_{\mathcal{D}}A(\rho - 1)}{D\,(r + (1-\alpha)\delta)} \tag{19}$$

*at each $t \geq 0$. Moreover, each platform $k$'s attention share is $x_{kt} = 1/D$ at each $t \geq 0$. The NPV of any given platform's flow profits is*

$$\Pi_{\mathcal{D}} = \frac{\pi_{\mathcal{D}}A/D - L_{\mathcal{D}}}{r}. \tag{20}$$

From equation (19) we see that platforms' incentives to invest are determined by ad revenues $\pi_{\mathcal{D}}A$ and platform interoperability $\rho$. The latter determines how easily a platform can steal attention away from its rivals. Neither of these are directly related to a consumer's utility from platform use. We study the implications of this for welfare in Section 6.

[17]This turn outs to be more tractable than applying Bellman's principle of optimality which requires solving for the optimal policy on the whole state space $\mathbb{R}_{+}$.

### 3.7 Income

The last part of the model to describe is how income $I$ is determined. We endogenize income to close the economy and establish a microfounded measure of social welfare. We assume that consumers own firms and platforms who distribute their flow profits uniformly across consumers. That is, $I = wL + rD\Pi_{\mathcal{D}} + rF\Pi_{\mathcal{F}}$. The relevant measure of social welfare is therefore consumer surplus.[18]

## 4 Unique Stationary Equilibrium

We compute the unique stationary equilibrium by finding the level of platform investment $L_{\mathcal{D}}$ that is consistent with income $I$ in that the labor market clears. Given income $I$, Propositions 1, 2, 3, 4, and 5 characterize the stationary equilibrium.

Since the labor market clears, the labor allocated to production is $L - DL_{\mathcal{D}}$. Then, since platforms extract profits from firms and all costs are labor costs, it follows that

$$I = \frac{\sigma}{\sigma - 1}(L - DL_{\mathcal{D}}). \tag{21}$$

This is one equation relating income $I$ and investment $L_{\mathcal{D}}$. A second equation is (19) (recall that income $I$ appears in $\pi_{\mathcal{F}}$). Define $\hat{\pi}_{\mathcal{D}}$ such that $\pi_{\mathcal{D}} = I\hat{\pi}_{\mathcal{D}}$. That is,

$$\hat{\pi}_{\mathcal{D}} = \frac{1}{\sigma M \mu_H} \int_0^\infty \frac{1 - NH^c(s)^{N-1} + (N-1)H^c(s)^N}{r + \lambda_f + \lambda_e(s)} ds. \tag{22}$$

There is an explicit solution to the system (21), (19) which determines the unique stationary equilibrium.

Theorem 1. *There exists a unique stationary equilibrium. For this equilibrium, the following properties hold.*

[18] An alternative would be to set social welfare to a linear combination of consumer surplus, firm profits, and platform profits. But, consumer surplus and profits are measured in different units. Moreover, it is unclear why social welfare should correspond to a linear combination.

*1. The size and composition of consideration sets are as in Proposition 3.*

*2. Each platform invests*

$$L_{\mathcal{D}} = \frac{\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}}A(\rho-1)}{r+\delta+\varphi\delta\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}}A(\rho-1)}\frac{L}{D} \tag{23}$$

*and has quality $q = L_{\mathcal{D}}^{\varphi}/\delta$ where $\hat{\pi}_{\mathcal{D}}$ is as in (22).*

*3. Income $I$ is given by (21).*

*4. Given income $I$, demands, prices, bidding, firm profits, and platform profits are as in Propositions 1, 2, 4, and 5.*

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

Theorem 1 summarizes the nearly explicit characterization of the equilibrium. Notice in Part 5, that final product consumption $C$ depends on consideration sets only through the cumulative match value of firms $M\mu_H$ within any given set.[19] Thus, overall welfare depends on only 1. the cumulative value $M\mu_H$ and 2. the quantity of labor $L_{\mathcal{D}}$ used for investment. Given the tractability and flexibility of the framework, there is room to explore many important extensions. We describe some in Section 7. These include endogenizing total platform attention, ad frequency, network effects, reserve prices, and allowing for heterogeneous data across platforms.

# 5 Comparative Statics

This section presents various comparative statics results. The goal of these is to understand how changes to model primitives affect welfare. In particular, we study changes to data informativeness, ad frequency, platform interoperability, firm substitutability, the measure

[19] This is also true in Melitz (2003) and is one advantage of CES when there is firm heterogeneity.

of platforms, and the measure of firms. These results may be relevant to regulators who can affect these market features with their policies.

In general, parameter changes have nonmonotone effects on overall welfare. We therefore derive comparative statics on final product consumption $C$ and final platform consumption $X$ separately. To build up to these results and better illustrate the intuition, we first derive comparative statics on $M$ and $H$. We then trace those through to comparative statics on ad revenue $\pi_{\mathcal{D}} A$, and then to $C$ and $X$. Later, in Section 6, we explain how these results can be used to derive comparative statics on the distance between equilibrium investment and first best.

## 5.1 Size and Composition of Consideration Sets

Proposition 6 summarizes comparative statics on the size and composition of consideration sets.

Proposition 6. *The following comparative statics with respect to the size $M$ of consideration sets and the distribution $H$ over expected values within them hold:*

1. *An increase in data informativeness (increase in $G$ with respect to $\succ_{MPS}$) causes the cumulative match value $M\mu_H$ to increase and the cdf $H^c$ to decrease in $\succ_{sosd}$.*
2. *An increase in ad frequency $A$ causes the cdfs $H$ and $H^c$ to decrease in $\succ_{fosd}$ but the cumulative match value $M\mu_H$ to increase.*
3. *An increase in the measure of firms $F$ causes the cdfs $H$ and $H^c$ to increase in $\succ_{fosd}$ and there $M\mu_H$ to increase.*
4. *A change to platform interoperability $\rho$, firm substitutability $\sigma$, or the measure of platforms $D$ has no effects on $M$ or $H$.*

The main takeaway is that an increase in data informativeness, ad frequency, and the measure of firms all lead to an increase in the cumulative match value $M\mu_H$. As we noted at the end of Section 4, $M\mu_H$ is one of the two main determinants of welfare.

While $H^c$ increases in $\succ_{fosd}$ when $F$ increases, it decreases in $\succ_{fosd}$ when $A$ increases, and decreases in $\succ_{sosd}$ when $G$ increases in $\succ_{MPS}$. Thus, increases in $A$, $F$, or $G$ have different effects on competition for platforms' intermediation services and therefore ad revenues, which we describe next.

## 5.2 Ad Revenues

Proposition 7 summarizes comparative statics on ad revenue.[20]

PROPOSITION 7. *The following comparative statics with respect to ad revenue $\pi_{\mathcal{D}}A$ hold:*

1. *An increase in ad frequency $A$ causes ad revenue $\pi_{\mathcal{D}}A$ to decrease. If $M$ and $H$ are fixed, then the reverse is true.*
2. *Ad revenue $\pi_{\mathcal{D}}A$ is maximal when data is uninformative. An increase in data informativeness (increase in $G$ with respect to $\succ_{MPS}$) may cause ad revenue $\pi_{\mathcal{D}}A$ to decrease.*
3. *An increase in $\sigma$ or $\rho$ causes ad revenue $\pi_{\mathcal{D}}A$ to decrease.*
4. *An increase in the measure $F$ of firms causes ad revenue $\pi_{\mathcal{D}}A$ to increase.*
5. *An increase in the measure $D$ of platforms has no effect on ad revenue $\pi_{\mathcal{D}}A$.*

Proposition 7 shows that factors which tend to ease search frictions reduce ad revenues while factors which exacerbate search frictions increase ad revenues. Part 1 states that when ad frequency increases, ad auction revenue $\pi_{\mathcal{D}}$ decreases sufficiently fast so that total ad revenue $\pi_{\mathcal{D}}A$ decreases. The intuition for this can be seen in the equation for the equilibrium bid (16). By Proposition 6 an increase in $A$ leads to an increase in $M\mu_H$ which appears in the denominator of $\pi_{\mathcal{F}}$. Bids decrease because firms internalize the higher match value of

[20]These can be traced through to investment $L_{\mathcal{D}}$ using (23). This is how we derive Proposition 8.

their competitors. This alone is not enough to offset the larger quantity of ads sold when $A$ increases. But, because $H^c$ decreases in $\succ_{fosd}$ so that there is more competition in the ad auction and because $\lambda_e$ increases so that the outside option is more valuable, ad revenue $\pi_{\mathcal{D}} A$ decreases.

This can be contrasted with an increase in $F$ as seen in Part 4. There, $M\mu_H$ also increases but because $H^c$ increases in $\succ_{fosd}$ and $\lambda_e$ decreases, ad revenue $\pi_{\mathcal{D}} A$ increases. Before discussing Part 2 (the intuition is more subtle), we first give intuition for Parts 3 and 5. An increase in $\sigma$ causes firms' profits from sales to decrease and in turn, bids and ad revenues. An increase in $\rho$ means that attention is more elastic. Platforms compete more aggressively for attention. Because more labor is allocated to investment, income decreases, bids decrease and in turn ad revenues. Since investment (23) is inversely proportional to the measure of platforms $D$, the quantity of labor allocated to production does not depend on $D$. As a result income, and therefore ad revenues are unaffected by changes to $D$.

Part 2 states that ad revenue is at it's maximal level when data is uninformative. In models of ad auctions where competition among bidders in the product market is not modeled, an increase in data informativeness typically leads to an increase in ad revenues (Hummel and McAfee, 2016; Board, 2009; Morris, Bergemann, Heumann, Sorokin, and Winter, 2021). This is because an increase in data informativeness increases bidder surplus. The auctioneer can extract a fraction of this surplus.

In this paper, an increase in data informativeness does not lead to an increase in bidder surplus. It merely shifts demand from one firm to another. The only beneficiaries are the consumers who are exposed to better matches. But the firms are the bidders. In addition, when data is more informative, bidders earn higher information rents which tends to reduce ad revenue. Moreover, an increase in data informativeness may lead to an increase the value of the outside option. When data is uninformative, the outside option has no value—the price in all future auctions is the same as that of the current auction. Because of these various factors, ad revenues typically decline when data informativeness improves.

Figure 1 illustrates.

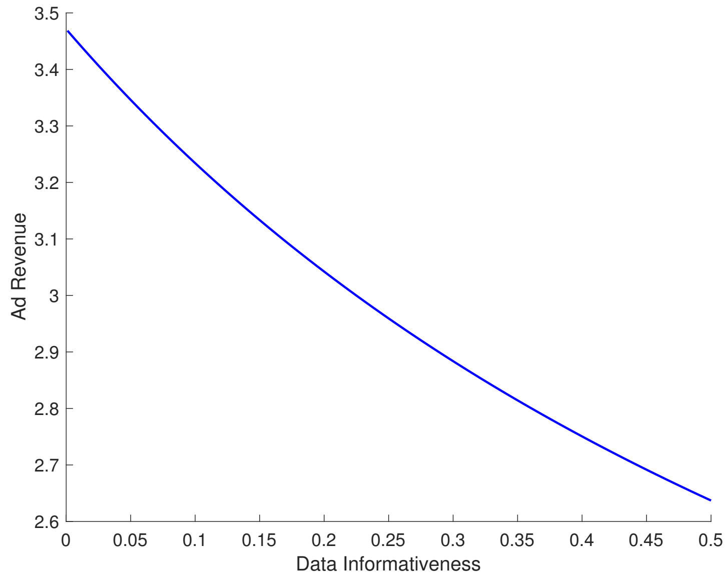


**Figure 1:** Ad revenue is plotted when $G$ is $U[.5-\epsilon, .5+\epsilon]$ as $\epsilon$ ranges from 0 to .5. Parameters are $M = 1$, $F = 10$, $N = 5$, $r = .1$, $A = .5$, $\sigma = 3.33$, $D = 1$, $\rho = 1.33$, $\alpha = .5$, $\delta = .1$, $\varphi = .5$ $L = 10$ .

Above, we plot equilibrium ad revenue as data informativeness increases. The parameters are given in the figure caption.[21] In this example, $P$ is the uniform distribution on $[0, 1]$ and $G$ is the uniform distribution on $[.5-\epsilon, .5+\epsilon]$. We parameterize data informativeness by $\epsilon$ ($G$ is increasing in $\succ_{MPS}$ as $\epsilon$ increases). As data goes from uninformative to fully informative, ad revenues drop by approximately 25 percent. Regulators often voice a concern that a reduction in data informativeness may lead to declining ad revenues and platform quality. However, our model shows that this need not be the case. In fact, the opposite may occur.

This result does not imply that more informative data is not valuable to an individual platform. While platforms would prefer uninformative data if they could collude, individually they typically benefit from more informative data. We can quantify the value of information to an individual platform $k$. Suppose that $\tilde{v}_{ij}$ denotes firm $j$'s expectation of

[21] In Figure 1 it may seem odd that $D = 1$ and $F = 10$ when we have normalized $|\mathcal{C}| = 1$. This is just for a more attractive graph. Given the model's homotheticity one can think of $|\mathcal{C}| >> 10$ but with low labor supplied by any individual worker so that total labor remains at $L = 10$.

$v_{ij}$ when supplied with data from platform $k$. Suppose that the other platforms all have the same data. As before, $\hat{v}_{ij}$ denotes firm $j$'s expectation using this data. $\hat{v}_{ij}$ may differ from $\tilde{v}_{ij}$. Given the joint distribution between $\hat{v}_{ij}$ and $\tilde{v}_{ij}$ we can compute the bidding function $\tilde{B}$ used on platform $k$:

$$\tilde{B}(\tilde{v}_{ij}) = \pi_{\mathcal{F}}\tilde{v}_{ij} + \left(\frac{\lambda_f}{\lambda_f + r}\frac{\lambda_a}{\lambda_a + r} - \frac{\lambda_a}{\lambda_a + r}\right)\mathbb{E}[V(\hat{v}_{ij})|\tilde{v}_{ij}, j \in \Omega^c_{it}]. \tag{24}$$

From this, we can compute the expected auction revenue on platform $k$. Since platform $k$ is infinitessimal, the same bidding function $B$ of (16) is used on the other platforms.

Figures 2 and 3 illustrate. Details concerning how these figures are constructed and the particular joint distribution between $\hat{v}_{ij}$ and $\tilde{v}_{ij}$ are given in Appendix B. In Figure 2 we plot the ad revenue of platform $k$ as we increase the informativeness of its data while holding that of the other platforms fixed. As illustrated, platform $k$'s ad revenues increase. In Figure 3 we fix platform $k$'s data but increase the data informativeness of the other platforms. As illustrated, platform $k$'s ad revenues decline.

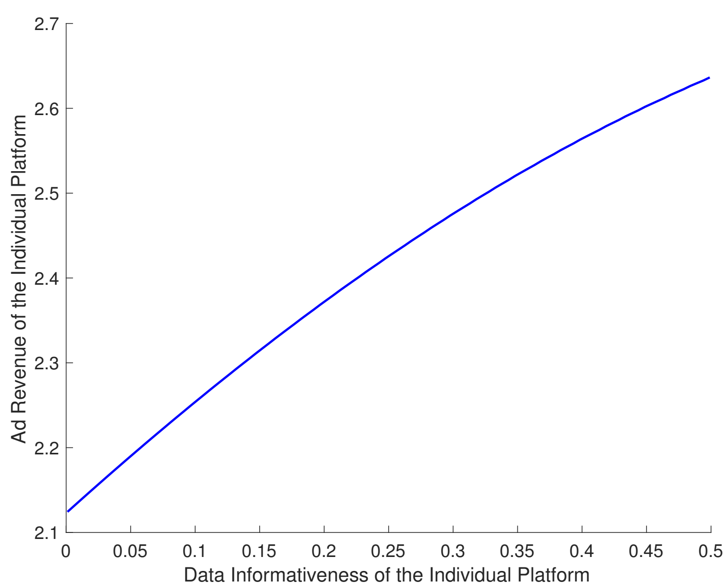


**Figure 2:** Ad revenue is plotted when $G$ is $U[0, 1]$, and $\tilde{G} \sim U[.5 - \tilde{\epsilon}, .5 + \tilde{\epsilon}]$ when $\tilde{\epsilon}$ ranges from 0 to .5. Additional details given in the Appendix B. All other parameters are as in Figure 1.

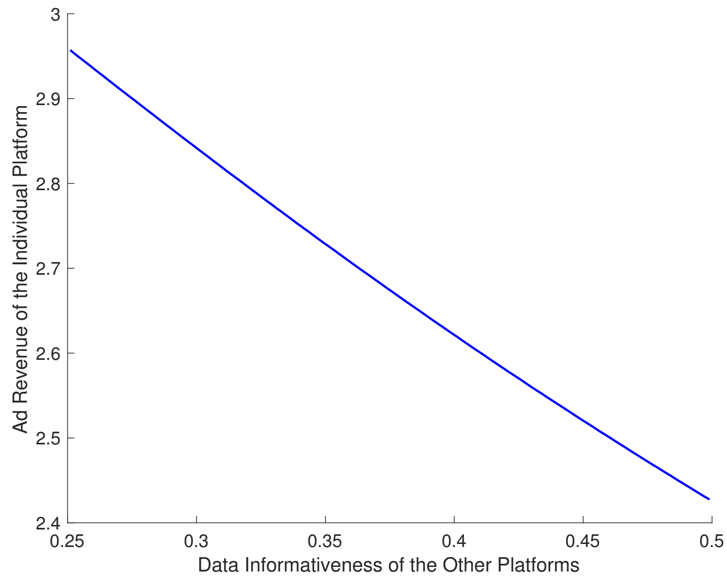


**Figure 3:** Ad revenue is plotted when $G$ is $U[.5-\epsilon, .5+\epsilon]$, and $\tilde{G} \sim U[.25, .75]$ when $\epsilon$ ranges from .25 to .5. Additional details given in the Appendix (under construction, available upon request).All other parameters are as in Figure 1.

We have presented comparative statics on ad revenues and on the size and composition of consideration sets. Using Theorem 1 we can trace these through to comparative statics on final consumption.

## 5.3 Final Consumption

Proposition 8 summarizes various comparative statics on final consumption $C$ and $X$.

Proposition 8. *The following comparative statics with respect to final product consumption $C$ and final platform consumption $X$ hold:*

1. *An increase in ad frequency $A$ causes $C$ to increase and $X$ to decrease.*

2. *When data is uninformative, $C$ is minimal and $X$ is maximal across all data $G$.*

3. *An increase in product substitutability $\sigma$ has an ambiguous effect on $C$ but causes $X$ to decrease.*

4. *An increase in platform interoperability $\rho$ causes $C$ to decrease and $q$ to increase.*[22]

5. *An increase in the measure of platforms $D$ has no effect on $C$ but causes $X$ to increase if $\frac{1}{\rho-1} > \varphi$ and decrease if $\frac{1}{\rho-1} < \varphi$.*

6. *An increase in the measure of firms $F$ has an ambiguous effect on $C$ but causes $X$ to increase.*

These results follow, almost immediately, from Propositions 6 and 7 and the expressions for $C$ and $X$ in Part 5 of Theorem 1. The main takeaway from Parts 1 and 2 is that factors which reduce search frictions tend to improve final product consumption and worsen final platform consumption. Part 3 states that an increase in product substitutability causes $X$ to decrease. This is because ad revenues decrease and so platform investment decreases. The effect on $C$ is ambiguous because, while income goes up due to decreased investment, the gains to variety decrease. Part 4 states that an increase in platform interoperability $\rho$ leads to lower production and more platform investment in quality. This is because attention in (6) is more elastic and so platforms have more incentives to invest to steal business from their rivals. Part 5 follows from (23) and Part 5 of Theorem 1. When substitutability is high, when $D$ increases, the gains to variety are too low to offset the decrease in platform quality. Part 6 states that an increase in the measure of firms $F$ leads to an increase in $X$ and has an ambiguous effect on $C$. $X$ increases because investment increases since ad revenue increases. The effect on $C$ is ambiguous because income goes down but $H$ goes up in $\succ_{fosd}$.

# 6 Equilibrium Relative to First Best

In this section, we compare the equilibrium to a first best benchmark for the special case of Cobb-Douglas utility. The main conclusion is that investment may be either too high

[22]We report the comparative static on $q$ rather than $X$ to hold the gains to variety fixed.

or too low. We derive a simple condition that can be used to determine which is the case. Using the condition, we discuss the main sources of inefficiency: appropriability, business stealing, and market power. We explain the extent to which the analysis generalizes to other utility functions in Appendix A.

## 6.1 Preliminaries

It will be useful, for the welfare analysis, to extend the baseline model by endogenizing the total attention spent on platforms. First, attention optimization links the marginal utility for platform consumption to that of product consumption. Because platforms charge zero prices to consumers, this link is missing in the baseline model. It will allow us to partially extend the analysis beyond Cobb-Douglas utility as we describe in Appendix A. Second, endogenizing total platform attention will allow us to more easily interpret why equilibrium deviates from first best.

We extend the baseline model in the following way. We assume that there are now two sectors in the product market. One sector is a leisure sector which requires attention as well as income to consume. We assume that 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 limit the need for any additional analysis, we assume that each sector is CES with the same parameter $\sigma$ as in the baseline. The total measure of products in the market remains $F$. But, the measure of leisure products is $\beta F$ while the measure of nonleisure products is $(1-\beta)F$.[23] Consumers' values for products of either type are drawn from $P$ and firms' expectations over these values are drawn from from $G$, just as in the baseline model.

A consumer's flow utility is now

[23]It will be clear later that if we were to endogenize the choice by firms concerning which type of product to produce, $\beta$ would be the fraction which choose to produce leisure products.

$$\tilde{u}(C_N, \tau C_L, (1-\tau)X) = \left(C_N^{1-\beta}(\tau C_L)^{\beta}\right)^{1-\gamma}((1-\tau)X)^{\gamma}$$

where $\tau$ is the fraction of attention devoted to leisure products, $C_N$ is the CES aggregate over nonleisure products, $C_L$ is the CES aggregate over leisure products, and $X$ is the CES aggregate over platform use. We interpret $C_L$ and $X$ as the rates at which leisure products and platforms are consumed per unit of attention respectively.[24]

We assume that the consumer chooses attention $\tau$ 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 $\gamma/(\gamma+(1-\gamma)\beta)$ of his attention on platforms. Moreover, the consumer spends a fraction $\beta$ of his income on leisure products. With this, let us define

$$C = C_N^{1-\beta} C_L^{\beta}.$$

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

$$u(C, X) = C^{1-\gamma} X^{\gamma}. \tag{25}$$

From here, it is easy to verify that the stationary equilibrium is characterized by the same equations as in the baseline model. Namely $C$ and $X$ are as in Part 4 of Theorem 1, and $M$, $H$, and $\pi_{\mathcal{D}}$ are defined by equations (12), (13), (11), and (17) except with $A\gamma/(\gamma + (1-\gamma)\beta)$ in place of $A$, since now only $\gamma/(\gamma + (1-\gamma)\beta)$ 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: $\beta$. Consequently, a firm's flow profit from selling to a consumer depends on the match value but not on the product type. Moreover, $H$ in

[24]Ex. You may purchase a guitar. But to consume it you must devote attention to playing it. We assume consumption scales linearly with attention (ex. you never get tired and each instant spent playing the guitar is like any other).

Proposition 3 is the empirical cdf over firms' expectations for both product types within the consideration set of any consumer. 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. For details, see the Appendix.

In short, this extended model reduces to the baseline model except with a change in the parameter $A$ to $A(1-\tau)$ where, with abuse of notation,

$$\tau = \frac{(1-\gamma)\beta}{\gamma + (1-\gamma)\beta}$$

is the share of attention spent on leisure products. We are thereby able to incorporate attention optimization into the model without any additional analysis. Moreover, the parameter $\gamma$ can be expressed in terms of $\tau$ and $\beta$ which represent the fraction of attention and income that individuals spend on leisure products.

## 6.2 First Best Benchmark

The social planner's problem is to allocate labor between production and investment. The planner can not alter the ad frequency or any other aspect of the matching technology. Later, we explain that this is without loss of generality when it comes to computing the efficient investment level, given our special choice of utility.

To write the social planner's problem we observe that 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 $\beta$ of production labor is used for leisure products.

With these observations made, the social planner's problem is

$$\max_{\{L_{\mathcal{D}t}\}_{t\geq 0}} \int_0^\infty e^{-rt} u(C_t, X_t) dt \tag{26}$$

subject to $C_t = (L - DL_{\mathcal{D}t})\,(M\mu_H)^{\frac{1}{\sigma-1}}$, $X_t = D^{\frac{1}{\rho-1}} q_t$, and

$$\dot{q}_{kt} = L_{\mathcal{D}t}^{\varphi} - \delta q_{kt}.$$

Above, we have assumed that the distribution $H_t$ is at its stationary level $H$ because we are interested only in the steady state which the planner eventually converges to. Using the Maximum Principle, we derive the social planner's steady state level of investment, stated below in Theorem 9.

Proposition 9. *The steady state level of investment by any given platform under the social planner is*

$$L_{\mathcal{D}}^{FB} = \frac{\varphi\delta^{\frac{1-\tau}{\tau}}\beta}{r + \delta + \varphi\delta^{\frac{1-\tau}{\tau}}\beta}\frac{L}{D}. \tag{27}$$

*The social planner's optimal level of investment does not depend on any features of the product market.*

A key takeaway from Theorem 9 is that first best investment does not depend in any way on the consideration set of the consumer. Moreover, it does not depend on ad frequency, data informativeness, platform interoperability, or product firm market power. This is a special consequence of Cobb-Douglas utility. But, it simplifies the comparison with equilibrium and allows us to later derive simple comparative statics on the distance of equilibrium investment from first best.

In Appendix A, we show that Theorem 9 extends to a much broader class of utility functions if the social planner does not internalize the impact of investment on $M$ and $H$ by affecting the attention consumers spend on platforms. We conjecture that (27) is a lower bound on first best investment when the social planner does internalize this impact.

### 6.3 Comparing Equilibrium and First Best

We compare the equilibrium level of investment (23) to the first best level (27).

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

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

*Moreover the distance between equilibrium and first best investment is increasing in the distance* $\left|\frac{\tau}{\beta}\frac{\sigma}{\sigma-1}\hat{\pi}_{\mathcal{D}}A(\rho-1)-1\right|$.

Theorem 2 is immediate from inspecting equations (23) and (27). It provides a simple condition which can be used to understand the efficiency of equilibrium investment. In the next subsection, we will explain the various sources of inefficiency and where they appear in the condition. But first, we make some basic observations.

COROLLARY 2.1. *The equilibrium level of investment may be excessive or insufficient depending on model parameters.*

When $\tau$ is near zero, investment is inefficient. This is because ad auction revenue per unit of income $\hat{\pi}_{\mathcal{D}}$ is decreasing in the effective ad frequency $A(1-\tau)$.[25] On the other had, when $\tau$ tends to 1, $\hat{\pi}_{\mathcal{D}}$ diverges since $M$ tends to zero because the effective ad frequency $A(1-\tau)$ tends to zero. Thus, investment is excessive in that case. In fact, it turns out that when $\tau$ is smaller so that consumers value platform use more, platform quality is actually lower. This follows from Part 1 of Proposition 8 since the effective ad frequency is lower. This stark inefficiency is because platforms' investment incentives are determined by ad revenues which do not fully internalize the surplus which platforms generate for consumers.

Using Theorem 2, we can compute the tax or subsidy on auction revenues that eliminates this inefficiency.

[25] In the Appendix, we used this fact to prove Part 1 of Proposition 7.

Corollary 2.2. *A proportional tax/subsidy on auction revenues which is paid out/financed by consumers equal to*

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

*can restore equilibrium investment to first best.*

Thus, there are parameter values when platforms should be taxed and others when platforms should be subsidized for their services.

Next, we observe that we can use Theorem 2 to derive comparative statics on the distance of equilibrium investment from first best.

Corollary 2.3. *The following comparative statics hold:*

1. *An increase in ad frequency $A$ causes equilibrium investment to be closer to (farther from) first best when it is excessive (insufficient).*

2. *An increase in data informativeness can cause equilibrium ivnestment to be closer to (farther from) first best when it is excessive (insufficient).*

3. *An increase in firm substitutability $\sigma$ causes equilibrium investment to be closer to (farther from) first best when it is excessive (insufficient).*

4. *An increase in platform interoperability $\rho$ causes equilibrium investment to be farther from (closer to) first best when it is excessive (insufficient).*

While we are able to derive comparative statics related to the efficiency of investment, we also note that changes to model parameters also affect match value and the gains to variety. It is therefore not always clear how welfare as a whole moves with parameters. In some cases, it can be deduced. For instance, an increase in ad frequency when investment is excessive will lead to both higher cumulative match value and more efficient investment which implies welfare should improve.

Finally, we observe that the condition in Theorem 2 suggests a sufficient statistic approach to assesing investment efficiency. In the US, consumers spend roughly 5 percent of their income on leisure goods. This implies that $\beta \approx .05$. Suppose that the typical individual spends 30 percent of leisure time on platforms so that $\tau = .3$. Since digital ad revenues are 1 percent of GDP, if product substitutability is 4, then

$$\frac{\tau}{\beta}\frac{\sigma}{\sigma - 1}\hat{\pi}_{\mathcal{D}} A \approx .622.$$

Then the condition implies that investment is insufficient if $\rho < 2.6$. An upper bound on $\rho$ is 2 which implies that investment is insufficient if we take the model literally.[26]

## 6.4 Sources of Inefficiency

We now describe, in more depth, the sources of inefficiency in the model. It turns out that inefficiencies arise due to issues with appropriability, business stealing externalities, and market power.

Firstly, because the platforms charge zero prices, they are unable to appropiate any of the surplus which they generate to consumers, either by exposing them to new products with ads or by providing them with services. This is a force leading to insufficient investment.

Instead, platforms appropriate surplus from firms who are not the direct beneficiaries of their services. In fact, platforms collectively appropriate more surplus from firms than they generate. Actually, platforms generate no surplus for firms since, by displaying ads, they merely shift demand among the different firms. The only reason firms purchase ads is to steal business from other firms. This force also appears in Grossman and Shapiro (1984) who show that it leads to excessive advertising. In the model of this paper, it is a force leading to excessive investment. The fact that investment incentives are determined by firms' business stealing externalities leads to some stark inefficiencies. For instance, as was highlighted earlier, incentives to steal business are higher when $\tau$ is higher since consumers

[26]One way to estimate $\rho$ would be to run the following experiment. Give subjects a fixed pool of money to spend on social media services. By varying the prices of the services, one can compute the elasticity of demand which can be used to estimate $\rho$.

discover fewer firms. But then, ad revenues are higher since search frictions are worsened. This in turn means that ad revenues and therefore quality are *lower* when consumers value quality *more*.

On its own, ad revenues are not enough to incentivize platforms to invest. Attention must be elastic (as determined by the parameter $\rho$) so that platforms can steal attention from each other. This externality is similar to that of Anderson and Coate (2005) who show that business stealing may lead to excessive entry by television programs. This second business stealing externality is another force leading to excessive investment.

Finally, not only do platforms fail to internalize the social value of their investments, they also fail to internalize the social costs. That is, there are distortions in the price of labor. One reason for this is that the market power of product firms leads them to have lower demand for labor than then the socially efficient level. This is a force leading to excessive investment. A second reason is that because platforms fail to internalize the social value of investments, their demand for labor is also distorted causing additional distortions in the price of labor.

# 7 Extensions

This section summarizes the results of several extensions.

## 7.1 Ad Frequency

In the first extension, we endogenize platforms' choices of ad frequency. We assume that consumers incur nuisance costs from viewing ads which reduces the effective quality levels of the platforms. We redefine the CES aggregate over platform consumption as

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

Above $\nu$ is a decreasing function of $A_{kt}$, the ad frequency on platform $k$. Under this extension, the share of attention devoted to platform $k$ is

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

To maximize profits, platform $k$ therefore selects $A_{kt}$ to maximize $\nu(A_{kt})^{\rho-1}A_{kt}$ (we assume $\nu$ decays sufficiently fast for the maximum to be obtained). Thus, the platform's optimal choice of ad frequency depends only on the nuisance cost $\nu$ and platform interoperability $\rho$. It does not depend on any other properties of the model such as ad revenue.

It turns out that ad frequency can be either too high or too low relative to first best. To see this, 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 $\gamma$. Suppose also that the social planner sets the same ad frequency $A$ for each platform. Then the planner chooses $A$ to maximize

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

Recall that $\mu_H$ is decreasing in $A$ by Proposition 6. Suppose that $\sigma$ and $\rho$ both equal 3/2. Then, if $\gamma = 1/2$, equilibrium advertising is too high. On the other hand, for $\gamma$ near 1, equilibrium advertising is too low as then the planner's choice of ad frequency diverges.

## 7.2 Network Effects

In the second extension, we incorporate network effects. We redefine the CES aggregate over platform consumption as

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

Above, $\eta$, which multiplies the quality of each platform $k$, is an increasing function of the total attention $x_{kt}$ paid to that platform. We consider a simple class of examples where $\eta(l) = l^{\zeta}$ for some $\zeta > 0$.

Under this extension, the share of attention devoted to platform $k$ is

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

This holds for all $i \in \mathcal{C}$. The equilibrium condition is that $x_{ikt} = x_{kt}$. There is a continuum of solutions to this fixed point. Each solution is of the following form.

There is a subset $\mathcal{E} \subseteq \mathcal{D}$ of platforms. If $k \in \mathcal{E}$ then

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

if the expression is well-defined. Otherwise $x_{kt} = 0$.

Under the equilibrium refinement that all platforms' in $\mathcal{D}$ receive a positive share of attention at all points in time, network effects are effectively equivalent to an increase in platform interoperability. This seems a natural refinement given the restriction to stationary equilibria. One can view the analysis in this paper as pertaining to a given subset of platforms (of positive measure) who remain active with stable market shares in the long run. However, the model has nothing to say about which subset prevails. Most results continue to hold with some minor alterations.

### 7.3 Data Which Differs Across Platforms

In the third extension, detailed in Appendix C, we extend the model to allow for the case when platforms have different consumer data except we take as given platforms' attention shares. We extend the analysis of equilibrium bidding strategies. Firms bid differently on platforms with different data. Under some conditions, we characterize the equilibrium bidding strategies by the fixed point of a contraction mapping. They can therefore be computed by interating the contraction map. However, computing the stationary distribution in consideration sets is more difficult and suffers from the curse of dimensionality. Numerically it is easiest to study the case with two groups of platforms with differing data. In a special case when one set of platforms has data and the other has no data, we derive explicit expressions for the equilibrium bidding strategies (given the stationary distribution). This extension may be useful for future work that studies the impact of policies such as GDPR on platforms with different data.

### 7.4 Reserve Prices

In the last extension, detailed in Appendix D, we allow platforms to choose reserve prices. Under regularity conditions, in the spirit of Myerson (1981), we characterize reserve prices nearly explicitly. We show that reserve prices are positive even though there is a continuum of platforms. Allowing for reserve prices therefore introduces another potential source of inefficiency into the model. This is because positive reserve prices lead to worsened search frictions since consumers view ads at a lower rate.

## 8 Conclusion

This paper has presented a general equilibrium model of the market for attention. The model formalizes the roles that platforms play in providing quality services to consumers and mitigating search frictions in product markets while extracting ad revenues. We have characterized search frictions, ad revenues, platforms' quality levels, and welfare in stationary equilibrium. We have shown how banning data tracking, capping ad frequency, and enforcing platform interoperability may lead to better platform consumption at the expense of worse product consumption. We have also shown that equilibrium investment and ad frequency may be too high or too low depending on parameters. The main sources of inefficiency concern appropriability, business stealing, and market power. Given the flexibility and tractability of the framework, we hope that future work may find it a useful starting point for further research.

There are many potential avenues to explore. One would be to study the case when platforms are large and there are strategic interactions among them. Intuition from this paper suggests that strategic interactions would lead to lower investment and ad frequency. Another would be to study entry by firms and platforms. The results of this paper can be shown to imply that there exists a unique stationary equilibrium with endogenous entry. A classic question (Spence, 1976; Dixit and Stiglitz, 1977) to explore in this new setting is whether the market supplies an efficient variety of products and platforms.

# A Omitted Proofs

*Proof of Proposition 5.* We conjecture that there exists a stationary equilibrium of the investment subgame where each platform has constant quality level $q$ and invests a constant amount $L_{\mathcal{D}}$ provided that the initial quality level is $q_0 = q$. Given the law of motion of platform quality, this implies that $q = \left(L_{\mathcal{D}}^{\varphi}/\delta\right)^{1/1-\alpha}$.

The (present-value) Hamiltonian for platform $k$'s optimization problem is

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

where $\lambda_t$, the costate variable, evolves according to

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

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

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

Under the conjectured stationary strategy, we have

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

This implies that $\lambda_t$ must be a constant $\lambda$. Namely, by the costate evolution equation,

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

Substituting into the earlier equation, we have

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

This implies that

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

In order to verify the optimality of setting $L_{kt} = L_{\mathcal{D}}$, by Arrow's Sufficiency Theorem, it suffices to check that the maximized Hamiltonian is concave in the state along the optimal trajectory.

By the earlier analysis, the maximized Hamiltonian is

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

Taking two derivatives of the maximized Hamiltonian, we see that it is concave in the state along the optimal trajectory if

$$\rho \leq 2 + \frac{\delta}{r + (1-\alpha)\delta}\left(1 - \frac{\alpha}{1-\varphi}\right)\alpha.$$

Moreover, if this condition is not satisfied, the maximized Hamiltonian is strictly convex in the state along the optimal trajectory. This can never be optimal. Thus the condition is both necessary and sufficient for the existence of a stationary equilibrium of the investment subgame.

First derivative:

$$\pi_{\mathcal{D}} A(\rho - 1)\frac{q_{kt}^{\rho-2}}{Dq^{\rho-1}} - \frac{\alpha}{1-\varphi}\left(\varphi^{\frac{1}{1-\varphi}} - \varphi^{\frac{\varphi}{1-\varphi}}\right)\lambda_t^{\frac{1}{1-\varphi}} q_{kt}^{\frac{\alpha}{1-\varphi}-1}$$

Second Derivative:

$$\pi_{\mathcal{D}} A(\rho - 1)(\rho - 2)\frac{q_{kt}^{\rho-3}}{Dq^{\rho-1}} - \frac{\alpha}{1-\varphi}\left(\frac{\alpha}{1-\varphi} - 1\right)\left(\varphi^{\frac{1}{1-\varphi}} - \varphi^{\frac{\varphi}{1-\varphi}}\right)\lambda_t^{\frac{1}{1-\varphi}} q_{kt}^{\frac{\alpha}{1-\varphi}-2}$$

Simplifying:

$$\pi_{\mathcal{D}} A(\rho-1)(\rho-2)\frac{1}{D} - \frac{\alpha}{1-\varphi}\left(\frac{\alpha}{1-\varphi}-1\right)\left(\varphi^{\frac{1}{1-\varphi}} - \varphi^{\frac{\varphi}{1-\varphi}}\right)\lambda^{\frac{1}{1-\varphi}} q^{\frac{\alpha}{1-\varphi}}$$

$$\Leftrightarrow$$

$$\pi_{\mathcal{D}} A(\rho-1)(\rho-2)\frac{1}{D} - \frac{\alpha}{1-\varphi}\left(\frac{\alpha}{1-\varphi}-1\right)\left(\varphi^{\frac{1}{1-\varphi}} - \varphi^{\frac{\varphi}{1-\varphi}}\right)\left(\frac{\pi_{\mathcal{D}A}(\rho-1)}{D(r+(1-\alpha)\delta)}\right)^{\frac{1}{1-\varphi}} q^{\frac{\alpha-1}{1-\varphi}}$$

Recall that

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

Using this, we see that

$$\pi_{\mathcal{D}} A(\rho-1)(\rho-2)\frac{1}{D} - \frac{\alpha}{1-\varphi}\left(\frac{\alpha}{1-\varphi}-1\right)\left(\varphi^{\frac{1}{1-\varphi}} - \varphi^{\frac{\varphi}{1-\varphi}}\right)\left(\frac{\pi_{\mathcal{D}A}(\rho-1)}{D(r+(1-\alpha)\delta)}\right)\delta(\varphi)^{-\varphi/(1-\varphi)}$$

□

*Details of the preliminaries in Section 6.* We take income $I$ as given. We conjecture that the fraction of products in the consideration set which are leisure products is $\beta$ in the stationary equilibrium. We also conjecture that both the distribution over expected values for leisure products and nonleisure products within consideration sets is $H$. Then the flow profit accruing to firm $j$ from selling a leisure product to consumer $i$ is

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

That is, the $\beta$ fraction of income cancels out with the $\beta$ fraction of the size of the consideration set. By similar logic, a nonleisure product which will therefore have the same flow profit as a leisure product. It follows immediately that the same bidding function as in the baseline model must apply for both product types. This further implies that a fraction $\beta$ of products in the consideration set are leisure products as conjectured. Moreover, by the same logic as in the baseline model, the distribution over expectations in the consideration set corresponding to leisure products must satisfy

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

Clearly, $\beta$ cancels out from both sides and we have the same condition as in the baseline model. By symmetry, the same condition applies also for non-leisure products. Thus all of the conjectures are verified. From here, it follows that the equations from the baseline model characterize the equilibrium of this extended model as desired. $\square$

We will now prove a more general version of Proposition 9. Suppose that utility is more generally of the form

$$u(\tau^\beta C, (1-\tau)X)$$

and that there exists, for each $C$, $X$, a unique interior optimal $\tau$ which satisfies

$$u_1(\tau^\beta C, (1-\tau)X)\beta\tau^{\beta-1}C = u_2(\tau^\beta C, (1-\tau)X)X. \tag{28}$$

Let $\tau_u(C, X)$ denote the solution.

*Proof of Proposition 9.* Throughout, we supress the arguments in $\tau_u$. The (present-value) Hamiltonian for the social planner's problem is

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

where $\lambda_t$ is the costate variable which evolves according to

$$r\lambda_t - \dot{\lambda}_t = -D^{\frac{1}{\rho-1}}(1-\tau_u)u_2 + \lambda_t\left(\alpha q_t^{\alpha-1} L_{\mathcal{D}t}^{\varphi} - \delta\right).$$

By Pontryagin's Maximum Principle, investment must maximize the Hamiltonian along the optimal trajectory. By the envelope theorem, using the first order condition from attention optimization (28), the first-order condition at the steady state $q$ is

$$\frac{\sigma}{\sigma-1}\hat{C}D\tau_u^{\beta}\frac{1}{\varphi}(\delta q^{1-\alpha})^{\frac{1-\varphi}{\varphi}}u_1 = \lambda_t q^{\alpha}.$$

Above, I have used the fact that $\delta q^{1-\alpha} = L_{\mathcal{D}}^{\varphi}$ in steady state. This condition implies that $\lambda_t$ is a constant $\lambda$ in steady state. By the costate evolution equation,

$$\lambda = \frac{D^{\frac{1}{\rho-1}}(1-\tau_u)u_2}{r+\delta(1-\alpha)}.$$

Substituting into the first order condition gives

$$\frac{\sigma}{\sigma-1}\hat{C}D\tau_u^{\beta}\frac{1}{\varphi}(\delta q^{1-\alpha})^{\frac{1-\varphi}{\varphi}}\frac{u_1}{u_2} = \frac{D^{\frac{1}{\rho-1}}(1-\tau_u)}{r+\delta(1-\alpha)}q^{\alpha}. \tag{29}$$

From (28), we have

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

Substituting into (29), we have

$$\frac{\sigma}{\sigma-1}D\frac{1}{\beta}\frac{\tau_u}{1-\tau_u}\frac{1}{I}L_{\mathcal{D}}\frac{1}{\varphi\delta} = \frac{1}{r+\delta(1-\alpha)}$$

which rearranges to

$$L_{\mathcal{D}} = \frac{\frac{\sigma-1}{\sigma}I\varphi\delta\frac{1-\tau_u}{\tau_u}\beta}{D(r+(1-\alpha)\delta)}.$$

Recall that

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

Using this relationship, we find that

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

$\square$

# B Details for Figures 1 and 2

To derive the expectation in (24) we must specify the joint distribution between the posterior expectations generated by the two types of data. We specify a joint distribution as follows.

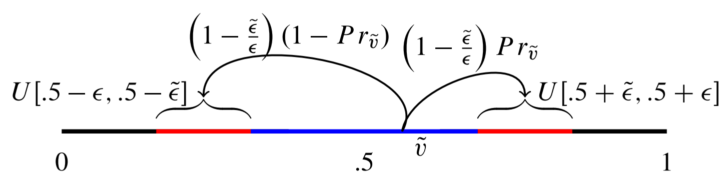


**Figure 4:** Construction of the conditional distribution over $\hat{v}$ given $\tilde{v}$.

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

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

Then conditional on receiving the signal realization $\tilde{v}$ the expectation over $v$ is in fact $\tilde{v}$. Moreover $\hat{v} \sim U[.5 - \epsilon, .5 + \epsilon]$. It is not necessary to specify how $\hat{v}$ is generated but it is easy to see that an analogous construction of $\hat{v}$ from $v$ (the true value) works. From this, we compute the joint distribution over $\hat{v}, \tilde{v}$ conditional on $j \in \Omega_{it}^c$, and then to use (24) to numerically compute the value of information.

Consider $\hat{v} \sim U[.5-\epsilon, .5+\epsilon]$ and $\tilde{v} \sim U[.5-\tilde{\epsilon}, .5+\tilde{\epsilon}]$ where $\tilde{\epsilon} < \epsilon$. Based on the construction described in the main text and illustrated by Figure 4, each point $\tilde{v}$ in $[.5-\epsilon, .5+\epsilon]$ stays with probability $\tilde{\epsilon}/\epsilon$. Otherwise, it jumps. Conditional on jumping, the probability it jumps up $Pr_{\tilde{v}}$ solves (30):

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

The probability that it jumps down conditional on jumping is $1 - Pr_{\tilde{v}}$. Let $\chi(\hat{v}|\tilde{v})$ denote the density for $\hat{v}$ given $\tilde{v}$. 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 the density for $\tilde{v}$ given $\hat{v}$. 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. Next, given $\tilde{v}$, what is the distribution over $\hat{v}$ conditional on knowing that the firm is not in the consideration set? It is

$$\frac{\chi(\tilde{v}|\hat{v})H^c(\hat{v})'}{\tilde{H}^c(\tilde{v})'}$$

whenever $\hat{v}$ is in one of the red regions and if $\hat{v} = \tilde{v}$ it is

$$\frac{H^c(\tilde{v})'}{\tilde{H}^c(\tilde{v})'}.$$

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

# C Extension: Data Which Differs Across Platforms

We investigate the case when there are two groups of platforms where data differs across groups but is the same within any given group. It will be clear that the analysis can be extended to an arbitrary number of such groups (though numerical tractability will suffer with more groups).

## C.1 More General Case

Let $x_1$ denote the share of attention spent by consumers on platforms in group 1. Let $\hat{v}_{ij}^1$ denote the signal sent to firm $j$ about $v_{ij}$ when bidding on platform 1. Let $\hat{v}_{ij}^2$ denote the signal on $v_{ij}$ sent to firm $j$ when bidding on platform 2. We assume that these are drawn from a joint distribution $G : [0, \overline{v}]^2 \to [0, 1]$ independently across consumers $i$ and firms $j$. The assumption that the signals are one dimensional is restrictive—we can not tractably analyze higher dimensional signal structures.

Proposition 10. *Suppose that bidding functions are monotone. The stationary distribution over values $H$ in the consideration set is determined by*

$$h(\hat{v}_1, \hat{v}_2) = x_1 N H^c(\hat{v}_1, 1)^{N-1} h^c(\hat{v}_1, \hat{v}_2) + x_2 N H^c(1, \hat{v}_2)^{N-1} h^c(\hat{v}_1, \hat{v}_2)$$

*where*

$$M h(\hat{v}_1, \hat{v}_2) + \frac{F - M}{M} h^c(\hat{v}_1, \hat{v}_2) = g(\hat{v}_1, \hat{v}_2).$$

We now derive the equilibrium bidding functions. IN what follows let $L = 2$. Below, all expectations are taken conditional on the information that firm $j$ is outside the consideration set of the consumer ($j \in \Omega_t^c$) though I do not indicate this to ease notation. Below, $x_l$ denotes the total attention share of group $l \in L$. Let $O_l$ denote the cdf of the first order statistic of $N - 1$ highest draws, from outside the consideration set, of the expected value given data on a platform in group $l$.

PROPOSITION 11. *Let $\Phi = (\Phi_l)_{l \in L}$ be the operator that takes as input a bidding profile $B = (B_l)_{l \in D}$ and sets*

$$\Phi_m(B)(\hat{v}_{mj}) = \frac{\mathbb{E}\left[\frac{1}{\sigma(A+r)}\frac{1}{M\mu_H}\left(v_j + \frac{K}{f^*+r}\sum_{l \in D} x_l O_l(\hat{v}_{lj})\left(\mathbb{E}[v_j|\hat{v}_{mj}] - v_j\right)\right) + x_l \int_0^{\hat{v}_{lj}} B_l(s) O_l'(s) \middle| \hat{v}_{mj}\right]}{1 + \frac{K}{f^*+r}\sum_{l \in D} x_l \mathbb{E}[O_l(\hat{v}_{lj})|\hat{v}_{mj}]}$$

*for each $\hat{v}_{mj} \in [0, \overline{v}]$ and $m \in L$. $\Phi$ takes as input a bidding profile and outputs another bidding profile. Moreover, $\Phi$ is a contraction mapping with respect to the sup-norm. Thus there exists a unique bidding profile $B^*$ such that $B^* = \Phi(B^*)$. $B^*$ is an equilibrium in bidding strategies given the stationary distribution in consideration sets.*

*Proof.* Let $V$ denote the NPV of payoffs from sales to the consumer at the time when the firm is about to enter an auction, but does not know which platform the auction is being run on, but does know all of the available signals. Let $\hat{v}_j := \mathbb{E}[v_j|(\hat{v}_{lj})_{l \in L}]$ and $c := \frac{1}{\sigma(f^*+r)}\frac{1}{M\mu_H}$. We have

$$V^*((\hat{v}_{lj})_{l \in L}) = \sum_{l \in L} x_l O_l(\hat{v}_{lj})\left(c\hat{v}_j - \int_0^{\hat{v}_{lj}} B_l(s) O_l'(s) ds + \frac{f^*}{f^*+r}\frac{K}{K+r}V^*((\hat{v}_{lj})_{l \in L})\right) + x_i(1 - O_l(\hat{v}_{lj}))\frac{K}{K+r}V^*((\hat{v}_{lj})_{l \in L})$$

This rearranges to

$$\frac{r}{K+r}V^*((\hat{v}_{lj})_{l \in L}) = \sum_{l \in L} x_l O_l(\hat{v}_{ij})\left(c\hat{v}_j - \int_0^{\hat{v}_{lj}} B_l(s) O_l'(s) ds - \frac{r}{f^*+r}\frac{K}{K+r}V^*((\hat{v}_{lj})_{l \in D})\right).$$

Let $B_l$ denote the equilibrium bidding function for platform in group $l$. Then

$$B_l(\hat{v}_{lj}) = c\mathbb{E}[\hat{v}_j|\hat{v}_{lj}] - \frac{r}{f^*+r}\frac{K}{K+r}\mathbb{E}[V((\hat{v}_{lj})_{l \in L})|\hat{v}_{lj}].$$

We therefore have

$$(31)\quad B_l(\hat{v}_{lj}) = c\mathbb{E}[\hat{v}_j|\hat{v}_{lj}] - \frac{K}{f^*+r}\left(x_l O_l(\hat{v}_{lj})B_l(\hat{v}_{lj}) - x_l\int_0^{\hat{v}_{lj}} B_l(s)O_l'(s)ds\right)$$
$$-\frac{K}{f^*+r}\sum_{m\neq l}\mathbb{E}\left[x_m O_m(\hat{v}_{mj})\left(c\hat{v}_j + B_l(\hat{v}_{lj}) - c\mathbb{E}[\hat{v}_j|\hat{v}_{lj}]\right) - x_m\int_0^{\hat{v}_{mj}} B_m(s)O_m'(s)\middle|\hat{v}_{lj}\right].$$

Rearranging,

$$B_l(\hat{v}_{lj}) = \frac{\mathbb{E}\left[c\hat{v}_j + \frac{K}{f^*+r}\sum_{m\in L} x_m O_m(\hat{v}_{mj})\left(c\mathbb{E}[\hat{v}_j|\hat{v}_{lj}] - c\hat{v}_j\right) + x_m\int_0^{\hat{v}_{mj}} B_m(s)O_m'(s)\middle|\hat{v}_{lj}\right]}{1+\frac{K}{f^*+r}\sum_{m\in L} x_m\mathbb{E}[O_m(\hat{v}_{mj})|\hat{v}_{lj}]}$$

That this holds for each $l$ is a necessary condition for $B$ to be an equilibrium bidding profile. By inspection, for any two bid profiles $B$ and $\hat{B}$,

$$\max_{l\in L}\max_s|\Phi_l(B_l(s)) - \Phi_l(\hat{B}_l(s))| \leq \max_{l\in L}\max_{\hat{v}_{lj}}\frac{\frac{K}{f^*+r}\sum_{m\in L} x_m\mathbb{E}[O_m(\hat{v}_{mj})|\hat{v}_{lj}]}{1+\frac{K}{f^*+r}\sum_{m\in L} x_m\mathbb{E}[O_l(\hat{v}_{lj})|\hat{v}_{ij}]}\max_s|B_l(s) - \hat{B}_l(s)|.$$

□

## C.2 Special Case

Though the general theory is best illustrated numerically, there are two important cases when the equilibrium can be characterized (nearly) explicitly. These are when 1. all platforms have the same data as in the baseline model and 2. there are two groups of platforms and one group has no data. I will derive the equations for the equilibrium in this second case now.

There are two groups of platforms, 1 and 2. Platforms in group 1 have data while platforms in group 2 have no data. For simplicity, I assume that each group has one half of the attention share, though the analysis may be easily generalized to cover asymmetric market shares.

Proposition 12. *The stationary distribution* $H_1^c$ *of expectations conditional on platform 1's data outside the consideration set* $\Omega_t$ *solves*

$$\frac{1}{2}H_1^c(\hat{v})^N + \left(\frac{F}{M} - \frac{1}{2}\right)H_1^c(\hat{v}) = \frac{F}{M}G(\hat{v})$$

*for each* $\hat{v} \in [0, \overline{v}]$. *There is a unique stationary distribution.*

*Proof.* The distribution must match inflows and outflows:

$$\frac{1}{2}H_1^c(\hat{v})^N + \frac{1}{2}H_1^c(\hat{v}) = H_1(\hat{v}).$$

By the accounting identity

$$MH_1(\hat{v}) + (F - M)H_1^c(\hat{v}) = FG(\hat{v})$$

we derive the equation

$$\frac{1}{2}H_1^c(\hat{v})^N + \frac{1}{2}H_1^c(\hat{v}) = \frac{F}{M}G(\hat{v}) - \frac{F-M}{M}H_1^c(\hat{v}).$$

Simplifying, gives

$$\frac{1}{2}H_1^c(\hat{v})^N + \left(\frac{F}{M} - \frac{1}{2}\right)H_1^c(\hat{v}) = \frac{F}{M}G(\hat{v}).$$

□

Proposition 13. *Given the stationary distribution* $H$, *the equilibrium bid functions* $B_1$ *and* $B_2$ *are characterized explicitly by the equations:*

$$B_1(\hat{v}_j) = \frac{\hat{v}_j}{\sigma(f^* + r)M\mu_H} - \frac{1}{2}\frac{r}{f^* + r}\frac{K}{K + r}\left(\frac{V_1^*(\hat{v}_j) + \frac{1}{N}\left(\frac{\hat{v}_j}{\sigma M\mu_H(f^*+r)} - B_2\right)}{1 + \frac{r}{f^*+r}\frac{K}{2(K+r)}}\right)$$

*and*

$$B_2 = \frac{\mu_{H^c}}{\sigma(f^*+r)M\mu_H} - \frac{r}{f^*+r}\frac{K}{K+r}\left(\frac{1+\frac{K}{K+r}}{2+\frac{K}{K+r}}\right)\mathbb{E}[V_1^*(\hat{v}_1)].$$

*Above,*

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

*where*

$$\psi(\hat{v}_j) = \frac{\frac{1}{\sigma(f^*+r)M\mu_H}\left(H_1^c(\hat{v}_1)^{N-1} + \frac{1}{N}\frac{\frac{K}{K+r}}{2+\frac{r}{f^*+r}\frac{K}{K+r}}\left(1 - H_1^c(\hat{v}_1)^{N-1}\frac{r}{f^*+r}\right)\right)}{1+\frac{K}{2(K+r)} + H_1^c(\hat{v}_1)^{N-1}\frac{r}{f^*+r}\frac{K}{2(K+r)} + \frac{1}{4}\left(\frac{K}{K+r}\right)^2\frac{\frac{r}{f^*+r}}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\left(1 - H_1^c(\hat{v}_1)^{N-1}\frac{r}{f^*+r}\right)},$$

$$a = \frac{-\frac{1}{N}\frac{\frac{K}{K+r}}{2+\frac{r}{f^*+r}\frac{K}{K+r}}\frac{\mu_{H^c}}{\sigma(f^*+r)M\mu_H}}{1+\frac{1}{2}\frac{f^*+2r}{f^*+r}\frac{K}{K+r} + \frac{1}{4}(\frac{K}{K+r})^2\frac{r}{f^*+r}\left(\frac{2+\frac{r}{f^*+r}\frac{K}{2(K+r)}}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\right)},$$

*and*

$$b = \frac{1}{N}\frac{\frac{K}{K+r}}{2+\frac{r}{f^*+r}\frac{K}{K+r}}\frac{\frac{r}{f^*+r}\frac{K}{K+r}\left(\frac{1+\frac{K}{K+r}}{2+\frac{K}{K+r}}\right)}{1+\frac{1}{2}\frac{f^*+2r}{f^*+r}\frac{K}{K+r} + \frac{1}{4}(\frac{K}{K+r})^2\frac{r}{f^*+r}\left(\frac{2+\frac{r}{f^*+r}\frac{K}{2(K+r)}}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\right)}.$$

*Proof.* Let $R$ denote the stationary distribution of values in the consideration set $\Omega_t$ in the unique stationary equilibrium. To ease notation, let $O(\hat{v}_1) = H_1^c(\hat{v}_1)^{N-1}$. The equilibrium bid by firm $j$ on platform 1 must satisfy

$$B_1(\hat{v}_j) = \frac{\hat{v}_j}{\sigma(f^*+r)M\mu_H} - \frac{r}{f^*+r}\frac{K}{K+r}\left(\frac{V_1^*(\hat{v}_j) + V_{12}^*(\hat{v}_j)}{2}\right). \tag{32}$$

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. It can be shown that $V_1^*$ satisfies the recursive equation

$$V_1^*(\hat{v}_1)\left(1+\frac{K}{2(K+r)}\right)=O(\hat{v}_1)\left(B_1-\mathbb{E}[B_1|\hat{v}_1\geq\hat{v}_{1(1)}]\right)+\frac{K}{K+r}\frac{V_{12}^*(\hat{v}_j)}{2}. \tag{33}$$

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 $\Omega_t$. Then

$$V_2^*\left(1+\frac{K}{2(K+r)}\right)=\frac{K}{K+r}\frac{\mathbb{E}[V_1^*(\hat{v}_1)]}{2}$$

and so rearranging,

$$V_2^*=\frac{\frac{K}{2(K+r)}}{1+\frac{K}{2(K+r)}}\mathbb{E}[V_1^*(\hat{v}_1)]. \tag{34}$$

Thus the optimal bid on platform 2 is

$$B_2=\frac{\mu_{H^c}}{\sigma(f^*+r)M\mu_H}-\frac{r}{f^*+r}\frac{K}{K+r}\left(\frac{V_2^*+\mathbb{E}[V_1^*(\hat{v}_j)]}{2}\right). \tag{35}$$

Similarly, $V_{12}^*$ solves

$$V_{12}^*(\hat{v}_j)=\frac{1}{N}\left(\frac{\hat{v}_j}{\sigma M\mu_H\,(f^*+r)}-B_2\right)-\frac{r}{f^*+r}\frac{K}{K+r}\frac{V_{12}^*(\hat{v}_j)+V_1^*(\hat{v}_j)}{2}.$$

Rearranging, we have

$$V_{12}^*(\hat{v}_j)\left(1+\frac{r}{f^*+r}\frac{K}{2(K+r)}\right)=\frac{1}{N}\left(\frac{\hat{v}_j}{\sigma M\mu_H\,(f^*+r)}-B_2\right)-\frac{r}{f^*+r}\frac{K}{2(K+r)}V_1^*(\hat{v}_j). \tag{36}$$

Substituting (32) into (33), I derive

$$V_1^*(\hat{v}_1)\left(1+\frac{K}{2(K+r)}\right)=O_1(\hat{v}_j)\frac{\hat{v}_1}{\sigma(f^*+r)M\mu_H}-O(\hat{v}_j)\frac{r}{f^*+r}\frac{K}{K+r}\left(\frac{V_1^*(\hat{v}_1)+V_{12}^*(\hat{v}_j)}{2}\right)$$
$$-\int_0^{\hat{v}_j}\left(\frac{s}{\sigma(f^*+r)M\mu_H}-\frac{r}{f^*+r}\frac{K}{K+r}\left(\frac{V_1^*(s)+V_{12}^*(s)}{2}\right)\right)O'(s)ds+\frac{K}{K+r}\frac{V_{12}^*(\hat{v}_j)}{2}$$

Combining like terms gives,

$$\begin{aligned}V_1^*(\hat{v}_1)\left(1+\frac{K}{2(K+r)}\right)&=O_1(\hat{v}_j)\frac{\hat{v}_1}{\sigma(f^*+r)M\mu_H}-O(\hat{v}_j)\frac{r}{f^*+r}\frac{K}{K+r}\frac{V_1^*(\hat{v}_1)}{2}\\&-\int_0^{\hat{v}_j}\left(\frac{s}{\sigma(f^*+r)M\mu_H}-\frac{r}{f^*+r}\frac{K}{K+r}\left(\frac{V_1^*(s)+V_{12}^*(s)}{2}\right)\right)O'(s)ds\\&+\frac{K}{K+r}\frac{V_{12}^*(\hat{v}_j)}{2}\left(1-O(\hat{v}_j)\frac{r}{f^*+r}\right).\end{aligned}$$

Taking a derivative of both sides I derive,

$$\begin{aligned}V_1^*(\hat{v}_1)'\left(1+\frac{K}{2(K+r)}\right)&=O_1(\hat{v}_j)'\frac{\hat{v}_1}{\sigma(f^*+r)M\mu_H}+O_1(\hat{v}_j)\frac{1}{\sigma(f^*+r)M\mu_H}\\&-O_1'(\hat{v}_j)\frac{r}{f^*+r}\frac{K}{K+r}\frac{V_1^*(\hat{v}_1)}{2}-O_1(\hat{v}_j)\frac{r}{f^*+r}\frac{K}{K+r}\frac{V_1^*(\hat{v}_1)'}{2}\\&-\left(\frac{\hat{v}_j}{\sigma(f^*+r)M\mu_H}-\frac{r}{f^*+r}\frac{K}{K+r}\left(\frac{V_1^*(\hat{v}_j)+V_{12}^*(\hat{v}_j)}{2}\right)\right)O'(\hat{v}_j)\\&-\frac{K}{K+r}\frac{V_{12}^*(\hat{v}_j)}{2}O'(\hat{v}_j)\frac{r}{f^*+r}+\frac{K}{K+r}\frac{V_{12}^*(\hat{v}_j)'}{2}\left(1-O(\hat{v}_j)\frac{r}{f^*+r}\right).\end{aligned}$$

Simplifying,

(37)
$$\begin{aligned}V_1^*(\hat{v}_1)'\left(1+\frac{K}{2(K+r)}\right)&=O_1(\hat{v}_j)\frac{1}{\sigma(f^*+r)M\mu_H}-O_1(\hat{v}_j)\frac{r}{f^*+r}\frac{K}{K+r}\frac{V_1^*(\hat{v}_1)'}{2}\\&+\frac{K}{K+r}\frac{V_{12}^*(\hat{v}_j)'}{2}\left(1-O(\hat{v}_j)\frac{r}{f^*+r}\right).\end{aligned}$$

Using (36), I find that

$$V_{12}^*(\hat{v}_j)'\left(1+\frac{r}{f^*+r}\frac{K}{2(K+r)}\right)=\frac{1}{N}\frac{1}{\sigma M\mu_H\,(f^*+r)}-\frac{r}{f^*+r}\frac{K}{2(K+r)}V_1^*(\hat{v}_j)'.$$

Rearranging

$$V_{12}^*(\hat{v}_j)'=\frac{\frac{1}{N}\frac{1}{\sigma M\mu_H(f^*+r)}-\frac{r}{f^*+r}\frac{K}{2(K+r)}V_1^*(\hat{v}_j)'}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}.$$

Substituting into (37) gives

$$\begin{aligned}V_1^*(\hat{v}_1)'\left(1+\frac{K}{2(K+r)}\right)&=O_1(\hat{v}_j)\frac{1}{\sigma(f^*+r)M\mu_H}-O_1(\hat{v}_j)\frac{r}{f^*+r}\frac{K}{K+r}\frac{V_1^*(\hat{v}_1)'}{2}\\&\quad+\frac{1}{2}\frac{K}{K+r}\frac{\frac{1}{N}\frac{1}{\sigma M\mu_H(f^*+r)}-\frac{r}{f^*+r}\frac{K}{2(K+r)}V_1^*(\hat{v}_j)'}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\left(1-O(\hat{v}_j)\frac{r}{f^*+r}\right).\end{aligned}$$

I next rearrange the equation so that $V_1^*(\hat{v}_1)'$ appears on the left hand side and all remaining terms appear on the RHS. I find that the coefficient on $V_1^*(\hat{v}_1)'$ is

$$1+\frac{K}{2(K+r)}+O_1(\hat{v}_j)\frac{r}{f^*+r}\frac{K}{K+r}\frac{1}{2}+\frac{1}{2}\frac{K}{K+r}\frac{r}{f^*+r}\frac{K}{2(K+r)}\frac{1}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\left(1-O(\hat{v}_j)\frac{r}{f^*+r}\right)$$

Simplifying we have

$$1+\frac{K}{2(K+r)}+O_1(\hat{v}_j)\frac{r}{f^*+r}\frac{K}{2(K+r)}+\frac{1}{4}\left(\frac{K}{K+r}\right)^2\frac{\frac{r}{f^*+r}}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\left(1-O(\hat{v}_j)\frac{r}{f^*+r}\right).$$

The term on the RHS is

$$\frac{1}{\sigma(f^*+r)M\mu_H}\left(O_1(\hat{v}_j)+\frac{1}{N}\frac{\frac{K}{K+r}}{2+\frac{r}{f^*+r}\frac{K}{K+r}}\left(1-O(\hat{v}_j)\frac{r}{f^*+r}\right)\right).$$

Define $\psi(\hat{v}_j)$ by

$$\psi(\hat{v}_j)=\frac{\frac{}{\sigma(f^*+r)M\mu_H}\left(O_1(\hat{v}_j)+\frac{1}{N}\frac{\frac{K}{K+r}}{2+\frac{r}{f^*+r}\frac{K}{K+r}}\left(1-O(\hat{v}_j)\frac{r}{f^*+r}\right)\right)}{1+\frac{K}{2(K+r)}+O_1(\hat{v}_j)\frac{r}{f^*+r}\frac{K}{2(K+r)}+\frac{1}{4}\left(\frac{K}{K+r}\right)^2\frac{\frac{r}{f^*+r}}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\left(1-O(\hat{v}_j)\frac{r}{f^*+r}\right)}.$$

Then we have

$$V_1^*(\hat{v}_j) = \int_0^{\hat{v}_j} \psi(s)ds + constant.$$

What remains is to derive $constant$. By (33) and (36), I see that

$$V_1^*(0)\left(1+\frac{K}{2(K+r)}\right) = \frac{K}{K+r}\frac{V_{12}^*(0)}{2}$$

and

$$V_{12}^*(0)\left(1+\frac{r}{f^*+r}\frac{K}{2(K+r)}\right) = -\frac{1}{N}B_2 - \frac{r}{f^*+r}\frac{K}{2(K+r)}V_1^*(0).$$

Rearranging this second equation gives

$$V_{12}^*(0) = \frac{-\frac{1}{N}B_2 - \frac{r}{f^*+r}\frac{K}{2(K+r)}V_1^*(0)}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}.$$

Substituting into the first equation gives

$$V_1^*(0)\left(1+\frac{K}{2(K+r)}\right) = \frac{K}{K+r}\frac{1}{2}\frac{-\frac{1}{N}B_2 - \frac{r}{f^*+r}\frac{K}{2(K+r)}V_1^*(0)}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}.$$

Therefore,

$$V_1^*(0)\left[\left(1+\frac{K}{2(K+r)}\right)\left(1+\frac{r}{f^*+r}\frac{K}{2(K+r)}\right) + \frac{K}{K+r}\frac{1}{2}\frac{r}{f^*+r}\frac{K}{2(K+r)}\frac{1}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\right]$$

$$= -\frac{1}{N}\frac{K}{K+r}\frac{1}{2}\frac{1}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}B_2.$$

Simplifying, we have

$$V_1^*(0)\left[1+\frac{1}{2}\frac{f^*+2r}{f^*+r}\frac{K}{K+r} + \frac{1}{4}\left(\frac{K}{K+r}\right)^2\frac{r}{f^*+r}\left(\frac{2+\frac{r}{f^*+r}\frac{K}{2(K+r)}}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\right)\right] = -\frac{1}{N}\frac{\frac{K}{K+r}}{2+\frac{r}{f^*+r}\frac{K}{K+r}}B_2.$$

Next, using (35) with (34) we find that

$$B_2 = \frac{\mu_{H^c}}{\sigma(f^*+r)M\mu_H} - \frac{r}{f^*+r}\frac{K}{K+r}\left(\frac{1+\frac{K}{K+r}}{2+\frac{K}{K+r}}\right)\mathbb{E}[V_1^*(\hat{v}_1)].$$

Therefore,

$$V_1^*(0)\left[1+\frac{1}{2}\frac{f^*+2r}{f^*+r}\frac{K}{K+r}+\frac{1}{4}(\frac{K}{K+r})^2\frac{r}{f^*+r}\left(\frac{2+\frac{r}{f^*+r}\frac{K}{2(K+r)}}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\right)\right]$$
$$= -\frac{1}{N}\frac{\frac{K}{K+r}}{2+\frac{r}{f^*+r}\frac{K}{K+r}}\left(\frac{\mu_{H^c}}{\sigma(f^*+r)M\mu_H} - \frac{r}{f^*+r}\frac{K}{K+r}\left(\frac{1+\frac{K}{K+r}}{2+\frac{K}{K+r}}\right)\mathbb{E}[V_1^*(\hat{v}_1)]\right).$$

Define $V_1^*(0) = a + b\mathbb{E}[V_1^*(\hat{v}_1)]$ where the coefficients are defined by the above equation. That is,

$$a = \frac{-\frac{1}{N}\frac{\frac{K}{K+r}}{2+\frac{r}{f^*+r}\frac{K}{K+r}}\frac{\mu_{H^c}}{\sigma(f^*+r)M\mu_H}}{1+\frac{1}{2}\frac{f^*+2r}{f^*+r}\frac{K}{K+r}+\frac{1}{4}(\frac{K}{K+r})^2\frac{r}{f^*+r}\left(\frac{2+\frac{r}{f^*+r}\frac{K}{2(K+r)}}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\right)}$$

and

$$b = \frac{1}{N}\frac{\frac{K}{K+r}}{2+\frac{r}{f^*+r}\frac{K}{K+r}}\frac{\frac{r}{f^*+r}\frac{K}{K+r}\left(\frac{1+\frac{K}{K+r}}{2+\frac{K}{K+r}}\right)}{1+\frac{1}{2}\frac{f^*+2r}{f^*+r}\frac{K}{K+r}+\frac{1}{4}(\frac{K}{K+r})^2\frac{r}{f^*+r}\left(\frac{2+\frac{r}{f^*+r}\frac{K}{2(K+r)}}{1+\frac{r}{f^*+r}\frac{K}{2(K+r)}}\right)}.$$

Then

$$\mathbb{E}[V_1^*(\hat{v}_1)] = \frac{1}{1-b}\left(a+\mathbb{E}\left[\int_0^{\hat{v}_j}\psi(s)ds\right]\right).$$

To conclude,

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

□

# D Extension: Reserve Prices

I extend the baseline model by allowing each platform to set reserve prices. The proposition below provides a nearly explicit characterization of the unique candidate equilibrium reserve price $R$ and stationary distribution $H$.

PROPOSITION 14. *A candidate competitive equilibrium in reserve prices is characterized by the following equations. The cutoff bidder expected value $Y$ solves*

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

*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 $s \leq Y$,*

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

*The reserve price is*

$$R = \frac{Y}{\sigma M \mu_H (f^* + r)}.$$

*The bidding function is*

$$B(\hat{v}_{ij}) = \frac{1}{\sigma(f^*+r)} \frac{1}{M\mu_H} \int_Y^{\hat{v}_{ij}} \frac{1}{1 + \frac{K}{f^*+r} H^c(s)^{N-1}} ds + R$$

*for $\hat{v}_{ij} \geq Y$. There is a multiplicity of equilibrium bids for $\hat{v}_{ij} < Y$. I restrict attention to an equilibrium where*

$$B(\hat{v}_{ij}) = \frac{1}{\sigma(f^*+r)} \frac{1}{M\mu_H} \hat{v}_{ij}$$

*for $\hat{v}_{ij} < Y$.*

*Proof.* Firm $j$'s optimal bid must satisfy

$$B(\hat{v}_{ij}) = \frac{1}{\sigma(f^*+r)}\frac{\hat{v}_{ij}}{M\mu_H} + \frac{f^*}{f^*+r}\frac{K}{K+r}V^*(\hat{v}_{ij}) - \frac{K}{K+r}V^*(\hat{v}_{ij}).$$

Let $Y := R\sigma M\mu_H(f^*+r)$. Above,

$$\begin{aligned}&V^*(\hat{v}_{ij}) = \left(1 - H(\hat{v}_{ij})^{N-1}\right)\frac{K}{K+r}V^*(\hat{v}_{ij}) + \\ &H^c(\hat{v}_{ij})^{N-1}\left(\frac{1}{\sigma(f^*+r)}\frac{\hat{v}_{ij}}{M\mu_H} + \frac{f^*}{f^*+r}\frac{K}{K+r}V^*(\hat{v}_{ij}) - \mathbb{E}\left[\max\{B(\hat{v}^{(1)}_{-j}), R\}|\hat{v}_{ij} > \hat{v}^{(1)}_{-j}\right]\right)\end{aligned}$$

when $\hat{v}_{ij} \geq Y$. Letting $O(\hat{v}_{ij}) = H^c(\hat{v}_{ij})^{N-1}$, we have

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

Then

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

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

$$B(\hat{v}_{ij}) = \frac{1}{\sigma(f^*+r)}\frac{\hat{v}_{ij}}{M\mu_H} - \frac{r}{f^*+r}\frac{K}{r}\left(O(\hat{v}_{ij})B(\hat{v}_{ij}) - RO(R) - \int_Y^{\hat{v}_{ij}} B(s)O'(s)ds\right).$$

Differentiating with respect to $\hat{v}_{ij}$, we obtain a first order linear differential equation for $B$. Using the boundary condition $B(Y) = R$, we find that

$$B(\hat{v}_{ij}) = \frac{1}{\sigma(f^*+r)}\frac{1}{M\mu_H}\int_Y^{\hat{v}_{ij}} \frac{1}{1 + \frac{K}{f^*+r}H^c(s)^{N-1}}ds + R$$

for $\hat{v}_{ij} \geq Y$. There is a multiplicity of equilibrium bids for $\hat{v}_{ij} < Y$. I will restrict attention to an equilibrium where $B$ is continuous on its entire domain. Specifically,

$$B(\hat{v}_{ij}) = \frac{1}{\sigma(f^* + r)} \frac{1}{M\mu_H} \hat{v}_{ij}$$

for $\hat{v}_{ij} < Y$. Next, we derive the stationary distribution over expected values $H$. Matching inflows with outflows gives,

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

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 $Y$ to derive an equation pinning down the cutoff value $Y$. First, we derive $h^c(Y)$ by observing that

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

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 I derive the first order condition. Consider the equation for profit:

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

Note that implicitly $\mu_H$ 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,

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

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 $Y$. From this, $\mu_H$ is obtained by first solving the polynomial equation for $H^c$ derived earlier. Then $R$ is obtained. I write the equation for $Y$ in terms of model primitives:

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

The solution to this equation gives a unique candidate for a symmetric competitive equilibrium cutoff. □

Corollary 2.4. *In the limit as $F$ tends to infinity, the candidate competitive equilibrium cutoff converges to the solution to*

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

Proposition 15. *A sufficient condition for the candidate equilibrium identified in the previous proposition to in fact be an equilibrium is that*

$$(N-1)(1 - H^c(s))\left(s - \int_Y^s \frac{1}{1 + \frac{K}{f^*+r}H^c(l)^{N-1}}dl - Y\right) - H^c(s)\left(s - \frac{1 - H^c(s)}{h^c(s)}\right)$$

*is decreasing for $s \in [Y, \overline{v}]$ and*

$$s - \frac{1 - \frac{F}{F-M}G(s)}{g(s)}$$

*is increasing for $s \in [0, Y]$. This holds provided $G$ is strictly regular in that the derivative of its virtual value is bounded from zero and $F - M$ is sufficiently large.*

*Proof.* Recall that in the region $s \leq Y$, the first derivative of the objective function is

$$NH(s)^{N-1}\left(-sh^c(s) + 1 - H^c(s)\right).$$

Since this is always positive at $s = 0$, it is never optimal to set a reserve price equal to 0.

Similarly it is never optimal to set a reserve price equal to $\overline{v}$. Thus the optimal reserve price must occur at an interior point when the first order condition is satisfied. Consider any such point, if it exists, in the region $s \leq Y$. Then

$$sh^c(s) = 1 - H^c(s).$$

Provided $s - \frac{1-H^c(s)}{h^c(s)}$ is monotone, there exists a unique solution to this equation. I now consider the first order condition when $s \geq Y$. Here, we have a couple additional terms.

$$\begin{aligned} - [(N-1)H^c(s)^{N-2}(1-H^c(s))h^c(s)] \left( \int_Y^s \frac{1}{1 + \frac{K}{f^*+r} H^c(l)^{N-1}} dl + Y \right) \\ + s\left((N-1)H^c(s)^{N-2} - NH^c(s)^{N-1}\right) h^c(s) \\ + (1-H^c(s))H^c(s)^{N-1} = 0. \end{aligned}$$

Simplifying and rearranging, we arrive at

$$(N-1)(1-H^c(s)) \left( s - \int_Y^s \frac{1}{1 + \frac{K}{f^*+r} H^c(l)^{N-1}} dl - Y \right) - H^c(s) \left( s - \frac{1-H^c(s)}{h^c(s)} \right) = 0.$$

Provided the above is also decreasing, the candidate equilibrium of Proposition 14 is in fact an equilibrium. This is the case if $F - M$ is sufficiently large and $G$ is strictly regular.

□

# References


Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2019): "Too Much Data: Prices and Inefficiencies in Data Markets," Discussion paper, National Bureau of Economic Research.

Ali, S. N., G. Lewis, and S. Vasserman (2020): "Voluntary Disclosure and Personalized

Pricing," in *Proceedings of the 21st ACM Conference on Economics and Computation*, pp. 537–538.

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

Athey, S., and G. Ellison (2011): "Position Auctions with Consumer Search," *The Quarterly Journal of Economics*, 126(3), 1213–1270.

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.

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

Bergemann, D., A. Bonatti, and T. Gan (2019): "The Economics of Social Data," .

Bergemann, D., B. Brooks, and S. Morris (2015): "The limits of price discrimination," *American Economic Review*, 105(3), 921–57.

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.

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., N. Gârleanu, and L. H. Pedersen (2005): "Over-the-Counter Markets," *Econometrica*, 73(6), 1815–1847.

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.

Evans, D. S. (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.

Grossman, G. M., and E. Helpman (1991): “Quality Ladders in the Theory of Growth,” *The Review of Economic Studies*, 58(1), 43–61.

Grossman, G. M., and C. Shapiro (1984): “Informative Advertising with Differentiated Products,” *The Review of Economic Studies*, 51(1), 63–81.

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.

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

Jones, C. I., and C. Tonetti (2020): “Nonrivalry and the Economics of Data,” *American Economic Review*, 110(9), 2819–58.

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

Melitz, M. J. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71(6), 1695–1725.

Morris, S., D. Bergemann, T. Heumann, C. Sorokin, and E. Winter (2021): "Selling Impressions: Efficiency vs. Competition," *CEPR Discussion Paper No. DP16507*.

Myerson, R. B. (1981): "Optimal auction design," *Mathematics of operations research*, 6(1), 58–73.

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

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.

Scott Morton, F. M., G. S. Crawford, J. Crémer, D. Dinielli, A. Fletcher, P. Heidhues, M. Schnitzer, and K. Seim (2021): "Equitable Interoperability: the "Super Tool" of Digital Platform Governance," *Available at SSRN*.

Spence, M. (1976): "Product selection, fixed costs, and monopolistic competition," *The Review of economic studies*, 43(2), 217–235.

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.