Data Sharing and Market Power with Two-Sided Platforms*

Rishabh Kirpalani† and Thomas Philippon‡

First version: November 2019; this version: November 2020

Abstract

We study an economy in which consumers and merchants (sellers) interact on a two-sided platform. Consumers can share data about their tastes for different varieties of a single good with the platform which in turn sells this data to merchants. Data sharing increases gains from trade by improving match quality but gives more market power to the platform relative to the merchants which can reduce entry and consequently consumer welfare. This leads to an externality not internalized by consumers thus leading to more data sharing than is efficient. We highlight two reasons why more precise information increases the market power of the platform. The first is a copycat (private label) externality that increases the outside option for the platform of selling the good directly to consumers. The second is a consumer access externality that reduces the outside option of the merchants when information gets more precise, as more buyers are able to find their desired variety on the platform. Our model explains the qualitative differences between traditional retail platforms (physical stores) and digital online platforms and why the latter are more likely to require regulatory interventions that the former.

*We are especially grateful to German Gutierrez for helping us understand the business model of Amazon, and to Luis Cabral, Steve Davis, Andres Drenik, Janice Eberly, Paolo Martellini, Diego Perez, Christopher Sullivan, Larry White, Rob Seamans, and seminar participants at NBER, University of Chicago, and New York University for stimulating discussions. We are grateful to the Smith Richardson Foundation for a research grant.

†University of Wisconsin-Madison, rishabh.kirpalani@wisc.edu.

‡New York University, CEPR and NBER, tphilipp@stern.nyu.edu

"They [Wall Street analysts] paid less attention to the thing that was attracting so many customers and keeping them coming back for more: the data. For Amazon wasn't just able to give its visitors the books they came to buy, but the books that they didn't even realize they wanted."

Margaret O'Mara, The Code, 2019.

O'Mara (2019)'s description of the rise of Amazon makes it clear that Amazon's founder, Jeff Bezos, thought from the beginning that data and technology would be the most important factor for the success of his new firm. For instance, an important reason for locating the firm in Seattle was the well established presence of Microsoft and thousands of tech engineers, and indeed, Amazon became a technological and logistical powerhouse. Using data to one's advantage is of course a perfectly legitimate strategy. In April 2020, however, the *Wall Street Journal* reported that Amazon employees on the private-label side of its business had used data about individual third-party sellers on its site to create competing products.[1] Similarly, in July 2020, the *Wall Street Journal* reported that Amazon appeared to use its investment and deal-making process [on its VC side] to help develop competing products. Amazon takes stakes in some startups and acquires others outright. Many investments are made through its Alexa Fund, an investment vehicle launched in 2015.[2] Is this business as usual, or does it reflect excessive market power by Amazon? Can data access create anticompetitive effects? Can data usage be excessive? If so, when, and why? These are some the questions we try to analyze.

Large internet platforms have changed the way market participants interact. One reason for this is the extraordinary ability of platforms such as Amazon and Google to gather and analyze large amounts of data. Platforms use this data to enable better matching between participants as well as for commercial purposes, including sale to third parties (Gutierrez, 2020b). In this paper, we focus on one set of benefits and costs and study the welfare implications of data sharing in the context of large platforms. In our model, information shared by buyers enables better matching between buyers and sellers on the platform. However, more information/data endogenously increases the market power of the platform relative to merchants. We ask if data sharing can be excessive and under what conditions. In doing so, we shed light on the qualitative differences between new platforms such as Amazon and Google and more traditional retail platforms.

In our model, consumers and sellers interact on a two-sided platform. Consumers can share information with the platform regarding their tastes for different varieties of a good. The platform then sells this data to sellers. Sellers and buyers interact in a directed search market on the platform. As the information gathered from the consumers becomes more precise, sellers can better predict the varieties desired by the consumers which improves overall match quality. On the other hand more precise information increases the market power of the platform relative to the sellers.

The platform and the sellers bargain over the price of information. As information becomes more precise, the outside option of the platform increases while that of the seller decreases. The former, which we term the copycat effect occurs because the platform could use its information to produce the product by itself, thereby undercutting the sellers. In equilibrium the platform does not need to implement this threat,

[1] https://www.wsj.com/articles/amazon-scooped-up-data-from-its-own-sellers-to-launch-competing-products-11587650015. Amazon said it was conducting an internal investigation into the practices described in the story.

[2] https://www.wsj.com/articles/amazon-tech-startup-echo-bezos-alexa-investment-fund-11595520249

but it increases its bargaining power.

In addition, the seller's outside option decreases as information gets more precise. We call this the consumer access effect. The seller has the option to sell the good at a store outside the platform, thereby matching with consumers who were unsuccessful at obtaining their desired variety on the platform. As information gets more precise, the fraction of such consumers decreases implying that such consumers no longer need to show up at stores outside the platform. Thus the outside option of the seller decreases.

These two effects combine to increase the price of information the platform can charge as information gets more precise, which lowers the seller's payoff. The externality comes from the fact that consumers do not internalize the impact of their information sharing on the price they face in the market.

We show that as information gets more precise, the latter effect dominates the benefits of better matches which in turn implies that the number of sellers participating on the platform decreases. This lowers consumers welfare. Since consumers are small they do not internalize the effect of information disclosure on seller entry. As a result, there may be excessive information revelation compared with that chosen by a social planner. Consequently regulation that restricts the information shared by consumers can increase overall welfare. A corollary of this result is that if information technology decreases the private cost of information disclosure, there is more likely to be too much information disclosure.

A natural question that arises is how platforms like Amazon differ from traditional retail platforms like Walmart and Target. For example, the latter also gather information about consumers tastes via sales, and produce their own in-house "private label" products. However, the key difference is the amount of data collected and processed by Amazon relative to Walmart, as well as the efficiency of their matching technologies. For example, physical stores can only observe the products actually purchased by consumers while Amazon can observe the consumers browsing history including the items in the cart (but not actually purchased). This enables Amazon to better predict consumers' tastes, price elasticities, susceptibility to advertising, etc. As we argue in the paper, it is exactly when the level of information is high that there is likely to be excessively high disclosure. This provides a rationale for greater regulation of online platforms relative to the physical stores.

Our mechanism is distinct from the ones analyzed in the literature and it sheds light on a central regulatory issue concerning large internet platforms, namely their market power vis-a-vis other merchants.

We present our baseline model in Section 1. We compare the planner and the decentralized equilibria in Section 2. We study extensions in Section 3.

Related Literature Katz and Shapiro (1985) provide an early analysis of network externalities in a Cournot model. They characterize equilibria with varying degree of compatibility across firms. Rochet and Tirole (2003) incorporate the insights of network economics and multi-product pricing to study competition among two sided platforms. Caillaud and Jullien (2003) model platforms as matchmakers and study competition with multi-homing, and price discrimination. They consider two-tier price systems with registration and transaction fees. Armstrong (2006) emphasizes the role of consumers who join all platforms and analyzes cross-group externalities. The study of single- versus multi-homing decisions has expanded rapidly since these early contributions, in particular regarding media markets. Ambrus et al. (2016) model consumers who spread their attention between several outlets. Park et al. (2020) model the entry of TV station in local news paper markets in the 1950s. Gutierrez (2020a) estimates a large scale model of consumers' choices on Amazon's platform and studies the impact of private labels.

A growing literature studies information disclosure by consumers. On the one hand, it is clear that the use of personal data can improve the allocation of online resources. Excessive information sharing can potentially harm consumers, however. Bergemann et al. (2015) analyze how a monopolist can use information about consumers' tastes to engage in third degree price discrimination. Intermediaries look for ways to monetize information. Bergemann et al. (2018) study how a data seller should optimally design and sell statistical experiments. Acemoglu et al. (2019) and Bergemann et al. (2019) study environments in which there can be excessive data sharing since consumers do not internalize that their data can reveal information about other consumers. See also Bergemann and Bonatti (2019) for a survey of the literature on data markets.

Our model of taste uncertainty is similar to Bergemann et al. (2015), Bergemann et al. (2018), and Ichihashi (2019). Compared to these papers, our contribution is to analyze information in a standard directed search environment where we can analyze the market power of a platform vis-a-vis its merchants. The key externality in our model operates through the outside option of merchants and their entry decisions.

To model the interaction between buyers and sellers we use the directed search framework developed by Shimer (1996) and Moen (1997) among others. See Wright et al. (2019) for an excellent survey. Directed search determines both trade on the platform as well as the outside market which in turn affects the sellers' outside option.

1 Baseline Model

We consider a trading model with an exogenous mass $\bar{N}_b$ of consumers (buyers, b, he), an exogenous number of legacy producers $\bar{N}_o$, an endogenous number of new producers N_s (sellers, s, she), and a market structure where buyers and sellers can match and trade. There is a single good which comes in several varieties and consumers are initially unsure about which variety suits their needs. Matching takes place either with a standard directed-search technology or on a platform (M). The key feature of the platform is that, in addition to directed search, it offers product recommendation by analyzing consumers' data.

1.1 Equilibrium without a platform

We start by describing the model without the information-gathering platform and then introduce the platform in the next subsection. This helps us understand exactly when the platform will be used by buyers and sellers. One can think of the model without the platform as a standard retail market. We refer to this market as the "outside" market, indexed by 'o' in the remainder of the paper. We model this outside market as a standard directed search environment with free entry. Each consumer wants to buy a fixed quantity (normalized to 1) of a good that comes in several varieties $i \in \mathcal{I} = [1, .., I]$. Each consumer has exactly one preferred variety which we refer to as the consumer's *taste*. Tastes are *i.i.d.* across varieties and across consumers, and all varieties are equally likely to be the preferred one ex-ante. Formally, consumption of variety i delivers utility u_i with $\max_{i\in\mathcal{I}}\{u_i\} = u > 0$, and $u_j = 0$ otherwise. All buyers and sellers are ex-ante identical. To simplify our notation, we present the model in terms of a representative buyer and a representative seller.[3]

The timing of the model is as follows. There is a single period with two stages:

1. There are $\bar{N}_o$ existing sellers. In stage 1, new sellers decide whether to pay the entry cost κ. Let N_s^e be the number of entrants.

2. In stage 2, consumers N_b and sellers $N_{s,o} = \bar{N}_o + N_s^e$ match in a directed search environment.

We solve the model by backward induction, starting from stage 2. Each seller has z units to sell and we normalize the marginal cost of production to zero. We assume a constant-elasticity matching function where the number of matches is $\bar{\alpha}_o \bar{N}_b^{\gamma}\left(z\bar{N}_o + zN_s^e\right)^{1-\gamma}$. Let $n_o \equiv \left(z\bar{N}_o + zN_s^e\right)/\bar{N}_b$ denote market tightness. The probability that a seller meets a buyer is given by $\alpha_o(n_o) = \bar{\alpha}_o n_o^{-\gamma}$. The Cobb-Douglass constant elasticity functional form simplifies the exposition but is not crucial. The important feature is that α_o is decreasing and convex.[4]

Sellers post prices and consumers direct their search after observing all the prices. Thus, there are potentially many sub-markets characterized by their price and tightness, (p, n). Following the search literature, we consider a market utility approach in which sellers maximize their payoffs by selecting (p, n) subject to participation by consumers. To sell the good the seller must clear two hurdles: first, she must match with a buyer – which happens with probability $\alpha_o(n_o)$; second, she must produce the correct variety of the good – which happens with probability $1/I$ in the outside market. The program of the seller for each unit is

$$v_{s,o} = \max_{n,p} \alpha_o(n)\frac{p}{I}$$

subject to

$$V_{b,o} = n\alpha_o(n)\frac{u-p}{I}.$$

This problem is equivalent to

$$\max_{n} \alpha_o(n)\frac{u}{I} - \frac{V_{b,o}}{n}. \tag{1}$$

The first order condition is $\alpha_o'(n)\,u/I + V_b/n^2 = 0$, which implies $p_o = \left(1 + \frac{n_o\alpha_o'}{\alpha_o}\right)u$. Given our Cobb-Douglass matching function we get $p_o = (1-\gamma)\,u$. The value for buyers is then

$$V_{b,o} = \gamma\frac{u}{I}\bar{\alpha}_o n_o^{1-\gamma}.$$

The value function per unit for sellers is $v_{s,o} = (1-\gamma)\frac{u}{I}\bar{\alpha}_o n_o^{-\gamma}$. Since each seller has z units to sell, the total value is $V_{s,o} = zv_{s,o}$ or

$$V_{s,o} = z(1-\gamma)\frac{u}{I}\bar{\alpha}_o n_o^{-\gamma}. \tag{2}$$

[3]Formally, "money" is a divisible good x that can be produced by both consumers and sellers at cost $C(x) = x$ and yields utility $U(x) = x$, and p is the price of the traded good in terms of "money".

[4]We also require – in all our matching functions – that the number of matches satisfies $\alpha N_b^{\gamma} z N_s^{1-\gamma} \leq \min(N_b, zN_s)$. If we denote by $n = zN_s/N_b$ the market tightness, we need to ensure that $\alpha n^{-\gamma} < 1$ and $\alpha n^{1-\gamma} < 1$, which of course is the same as checking that the probabilities of matching are less than one.

The free entry condition requires

$$V_{s,o} \leq \kappa. \tag{3}$$

The equilibrium in this model is fully characterized by the market tightness n_o. Assuming positive entry, the equilibrium tightness solves

$$n_o = \left((1-\gamma) \frac{z}{\kappa} \frac{u}{I} \bar{\alpha}_o \right)^{\frac{1}{\gamma}}. \tag{4}$$

The number of entrants is therefore $N_s^e = \max\left(0; z^{\frac{1-\gamma}{\gamma}} \bar{N}_b \left((1-\gamma) \frac{u\bar{\alpha}_o}{\kappa I}\right)^{\frac{1}{\gamma}} - \bar{N}_o\right)$.

A well known result is that the directed search environment is efficient. This is readily apparent from the planner problem:

$$\max_{N_s^e} \bar{\alpha}_o \bar{N}_b^{\gamma} z^{1-\gamma} \left(\bar{N}_o + N_s^e\right)^{1-\gamma} \frac{u}{I} - \kappa N_s^e.$$

The first order condition of the planner's problem

$$(1-\gamma) \bar{\alpha}_o z^{1-\gamma} \left(\frac{\bar{N}_b}{\bar{N}_o + N_s^e} \right)^{\gamma} \frac{u}{I} = \kappa$$

is the same as the free entry condition (3). As we will show this efficiency result will no longer hold when we introduce the platform and endogenous information disclosure by consumers.

1.2 Directed Search on the Platform

We now introduce a platform where buyers and sellers can match. The platform M combines two technologies: a matching technology and an information technology. If N_b buyers and N_s sellers are active on the platform, the number of matches is $\bar{\alpha} N_b^{\gamma} (z N_s)^{1-\gamma}$. The information technology depends on the disclosure policy chosen by the consumers. Consumers can share their personal data with the platform in order to get product recommendations. Formally, the consumer chooses the precision δ of a signal $\sigma \in \mathcal{I}$ about his taste. This signal is observed by the platform. If $u_i = u$, the signal realization is $\sigma = i$ with probability δ, i.e., $\Pr(\sigma = i \mid u_i = u) = \delta$. Since we assume flat priors, Bayes' law implies

$$\Pr(u_i = u \mid \sigma = i) = \frac{\Pr(\sigma = i \mid u_i = u) \Pr(u_i = u)}{\Pr(\sigma = i)} = \frac{\delta \times {}^1\!/_I}{\delta \times {}^1\!/_I + (1 - {}^1\!/_I)(1-\delta) \frac{1}{I-1}} = \delta.$$

The signal precision δ is therefore also the posterior probability that $u_\sigma = u$. Without loss of generality we assume that $\delta \in \left[{}^1\!/_I, \bar{\delta}\right]$.

In our baseline model, consumers can access both the platform and the outside market for free. Sellers, on the other hand, need to pay a fee to use the platform. In exchange for the fee m, the seller gets access to the matching technology of the platform and to the the "data" (σ, δ). The timing of the model is now:

1. Buyers set the disclosure parameter $\delta \in \left[{}^1\!/_I, \bar{\delta}\right]$ subject to a personal cost $\Lambda(\delta)$.

2. Sellers decide whether to pay the entry cost κ. Let N_s^e be the number of entrants

3. The sellers and the platform bargain. If they agree, the seller pays m and gets the chance to trade on the platform and receive information. If not, the seller can sell on the outside market.

4. Directed search and matching take place in two venues
    (a) Platform consumers and platform sellers match on the platform. The sellers receive information about consumers's tastes.
    (b) Other consumers and sellers match outside the platform.

Suppose that N_s sellers have paid the fee m to purchase the data (σ, δ) and participate in the directed search market. Since $\delta \geq 1/I$, it is optimal for the seller to produce the variety $j = s$ after observing signal $\sigma = s$. Conditional on (N_b, N_s, δ) the platform is a conventional directed search market. Therefore the price is $p = (1-\gamma)\, u$, and the value for the seller is

$$V_s = z\,(1-\gamma)\,\delta u \bar{\alpha} n^{-\gamma}, \tag{5}$$

where market tightness is $n \equiv \frac{zN_s}{N_b}$, while the value for the buyer is

$$V_b\,(\delta, n) = \gamma \delta u \bar{\alpha} n^{1-\gamma}. \tag{6}$$

Two endogenous variables play a crucial role in our model: the precision of disclosure δ, and the number of sellers N_s. Disclosure is chosen by buyers while the number of sellers is determined by free entry. Disclosure creates gains from trade and therefore increases both V_s and V_b. The number of sellers, on the other hand, increases V_b and decreases V_s.

1.3 Equilibrium with a Platform and Outside Market

Suppose that we start from an economy with only the traditional retail market. When will the platform be used? This depends on the efficiency of the platform captured by δ and $\bar{\alpha}$. Consider the outside equilibrium with free entry $n_o = \left((1-\gamma)\frac{z}{\kappa}\frac{u}{I}\bar{\alpha}_o\right)^{\frac{1}{\gamma}}$ and $V_{s,o} = \kappa$. Suppose a mass of buyers and a mass of sellers migrate to the platform keeping n_o constant. The outside market would not change. On the platform sellers would then get $\hat{V}_s = \delta I \frac{\bar{\alpha}}{\bar{\alpha}_o}\kappa$, which is more than κ if and only if $\delta I \bar{\alpha} > \bar{\alpha}_o$.

Lemma 1. *If $\delta I \bar{\alpha} < \bar{\alpha}_o$ the platform is inactive and the equilibrium is the one in Section 1.1. If $\delta I \bar{\alpha} > \bar{\alpha}_o$ the platform attracts all the new entrants and only legacy sellers remain in the outside market.*

Proof. First, suppose we have an equilibrium in which both the platform and outside option are active. By this we mean that there are new entrants on both the platform and in the outside market. Then it must be that $V_s = V_{s,o} = \kappa$ and $V_b = V_{b,o}$. These two equations imply that $n = n_o$ and thus it must be that

$$\frac{1}{I}\bar{\alpha}_o = \delta \bar{\alpha}.$$

Therefore, if this equality does not hold then only one of the two can be active (i.e. attract new entrants) in equilibrium. Suppose now that $\delta I \bar{\alpha} < \bar{\alpha}_o$. Suppose for contradiction that only the platform is active. Since buyers are always indifferent we must have that

$$V_b = \delta\gamma u\bar{\alpha}n^{1-\gamma} = V_{b,o} = \frac{1}{I}\bar{\alpha}_o\gamma u n_o^{1-\gamma}$$

which implies that

$$n = \left(\frac{I\delta\bar{\alpha}}{\bar{\alpha}_o}\right)^{-\frac{1}{1-\gamma}} n_o.$$

Since sellers can freely enter the outside market, it must be that

$$z\left(1-\gamma\right)\frac{u}{I}\bar{\alpha}_o n_o^{-\gamma} \leq \kappa$$

but then

$$z\left(1-\gamma\right)u\delta\bar{\alpha}n^{-\gamma} < \left(\frac{I\delta\bar{\alpha}}{\bar{\alpha}_o}\right)^{\frac{\gamma}{1-\gamma}} z\left(1-\gamma\right)\frac{u}{I}\bar{\alpha}_o n_o^{-\gamma} < \kappa$$

which is a contradiction. Thus, the platform can never be active. Finally, suppose that $\delta I\bar{\alpha} > \bar{\alpha}_o$. Suppose for contradiction that only the outside market is active. Then it must be that

$$z\left(1-\gamma\right)u\delta\bar{\alpha}n^{-\gamma} \leq \kappa$$

and so

$$z\left(1-\gamma\right)u\frac{1}{I}\bar{\alpha}_o n_o^{-\gamma} \leq \left(\frac{\bar{\alpha}_o}{I\delta\bar{\alpha}}\right)^{\frac{\gamma}{1-\gamma}} z\left(1-\gamma\right)u\delta\bar{\alpha}n^{-\gamma} < \kappa$$

which is a contradiction. □

To be active, the platform must have a technological advantage in either information or matching. When the platform is active all entrants go to the platform, $N_s = N_s^e$. Consumers, on the other hand, must be indifferent between searching on the platform or on the outside market. The indifference condition $V_b = V_{b,o}$ then implies

$$\delta\bar{\alpha}n^{1-\gamma} = \frac{\bar{\alpha}_o}{I}n_o^{1-\gamma}, \tag{7}$$

Sellers must pay a fee m to the platform, therefore the entry condition becomes

$$V_s - m = \kappa. \tag{8}$$

Market clearing for consumers requires

$$N_b + N_{b,o} = \bar{N}_b, \tag{9}$$

with $n_o \equiv \frac{zN_o}{N_{b,o}}$. To complete the model we need to derive m from the bargaining game between the platform and the sellers.

1.4 Bargaining between Sellers and the Platform

Let V_M denote the outside option of the platform, which we discuss below The outside option of the seller is to join the outside market. Let θ be the bargaining power of the platform vis-a-vis the sellers. The Nash bargaining problem is then:

$$\max_m (m - V_M)^\theta (V_s - m - V_{s,o})^{1-\theta},$$

assuming that $V_s \geq V_M + V_{s,o}$. The negotiated fee is therefore $m = \theta(V_s - V_{s,o}) + (1-\theta) V_M$, and the total payoff for the seller if she buys access to the platform is $V_s - m = (1-\theta) V_s + \theta V_{s,o} - (1-\theta) V_M$. Thus, the free entry condition is

$$(1-\theta) V_s + \theta V_{s,o} = \kappa + (1-\theta) V_M. \tag{10}$$

Outside Option of the Platform Instead of being an intermediate between buyers and sellers, the platform can produce the good itself at cost c, and then sell it to consumers directly. For instance, Amazon has its AmazonBasics private-label brand (see for example Khan (2016) and Mattiolo (2020)). Mattiolo (2020) documents how Amazon uses data to "decide how to price an item, which features to copy or whether to enter a product segment based on its earning potential". The platform controls the search environment and thus can ensure that it matches with a consumer and it is subject to the same information requirements. On the other hand, it must incur a production cost, $c > 0$. The interpretation of this cost is that the platform is less efficient than the sellers at producing the good. The expected profit of the platform is $\delta p - c$, therefore

$$V_M(\delta) = \max\{0; \delta(1-\gamma) u - c\}. \tag{11}$$

The important feature is that V_M is a (weakly) increasing function of information precision δ.

1.5 Decentralized Equilibrium

We describe the equilibrium in two steps. First given disclosure level δ, we characterize an equilibrium on the platform (platform equilibrium). Next, we endogenize δ and define an equilibrium for the environment (decentralized equilibrium). Given any level of disclosure δ chosen by consumers ex-ante, the platform-equilibrium is characterized by the two market tightness variables (n, n_o) which solve the free entry condition (10)

$$z(1-\gamma) u \left((1-\theta)\delta\bar{\alpha} n^{-\gamma} + \theta \frac{\bar{\alpha}_o}{I} n_o^{-\gamma}\right) = \kappa + (1-\theta) V_M(\delta),$$

and the consumer indifference condition (7)

$$n_o = \left(\frac{\bar{\alpha}_o}{I\delta\bar{\alpha}}\right)^{\frac{-1}{1-\gamma}} n.$$

Combining the two conditions we get market tightness as a function of information disclosure

$$n^{\gamma} = z\bar{\alpha}\left(1-\gamma\right)u\frac{G\left(\delta\right)}{\kappa+\left(1-\theta\right)V_{M}\left(\delta\right)}, \tag{12}$$

where

$$G\left(\delta\right) \equiv \left(1-\theta\right)\delta+\theta\left(\frac{\bar{\alpha}_{o}}{I\bar{\alpha}}\right)^{\frac{1}{1-\gamma}}\delta^{-\frac{\gamma}{1-\gamma}}.$$

Equation (12) shows that the effect of δ on n is ambiguous. The denominator is increasing in δ. This captures the *copycat* effect in the outside value of the platform. When δ is high, the copycat option is valuable, the platform has a stronger bargaining position and seller entry decreases. The function $G\left(\delta\right)$ capture gains from trade and customer access. The first term, $\left(1-\theta\right)\delta$, is increasing in δ since gains from trade encourage entry and increases market tightness. The second term is decreasing in δ and captures the *customer access* effect. When information improves, consumers flock to the platform, which depresses the outside market and therefore the outside option of the sellers. Taking log-derivatives we obtain

$$\gamma\frac{n'}{n} = \frac{G'\left(\delta\right)}{G\left(\delta\right)}-\frac{V_{M}'}{\frac{\kappa}{1-\theta}+V_{M}}$$

A key insight of the model is that an increase in δ can lead to an decrease in tightness.

Lemma 2. *Market tightness n is a decreasing function of δ if and only if*

$$\frac{G'\left(\delta\right)}{G\left(\delta\right)}-\frac{V_{M}'}{\frac{\kappa}{1-\theta}+V_{M}}<0.$$

The condition in Lemma 2 is likely to hold when the copycat effect V_{M}' is large, or when the customer access effect is high. A sufficient condition for n to be decreasing in δ is $G'\left(\delta\right)<0$ or

$$\frac{\gamma}{1-\gamma}\frac{\theta}{1-\theta}\left(\frac{\bar{\alpha}_{o}}{I\delta\bar{\alpha}}\right)^{\frac{1}{1-\gamma}}>1,$$

which holds when either θ (bargaining power of the platform) or γ (importance of consumer access) is high. As we will see, this can potentially justify regulations of data gathering by the platform.

To close the model we need to solve for δ. The program of a buyer is

$$\max_{\delta} V_{b}\left(\delta,n\right)-\Lambda\left(\delta\right),$$

subject to (6), and where Λ is an increasing and convex function that captures the personal cost of disclosing precise information. Note that $\Lambda\left(1/I\right)=0$ since this corresponds to no information. Since each consumer is of measure zero, he takes n as given. Assuming an interior solution the disclosure choice δ^{b} by a consumer satisfies

$$\frac{\partial V_{b}}{\partial\delta}\left(\delta^{b},n\right)=\Lambda'\left(\delta^{b}\right) \tag{13}$$

The equilibrium is defined by the tightness equation (12) and the optimal disclosure choice (13). Without loss of generality we can consider the case where $\lim_{\delta\to\bar{\delta}}\Lambda' = \infty$ so we can fold the technological constraint $\delta \leq \bar{\delta}$ into the cost function Λ. The particular case of no private disclosure costs simply corresponds to $\Lambda = 0$ over $\left(1/I, \bar{\delta}\right)$. In that case, since the value functions are all increasing in δ, it is clear that consumers' private choice is $\delta = \bar{\delta}$ and we can treat δ as a technological parameter in the decentralized equilibrium.

Proposition 1. *The decentralized equilibrium* $\left(n, \delta^b\right)$ *is characterized by the information disclosure choice and the free-entry market tightness condition*

$$\gamma u\bar{\alpha}n^{1-\gamma} = \Lambda'\left(\delta^b\right),$$
$$n^{\gamma} = z\bar{\alpha}\left(1-\gamma\right)u\frac{G\left(\delta\right)}{\kappa+\left(1-\theta\right)V_M\left(\delta\right)}.$$

2 Social Planner

2.1 Planner's Solution

Consider the problem of a social planner who cares about the welfare of consumers and sellers. We ignore the welfare of the platform in this baseline calculation since we did not model the platform entry decision in the baseline model. Since the expected profits of the sellers is pinned down by free entry the planner only needs to estimate the welfare of the consumers. The problem for the planner is thus

$$\delta^* = \arg\max_{\delta} \bar{N}_b W\left(\delta\right),$$

where $W\left(\delta\right) = V_b\left(\delta, n\left(\delta\right)\right) - \Lambda\left(\delta\right)$, subject to the consumer value function (6) and the free entry of sellers (10). Thus the planner solves $\max_{\delta}\gamma u\bar{\alpha}\delta n^{1-\gamma} - \Lambda\left(\delta\right)$, and the marginal value of information from the perspective of the planner is

$$\frac{1}{\gamma u\bar{\alpha}}\frac{dW}{d\delta} = n^{1-\gamma} - \frac{\Lambda'\left(\delta\right)}{\gamma u\bar{\alpha}} + \left(1-\gamma\right)\delta n' n^{-\gamma}.$$

The first two terms capture the private partial equilibrium trade-off that consumers take into account. It is zero at the decentralized equilibrium $\delta = \delta^b$. The planner has a lower marginal value of information than the private sector if and only if $n' < 0$.

Proposition 2. *The social planner wants to restrict information disclosure if and only if* $n'\left(\delta^b\right) < 0$.

From Lemma 2 we know that a sufficient condition is $\frac{\gamma}{1-\gamma}\frac{\theta}{1-\theta}\left(\frac{\bar{\alpha}_o}{I\delta\bar{\alpha}}\right)^{\frac{1}{1-\gamma}} > 1$. Note that when n' is large in absolute value, welfare can be decreasing in δ even if private disclosure costs $\Lambda\left(\delta\right)$ are zero. This implies that the planner's problem can have an interior solution even without private costs of information disclosure. It also shows that private costs are not crucial for our analysis; they are simply convenient to obtain an interior solution to the private sector problem.

2.2 Implications for Regulation and Entry

The most direct implication of our model is that consumers do not usually choose the socially optimal degree of information sharing because they ignore the impact of their disclosure on the degree of competition among sellers.

It is easy to see how disclosure can be inefficiently low. When disclosure is privately costly consumers disclose little information and this lowers the gains from trade. Sellers do not expect much profit, and entry is limited. As a result competition is low and prices are high. This case is likely when the bargaining power of the platform θ is low.

It is less obvious how disclosure can be too high since, in our model, information increases the gains from trade. This stands in contrast to papers that assume that one consumer's privacy is directly hurt by the disclosure of other consumers' information, as in Acemoglu et al. (2019). The key for the excessive disclosure result lies in the market power of the platform vis-a-vis the sellers. This can happen either because better disclosure allows the platform to create copycat products that compete the sellers' products, and/or because the platform attracts many consumers and depresses the value of the outside market for the sellers. This case is likely when either θ (bargaining power of the platform) or γ (importance of consumer access) is high.

Note that the state variable in our model is market tightness on the platform defined as $n \equiv \frac{zN_s}{N_b}$. Thus, tightness decreases when N_b increases faster than N_s. In numerical simulations (see Figure 1), we find that N_b is usually an increasing function of δ while N_s is increasing in δ for low values of δ and decreasing afterwards. To see why start from (12) which defines the function $n(\delta)$. We can then find from the indifference condition (7): $n_o \equiv \frac{z\bar{N}_o}{N_{b,o}} = \left(\frac{\bar{\alpha}_o}{I\delta\bar{\alpha}}\right)^{\frac{-1}{1-\gamma}} n$. Then we use the market clearing condition $N_b + N_{b,o} = \bar{N}_b$ to write

$$N_b = \bar{N}_b - \left(\frac{\bar{\alpha}_o}{I\delta\bar{\alpha}}\right)^{\frac{1}{1-\gamma}} \frac{zN_o}{n}$$

which shows that N_b is increasing in δ as long as $\frac{dn}{n} > -\frac{1-\gamma}{\delta}$. This condition holds for most parameter values, but we can find counter-examples. For N_s, on the other hand, we have

$$N_s = \frac{n}{z} N_b = \frac{n}{z}\bar{N}_b - \left(\frac{\bar{\alpha}_o}{I\delta\bar{\alpha}}\right)^{\frac{1}{1-\gamma}} N_o$$

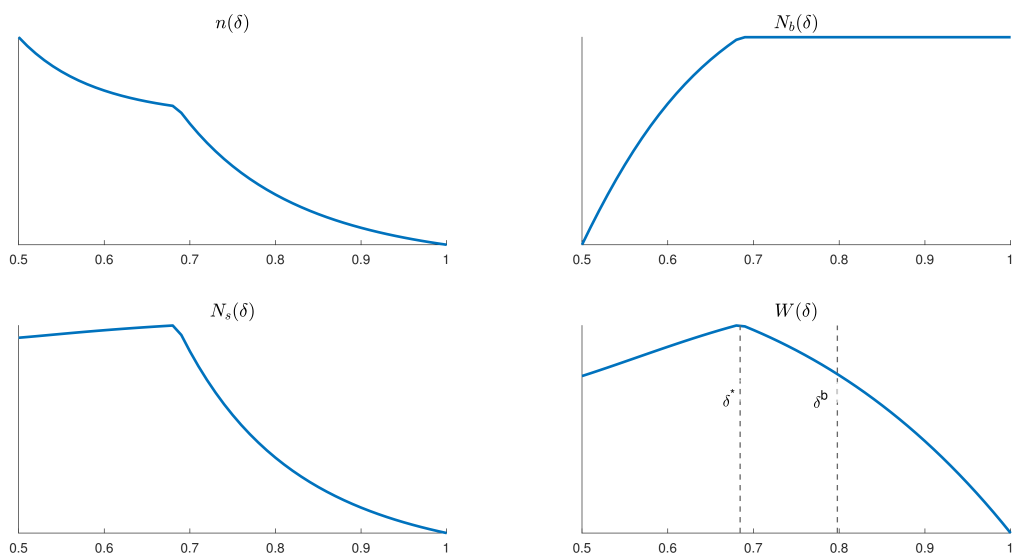
The second term increases N_s when δ is small while the first term decreases N_s if δ is large enough. Therefore when δ is high, we are more likely to be in the region where N_s decreases with δ. To the extent that an IT platform operates a more efficient data gathering technology, they are more likely to enter the parameter space where disclosure reduces the entry of new sellers. This leads to the following result.

Proposition 3. *Information technologies (IT) that improve data gathering increase the risk of excessive disclosure.*

Proposition 3 helps us understand the qualitative difference between a modern platform such as Amazon and traditional retail store. Retail stores are two-sided and operate a matching technology. They typically operate their own brands that compete with their outside sellers. They also gather some information about their customers. What then, is the difference between CVS and Amazon? Why would regulators be worried about the excessive information gathering in one case but not in the other? Some commentators want to prevent platforms from also being able to sell private label products on their site. Why would this policy not apply similarly to traditional retail stores?

Our model provides a precise theoretical argument. Disclosure that requires filling paper work or active action by users is costly. Gathering digital information, on the other hand, is cheap and often passive (e.g., cookies, browsing history). One way to think about Amazon versus traditional retail is that $\bar{\delta}$ for Amazon is a lot higher than for brick-and-mortar retailers. Proposition 3 says that this increases the risk of a negative impact on entry.

Figure 1: *Numerical Example*



Parameters: $I = 2$, $\gamma = 0.8$, $u = 10$, $\theta = 0.8$, $\bar{\alpha} = 0.3$, $\bar{\alpha}_o = 0.2$, $\kappa = 0.05$

3 Extensions

We now study some extensions of our baseline model.

3.1 Monopolist Platform

In our baseline model we assume Nash bargaining between the platform and the sellers. A natural question is to understand what happens if the platform behaves as a monopolist instead. The platform maximizes its revenues subject to the participation constraints of the sellers. The problem for the platform is then

$$\max_{m} mN_s,$$

subject to the sellers' participation constraints

$$V_s - m \geq \max\left(V_{s,o}, \kappa\right),$$

and the consumer indifference condition

$$V_b = V_{b,o}.$$

The first set of constraints captures both the entry condition $V_s - m > \kappa$ and the choice of the platform over the outside market $V_s - m \geq V_{s,o}$. The entry condition must hold with equality otherwise there would be infinite entry: $V_s - m = \kappa$. The other condition may be tight or slack, so there are two cases to consider.

Case $V_s - m = V_{s,o}$. Let us start with the case in which the condition is tight. Then $V_{s,o} = \kappa$ implies

$$n_o = \left(\frac{z}{\kappa}\left(1-\gamma\right)\frac{u}{I}\bar{\alpha}_o\right)^{1/\gamma}.$$

The consumer indifference condition $n = \left(\frac{\delta I \bar{\alpha}}{\bar{\alpha}_o}\right)^{\frac{-1}{1-\gamma}} n_o$ then implies

$$n = \left(\frac{\delta I \bar{\alpha}}{\bar{\alpha}_o}\right)^{\frac{-1}{1-\gamma}} \left(\frac{z}{\kappa}\left(1-\gamma\right)\frac{u}{I}\bar{\alpha}_o\right)^{1/\gamma}.$$

This shows that n is decreasing in δ. The fee is pinned down by the free entry condition $n = \left(\frac{z(1-\gamma)\delta u \bar{\alpha}}{\kappa+m}\right)^{\frac{1}{\gamma}}$ hence

$$m = \kappa\left(\left(\frac{\delta I \bar{\alpha}}{\bar{\alpha}_o}\right)^{\frac{1}{1-\gamma}} - 1\right) \tag{14}$$

is also increasing in δ. Note that entrants join the outside market as well as the platform. This constitutes an equilibrium as long as $N_{s,o} \geq \bar{N}_o$. With constant returns this happens only in a knife edge case $\delta I \bar{\alpha} = \bar{\alpha}_o$ since the platform can scale up infinitely quickly to accommodate new entrants and customers. With decreasing returns the limit pricing equilibrium would exist for a range of parameters.

Case $V_s - m > V_{s,o}$. In this case, all entrants join the platform and the number of sellers in the outside market is the exogenous number of legacy sellers $N_{s,o} = \bar{N}_o$. Using the definition of market tightness $n \equiv \frac{zN_s}{N_b}$ and $n_o \equiv \frac{zN_{s,o}}{\bar{N}_b - N_b}$ we can write the revenues of the platform as

$$mN_s = m\left(\frac{n}{z}\bar{N}_b - \frac{n}{n_o}\bar{N}_o\right)$$

We can then use the consumer indifference condition $n_o^{1-\gamma} = \frac{\delta I \bar{\alpha}}{\bar{\alpha}_o} n^{1-\gamma}$ to get

$$mN_s = m\left(n\frac{\bar{N}_b}{z} - \left(\frac{\delta I \bar{\alpha}}{\bar{\alpha}_o}\right)^{-\frac{1}{1-\gamma}}\bar{N}_o\right).$$

Before taking the first order condition we need to take a stand on the choice of δ. As explained in Katz and Shapiro (1985), there are two equilibrium concepts in network formation. In one , consumers form expectations about the size of the network and firms then play a game, taking consumers expectations as given. The key point in this case is that the pricing decision of firm does not directly influence the expectations of consumer. The other concept is when expectations depend explicitly on firms' pricing decisions. In our model consumers make two choices: where to shop and how much information to disclose. Consistent with the literature and our timing assumption, we assume in our benchmark model that the platform takes δ as given when it chooses m.[5] The FOC for the platform is then

$$n\frac{\bar{N}_b}{z}-\left(\frac{\delta I\bar{\alpha}}{\bar{\alpha}_o}\right)^{-\frac{1}{1-\gamma}}\bar{N}_o+m\frac{\partial n}{\partial m}\frac{\bar{N}_b}{z}=0,$$

which implies that $\frac{\partial n}{\partial m}<0$ at any interior solution. The monopolist always raises its price until tightness decreases.

We can compute $\frac{\partial n}{\partial m}$ from the free entry condition[6], $V_s=\kappa+m$, to get $n=\left(\frac{z(1-\gamma)\delta u\bar{\alpha}}{\kappa+m}\right)^{\frac{1}{\gamma}}$ and

$$\frac{1}{n}\frac{\partial n}{\partial m}=-\frac{1}{\gamma}\frac{1}{\kappa+m}.$$

We can therefore write the FOC as $\left(1-\frac{1}{\gamma}\frac{m}{\kappa+m}\right)\frac{\bar{N}_b}{z\bar{N}_o}\left(\frac{\delta I\bar{\alpha}}{\bar{\alpha}_o}\right)^{\frac{1}{1-\gamma}}=n^{-1}$, which we can write as

$$\frac{(\kappa+m)^{\frac{1}{\gamma}+1}}{\kappa-\frac{1-\gamma}{\gamma}m}=z^{\frac{1-\gamma}{\gamma}}\left((1-\gamma)\,u\right)^{\frac{1}{\gamma}}\left(\delta\bar{\alpha}\right)^{\frac{1}{\gamma}+\frac{1}{1-\gamma}}\left(\frac{I}{\bar{\alpha}_o}\right)^{\frac{1}{1-\gamma}}\frac{\bar{N}_b}{\bar{N}_o}. \tag{15}$$

The LHS is increasing in m. The RHS is increasing in $\frac{\bar{N}_b}{\bar{N}_o}$. M charges a higher fee when there are many consumers, as expected, and a lower fee when there are many outside sellers who attract consumers outside the platform. The impact of z and u is intuitive: higher gains, higher fees. The impact of $\frac{I}{\bar{\alpha}_o}$ is also intuitive: a higher value means that the platform is more valuable and can charge a higher fees. Also as expected, we have $\frac{\partial m}{\partial \delta}>0$. It is worth noting that δ and $\bar{\alpha}$ appear together in the expression: the fee depends on the product of information and matching services offered by the platform and not on each individually. Finally, note than m is bounded from above : $m<\frac{\gamma}{1-\gamma}\kappa$.

We can summarize our results as follows

Proposition 4. *Monopoly pricing. When $\delta I\bar{\alpha}<\bar{\alpha}_o$ the platform is inactive. When $\delta I\bar{\alpha}=\bar{\alpha}_o$ the platform engages in limit pricing (14), keeping sellers indifferent between the platform and the outside market. When $\delta I\bar{\alpha}>\bar{\alpha}_o$ the platform attracts all the new entrants and charges the monopoly price (15). The platform fee m is always increasing in information quality δ.*

The impact of δ on n is ambiguous in the monopoly pricing case. Note that $m=z\,(1-\gamma)\,\delta u\bar{\alpha}n^{-\gamma}-\kappa$

[5] One can show that the results are similar in the other case.

[6] Recall that the sellers value functions are $V_s=z\,(1-\gamma)\,\delta u\bar{\alpha}n^{-\gamma}$ and $V_{s,o}=z\,(1-\gamma)\,\frac{u}{I}\bar{\alpha}_o n_o^{-\gamma}$. The buyers value functions are $V_b=z\gamma\delta u\bar{\alpha}n^{1-\gamma}$ and $V_{b,o}=z\gamma\frac{u}{I}\bar{\alpha}_o n_o^{1-\gamma}$.

therefore

$$\gamma n^{-1} = \left(\frac{\kappa n^{\gamma}}{z(1-\gamma)\delta u\bar{\alpha}} + \gamma - 1\right)\frac{\bar{N}_b}{z\bar{N}_o}\left(\frac{\delta I\bar{\alpha}}{\bar{\alpha}_o}\right)^{\frac{1}{1-\gamma}}$$

When $\kappa n^{\gamma} >> z(1-\gamma)^2\delta u\bar{\alpha}$, n is decreasing in δ. In other cases n can be increasing.

3.2 Buyer Entry with Observable Types

Prime membership is an important feature of Amazon. The history of the company highlights two crucial turning points: one is the introduction of Prime and adjacent services, the other is Amazon Cloud. To think about Prime membership, however, we need to introduce heterogeneity in consumers. Prime members are wealthier than the average consumer. According to Statista "*as of December 2019, there were an estimated 112 million U.S. Amazon Prime subscribers, up from 95 million in June 2018. On average, Amazon Prime members spent 1,400 U.S. dollars on the e-retail platform per year. March 2019 data also states that non-Prime members only spent 600 U.S. dollars annually.*"[7]

Let us assume that buyers are heterogeneous in their utilities $u_j \in [\underline{u}, \overline{u}]$ and face a fixed entry cost η. Conditional on entry they either participate on the platform or in traditional retail. We consider first the case where u is observable. Conditional on u the directed search market is as before. There exists some threshold type u^* so that all types above enter. The sellers' free entry condition is then

$$z(1-\gamma)U(u^*)\left((1-\theta)\delta\bar{\alpha}n^{-\gamma} + \theta\frac{\bar{\alpha}_o}{I}n_o^{-\gamma}\right) = \kappa + (1-\theta)V_M(\delta),$$

where

$$U(u^*) \equiv \mathbb{E}[u \mid u \geq u^*] = \frac{\int_{u^*} uf(u)\,du}{1-F(u^*)}.$$

Note that U is an increasing function. When u is observable, the indifference condition holds type by type so we still have $\bar{\alpha}_o n_o^{1-\gamma} = \delta I\bar{\alpha}n^{1-\gamma}$. Combining these conditions we get

$$n^{\gamma} = z\bar{\alpha}(1-\gamma)U(u^*)\frac{G(\delta)}{\kappa + (1-\theta)V_M(\delta)},$$

with the same function $G(\delta) \equiv (1-\theta)\delta + \theta\left(\frac{\bar{\alpha}_o}{I\bar{\alpha}}\right)^{\frac{1}{1-\gamma}}\delta^{-\frac{\gamma}{1-\gamma}}$ as in the benchmark case. Taking log differentials, we get

$$\gamma\frac{n'}{n} = \varepsilon^*\frac{(u^*)'}{u^*} + \frac{G'(\delta)}{G(\delta)} - \frac{V_M'}{\frac{\kappa}{1-\theta} + V_M}$$

where $\varepsilon^* = \frac{u^*U'(u^*)}{U(u^*)} > 0$ is the elasticity of the conditional expectation.

The marginal type is defined by the free entry condition of consumers

$$\gamma\delta u^*\bar{\alpha}n^{1-\gamma} = \eta,$$

[7] https://www.statista.com/statistics/304938/amazon-prime-and-non-prime-members-average-sales-spend/

so

$$\frac{(u^*)'}{u^*} = -(1-\gamma)\frac{n'}{n} - \frac{1}{\delta},$$

and combining these conditions we have the following Lemma

Lemma 3. *With free entry by heterogenous buyers and perfect price discrimination, the impact of information disclosure on market tightness is given by*

$$\frac{n'}{n}\left(\gamma + (1-\gamma)\varepsilon^*\right) = \frac{G'(\delta)}{G(\delta)} - \frac{V_M'}{\frac{\kappa}{1-\theta} + V_M} - \frac{\varepsilon^*}{\delta}.$$

Note that our previous model with homogeneous consumers simply corresponds to $\varepsilon^* = 0$, and the sufficient condition in Lemma 2 is unchanged. When consumers are heterogenous, there is an extra factor than can lead to decreasing tightness: the quality of the marginal consumer decreases and thus, the gains from trade decrease.

3.3 Private Types

We now consider the case where sellers do not observe u. When a seller matches with a consumer, she does not know the consumer's type and thus must post the same price. Consider a consumer of type u. On the outside market he gets

$$V_{b,o}(u) = n_o^{1-\gamma}\frac{u - p_0}{I}$$

and on the platform he gets

$$V_b(u) = n^{1-\gamma}\delta(u-p)$$

It turns out that even if types are private, one can construct a separating equilibrium identical to the case in the previous subsection so that buyers self select into the right submarket. To do this we consider a slightly different problem of buyers posting prices and then sellers directing their search. The buyer of type u solves

$$V_b(u) = \max_{n,p} n\alpha(n)\delta(u-p)$$

subject to $v_s = \alpha(n)p$ where n and p are the tightness and price associated with this submarket. The first order condition of this problem implies that $p(u) = (1-\gamma)u$. Suppose that the outside market is also separating and thus the equilibrium tightness levels for submarket u, $(n(u), n_o(u))$ are determined by the indifference condition for the buyers

$$\delta\bar{\alpha}n(u)^{1-\gamma} = \frac{\bar{\alpha}_o}{I}n_o(u)^{1-\gamma}$$

and the free entry condition for the sellers. Finally, as in the previous subsection there is a threshold type u^* so that all types above this enter. By construction this constitutes an equilibrium since sellers are indifferent between all submarkets and buyers are optimizing. Notice that the same allocation constitutes an equilibrium even if sellers post prices. Consider a submarket where sellers post price $p(u) = (1-\gamma)u$. Then the previous formulation implies that buyers of type u prefer to self select into submarket u and thus we have an equilibrium. As before the equilibrium tightness is determined by the buyers indifference condition and the sellers free entry condition.

One way to interpret this separating equilibrium is to assume that the buyer types represent their preferences for quality. Thus higher u types prefer higher quality goods. Consequently, we can think of the sellers as producing different qualities of the good and charging higher prices for higher quality.

4 Conclusion

Our model helps explain why excessive data sharing might not have been a problem with traditional retail brick-and-mortar stores but could become one with internet platforms. We show that there can be both under-sharing and over-sharing depending on the information and matching technologies. When matching is efficient and information is cheap, then excessive sharing is more likely.

Our model focuses on the case of one good. A natural extension is to introduce several goods and to study the decision by consumers and merchants to single-home or to multi-home.

References

Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2019). Too much data: Prices and inefficiencies in data markets. Working Paper.

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

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

Bergemann, D. and A. Bonatti (2019). Markets for information: An introduction. *Annual Review of Economics 11*, 85–107.

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

Bergemann, D., A. Bonatti, and A. Smolin (2018, January). The design and price of information. *American Economic Review 108*(1), 1–48.

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

Caillaud, B. and B. Jullien (2003). Chicken and egg: Competition among intermediation service providers. *The RAND Journal of Economics 34*(2), 309–328.

Gutierrez, G. (2020a). The seller welfare standard. NYU Stern Working Paper.

Gutierrez, G. (2020b). Ten facts about Amazon. NYU Stern Working Paper.

Ichihashi, S. (2019). Online privacy and information disclosure by consumers. *American Economic Review*.

Katz, M. L. and C. Shapiro (1985). Network externalities, competition, and compatibility. *The American Economic Review 75*(3), 424–440.

Khan, L. M. (2016). Amazon's antitrust paradox. *Yale lJ 126*, 710.

Mattiolo, D. (2020). Amazon scooped up data from its own sellers to launch competing products; contrary to assertions to congress, employees often consulted sales information on third-party vendors when developing private-label merchandise. *Wall Street Journal (Online)*.

Moen, E. R. (1997). Competitive search equilibrium. *Journal of political Economy 105*(2), 385–411.

O'Mara, M. (2019). *The Code: Silicon Valley and the Remaking of America*. Penguin Press.

Park, K. F., R. Seamans, and F. Zhu (2020). Homing and platform responses to entry: Historical evidence from the u.s. newspaper industry. *Strategic Management Journal n/a*(n/a).

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

Shimer, R. (1996). *Essays in search theory*. Ph. D. thesis, Massachusetts Institute of Technology.

Wright, R., P. Kircher, B. Julien, and V. Guerrieri (2019). Directed search and competitive search: A guided tour. *Journal of Economic Literature*.