WHAT'S SO SPECIAL ABOUT TWO-SIDED MARKETS?

Benjamin E. Hermalin and Michael L. Katz

Haas School of Business
University of California, Berkeley

1 May 2016

Abstract

An unusual feature of two-sided markets is that there is no consensus regarding what they are. Our approach to deriving a definition is to identify examples that have been found to represent an interesting phenomenon in common and then reverse engineer the outcomes to determine the drivers of what are perceived to be the distinguishing or interesting features of equilibrium. In our view, the central focus of the two-sided-markets literature has been on identifying and analyzing cross-platform externalities. We identify two critical features that give rise to such externalities at the margin: *idiosyncratic matching* and *inefficient rationing*.

Forthcoming in
ECONOMIC THEORY AND PUBLIC POLICIES:
JOSEPH STIGLITZ AND THE TEACHING OF ECONOMICS,
Columbia University Press.

I. INTRODUCTION

A man walks into a bar… What happens next is the subject of many bad jokes and good economics papers. In the past decade or so, economists have devoted considerable attention to understanding markets in which an enterprise (a "platform") facilitates exchange between two or more parties ("platform users"). Singles bars and hook-up apps (*e.g.,* Tinder), which facilitate romantic and/or sexual encounters among their users, are examples. So are payment-card networks (*e.g.*, Visa), which facilitate exchange between merchants and customers. Such markets have come to be called *two sided*, in part because a platform's strategy, if well designed, must concern itself with users on both sides simultaneously.

An unusual feature of two-sided markets is that there is no consensus regarding what they are. There have been many attempts to offer precise definitions of two-sided markets, but none is fully accepted. The name itself is unhelpful: aren't all markets two-sided, bringing buyers and sellers together? Such an expansive conception would provide no guidance for either scholarship or public policy (parties in antitrust litigation and regulatory proceedings frequently invoke two-sidedness as a reason to deviate from traditional antitrust principles).

One might argue that no formal definition is needed and that two-sidedness is like pornography: you know it when you see it. However, like pornography, two people may look at the same thing differently. There is consensus that advertising-supported media (*e.g.,* newspapers and Internet search engines), payment-card networks, and online commerce and social networking platforms (*e.g.,* eBay and Facebook, respectively) constitute two-sided markets. But there is less agreement that communications networks (*e.g.,* telephone networks and Internet service providers) and health insurance plans (which bring together care providers and patients) are two-sided, even though their structures are equivalent to many that are regarded as two-sided. And the literature has excluded manufacturers (even though a manufacturer facilitates exchange between input suppliers and final-product consumers) and tended to exclude retailers.[1] There is little value in having a signifier if there is no agreement what it signifies. Moreover, deriving a meaningful definition may generate insight into the underlying economic phenomenon and clarify its relationship with other distinct phenomena, such as network effects.

Our approach to deriving a definition is to identify examples that have been found to represent an interesting phenomenon in common and then reverse engineer the outcomes to determine the drivers of what are perceived to be the distinguishing or interesting features of equilibrium. In our view, the central focus of the two-sided-markets literature has been on identifying and analyzing settings in which the actions of one set of participants affect the surplus of another set, so that the possibility of externalities arises. The goal of our definition thus is to isolate what factors give rise to externalities that are not internalized by end users when the platform engages in "conventional" pricing. Below, we identify two critical features: *idiosyncratic matching* and *inefficient rationing*.

II. PAST DEFINITIONS OF TWO-SIDED MARKETS

Before offering our own (implicit) definition, we first review existing definitions. There is agreement that the central role of a platform is to facilitate transactions among its users, and that it may undertake a variety of actions to do so. However, this is true of any intermediary. So what—if anything—distinguishes a platform?

---

[1] There are exceptions. For example, Armstrong (2006, p. 684) considers a supermarket to be a platform operating in a two-sided market.

Rochet and Tirole (2003) identify the importance of the externalities that can arise when a user on one side of a platform cares about the number of users on the other side and how they behave (*e.g.,* how intensively they utilize the platform):

> The interaction between the two sides gives rise to strong complementarities, but the corresponding externalities are not internalized by end users, unlike in the multiproduct literature (the same consumer buys the razor and the razor blade). In this sense, our theory is a cross between network economics, which emphasizes such externalities, and the literature on (monopoly or competitive) multiproduct pricing, which stresses cross-elasticities.

Two types of external effects potentially arise: *access externalities* and *usage externalities*. An access externality (referred to as a "membership externality" by Rochet and Tirole, 2006) is the benefit a user on one side generates for users on the other by joining the platform, thereby making himself accessible or available for transactions. A usage externality is the benefit a user on one side generates for a user on the other by actually engaging in transactions on the platform. As an example, your acquisition of a mobile phone generates an access externality for others because they find you easier to reach. Your calling someone on that phone generates an externality (positive or negative) for that person depending on her value from hearing from you.

A common view is that two-sided markets are those in which network effects or access externalities are present. A network effect is said to arise when the value of a good or service to a given user rises as the number of other users of that good or service rises. For example, a given subscriber's value of belonging to a communications network depends on the number of other subscribers on the network. This is a direct effect in that users directly interact with one another. Network effects can also be indirect: the value a user places on an operating system for a computer or smart phone, for example, tends to rise with the number of users because the greater that operating system's total number of users, the greater the incentive for application providers to produce for that operating system. With both direct and indirect network effects, the literature on network effects has focused on how a network effect drives members of one side (*e.g.,* communicators or device users) to cluster together.[2]

The two-sided market literature offers a somewhat different perspective on network effects. Instead of focusing on how the presence of members of one group makes the network more desirable to other members of that same group, the focus has been on inter-group network effects. That is, the two-sided markets literature focuses on how the presence of members of group *A* attracts members of group *B*, and *vice versa*. For example, how does having more merchants willing to accept Visa cards affect consumers' acquisition and use of Visa cards? Because members of the two groups are on opposite sides of the platform this phenomenon has come to be known as a *cross-platform* network effect.

Some researchers have suggested that inter-group or cross-platform network effects are the defining feature of two-sided markets. Indeed, Parker and Van Alstyne's (2005) seminal paper was titled "Two-Sided Network Effects: A Theory of Information Product Design." However, as Rochet and Tirole (2006) illustrate, there is more to the literature than the study of cross-platform network effects: for instance, a cross-group-network-effects definition fails to capture usage externalities.[3] Rochet and Tirole make a platform's pricing to the two sides a central part of their definition: they identify two-sided markets by the presence of cross-platform

---

2 See, for example, Rohlfs (1974) and Katz and Shapiro (1985). For a review of the network effects literature, see Farrell and Klemperer (2007).

3 Rochet and Tirole (2006, p. 657) also observe that, because the effects are mediated through prices, the statement that members of group *A* benefit from an increase in the number of members of group *B* is a statement about the relative utility levels in two different equilibria. Whether or not members of group *A* benefit from additional members of group *B* will depend, in part, on what price changes accompanied the increase. We observe that this critique applies to any model of indirect network effects, and calls for care in interpretation as opposed to undermining the utility of the concept.

welfare effects that the users would not internalize absent pricing by the platform, and these welfare effects can arise from membership or usage decisions or both. Rochet and Tirole (2006, p. 648) offer the following definition for the case of pure usage externalities:[4]

> Consider a platform charging per-interaction charges p_1 and p_2 to the two sides. The market is *not* two sided if the volume of transactions realized on the platform depends only on the aggregate price level, $P \equiv p_1 + p_2$; that is, if it is insensitive to reallocations of the total price P between the two sides. If, in contrast, volume varies with p_i holding P constant, the market is said to be two-sided.

They extend their definition to encompass membership decisions and the corresponding access, subscription, or membership fees. To paraphrase this definition, a two-sided market is one in which the *structure* of prices, as well as their levels, matter.

Although this definition has much to recommend it, it is not perfect. First, it should—and readily can—be extended to consider non-price strategies targeted at the two sides of the market; the structure of these non-price strategies can matter even in situations where the structure of a platform's pricing does not. Second, it may be too broad: any firm can be viewed as facilitating "transactions" between its input suppliers and output buyers. Specifically, consider a firm that sets both input and output prices. The difference between those prices is the sum of the fees the firm charges its input suppliers and output buyers for facilitating transactions between them. Holding that difference constant, the volume of transactions is affected by raising or lowering the two prices by equal amounts. In other words, the pricing structure matters holding the net level constant. Hence, this example falls within Rochet and Tirole's definition. Yet we suspect that

4 We have slightly edited the stated definition in Rochet and Tirole (2006) so it better accords with the conventions and notation of this paper.

few authors would consider the theory of two-sided markets to be equivalent to the theory of the firm.[5]

Rysman (2009) suggests that the focus of a definition of a two-sided market should be on the strategies employed rather than prices. Specifically, if the platform's optimal strategy *vis-à-vis* users on one side of the platform is independent of the strategy it adopts *vis-à-vis* users on the other side, then this is *not* a two-sided market. Otherwise, it is. This definition encompasses a key aspect of a two-sided market, namely that the platform must consider both sides simultaneously. At the same time, we believe Rysman's definition—like Rochet and Tirole's—is too broad, capturing any firm that has market power in both the input and output markets in which it participates. For example, a manufacturer with market power cannot decide its willingness to pay for inputs without estimating the demand for its output, and it cannot set the price of its output without estimating input prices.

Rather than offer a definition that creates bright-line boundaries, Weyl (2010, p. 1644) identifies a set of market characteristics that lead to interesting phenomena. One characteristic is that a platform has market power with respect to users on both sides of the market. However, Weyl narrows the set of firms considered to be platforms by also requiring that: (a) the platform provides distinct services—and can charge distinct prices—to the two sides of the market, and (b) there are cross-platform network effects. We believe this is a very useful perspective but that it is also useful to derive a sharper delineation of what distinguishes a platform from other intermediaries.

---

[5] A third shortcoming is somewhat more technical in nature. Proposition 2 of Hermalin and Katz (2004) shows that, whether only the level P matters for the volume of message exchange in a telecommunications setting depends on the specific distribution of the benefits the sender and receiver derive from such exchange. Many authors, including Rochet and Tirole, would consider platform-mediated communication a two-sided market and there is, thus, something unsatisfactory about a definition that holds except for certain distributions of benefits.

Hagiu (2007) seeks to develop a definition that would expressly rule out retailers and would, by extension, rule out many other types of firms as well. In his view, the key differences between a retailer and a platform are that: (a) the former takes possession (assumes ownership) of the manufacturer's product and has control over its sale to consumers, whereas the latter does not, and (b) the latter, unlike the former, shares control rights over the implementation of final sale to the consumer with the manufacturer. For example, Walmart has very considerable discretion over the prices and other aspects of the transactions with consumers *vis-à-vis* products it previously purchased from manufacturers. In contrast, on eBay, sellers retain considerable control over prices and other aspects of the transaction, such as mode of shipment (although eBay does set policies with which sellers must abide). Furthermore, in many instances, a dissatisfied consumer who buys from Walmart tends to deal with Walmart, whereas a dissatisfied consumer who buys on eBay often deals directly with the seller.

In our view, lack of possession is not the defining feature of a two-sided market. For example, US bookstores traditionally had contractual rights to return, for a refund, unsold books to their publishers. Yet bookstores would presumably not be seen by many as examples of two-sided markets. Moreover, if possession were the defining feature, then it could be difficult to distinguish a two-sided market from agency. For example, a realtor does not take possession of the seller's house, yet many aspects of what a realtor does to facilitate a transaction between a buyer and seller (*e.g.,* staging a house, holding an open house) seem better modeled using the tools of agency theory than those from the two-sided-markets literature.

In addition, there are certain types of retailing that would seem to encompass, in part, aspects of a two-sided market. Consider, for instance, Apple's iTunes (an example also given by Hagiu, 2007). At one point, iTunes set the price at 99¢ per song, which in Hagiu's view makes iTunes a retailer, not a platform. Possession, though, is not a clear concept with respect to digital goods—it is not obvious that iTunes has taken possession of any music in the same way a record store of yore took possession of LPs. Furthermore, there are many aspects of iTunes that appear to encompass two-sidedness, in particular important network effects. For instance, attracting music companies to iTunes was critical to Apple's ability to sell iPods and the number of iPods sold influenced music companies' willingness to license their music to iTunes.

Hagiu and Wright (2011) offer a definition that builds on the notion that platform users retain considerable autonomy with respect to their transactions with one another while also accounting for network effects: "[a multi-sided platform is] an organization that creates value primarily by enabling direct interactions between two (or more) distinct types of affiliated customers." By "affiliated," Hagiu and Wright mean the customers—end users—must deliberately choose to join the platform. The notion of user autonomy is captured by the reference to "direct" interactions (as opposed to the case of a retailer, whose suppliers may never directly interact with its customers). Although network effects are not explicit in this definition, the principal way—at least in modeling—the platform creates value is by internalizing, in part, the relevant externalities when it sets (negotiates) terms with the end users.

Below, we will offer a definition that also focuses on internalizing potential cross-platform externalities but that more fully isolates the relevant factors that make platforms an interesting, distinct phenomenon. We do so by examining the pricing decision of a monopoly platform and identifying market characteristics that give rise to unique market outcomes.

III. ANONYMOUS PAIRING AND EFFICIENT RATIONING

We begin by examining an important special case in which platform pricing corresponds to that of conventional firms. Suppose that a user cares about completing a transaction, but not about with whom. Specifically, the user derives utility *v* from completing a transaction independent of the identity of the user on the other side of the transaction. We call this property *anonymous pairing*. Because exchange is voluntary, a transaction takes place if and only if a pair of users, one from each side, desires it. Let t_i denote the number of transactions that users on side *i* of the platform, *i* =1 or 2, would collectively like to undertake. The fact that it takes two to tango is captured by assuming the total number of transactions actually completed equals $\min\{t_1, t_2\}$. Observe that, when $t_i > t_j$, it will be necessary to ration side-*i* users. We say that there is *efficient rationing* if the *i*-side users who transact are the t_j of them with the highest values of transacting. As we will demonstrate in the remainder of this section, in situations satisfying efficient rationing and anonymous pairing, a platform's pricing is like that of any other firm.

A platform may set both per-exchange, or transaction, prices and membership, or access, prices. For expositional ease, we follow much of the literature in examining the two types of pricing separately even though they can interact.

A. Transaction Pricing

Consider a platform that employs a linear tariff with prices p_1 and p_2 to sides 1 and 2, respectively. The realized utility of an individual on side *i* with value *v* is $v - p_i$ if he transacts and 0 otherwise. Hence, an individual agrees to transact only if $v \geq p_i$. Note that *v* and p_i can be positive or negative. Let $P_i(t)$ denote the t^{th}-highest value of *v* for users on side *i* of the market (*i.e.,* the inverse demand curve of type-*i* users).

Total surplus is

$$W(t_1, t_2) \leq \int_0^{\min\{t_1, t_2\}} \{P_1(z) + P_2(z)\} dz - c \min\{t_1, t_2\} \ ,$$

where c denotes the incremental cost of a transaction. The inequality is an equality under efficient rationing, and rationing is necessarily efficient at prices $P_1(t)$ and $P_2(t)$ because trade is desired by the same number on each side and the value of any completed transaction exceeds the value of any transaction forgone. Because $W(t_1,t_2) \le W(\min\{t_1,t_2\},\min\{t_1,t_2\})$, there is no loss in welfare from restricting attention to prices that induce $t_1 = t_2$. Therefore, the welfare maximizer's program is equivalent to choosing $t \in \arg\max W(t,t)$, which has the first-order condition

$$P_1(t) + P_2(t) = c \quad .[6] \tag{1}$$

Equation (1) states that the sum of the prices equals marginal cost. With efficient rationing and anonymous pairing, the problem is equivalent to an excludable public good problem: the optimal quantity is the one at which the sum of marginal benefits equals the marginal cost of completing another pairing. Observe that the efficient prices are not personalized—all users on a given side of the platform face the same price.

Next, consider a profit-maximizing platform. The platform's profits are

$$\pi(t_1,t_2) \equiv \left(P_1(t_1) + P_2(t_2) - c\right)\min\{t_1,t_2\} \quad .$$

A profit-maximizing platform will also induce $t_1 = t_2$: if, counterfactually, $t_i > t_j$, then reducing t_i would raise the price paid by side-*i* users without affecting the set of completed transactions. The first-order condition for the common value, t, is

$$t\left(P_1'(t) + P_2'(t)\right) + P_1(t) + P_2(t) = c \quad . \tag{2}$$

Comparing (1) and (2), one sees that the platform chooses higher overall prices, which induce a lower level of transactions. *Given the level of transactions,* a profit-maximizing platform

[6] The second-order conditions hold because demand curves slope down.

chooses the same *structure* of prices as would a welfare maximizer: each chooses prices that equalize the two sides' desired levels of transactions.[7] This property is similar to Joe's observation in Stiglitz (1977, p. 410) that profit-maximizing multi-product monopolist sets Ramsey prices conditional on the resulting profit level.

Examination of equations (1) and (2) reveals the following:

- Efficient prices do not rely on assigning some or all of the marginal costs to one side of the market or the other based on notions of which side caused or triggered the costs. Instead, the costs are common costs of facilitating a transaction.
- Profit and welfare maximization depend on demand conditions on both sides of the market simultaneously.
- Because of the usage externality, either a welfare- or profit-maximizing platform may find it optimal to charge one of side of the market a zero or even negative price. Negative prices can arise when completing a transaction is beneficial to one side of the market but costly to the other, so that the latter must be subsidized in order to transact voluntarily.
- The socially optimal prices for the platform's services sum to marginal cost.

In our experience, the first three findings above are often key results highlighted by researchers studying two-sided markets. We note, however, that these results apply to any firm! To see this fact, interpret $P_1(t)$ as the firm's customers' willingness to pay for *t* units of the firm's

---

7 Rochet and Tirole (2003) demonstrate that, holding the sum of the prices to the two sides constant, a profit-maximizing platform may not choose the same price structure as a welfare-maximizing one.

output, and $-P_2(t)$ as the firm's input suppliers' price for supplying the inputs needed to produce the t^{th} unit of output. Then the firm chooses its output level, t, to maximize

$$t\big(P_1(t)-\big(c-P_2(t)\big)\big) \ ,$$

where $c-P_2(t)$ is the firm's marginal cost of supplying output. Although the literature often highlights the "unusual" possibility of levying negative prices on one side of the market, there is nothing unusual about such prices when that side comprises input suppliers (*i.e.*, that a supplier is paid—not charged—for what it supplies).[8] In our view, it is the violation of the fourth finding that distinguishes platforms from other firms, at least as objects of economic analysis: in the key cases of interest in the two-sided markets literature, the socially optimal platform prices do *not* sum to marginal cost.

B. ACCESS PRICING

Now suppose access to the platform is priced, but transactions are free to platform members. Users decide whether to join the platform and then how much to transact. As we will see, the mapping from membership decisions to transaction levels is critical for the results obtained. To help fix intuition, suppose the platform is a singles bar or a hook-up app, and the two sides are heterosexual men and women. A type-v individual receives utility v from being paired off regardless of the partner. The utility of an individual on side i with value v who pays admission fee f_i to join the platform is $v-f_i$ if paired off and $-f_i$ if not. Suppose that the transaction-determination technology is the following: each user can transact with at most one user of the

---

8 This is not to say that these insights are unimportant. That the users on one side of a market can be seen as suppliers of services to users on the other side can be a useful change in perspective. For example, in telecommunications, it has proved useful to recognize that the calling party is not the sole cost causer and, hence, that it can be efficient to charge the calling party less than the cost of a call and have the receiving party pay a positive price.

other side of the platform, and transactions are determined by random pairing. If there are n_1 side-1 users and n_2 side-2 users, with $n_i \geq n_j$, then n_j transactions occur; each side-*j* user transacts with probability 1, but each side-*i* user transacts with probability $n_j / n_i < 1$. Faced with admission fee f_i, an *i*-side user who obtains *v* from transacting will join the platform if and only if $v \min\{n_j / n_i, 1\} \geq f_i$.

Consider the welfare-maximizing prices. First note that $n_1 = n_2$ is necessary for welfare optimization: if one started with unequal numbers and reduced the number of people on the more-populous side by raising the access fee to that side, then the total number of transactions would not change, but costs would fall and the average value of a transaction would rise, where we have used the fact that those users with the highest values of *v* are the ones that will choose to pay the admission fee and join the platform (*i.e.,* rationing is efficient). With equal numbers of members on the two sides, total welfare (surplus) is equal to

$$W(n,n) \equiv \int_0^n \left(P_1(z) + P_2(z)\right)dz - \left(k_1 + k_2\right)n ,$$

where k_i is the marginal cost of a side-*i* user's joining the platform and $P_i(n)$ is the n^{th} highest value of *v* among side-*i* individuals. The first-order condition for welfare maximization is

$$P_1(n) + P_2(n) = k_1 + k_2 \quad . \tag{3}$$

Equation (3) closely parallels the finding for transaction pricing, equation (1). In particular, note the efficient prices sum to the marginal cost of facilitating a transaction through membership.

Next, consider a profit-maximizing platform. Without loss of generality assume $n_1 \leq n_2$. The marginal side-1 user is willing to pay $P_1(n_1)$, and the marginal side-2 user (who is not certain to transact) is willing to pay $P_2(n_2)\, n_1 / n_2$, so the platform's profit is

$$\pi(n_1, n_2) \equiv \left(n_1 P_1(n_1) + n_2 \tfrac{n_1}{n_2} P_2(n_2)\right) - k_1 n_1 - k_2 n_2 \ .$$

If $n_1 < n_2$, then—given demand curves slope down and costs are increasing—the "solution" to maximizing profit would entail $n_2 = 0$, which is nonsense. The platform must choose $n_1 = n_2 = n$. The associated first-order condition is

$$n\left(P_1'(n) + P_2'(n)\right) + P_1(n) + P_2(n) = k_1 + k_2 \ , \qquad (4)$$

which closely parallels the finding for transaction pricing in equation (2).

IV. INEFFICIENT RATIONING

As we will see, the assumptions of efficient rationing and anonymous pairing made above are critical to the result that the welfare-maximizing prices for the platform's services sum to marginal cost. In the present section, we first examine the usage and then access externalities that arise when efficient rationing is violated, which is the case examined by most of the two-sided markets literature. In the section following this one, we present an example in which there is idiosyncratic matching rather than anonymous pairing.

A. TRANSACTION GAMES AND USAGE EXTERNALITIES

Above, we considered a particular technology that generated potential transactions among platform users. We now consider alternative technologies and demonstrate that they can give rise to externalities that affect socially and privately optimal pricing.

As above, assume users do not care about the identity of the user with whom they transact. Now, however, assume a user on one side can transact with *every* member of the other side willing to transact with him or her. For example, if we divide phone users into call initiators and call recipients, with each of the former potentially calling each of the latter, then the number of completed calls is the product of the number of those willing to initiate times the number willing to answer.[9] As we will now show, this multiplicative transaction technology has a profound effect on the nature of privately and socially optimal platform pricing.

Suppose that a *given* user on one side of the market gets a constant benefit per transaction, v, but this value varies across the population on each side. Without loss of generality, normalize the number of users on each side to one. Let $1 - D_i(v)$ denote the distribution function of benefits across side-i users. The use of the letter "D" is not accidental—the survival function, $D_i(v)$, is like a demand curve: $D_i(p_i)$ is the number of side-i users who want to transact when side-i users are charged p_i per transaction. Each user on the *other* side, side j, who wants to transact will thus complete $D_i(p_i)$ transactions, and the total number of transactions is $D_i(p_i)D_j(p_j)$.[10]

The aggregate surplus that members of side i gain from their transactions with any given member of side j is, as usual, the area beneath their demand curve for transactions with that individual and above price; that is, it is

$$S_i(p_i) = \int_{p_i}^{\infty} D_i(p)dp \quad .$$

Because $S_i(p_i)$ is the aggregate surplus side-i users obtain from each side-j user with whom they transact, the total side-i surplus over *all* transactions is $D_j(p_j)S_i(p_i)$. The platform's cost per exchange is c, so its profit (surplus) is $(p_1 + p_2 - c)D_1(p_1)D_2(p_2)$. Summing the platform's and the two sides' surpluses, total surplus is

---

9 See Hermalin and Katz (2004) for a discussion of the complications that arise for the analysis if there is correlation in who calls whom.

10 Alternatively, one could interpret the analysis as being conducted pairwise across the population. In this interpretation, $D_i(p_i)$ is the probability that a given individual on side i obtains a benefit greater than p_i from a transaction with a given individual on side j, $j \neq i$.

$$W(p_1, p_2) \equiv D_2(p_2)S_1(p_1) + D_1(p_1)S_2(p_2) + (p_1 + p_2 - c)D_1(p_1)D_2(p_2) \ .$$

We have expressed this model in terms of prices because that is how it has been developed in the literature.[11] However, it is useful to express it in the same terms we used above. Let $P_i(n)$ denote the n^{th}-highest per-transaction value for users on side *i* of the market. Faced with a per-transaction price of $P_i(n)$, *n* side-*i* users will make themselves available for transactions (*e.g.*, *n* merchants will accept a credit card or *n* consumers will wish to pay using that card). Hence, we can write

$$W(n_1, n_2) \equiv n_2 S_1(P_1(n_1)) + n_1 S_2(P_2(n_2)) + \big(P_1(n_1) + P_2(n_2) - c\big) n_1 n_2 \ .$$

Because $S_i'(p_i) = -n_i$, the first-order conditions for maximizing welfare can be given as

$$P_1(n_1) + P_2(n_2) = c - \frac{S_1\big(P_1(n_1)\big)}{n_1} = c - \frac{S_2\big(P_2(n_2)\big)}{n_2} \ .$$

Observe that the welfare-maximizing prices sum to less than marginal cost. Specifically, the sum of the prices is less than marginal cost by an amount equal to the average surplus that transactions generate for side-*i* users.[12]

An obvious question is: why do efficient prices sum to less than marginal cost here, but summed to marginal cost above? The answer is the lack of efficient rationing. In the present model, the transactions forgone due to a reduction in n_i include transactions that are, from side-*j*'s perspective, inframarginal: they generate positive average surplus for side *j*. This fact is relevant for pricing to side *i* because, when users on side *i* reduce their willingness to transact by

---

11 See, for example, Rochet and Tirole (2003 and 2006), Hermalin and Katz (2004), and Bolt and Tieman (2006).

12 Notice, too, that the welfare-maximizing prices equilibrate the average surplus members of one side derive from transacting with the average surplus members of the other side derive from transacting.

δ, users on the other side of the market lose surplus $\delta S_j(p_j) > 0$. By contrast, with efficient rationing, the transactions forgone due to a reduction in n_i are solely ones that were marginal from side-*j*'s perspective; that is, they yielded side *j* no consumer surplus.

Now consider a platform that seeks to maximize its profit,

$$\pi(n_1, n_2) \equiv \big(P_1(n_1) + P_2(n_2) - c\big) n_1 n_2 \ .$$

The corresponding first-order conditions can be expressed as

$$P_1(n_1) + P_2(n_2) = c - n_1\, P_1'(n_1) = c - n_2\, P_2'(n_2) \ . \tag{5}$$

Because demand curves slope downward, (5) implies that profit-maximizing prices sum to more than marginal cost. Equation (5) also implies

$$\frac{-D_1'(p_1)}{D_1(p_1)} = \frac{-D_2'(p_2)}{D_2(p_2)} \quad . \tag{6}$$

The left-hand (respectively, right-hand) side of (6) is the hazard rate associated with the distribution of benefits for side-1 (respectively, side-2) users. Expression (6) thus states that, at the profit-maximizing prices, the probabilities a random user is indifferent between transacting and not conditional on his being a user willing to transact must be the same for the two sides.

An additional issue is whether, *conditional* on the sum of its prices, the platform chooses an efficient price structure. Rochet and Tirole (2003) and Hermalin and Katz (2004) show that the answer may be yes or no. When no, a second distortion arises: for a given margin, a profit-maximizing platform chooses prices to maximize transaction volume, while the social planner takes into account the surplus those transactions generate for users. In essence, a profit-maximizer is concerned with generating value on the margin, whereas the social planner is also concerned with inframarginal value, a familiar source of distortion (see, *e.g.,* Spence, 1975).

B. MEMBERSHIP GAMES AND ACCESS EXTERNALITIES

We now return to a setting in which transaction prices are taken as given (for convenience, equal to zero) and focus on membership pricing. We consider a platform having a matching technology that results in the number of transactions being $n_1 n_2$ rather than $\min\{n_1, n_2\}$ as above. Interpreted in terms of a singles' bar, every transaction is a conversation, and every man in the bar converses with every woman there, and vice versa. Under this interpretation, $P_i(n)$ is the valuation placed on a transaction (*e.g.*, a conversation) by the member of side *i* with the n^{th} highest valuation. Note the user obtains this value from each transaction in which s/he engages. Hence, the n^{th} user on side *i* derives total benefits $n_j P_i(n)$.

With this pairing technology, total surplus is equal to

$$W(n_1, n_2) \equiv n_2 \int_0^{n_1} P_1(z)dz + n_1 \int_0^{n_2} P_2(z)dz - k_1 n_1 - k_2 n_2 \ .$$

Hence, the first-order condition for welfare maximization with respect to n_i is

$$n_j P_i(n_i) + \int_0^{n_j} P_j(z)dz = k_i \ .$$

Notice that, faced with admission fee f_i, a user on side *i* who values a transactions at *v* will choose to join the platform if and only if $n_j v \geq f_i$. This fact implies that the welfare-maximizing admission fee satisfies

$$f_i = k_i - \int_0^{n_j} P_j(z)dz < k_i \ .$$

Here, too, the sum of the socially optimal prices is less than (total) marginal cost.

As Armstrong (2006) observes, this analysis yields the expected result that, in an efficient equilibrium, each member of side *i* should pay his or her social cost of entering, where the social cost is physical cost, k_i, less the social benefit his or her entering provides to members of the other side, $\int_0^{n_j} P_j(z)dz$. Unlike the special case considered above, the marginal benefit to the other side is not zero: a change in the decision by the marginal person on one side of the market has effects on the welfare of infra-marginal users on the other side.

Recalling that the entry fee on side *i* is $n_j P_i(n_i)$, a profit-maximizing platform acts to maximize:

$$\pi(n_1, n_2) \equiv n_1 n_2 \left(P_1(n_1) + P_2(n_2)\right) - k_1 n_1 - k_2 n_2 .$$

The corresponding first-order conditions can be expressed as

$$f_i = n_j P_i(n_i) = k_i - n_j P_j(n_j) - n_1 n_2 P_i'(n_i) \ .$$

The profit-maximizing price is equal to marginal cost adjusted downward by the effect on the other side's willingness to pay, $-n_j P_j(n_j)$, and upward by the market-power effect, $-n_1 n_2 P_i'(n_i) > 0$.

V. IDIOSYNCRATIC MATCHING

So far, we have assumed a user cares about the number of transactions with users on the other side of the market, but not the identities or types of those users. For instance, in the singles bar examples above, one partner (for the evening or just a conversation) was assumed to be as good as any other. In this section, we will suppose that the partner does matter. We will consider an access-pricing example, but similar considerations arise with transaction pricing as well.

Specifically, return to a model of a singles' bar in which a transaction is pairing (*i.e.*, an individual who enters the bar engages in at most one transaction) and suppose that the bar charges only for admission. Further, as before, suppose that all women are equally good matches, but now suppose there are two types of men: cads and princes. As the names suggest, a woman derives greater value from being matched with a prince than a cad. Consider the *n* men with the highest willingnesses to pay and let $A(n)$ denote the fraction among them who are princes. Let $P_1(n, A)$ denote the value that the n^{th} woman places on being matched with a prince with probability A and a cad with probability $1-A$. Suppose that value is additively separable and the premium placed on princes (normalized to one) is the same for all women; that is,

$$P_1(n, A) = u(n) + A \ , \tag{7}$$

where $u(\cdot)$ is decreasing. As above, the n^{th} man's value of being matched with certainty is $P_2(n)$. Observe the women's side is indexed by 1 and the men's by 2.

Weakly more men than women must patronize the bar at the social optimum. To see why, suppose instead that there were more women than men and, thus, some unmatched women. By raising the admission fee charged to women, the platform could reduce the number of women slightly without affecting the behavior or wellbeing of men. The women who would stop patronizing the bar would be the ones with the lowest values of matching. Hence, costs would fall, the number of matches would remain constant, and the average value of a match would rise. In other words, benefits would rise and costs would fall.

Restricted to values for which $n_1 < n_2$ (*i.e.,* there are more men than women) expected welfare is

$$W(n_1, n_2) \equiv \int_0^{n_1} P_1(z, A(n_2))dz + \frac{n_1}{n_2}\int_0^{n_2} P_2(z)dz - k_1 n_1 - k_2 n_2 \ . \tag{8}$$

Using (7) and differentiating (8) with respect to n_2 yields

$$\frac{\partial W(n_1, n_2)}{\partial n_2} = n_1 A' + \frac{n_1}{n_2}\left[P_2(n_2) - \frac{1}{n_2}\int_0^{n_2} P_2(z)dz \right] - k_2 \, . \tag{9}$$

The term in square brackets is the difference between the marginal and average valuations men have for being paired with a woman, which is negative. Initially assume that the platform is characterized by beneficial selection in that inducing more men to participate lowers the average quality (*i.e.,* $A' < 0$). In this case, it readily follows that (9) is negative for all $n_1 < n_2$. In other words, the social optimum entails equal numbers of men and women patronizing the bar.

Differentiating $W(n,n)$ with respect to *n* yields the first-order condition

$$P_1(n, A(n)) + P_2(n) = k_1 + k_2 - nA'(n) \quad . \tag{10}$$

Given the assumption that $A' < 0$, equation (10) implies that the socially optimal platform prices sum to *more* than marginal cost. Intuitively, charging men a high admission fee leads to a better-quality pool of men, and the benefits of the higher average quality of matches outweighs the loss in benefits from forgone matches. Given the reduced number of men, it is efficient to raise the admission fee to women in order to equalize the numbers of users on the two sides.

Now, suppose that the platform is characterized by adverse selection, so that inducing more men to participate men raises average quality (*i.e.,* $A' > 0$). Assume, however, that effect is sufficiently weak that

$$A'(n_2) + \frac{1}{n_2}\left[P_2(n_2) - \frac{1}{n_2}\int_0^{n_2} P_2(z)dz\right] < 0 \tag{11}$$

for all positive n_2. Then (9) remains negative for all $n_1 < n_2$, so the social optimum again entails $n_1 = n_2$. Correspondingly, (10) remains the first-order condition for the optimal *n*. In this case, (10) implies that the socially optimal platform prices sum to strictly less than marginal cost, a result more typically found in models of two-sided pricing.

For some parameter values, it is socially optimal to have more men than women patronize the singles bar. This is readily seen by considering a slightly modified version of the model: there are *n* women and *n* + *m* men, where the *n* men with the highest values for matching are all cads. Assume, unlike the previous analysis, that women derive no utility if matched with a cad (*i.e.*, $u(\cdot) \equiv 0$), but considerable utility if matched with a prince. Then unless prices induce more than *n* men to patronize the bar, there is no value to women from entering the bar. If women gain enough from matches with princes, and the marginal costs of admission are low enough, it will be efficient to induce more than *n* men to patronize the bar. In other words, it will be efficient to set prices that lead women to be rationed. This result is an example of a general phenomenon identified by Joe and Andrew Weiss in Stiglitz and Weiss (1981): prices play two roles—screening the market participants to induce high-quality users to participate and clearing the market. In the presence of adverse selection, these roles can be in conflict.

In our view, the case of idiosyncratic matching is under-explored in the two-sided markets literature, and there are several directions in which the analysis could be extended.[13] In the analysis considered here, all women agreed on which type of man was the more desirable match: men varied along a vertical quality dimension. One could also imagine horizontal differentiation: users on side *i* of the market have heterogeneous preferences with respect to the types of users on side *j* of the market. In such a market, the quality of one's match may increase with the number of potential partners. In other words, a network effect arises because of the prospect of a superior idiosyncratic match.

---

[13] Within the two-sided markets literature, Damiano and Li (2007 and 2008) examine platforms' use of pricing strategies to induce sorting through self-selection. Following the seminal paper of Gale and Shapley (1962), there is a large literature that examines institutions to facilitate matching when users have idiosyncratic benefits. However, this literature tends to focus on mechanism design rather than taking the matching technology as given and examining how admissions fees and/or transaction prices affect welfare and profits, which has been the focus of most of the two-sided markets literature. For surveys of the matching literature, see Roth and Sotomayor (1990) and Sönmez and Ünver (2010).

VI. CONCLUSION

The literature on two-sided markets has led to many important insights. In our view, these insights arise from two primary factors. One is the recognition that users on one side of a platform can usefully be viewed as inputs to the production of benefits for users on the other side of the platform. Although this recognition does not give rise to entirely new phenomena, it has proven to be a very useful framework for examining a wide variety of business and public policy issues. In this chapter, we have focused on the second factor: the literature has concentrated on situations that give rise to unique features of equilibrium. Specifically, the literature has examined settings in which the marginal decision made by a user on one side of a platform affects the surplus enjoyed by users on the other side. As we have demonstrated, these surplus effects arise when there is inefficient rationing and/or idiosyncratic matching.

VII. REFERENCES

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

Bolt, W. and A.F. Tieman, 2006, "Social Welfare and Cost Recovery in Two-Sided Markets," *Review of Network Economics,* **5**(1), Article 7. Available at http://www.bepress.com/rne/vol5/iss1/7.

Damiano E. and H. Li (2008) "Competing Matchmaking," *Journal of the European Economic Association,* **6**(4), 789-818.

Damiano E. and H. Li (2007) "Price Discrimination and Efficient Matching," *Economic Theory,* **30**(2), 243-263.

Farrell, J., and P. Klemperer (2007) "Coordination and Lock-In: Competition with Switching Costs and Network Effects," in *Handbook of Industrial Organization, Vol. 3*, M. Armstrong and R. Porter (eds.), Amsterdam: Elsevier.

Gale, D., and L. Shapley (1962) "College Admissions and the Stability of Marriage," *American Mathematical Monthly,* **69**: 9-15.

Hagiu, A. (2007) "Merchant or Two-Sided Platform," *Review of Network Economics*, 6(2), 115-13.

Hagiu, A., and J. Wright (2011) "Multi-Sided Platforms," unpublished manuscript, Harvard Business School.

Hermalin, B.E., and M.L. Katz (2004) "Sender or Receiver: Who Should Pay To Exchange An Electronic Message?" *RAND Journal of Economics,* **35**(3): 423–448.

Katz, M.L., and C. Shapiro (1985) "Network Externalities, Competition, and Compatibility," *American Economic Review*, 75(June): 424-440.

Parker, G., and M. Van Alstyne (2005) "Two-Sided Network Effects: a Theory of Information Product Design," *Management Science,* **51**(10): 1494-1504.

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

Rochet, J-C, and J. Tirole (2006) "Two-sided markets: a progress report," *RAND Journal of Economics,* **37**(3): 645-667.

Rohlfs, J. (1974) "A Theory of Interdependent Demand for a Communications Service," *Bell Journal of Economics,* **5**(1), 16-37.

Roth, A. E., and M. Sotomayor (1990), *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis,* Econometric Society Monograph Series, Cambridge University Press, 1990.

Rysman, M. (2009) "The Economics of Two-Sided Markets," *Journal of Economic Perspectives,* **23**(3): 125–43.

Sönmez, T. and M. U. Ünver (2010), Matching, Allocation, and Exchange of Discrete Resources, in J. Benhabib, A. Bisin, and M. Jackson (eds.), *Handbook of Social Economics,* Elsevier.

Stiglitz, J.E. (1977) "Monopoly, Non-linear Pricing, and Imperfect Information: The Insurance Market," *Review of Economic Studies* **44:**407-430.

Stiglitz, J.E., and A. Weiss (1981) "Credit Rationing in Markets with Imperfect Information," *American Economic Review,* **71**(3): 393-410.

Weyl, E.G. (2010) "A Price Theory of Multi-Sided Platforms," *American Economic Review,* **100**(4): 1642-1672.