Anatomy of a Service Marketplace: Liquidity and Growth

Pavel Izhutov and Haim Mendelson*

November 2016

Abstract

The paper is an inquiry into the internal structure of an online marketplace for offline professional services. We develop a structural model of the marketplace and estimate it using data from the 2006-2014 period. Rather than view the marketplace as a "black box" that matches demand and supply, in this paper we take a *market microstructure* approach and explicitly model the matching and operating rules of the marketplace to determine the laws of marketplace evolution and the effectiveness of the matching process. The marketplace has two types of customers: consumers and suppliers. We carefully separate the liquidity of the marketplace, which focuses on the probability of conversion for existing customers, from the adoption process, which focuses on the acquisition of new customers of each type. Adoption reflects both the dissemination of information and prospective customers' expectations of conversion or value creation. We find that geographic location plays an important role in both conversion and adoption. We also find a nonlinear critical mass effect that depends on the pricing policy of the marketplace. We illustrate how a short-term revenue-optimizing strategy for the marketplace may hurt its long-term sustainability.

1 Introduction

Advances in information and communication technologies have resulted in the proliferation of online marketplaces that facilitate the exchange of goods and services. For example, eBay, the Amazon Marketplace and Etsy facilitate the exchange of goods between buyers and sellers and Upwork, Thumbtack and TaskRabbit facilitate the matching of service providers to consumers or businesses that need to get things done. Other examples of marketplaces are education platforms such as Coursera, funding marketplaces such as LendingClub and Prosper, and marketplaces for the use of assets such as living space (Airbnb) or cars (Turo). All of these marketplaces are examples of two-sided platforms (Rochet and Tirole (2006)). According to the Pew Research Center, 72% of Americans have used shared or on-demand online services (Smith (2016)), most of which are marketplace platforms and some venture capital firms dedicate large funds specifically to the marketplace sector (McBride (2013))

The growth of marketplaces is due to the the growth of the world wide web that reduced the cost of interaction, and the development of software that enabled the efficient and effective matching of heterogeneous demand and supply.

*Graduate School of Business, Stanford University, Stanford, California 94305 USA. Email: izhutov@stanford.edu, haim@stanford.edu

Marketplace platforms can sense, store and process vast amounts of demand and supply information and generate valuable transactions quickly and efficiently. Further, some marketplace platforms provide value added services that either help users navigate the ocean of available opportunities or facilitate the execution of transactions and the provision of services following the matching process.

The equilibrium market outcomes obtained in a marketplace are largely influenced by its operating rules which determine the way actual trades take place. Rather than view the marketplace as a “black box” that matches demand and supply, in this paper we take a *market microstructure* approach and explicitly model the matching and operating rules of the marketplace to determine the laws of marketplace evolution and the effectiveness of the matching process. We consider the operation of a marketplace as a two-stage process: first, *adoption*, whereby users on both sides of the platform join it to become customers, followed by *conversion*, whereby demand and supply are matched and exchanges take place. The effectiveness of the conversion process, often called liquidity, depends on the result the adoption process. On the other hand, adoption is dependent on the liquidity of the marketplace.

If we were to view the marketplace as a "black box" or “invisible hand”, it would be appropriate to study adoption using theories of diffusion developed in the Marketing domain. Further, marketplaces are characterized by two-sided (indirect) network effects (Rochet and Tirole (2006)). For example, in a marketplace involving buyers and sellers, more buyers make the marketplace more attractive to sellers, and more sellers make the marketplace more attractive to buyers. It is common in the literature to model network effects via a linear relationship (cf. Caillaud and Jullien (2003), Armstrong (2006)), which again may be appropriate when we view the marketplace as a "black box." In this paper, however, we aim to study marketplace operations at the micro level. We thus need to examine the interaction between the adoption and conversion processes, modeling each explicitly. In particular, both the dissemination of information about the marketplace and its liquidity influence the adoption decision. Thus, rather than assume a specific structure of network effects, their structure is derived from our model.

We develop a framework for analyzing online marketplaces and apply it to a specific service marketplace. In our model, the adoption decisions of users on both sides reflect the value they expect to receive through the platform. Whereas the extant literature addresses different aspects of online marketplaces from specific angles, in this paper we develop an operating model that integrates the adoption and conversion processes and allows us to identify the drivers of liquidity and growth. We find strong empirical support for our model, combining the interdependence of the adoption and conversion processes. Thus, the microstructure of the market affects the structure of network effects. We identify a critical mass phenomenon, suggesting strongly nonlinear network effects, and study its implications. Further, the critical mass effect depends on the pricing policy of the marketplace: early on, a short-term revenue-optimizing strategy will hurt the long-term sustainability of the marketplace. However, such a policy is close to optimal as the marketplace approaches its liquidity limit. We find that geographic location has a paramount effect on both adoption and conversion, implying a further substructure for both processes. For example, liquidity is driven not by the number of suppliers, but rather by the *effective* number of suppliers, determined by their location relative to the consumer. The diffusion of information among users is similarly highly location-dependent.

The following sections review relevant literature (Section 2), describe the model (Section 3), our data (Section 4) and empirical specification and results (Section 5). Implications for marketplace pricing are outlined in Section 6, followed by robustness checks (Section 7) and the concluding remarks (Section 8).

2 Related literature

Marketplaces that are structured as two-sided platforms have become ubiquitous as the Internet became mainstream. The first major marketplace of this type was eBay, launched in 1995, which was followed by hundreds of marketplaces of different types; today's leading marketplaces include AirBnB, Upwork, LendingClub, and eBay itself. Different aspects of these new markets have attracted researchers on both the theory side (e.g., Caillaud and Jullien (2003), Jullien (2005), Parker and Alstyne (2005), Eisenmann et al. (2006), Boudreau and Hagiu (2008), Weyl (2010), Gawer (2011), Hagiu (2014)) and the empirical side (see below). For a review of the theoretical literature on two-sided platforms, see Rochet and Tirole (2006), Rysman (2009) and Evans and Schmalensee (2013). On the empirical side, eBay has been and continues to be a leading object of study with a focus on its reputation system, which establishes a measure of trust between buyers and sellers (cf. Resnick and Zeckhauser (2002), Resnick et al. (2006), Nosko and Tadelis (2015)).

Dozens of papers study other marketplaces in a variety of areas. For example, Fradkin (2015) studied how the search ranking algorithm affects matching efficiency in AirBnB, Fraiberger and Sundararajan (2015) investigated welfare and distributional effects following the introduction of car rental marketplace Getaround, and Cullen and Farronato (2016) examined the TaskRabbit platform. Much of the research on two-sided platforms focuses on the implications of network effects (see, e.g., Katz and Shapiro (1994) for a review of network effects and Nair et al. (2004), Nair (2007) for structural models of indirect network effects in the PDA and video game industries).

In this paper, we open the marketplace "black box" to consider the processes of adoption and conversion. Both processes have been addressed in the literature, albeit separately and in different contexts.

2.1 Adoption

Our approach to customer acquisition or adoption focuses on the communication structure between existing customers and prospective new customers on both sides of the marketplace.

Early models in the marketing literature (e.g., Rogers (1962), Bass (1969)) focused on diffusion-based imitation or adoption based on word of mouth. In Bass (1969) and follow-up work (see, e.g., Mahajan et al. (1990) for a review), the higher the number of customers who are already using the product, the higher the probability that imitators will further adopt it. Models of viral adoption, based on word of mouth communications, are reviewed in Godes and Mayzlin (2004); see also Granovetter (1978) for a threshold-type model and Goldenberg et al. (2001) for a cascade-type model. Aral et al. (2009) separated the effects of viral adoption from those of homophily, and Aral and Walker (2011), using a randomized experiment on the Facebook platform, showed how a product's viral features improve its distribution. Katona et al. (2011) and Aral and Walker (2014) study the effects of network structure and social influence on the adoption of social networks. Mendelson and Moon (2016) study the adoption and survival of Facebook apps, identifying a strong viral effect.

Leskovec et al. (2007) analyzed a product recommendation network, studying the relation between purchases and recommendations and identifying effective viral distribution strategies. Bakshy et al. (2012), using a randomized experiment with the Facebook news feed, characterized the viral nature of sharing behavior. In addition, Leskovec et al. (2008) presented a network formation model with a focus on edge formation. Based on data from Flickr, Delicious, LinkedIn and Yahoo! Answers, they note that the node arrival process is network-specific (exhibiting exponential, super-linear, quadratic and sub-linear growth patterns, respectively) and therefore hard to endogenize. Tucker and Zhang (2010), using a randomized experiment in a B2B marketplace, find a negative effect on new seller listings of a large number of sellers displayed along with the number of buyers. The effect disappears when only the number of sellers is shown. A review and comparison of such experiments can be found in Aral and Walker (2014).

"Virality" is naturally studied in epidemiology, and the associated SIRS-type models may be applied to study viral product growth. For instance, Ribeiro (2014) employed a model of this sort for studying the dynamics of user engagement in a membership-based web platform, and Young (2009) provides a more general comparison between "contagion", "social influence" and "social learning" classes of adoption models.

2.2 Conversion

The literature in this domain has its roots in general equilibrium theory, dating back to (Walras (1874)) and the trial-and-error "Walrasian auction" mechanism that served as the "invisible hand" of the market. Although Walras did not model the actual exchange mechanism, the gap was filled by multiple successors in the mechanism design and auction and related search theory literature.

Matching

The literature on the optimal assignment of resources (goods or people) of different types dates back to Koopmans and Beckmann (1957), Gale and Shapley (1962), Shapley and Shubik (1971), and Becker (1973). The main focus is typically on characterizing the optimal allocation and designing a decentralized mechanism that implements it, which led to the deferred acceptance algorithm (Shapley and Shubik (1971)) and the concepts of positive and negative "assortative matching" (Becker (1973); for further details, see Roth and Sotomayor (1992)).

Another group of researchers focused on allocating goods among privately informed buyers, which formed the basis of auction theory and, more generally, mechanism design (cf. Vickrey (1961), Myerson (1979), Harris and Raviv (1981), Myerson (1981), Riley and Samuelson (1981), Wilson (1967), Milgrom and Weber (1982)).

Search and market design

The problem of matching workers and hiring firms motivated researchers to study the mechanics of matching. This literature can be classified along two dimensions. First, the search may be *non-sequential*, i.e., search is performed only once and the best option among those identified is selected (e.g., Stigler (1961)), or *sequential*, where the decision to continue the search is made repeatedly as options arrive (e.g., McCall (1965)). Second, the search may be *random*, i.e., potential options arrive randomly (e.g., McCall (1965)), or *directed,* i.e., there is some additional information that "directs" the search process; this information is typically contained in the prices of different options (e.g., Butters (1977)). The directed and sequential search settings are more relevant to the model studied here. In particular, a customer waiting for the best available supplier match is loosely similar to a job seeker looking for the employer offering the highest wage. McCall (1965) and McCall (1970) showed that having a reservation wage and waiting for an employer offering a wage above this threshold is the optimal strategy. Directed search models such as Butters (1977), Peters (1984), McAfee (2008), Acemoglu and Robert Shimer (1999), Kenneth Burdett et al. (2001) and Eeckhout and Kircher (2010) are also somewhat conceptually similar to a marketplace "directing" suppliers to different consumers by posting prices.

In addition to the papers cited above, applied research on *market design* dates back to Roth (1984) (see also Roth (2008)). While most of the applied market design results were derived for a static setting, some researchers developed different forms of dynamics (cf. Mendelson (1982), Ashlagi et al. (2013), Akbarpour et al. (2014), Arnosti (2015)). Roth (2008) provides a review of the importance of thickness, congestion and trust in market design. In the finance literature, "market microstructure" refers to the design of financial markets with a focus on liquidity (see, e.g., Amihud et al. (2006) for a review). Mendelson and Tunca (1999) have applied the notion of liquidity to a business-to-business marketplace.

Summary

Although there is a vast literature on adoption, matching and market design, none of these works is directly applicable to our setting, where we model the actual operation of a service marketplace. The reason is the salient features of a typical online marketplace. First, online marketplace resources are dynamic and perishable: consumers are impatient and have an opportunity to find a match outside the marketplace. Similarly, suppliers do not continuously stand ready for the best match. Second, user heterogeneity affects the speed of matching and conversion. Third, the search process in a marketplace is directed, sequential and two-sided, with both sides actively participating in the process. Fourth, marketplaces use simpler allocation rules and algorithms than those used in the market design literature. Fifth, online marketplaces are typically characterized by a viral adoption process which is not present in matching or network effect models of markets.

In this paper, we develop an operating model of an online service marketplace which incorporates these features, we empirically estimate it, test the predictions of the model, identify the relative strengths of the different forces that drive marketplace performance and growth, and study the implications of the results.

3 Model

We consider two types of potential marketplace users: *suppliers,* professionals who may provide services, and *consumers* who may choose to receive service from them. There are V types of services $v = 1, 2, \ldots, V$ which we call "verticals". Suppliers and consumers live on a discrete geographic grid X that evolves over time t. When a consumer ("she") wishes to find a suitable supplier ("he"), she places a service *request* for service at the marketplace, which acts as an intermediary and searches for a suitable supplier. If a suitable supplier is found and service takes place, the request is considered successful ("conversion") and the marketplace charges a supplier fee f (as discussed later, the fee depends on the state of the marketplace). Suppliers register on the marketplace and wait for service requests. We say that a consumer adopts the marketplace when she submits her first service request. Suppliers adopt by registering on the marketplace, meaning that they are willing to receive new service requests. Suppliers and consumers cannot directly identify each other.

The number of consumers in location x, vertical v and time period t who adopted the marketplace is denoted D_{xvt} ("D" for demand), and the corresponding number of suppliers by S_{xvt} ("S" for supply). Variables with no superscripts (e.g., S_{xvt}) indicate the number of adopters. Possible superscripts include: "reg" denoting the number of registrations and "req" denoting the number of consumer requests in a particular time period. While number of adopters are stock variables, "reg" and "req" are flow variables. In addition to superscripts, we use subscripts. In particular, subscript "$-x$" is used to denote the summation of a variable across all locations except location x (e.g., $S_{-xvt} = \sum_{y \neq x} S_{yvt}$). Similarly, subscript "$-v$" indicates summation across all verticals except vertical v (e.g., $D_{x,-v,t} = \sum_{w \neq v} D_{xwt}$) Relative to customers in neighborhood x, suppliers located in the same neighborhood x are called "local". Suppliers located in other neighborhoods are called "remote".

3.1 Conversion process

Once the marketplace receives a service request i, it sets a fee f to be charged to any supplier who will serve the request. Any registered supplier j exposed to request i evaluates his willingness to serve this request. If the supplier is wiling to serve the request, he communicates this willingness to the marketplace via a reply. On behalf of the consumer, the marketplace collects all the replies to request i as they arrive until there is a reply from a supplier j^* who delivers value to the consumer above a threshold u_0. However, the service request may expire by that point. If the consumer service request has not expired, she evaluates her willingness to be served by supplier j^*. Further, if the consumer request has not expired *and* she is wiling to be served by supplier j^*, service takes place, request i becomes successful and the marketplace collects its fee f. The decision of a consumer to place a request depends on the probability the request will result in actual service.

We begin the analysis of the conversion process by discussing the marketplace pricing (fee-setting) policy. We then proceed to analyze the response of a randomly chosen supplier j to a randomly chosen request i. After that, we describe how the marketplace decides when to stop collecting replies from suppliers. We then specify the determinants of success following a supplier assignment. In particular, we describe the request expiration process and the associated request "survival" curve. Then, we proceed to the consumer's decision to place a request. We consider this decision last, because it depends on the probability of success which in turn depends on the stages listed above.

Marketplace pricing policy

The marketplace we study fixes a fee for each service request, regardless of which supplier will fulfill it. The fee depends on a number of variables. First, it depends on the request parameters – e.g., service duration, location, nature of service, vertical, consumer's gender, etc. Second, it depends on market conditions, i.e. the number of suppliers and incoming request. All suppliers face the same matching fee for a given service request.

We model the fee associated with consumer request i as

$$f_{ixvt} = \theta_M A_i + \vec{\omega} \cdot \left[S_{xvt}, S_{-xvt}, D^{req}_{xvt}, D^{req}_{-xvt}\right] + \epsilon_{ixvt}, \tag{1}$$

where A_i is the expected revenue that a generic supplier will generate from request i, θ_M is the share of transaction revenue charged by the marketplace, v_{xv} is a fee correction related to the location and vertical, $\vec{\omega} \cdot \left[S_{xvt}, S_{-xvt}, D^{req}_{xvt}, D^{req}_{-xvt}\right]$ is correction related to market conditions, and ϵ_{ixvt} is random noise. The expected revenue A_i is drawn from an arbitrary distribution with mean A_0, independent of any other random variables. After the marketplace assigns a fee to a service request, it can be evaluated by suppliers.

Supplier response

Consider a consumer request i placed at location x, vertical v and time period t. There are S_{xvt} local suppliers and S_{-xvt} remote suppliers who can potentially serve it. Any of these suppliers may discover the request at rate $\delta \in (0,1)$ per unit time, so the number of suppliers discovering the request over a time period of length τ is Poisson-distributed with mean $\delta\tau$. Upon discovery, supplier j finds the request valuable if and only if $v_j\left(ij\right) = A_i - f_{ixvt} - \beta^S I\left(remote\right) + \epsilon_j\left(ij\right) > v_j\left(0\right)$, where $v_j\left(ij\right)$ is the utility that supplier j gets upon serving request i, A_i is the revenue that the supplier expects to receive upon serving the order, f_{ixvt} is the fee charged by the marketplace. β^S is a "distance" penalty, incurred whenever supplier serves a remote order (e.g., transportation cost) and $\epsilon_j\left(ij\right) \sim Exp\left(1/v_j^0\right)$ is a random shock that depends on the specific supplier-consumer pair, where $v_j\left(0\right)$ is supplier's outside option assumed common to all suppliers. All the components of $v_j\left(ij\right)$ are observed by supplier j before he replies.

The probability that a random supplier j replies to request i upon its discovery is

$$exp\left(\frac{A_i - v_j\left(0\right) - f_{ixvt} - \beta^S I\left(remote\right)}{v_j^0}\right). \tag{2}$$

The total reply rate to order i is thus

$$\delta exp\left(\frac{A_i - f_{ixvt} - v_j\left(0\right)}{v_j^0}\right)\left(S_{xvt} + S_{-xvt}exp\left(-\frac{\beta^S}{v_j^0}\right)\right)$$

The number of local and remote replies received over time τ are then Poisson-distributed with rates $\delta\tau exp\left(\frac{A_i - f_{ixvt} - v_j(0)}{v_j^0}\right)S_{xvt}$ and $\delta\tau exp\left(\frac{A_i - f_{ixvt} - v_j(0)+\beta^S}{v_j^0}\right)S_{-xvt}$ respectively. It follows that the expected inter-arrival time for suppliers' replies to request i is

$$E\left[\xi_{ixvt}\right] = \left[\delta exp\left(\frac{A_i - f_{ixvt} - v_j\left(0\right)}{v_j^0}\right)\left(S_{xvt} + S_{-xvt}exp\left(-\frac{\beta^S}{v_j^0}\right)\right)\right]^{-1}. \tag{3}$$

Marketplace stopping policy

The utility of the consumer who placed request i equals $U_{ixvt} = \theta_c A_i u_{ixvt}$, where $\theta_c A_i$ is the value exchanged between consumer and supplier (it is proportional to the supplier's revenue A_i that the supplier gets) and u_{ixvt} is the quality of the match that the marketplace delivers to the consumer. The marketplace aims to provide service quality above threshold u_0, common to all consumers, which corresponds to the consumer's reservation utility $U_i\left(0\right) = \theta_c A_i u_0$. Thus, for each service request the marketplace stops when it receives a reply from a supplier with utility above level $A_i u_0$. Put differently, the marketplace stops the search at τ if $\theta_c A_i u_{ixvt}\left(\tau\right) \geq \theta_c A_i u_0$. Or, equivalently, $u_{ixvt}\left(\tau\right) \geq u_0$, where $u_{ixvt}\left(\tau\right)$ is the highest quality that the marketplace can deliver for request i in time τ. The optimal stopping time τ^*_{ixvt} is thus $\tau^*_{ixvt} = min\left(\tau\right)\ \ s.t.\ u_{ixvt}\left(\tau\right) \geq u_0$, where u_0 is the matching quality threshold.

In order to compute the stopping time distribution and request success probability we need to define the matching quality.

Matching quality

The marketplace acts on behalf of the consumer and picks the best supplier $j^*\left(\tau\right)$ who delivers maximum service quality $u_{ixvt}\left(\tau\right)$. The service quality provided by supplier j to consumer i equals $v_i\left(ij\right) = v_i^0\left(ij\right) - \beta^D I\left(j\ remote\right)$, where $v_i^0\left(ij\right) \sim Exp\left(v_i^0\right)$ is the service quality that supplier j can deliver to consumer i, v_i^0 is a constant parameter common to all consumers, and β^D is the consumer "distance" cost incurred whenever the service is performed by a remote supplier. The marketplace observes the realization of service quality $v_i^0\left(ij\right)$ before choosing the best match.

The supplier that best fits order i is $j^* = argmax_{j\in J(\tau)}\left(v_i\left(ij\right)\right)$, where $J\left(\tau\right) = \{j:\ v_j\left(ij\right) \geq v_j\left(0\right),\ j \text{ discovered request } i \text{ by time } \tau\}$ is the set of suppliers who replied to request i over time interval τ since the request was placed. The quality that the consumer gets upon his assignment is $u_{ixvt}\left(\tau\right) = v_i\left(ij^*\right) = max\left(u^x_{ixvt}\left(\tau\right),\ u^{-x}_{ixvt}\left(\tau\right)\right)$. where $u^x_{ixvt}\left(\tau\right)$ is the highest quality that can be derived from the pool of local suppliers and $u^{-x}_{ixvt}\left(\tau\right)$ is the highest quality that can be derived from the pool of remote suppliers who responded over time τ. Given the optimal stopping policy and stopping time τ^*, we can derive that $u_{ixvt} \equiv u_{ixvt}\left(\tau^*\right) = max_{j\in J_{ixvt}(\tau^*)}\left(u_i\left(ij\right)\right)$. We can now compute a few matching process-related parameters.

Probability that the assigned supplier is local

In the Appendix we show that the probability that the supplier assigned to service request i is local is

$$Pr\left(u_{ixvt}^{x} > u_{ixvt}^{-x} \middle| u_{ixvt} \geq u_0\right) = \frac{Pr\left(u_{ixvt}^{x} > u_{ixvt}^{-x},\ u_{ixvt} \geq u_0\right)}{Pr\left(u_{ixvt} \geq u_0\right)} = \frac{S_{xvt}}{S_{xvt}^{eff}} \tag{4}$$

where

$$S_{xvt}^{eff} = S_{xvt} + S_{-xvt} exp\left(-\frac{\beta^S}{v_j^0} - \frac{\beta^D}{v_i^0}\right)$$

is the "effective" supply base (discussed below).

The probability of local assignment depends not only on the supply-side parameters $\beta^S,\ v_j^0, v_j(0)$, but also on the demand-side distance cost parameter β^D. The higher the distance cost, the higher the probability of the assignment being local. This probability does not depend on the stopping time τ^*.

Stopping time distribution

The optimal stopping time τ^* is distributed exponentially because

$$Pr\left(\tau^* < \tau\right) \quad = \quad Pr\left(u_{ixvt}\left(\tau\right) \geq u_0\right) \tag{5}$$

In the Appendix we show that

$$Pr\left(u_{ixvt}\left(\tau\right) \leq u\right) = exp\left(-\delta t exp\left(\frac{A_i - f_{ixvt} - v_j\left(0\right)}{v_j^0} - \frac{u_0}{v_i^0}\right) S_{xvt}^{eff}\right),$$

hence $Pr\left(\tau^* < \tau\right) = 1 - exp\left(-\mu_{ixvt}\tau\right)$, where $\mu_{ixvt} = \delta exp\left(\frac{A_i - f_{ixvt} - v_j(0)}{v_j^0} - \frac{u_0}{v_i^0}\right) S_{xvt}^{eff}$ is the "matching rate" for request i.

In addition, we can easily derive the expected stopping (matching) time that equals the inverse of the matching rate:

$$E\tau_{ixvt} = \frac{1}{\mu_{ixvt}} = \frac{exp\left(\frac{-A_i + f_{ixvt} + v_j(0)}{v_j^0} + \frac{u_0}{v_i^0}\right)}{\delta S_{xvt}^{eff}}. \tag{6}$$

The higher the effective supply base S_{xvt}^{eff} and discovery rate δ, the smaller the matching (stopping) time. The higher the fee f and quality threshold u_0, the longer the matching (stopping) time.

The effective supply base S_{xvt}^{eff} reflects the reduced value of remote supplier-consumer pairs, that is pairs in which a consumer and supplier live in different locations. Thus, we conclude that it is not the total supply base that matters for conversion but rather the effective supply base that takes into account the *heterogeneity* among users.

The higher the effective supply base S_{xvt}^{eff} and discovery rate δ, the higher the matching rate μ and the lower the matching (stopping) time τ^*. The higher the fee f and consumer reservation value u_0, the lower the matching rate. That is, the outside options available to suppliers and consumers, as well as the marketplace fee contribute to the matching frictions, whereas the effective supply base reduces these frictions.

In order to derive the probability that the matching process results in a service, i.e., the probability of request success, we need to specify the request expiration process.

Request expiration and request conversion to service

Each request "expires" according to survival function $S(\tau)$. That is, the probability that the request has not expired by time τ is $S(\tau) = Pr(\tau_{exp} > \tau)$, where τ_{exp} is the expiration time. The necessary condition for the request to be successful is $\tau^* < \tau_{exp}$, that is the optimal stopping time τ^* is less than the request expiration time τ_{exp}. In other words, the search process results in an acceptable match faster than the consumer finds a supplier on her own or changes her mind.

Even if the search completes before the request expires, it does not guarantee that the request will result in a service. There is still a "remaining" probability $(1-\pi_0)$ that the supplier and consumer will not find each other attractive enough to proceed to the service. Thus, the overall success probability of a request is given by

$$\phi_{ixvt} = \pi_0 Pr\left(\tau^*_{ixvt} < \tau_{exp}\right) = \pi_0 S\left(\tau^*_{ixvt}\right) \tag{7}$$

Assuming that the time to request expiration is distributed exponentially with mean $1/\tau_e$ we get

$$\phi_{ixvt} = \pi_0 exp\left(-\tau^*_{ixvt}/\tau_e\right) \tag{8}$$

Since $\tau^*_{ixvt} \sim Exp\left(\frac{\delta S^{eff}_{xvt}}{exp\left(\frac{-A_i+f_{ixvt}+v_j(0)}{v^0_j}+\frac{u_0}{v^0_i}\right)}\right)$ the request success probability equals (see Appendix)

$$\phi_{ixvt} = \pi_0\left(1+\left(\mu_{ixvt}\tau_e\right)^{-1}\right)^{-1}$$

where μ_{ixvt} is the matching rate.

Liquidity

In our service marketplace, each consumer request that expires at rate $1/\tau_e$ "meets" suppliers who discover service requests at rate δ. The discovery process is impeded by the matching fee and the distance, which both reduce suppliers' willingness to respond to the request. The resulting supplier response process can be described as two-dimensional marked Poisson processes that represents independent reply streams from local and remote suppliers respectively[1]. The parameters of these Poisson processes depend on the expected transaction revenue, marketplace matching fee and supply base. A larger

[1]for more details on the marked Poisson process, see Taylor and Karlin (2014)

supply base speeds up the reply process and increases the choice set for the marketplace to assign the best match from. Both of these effects increase the probability of conversion of a consumer request into actual service. Specifically, the conversion probability of request i is given by

$$\phi_{ixvt} = \pi_0 \left(1 + \left(exp\left(\frac{A_i - f_{ixvt} - v_j(0)}{v_j^0} - \frac{u_0}{v_i^0}\right)\delta\tau_e S_{xvt}^{eff}\right)^{-1}\right)^{-1} \tag{9}$$

In (9), if we increase the number of local and remote suppliers, the conversion probability of a consumer request increases, making the consumer better off. Thus, it is the effective supply base $S_{xvt}^{eff} = S_{xvt} + S_{-xvt}exp\left(-\frac{\beta^S}{v_j^0} - \frac{\beta^D}{v_i^0}\right)$ that drives the liquidity of the marketplace. The effective supply base weighs remote suppliers using the remote value depreciation factor $exp\left(-\frac{\beta^S}{v_j^0} - \frac{\beta^D}{v_i^0}\right)$. Thus, similar to findings in the finance literature (cf. Amihud et al. (2006)), market liquidity depends on the micro-foundations of the marketplace. To understand the drivers of market liquidity we need to analyze the marketplace and derive the probability of conversion.

The conversion probability is an increasing and concave function of the effective supply base with asymptotic value π_0, which is reached as $S_{xvt}^{eff} \to \infty$. The conversion probability ϕ is given by

$$\phi(S) = \frac{\pi_0}{1 + \frac{1}{kS^{eff}}} \tag{10}$$

where S^{eff} is the effective supply base. When the effective supply base is 0, conversion probability is 0. As the effective supply base increases, the conversion probability reaches the asymptotic limit π_0 (see Figure 1). The asymptotic conversion probability π_0 is outside of direct control of the marketplace and is limited by the data available for matching.

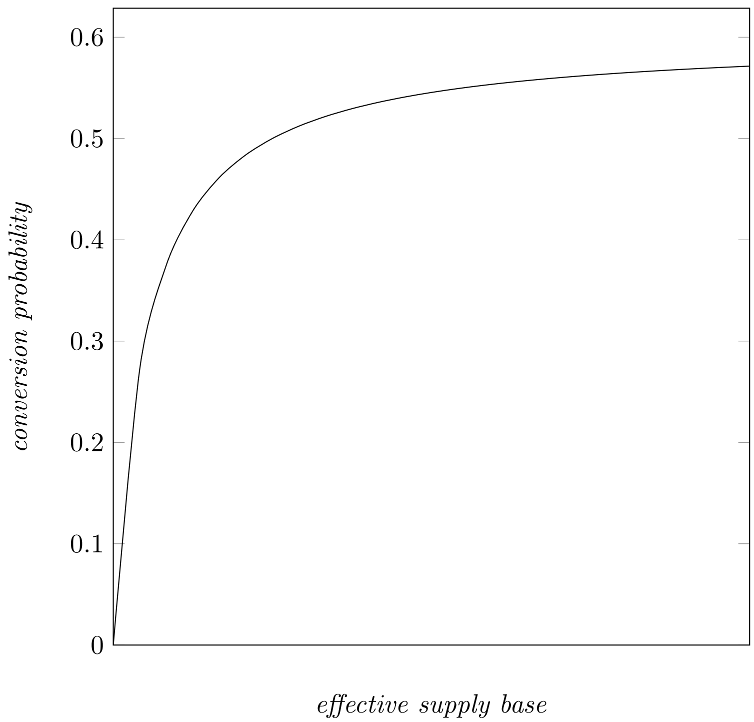


Figure 1: Conversion probability as a function of effective supply base

By (10), for a finite supply base, the coefficient

$$k = \delta\tau_e exp\left(\frac{A_i - f_{ixvt} - v_j\left(0\right)}{v_j^0} - \frac{u_0}{v_i^0}\right)$$

increases the conversion probability. A higher discovery rate δ, longer request expiration time τ_e, higher revenue A_i, lower marketplace fee f_{ixvt}, lower suppliers' and consumers' outside options $v_j\left(0\right)$ and u_0 – all increase the probability of success.

Conversion probability can be also expressed as

$$\phi\left(S\right) = \frac{\pi_0}{1 + exp\left(-\left(logS - log\left(1/k\right)\right)\right)}$$

that is, the conversion probability is a logistic function of the log-supply base ("log-network effects"). That is, the network effect on the supply side is nonlinear and follows a logistic shape.

3.2 Adoption

We next turn to the adoption process, which focuses on customer acquisition. There are different adoption processes for consumers and suppliers.

Consumers

A consumer's life cycle starts when she places her first order (service request). She learns about the marketplace either directly or by hearing about it from other consumers; the latter form is called "viral effects." For some services, which are inherently seasonal (e.g., travel, education, or shopping-related services) this dissemination of information is seasonal as well. The consumer may then place other orders (something that rarely happens in our marketplace). The consumer's life cycle ends when she churns, which happens after she did not place any orders over a long enough period of time. If D_{xvt} and D_{-xvt} are the numbers of local and remote consumers in the system, the flow of new service requests per unit of time in neighborhood x is given by

$$D_{xvt}^{req} = \pi_{xvt}^{D} \vec{V}^{D} \cdot [D_{xvt}, D_{-xvt}, 1]$$

where $\vec{V}^D = \left[V_x^D, V_{-x}^D, V_0^D\right]$ is the vector of viral coefficients and π_{xvt}^D defines the probability that a consumer who evaluates the decision to join a marketplace, actually joins it. As a result, $\pi_{xvt}^D V_x^D D_{xvt}$ and $\pi_{xvt}^D V_{-x}^D D_{-xvt}$ are the numbers of consumers who joined due to the effect of local and remote consumers respectively, and $\pi_{xvt}^D V_0^D$ – consumers who joined by non-viral means (e.g., search). The value of π_{xvt} depends on the probability that a consumer will ultimately receive service, i.e., convert. We assume that to place a request, consumers require a conversion probability which is higher than $(\phi_0 + \phi_i)$ – the sum of the baseline conversion ϕ_0 and an idiosyncratic threshold ϕ_i, drawn from an exponential distribution with mean ϕ, which is common for all consumers:

$$\phi_i \sim Exp\left(1/\phi\right).$$

Thus,

$$\begin{aligned}
\pi_{xvt}^D &= 1 - exp\left(-\left(\phi_{xvt} - \phi_0\right)/\phi\right) = \\
&= \frac{\phi_{xvt} - \phi_0}{\phi} + o\left(\frac{\phi_{xvt} - \phi_0}{\phi}\right) \\
&\simeq \frac{\phi_{xvt} - \phi_0}{\phi}.
\end{aligned}$$

Suppliers

A supplier's life cycle starts when he registers on the system. Supplier registrations may result from viral means (i.e., learning about the marketplace from other suppliers) or from non-viral means such as search. Suppliers churn when they have not replied to any service request for a long-enough period of time. If S_{xvt} and S_{-xvt} is the number of local and remote registered suppliers in the system, the number of new suppliers registering per unit of time in neighborhood x is given by

$$S_{xvt}^{reg} = \pi_{xvt}^{S} \vec{V}^{S} \cdot [S_{xvt}, S_{-xvt}, 1]$$

where $\vec{V}^S = \left[V_x^S, V_{-x}^S, V_0^S\right]$ is the vector of viral coefficients and π_{xvt}^S defines the probability that a suppliers who evaluates the decision to join the marketplace, does so. As a result, $\pi_{xvt}^S V_x^S S_{xvt}$ and $\pi_{xvt}^S V_{-x}^S S_{-xvt}$ are the numbers of suppliers who joined due to local and remote suppliers, respectively, and $\pi_{xvt}^S V_0^S$ – suppliers who joined by non-viral means. The value of π_{xvt}^S depends on the revenue a registered supplier is expected to get, which is in turn proportional to the rate of successful service requests $D_{xvt}\phi_{xvt}^S$. A supplier who learned about the marketplace thus decides to register if and only if the per-period value rate in a particular location and vertical, per local supplier in the vertical, is greater than $\left(\theta_0 + \theta_j^S\right)$ – the sum of the baselines rate θ_0 and an idiosyncratic threshold θ_j^S, drawn from an exponential distribution with mean θ, which is common for all suppliers:

$$\theta_j^S \sim Exp\left(1/\theta\right).$$

Thus,

$$\begin{aligned}
\pi_{xvt}^S &= 1 - exp\left(-\left(\theta_{xvt} - \theta_0\right)/\theta\right) = \\
&= \frac{\theta_{xvt} - \theta_0}{\theta} + o\left(\frac{\theta_{xvt} - \theta_0}{\theta}\right) \\
&\simeq \frac{\theta_{xvt} - \theta_0}{\theta},
\end{aligned}$$

where

$$\theta_{xvt} = \frac{D_{xvt}^{req}\phi_{xvt}^x}{S_{xvt}} + exp\left(-\frac{\beta^S}{v_j^0}\right)\sum_{-x}\frac{D_{yvt}^{req}\phi_{yvt}^{-y}}{S_{-yvt}}$$

is the expected value rate suppliers are facing when the value of a local request is normalized to 1, D_{xvt}^{req} is the number of consumer requests, $exp\left(-\frac{\beta^S}{v_j^0}\right)$ is the value of a remote request relative to a local one (see Appendix), ϕ_{yvt}^x is the probability that the order originated in location y is successfully matched to a supplier in location x,

In particular, ϕ_{xvt}^x is the probability that the request originating in location x will be served in location x, and ϕ_{yvt}^{-y} is the probability that the request originated in location y will be served outside of location y. In other words, a supplier is facing the with S_{xvt} suppliers for local requests and competition with S_{-yvt} for (remote) requests originated in each location $y \neq x$. Each remote request's value is discounted by the distance factor. We show in Appendix that $\phi_{xvt}^x = \frac{S_{xvt}}{S_{xvt}^{eff}}\phi_{xvt}$ and $\phi_{xvt}^{-x} = \phi_{xvt} - \phi_{xvt}^x = \phi_{xvt}\left(1 - \frac{S_{xvt}}{S_{xvt}^{eff}}\right)$.

Summary

It follows that the number of consumer service requests and supplier registrations is given by

$$\begin{cases} D_{xvt}^{req} &= \frac{\phi_{xvt}-\phi_0}{\phi}\vec{V}^D \cdot [D_{xvt}, S_{xvt}, D_{-xvt}, S_{-xvt}] + \epsilon_{xvt} \\ S_{xvt}^{req} &= \frac{\theta_{xvt}-\theta_0}{\phi}\vec{V}^S \cdot [D_{xvt}, S_{xvt}, D_{-xvt}, S_{-xvt}] + \epsilon_{xvt} \end{cases} \tag{11}$$

4 Data

We have obtained a rich dataset from an online marketplace for offline professional services. The marketplace aims to match consumer requests for service with suitable suppliers. Suppliers pay a fee for each successful match, whereas consumers get the matching service for free and pay directly to the suppliers. 93% of requests arrive through a web interface, while the rest arrive by phone. A typical request specifies the service requested, consumer location and other request parameters (e.g., gender preferences). Once an order is placed, most of the communication between the marketplace and the consumer takes place via a call center. The matching process typically entails a few calls: one to confirm request details, one to secure the consumer's agreement on the matched supplier and exchange contact details, and one within a month to have the consumer review the supplier. If the matched supplier did not satisfy the consumer, the consumer may decide to go through another the matching cycle; this happens in 20% of the cases.

Our sample includes weekly data on marketplace operations in one large city over the 2006-2014 period, corresponding to 260,000 service requests. The data for each request include (1) features of the consumer who placed the order: location, vertical, age, and gender; (2) request features: when placed, when matched, suppliers replied to request, assigned supplier (if any), matching fee and (3) request outcome: whether the match was successful (i.e., whether the actual service was performed). In addition, the dataset includes each supplier's attributes: location, skills, degree, experience, date of registration, wage rate and gender.

Our analysis focuses on the eight most popular market verticals. There are 155,000 consumers and 78,000 suppliers in the dataset. Summary patterns and trends for the main data variables are presented in Exhibits 2-3. Summary statistics of the key variables are in Exhibit 1. From Exhibits 2 and 3 one can observe that the marketplace follows a fast growth trajectory. The rate of growth is 1,000% in the first year; and it declines over time and stabilizes around 30-40% a year.

In addition, we also obtained data on the marketplace's marketing expenses. Most of the marketing budget was spent on web marketing, 99% of which was on cost-per-click search ads. Direct acquisition of suppliers constituted about 1% of overall adoption. The location data is available with about 1mi accuracy, which allowed us to create a discrete grid of 40x40 =1600 locations in total, with an average area of 1 sq. mile each. There are nine verticals defined by the marketplace which we kept intact.

The vast majority (99%) of requests had location data; the remaining 1% were removed from the sample. Only a half of suppliers had geographical information. In computing the aggregate metrics, suppliers with missing location were allocated to each location in proportion to the measured supplier count for that particular location, so that the total number of suppliers (aggregated across the city) was kept constant. Only .1% of previously registered consumers placed requests in a given week, so we classified each request to be from a new consumer. Suppliers that performed services in multiple verticals were counted as multiple suppliers. Consumers churned if they did not place a request within past 6 months. A supplier churned if he did not reply to any service request within 12 months.

5 Empirical specification and Results

Below we describe how we estimated our model and the results of each estimation.

Estimated equations and methods

To make sure our results are robust we used two different estimation techniques. First, we use the standard approach of pooling the panel data. In order to deal with potential correlations among the equations, we employ conservative assumptions on the variance structure in the form of error clustering by location (Wooldridge (2002), sec. 13.8.2). As an alternative, we performed weekly cross-sectional estimations of our equations and then averaged the resulting coefficients to compute the final coefficient (see Section 7). Details of our methodology and the results of the pooled estimations follow.

5.1 Pricing policy

Based on equation (1) we have

$$f_{ixvt} = \theta A_0 + \vec{\omega} \cdot \left[S_{xvt},\ S_{-x,vt},\ D^{orders}_{xvt},\ D^{orders}_{-x,vt}\right] + \epsilon_{ivt}.$$

This equation is estimated by OLS. There are two goals in this estimation. First, to understand whether the system-state variables are relevant instruments for the pricing (i.e., matching fee) function. Second, to use the predicted model of pricing behavior as the model employed by users before placing the service requests or registering in the system.

The results are presented in Exhibit 4. The fee assigned by the marketplace to each request is highly dependent on the number of suppliers and incoming demand requests. Adding one supplier to the local neighborhood increases the fee by \$0.0058, and adding a remote supplier increases the fee by \$0.0014. Incoming consumer requests have a negative, but statistically insignificant influence on the fee. Thus, the supply state vector $[S_{xvt}, S_{-xvt}]$ is identified as a relevant instrument for the matching fee. Its exogeneity is not rejected by the test for over-identification in the supply side estimations below.

5.2 Conversion – Supply side

Supply sensitivity to matching fee

The supply sensitivity links the matching fee to the probability that a random supplier would reply to a random request i after being exposed to it. Using the linear approximation to expression (2) we derive the indicator representing supplier j's reply to request i upon its discovery:

$$\begin{cases} I\left(replied_{ij}\right) & = 1 - \frac{v_j(0)-A_0}{v_j^0} - \frac{f_{ixvt}}{v_j^0} + \epsilon_{ixvt} \text{ if } ij \text{ local} \\ I\left(replied_{ij}\right) & = exp\left(-\frac{\beta^S}{v_j^0}\right)\left(1 - \frac{v_j(0)-A_0}{v_j^0}\right) - exp\left(-\frac{\beta^S}{v_j^0}\right)\frac{f_{ixvt}}{v_j^0} + \epsilon_{ixvt}, \text{ if } ij \text{ remote} \end{cases} \tag{12}$$

when $replied_{ij} = 1$ if supplier j replied to request i upon its discovery and $replied_{ij} = 0$ otherwise. Due to (1), the component $\theta_M A_i$ of the fee f_{ixvt} is correlated with the request parameters and likely correlated with ϵ_{ixvt}, hence the fee is endogenous and should not be estimated using ordinary least squares. We thus instrument the fee using $z_{xvt} = [S_{xvt}, S_{-xvt}]$ and assuming $E\left[[S_{xvt}, S_{-xvt}], \epsilon_{ixvt}\right] = 0$, that is, a supplier's decision to reply to an observed request is not a function of the size of the marketplace. We performed F-test and Hansen test for over-identification for relevance and exogeneity of instruments. The results are presented in Exhibit 5. The unit of observation is a request-supplier pair. For each request, we constructed the set of suppliers who likely discovered it. In particular, if supplier j replied to at least one request in the marketplace within the lifetime of request i, he is considered to have been exposed to request i. The lifetime of request i is the time interval between request initiation and its expiration. In addition, if supplier j was notified by the marketplace about request i, that supplier was also assumed to have been exposed to the request.

As seen in Exhibit 5, a higher matching fee assigned to a request leads to a lower probability of reply by the supplier. Each dollar increase in the fee decreases the reply probability by 1% (in relative terms). Local requests are four times more likely to receive a reply than remotes ones. Test for over-identification does not reject the exogeneity assumption, with a p-value of the J statistic being 0.33 for local replies and 0.19 for remote replies.

Discovery rate

Equation (3) yields the following expression for the reply inter-arrival times:

$$\xi_{ixvt} = \left[\delta\left(Pr\left(\text{local reply}\right)S_{xvt} + Pr\left(\text{remote reply}\right)S_{-xvt}\right)\right]^{-1} + \epsilon_{ixvt}.$$

Thus, the pooled discovery rate δ can be computed as follows:

$$\hat{\delta}_{pooled} = \frac{1}{\sum_i \xi_{ixvt}} \sum_i \left[\left(\rho_{xvt} S_{xvt} + \rho_{-xvt} S_{-xvt}\right)\right]^{-1},$$

where ρ_{xvt} and ρ_{-xvt} are the sample means of local and remote reply probabilities. Standard errors for the pooled estimate are computed using 1000 bootstrap replications with requests (*not* inter-arrival times) being sampled.

The daily discovery rate point estimate is 0.031, with a standard error of 0.0004. This implies that a randomly chosen supplier discovers a randomly chosen request within a day with probability 3% (if the request has not been assigned to anyone else within that day).

5.3 Conversion – Demand side

Network effects geographic structure

Expression (4) leads to the following specification:

$$I\left(\text{assigned supplier is local}\right)=\frac{S_{xvt}}{S_{xvt}+S_{-xvt}exp\left(-\left(\frac{\beta^S}{v_j^0}+\frac{\beta^D}{v_i^0}\right)\right)}+\epsilon_{ixvt}.$$

We estimate this equation by the method of moments on the subsample of requests to which a supplier was assigned. The results of the estimation are presented in Exhibit 7. We observe that in terms of market liquidity a local supplier is "worth" 30 remote suppliers. Thus the network effect structure strongly depends on location. By Exhibit 5, value of remote requests are deflated by approximately a factor of 4 compared to local requests. Thus, value of a remote supplier by consumers are deflated by a factor of 8 compared to the value of a local supplier.

Matching quality threshold

Equation (6) provides us with the expected stopping time:

$$E\tau_{ixvt}=\frac{exp\left(\frac{v_j(0)-A_i+f_{ixvt}}{v_j^0}+\frac{u_0}{v_i^0}\right)}{\delta S_{xvt}^{eff}}$$

from which the probability of a local match resulting in an assignment, $exp\left(-\frac{u_0}{v_i^0}\right)$, is estimated as

$$\left[exp\widehat{\left(-\frac{u_0}{v_i^0}\right)}\right]=\left(\sum\tau_{ixvt}\right)^{-1}\sum_i\frac{1}{\hat{\delta}\rho_{xvt}\left(S_{xvt}+S_{-xvt}exp\left(-\left(\widehat{\frac{\beta^S}{v_j^0}}+\frac{\beta^D}{v_i^0}\right)\right)\right)}$$

where ρ_{ixvt} and ρ_{-xvt} are the sample means of local and remote reply probabilities. Standard errors are computed using bootstrap with 1000 replications. The estimates from the above steps were used in constructing this estimator.

The resulting point estimate of $exp\left(-\frac{u_0}{v_i^0}\right)$ is 0.27, with a standard error of 0.003. Thus, the probability that a local reply results in an assignment is 27%.

Request success and expiration

Equation (8) yields

$$\phi_{ixvt}=\pi_0 exp\left(-\tau_{ixvt}^*/\tau_e\right)$$

In practice, following an unsuccessfully matched request, the consumer may repeat the matching process again, which happens in 20% of cases. We focus only on the first matching attempt and therefore estimate

$$
\begin{aligned}
Pr\left(success \middle| \text{one attempt}\right) &= \frac{Pr\left(success,\ \text{one attempt}\right)}{Pr\left(\text{one attempt}\right)} \\
&= \frac{exp\left(-\frac{\tau}{\tau_e}\right)\pi_0}{Pr\left(\text{one attempt}\right)}
\end{aligned}
$$

Let the sample estimate for $Pr\left(\text{one attempt}\right)$ equal $\hat{\pi}_1$, then, employing a linear approximation for the exponent, we get

$$
I\left(success_{ixvt} \middle| \text{one attempt}\right) = \frac{1}{\hat{\pi}_1}\left(\pi_0 - \frac{\pi_0}{\tau_e}\tau^*_{ixvt}\right) + \epsilon_{ixvt}
$$

which is estimated by OLS on the subsample of requests matched from the first attempt. The equation is estimated on the subsample of requests for which matching time was less than 33 days(.99 percentile) (requests that took longer to match are outliers that may skew the results, since the mean time is 5 days.)

The maximum probability of being matched from first time π_0 equals 58%, while the average request expiration time is 44 days, which also means that each day increase the chance of the consumer rejecting the assignment by 2%.

5.4 Adoption

Our linear dynamic system (11) yields

$$
\begin{cases}
\frac{D^{req}_{xvt}}{\gamma_t} & = \frac{\phi_{xvt}-\phi_0}{\phi}\vec{V}^D \cdot \left[D_{xvt}, D_{-xvt}, 1\right] + \mu^D M_{vt} + \epsilon_{xvt} \\
\frac{S^{req}_{xvt}}{\gamma_t} & = \frac{\theta_{xvt}-\theta_0}{\phi}\vec{V}^S \cdot \left[S_{xvt}, S_{-xvt}, 1\right] + \mu^S M_{vt} + \epsilon_{xvt}
\end{cases}
$$

where $\vec{V}^D$ and $\vec{V}^S$ are the viral coefficients, D^{reg}_{xvt}/γ_t and S^{reg}_{xvt}/γ_t are the de-seasonalized series (see Appendix for methodology). To eliminate noise in the conversion probability, we grouped conversion probability variable into 6 equally split groups: (0.0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, 0.5-0.6) and computed the mean conversion in each, thus obtaining a transformed variable with only 6 levels. We estimate each equation by least squares.

The estimation results are presented in Exhibits 9 (demand side) and 10 (supply side). First, we observe the strong presence of viral effects on both the demand and supply sides. 100 active users bring, on average, one more user in a week. In addition to that, we also see the presence of liquidity effects. Increase in conversion from 0 to 1 leads to 100 active consumers bringing in 5 more people per week which appears – quite a large effect. Interestingly, remote consumers do not matter in a significant way. Neither do network effects per se, i.e. network effects without consumers that trigger other consumers to evaluate the marketplace and take into account value create by network effects. In other words, it is the *interaction* of viral and network effects that influences adoption on the consumer side.

On the supply side, we also see the presence of network effects. One successful request per week delivered to a supplier increases the viral coefficient by a factor of 10, relative to situation when there are 0 successful requests delivered per week. In addition, value from the marketplace per se has significant impact on supply adoption (e.g., for those suppliers who find the platform through search). One consumer request served in a local neighborhood per week will bring 1 more supplier in that neighborhood over 5 weeks. Similarly to consumer side, remote suppliers have 4 orders of magnitude less impact on adoption than local suppliers. We also observe that on the supply side remote suppliers may detract suppliers from joining. This seems to be very natural since each additional supplier reduces the profit that an average supplier gets. Also, it may be interpreted as a "negative network effect". Namely, the higher the supplier base with established reputation, the harder it is to compete with them and thus the smaller the incentives to join the platform.

The effect of marketing is meaningful only on the consumer side. In particular, the viral effect of one active consumer can be replicated by spending $3 per neighborhood on marketing.

6 Implications for marketplace pricing

As discussed above, the functioning of the marketplace and in particular – its critical mass – depend on its pricing policy, i.e., how much a supplier pays for a successful match. The marketplace pricing policy determines the probability that a supplier who discovered a service request will actually reply to it. If the matching fee is high, then only few suppliers will reply, which will in turn reduce the market liquidity. A lower fee stimulates supplier response and increases liquidity but results in a lower revenue for the marketplace. In what follows, we study this tradeoff using a few examples.

Conversion optimization

Figure 3 depicts the conversion probability of an average request as a function of the effective supply base under three alternative fee scenarios: (a) a zero-fee policy, whereby the marketplace does not charge anyone for service; (b) the current pricing policy of our marketplace (averaging $30 per match); and (c) a higher fee of $45 per match.

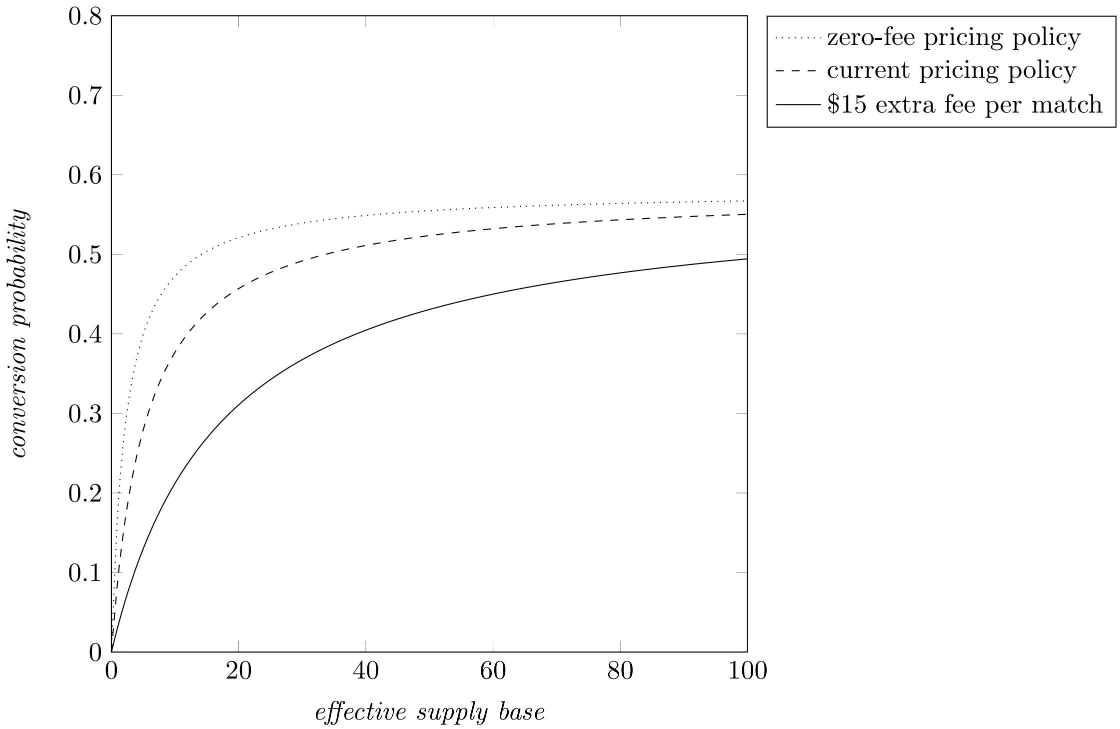


Figure 2: Average conversion probability as a function of the effective supply base under three alternative pricing policies: (a) a zero-fee policy; (b) current pricing policy ($30 per match on average), and (c) a price of $45 per match.

If the marketplace only eliminates the matching fee, and everything else stays the same, the average conversion probability will increase from the current 53% to 56%, and the average time to find a match would then be reduced by a factor of 1.8 to a little over a day. The conversion probability would be very close to its upper bound (58%) of the asymptotic conversion probability. If the marketplace adds $15 to the current fee, the conversion probability will be reduced by about 10% and the average matching time will increase to 5-10 days.

Critical mass For the current analysis, we define *critical mass* as the effective supply base required to bring the conversion probability to half of its asymptotic value, or 29%. Clearly, the critical mass depends on the parameters of the marketplace and is an increasing function of the matching fee. Under the current pricing policy, the critical mass in our marketplace is 5.3 effective suppliers. In other words, it takes a little more than five local suppliers to exceed the critical mass. If the marketplace prices each match at $45, the critical mass increases to 17, and if the matching fee is zero, critical mass declines to 2.3. Thus, lowering the marketplace matching fee increases the conversion rate by decreasing the critical mass of suppliers. However, at a given point in time, there is a tradeoff between conversion and revenue: higher fees may increase revenue while lowering conversions. In particular, conversion is maximized at the zero-fee policy which generates no revenue for the marketplace.

Short-term revenue optimization

We can find the revenue-maximizing matching fee for the marketplace by making a few simplifying assumptions. Assume $E\left[D_{xvt}^{req}\right] = \vec{v}^{D}\frac{\phi_{xvt}}{\phi}\left[D_{xvt}, S_{xvt}, D_{-xvt}, S_{-xvt}\right]$, that is, viral adoption is proportional to the probability of conversion, and assume that the state vector $\left[D_{xvt}, S_{xvt}, D_{-xvt}, S_{-xvt}\right]$ is stationary. Then, the revenue rate of the marketplace at a given location x is proportional to $f_{xvt}\phi_{xvt}^{2}\left(f_{xvt}\right)$.

By (9), the success probability of a request is given by

$$\pi_0\left(1+\left(exp\left(\frac{A_i - f_{ixvt} - v_j\left(0\right)}{v_j^0} - \frac{u_0}{v_i^0}\right)\delta\tau_e S_{xvt}^{eff}\right)^{-1}\right)^{-1}.$$

Thus, the problem of finding an optimal fee for a particular request can be formulated as maximizing

$$f\left(\frac{1}{1+kexp\left(\alpha f\right)}\right)^{2} \tag{13}$$

over f, where $k = exp\left(-\left(\frac{A_i - v_j(0)}{v_j^0} - \frac{u_0}{v_i^0}\right)\right)/\left(\delta\tau_e S_{xvt}^{eff}\right)$ and $\alpha = \frac{1}{v_j^0}$. We show in the Appendix that (13) is maximized by

It yields the following solution (see Appendix):

$$f^{*} = \frac{1}{\alpha}\left(\frac{1}{2} + W_0\left(\frac{1}{2k\sqrt{e}}\right)\right), \tag{14}$$

where $W_0\left(\cdot\right)$ is the principal branch of Lambert W-function. $W_0\left(\cdot\right)$ is a smooth monotone increasing function so that the optimal fee is also a smooth monotone increasing function of the supply base, discovery rate, average request expiration time and probability of a reply resulting in a match. With the linear approximation $Pr\left(reply_i\right) = \alpha_0 - \alpha\Delta f_i$, where $\alpha = \frac{1}{v_j^0}$ and $\alpha_0 = 1 - \frac{v_j(0) - A_0}{v_j^0}$, we show in the Appendix that $f^{*} \simeq \frac{\alpha_0}{\alpha}\left(1+\frac{3}{k_1}\right)\left(1-\sqrt{\frac{2}{\alpha_0 k_1}}\right)$, where $k_1 = 2\delta\tau_e S_{xvt}^{eff} exp\left(-\frac{u_0}{v_i^0}\right)$. Further, using estimated parameters, Figure 3 shows the revenue rate as a function of the average matching fee for different levels of the effective supply base, with all incoming requests having the parameters of the average request.

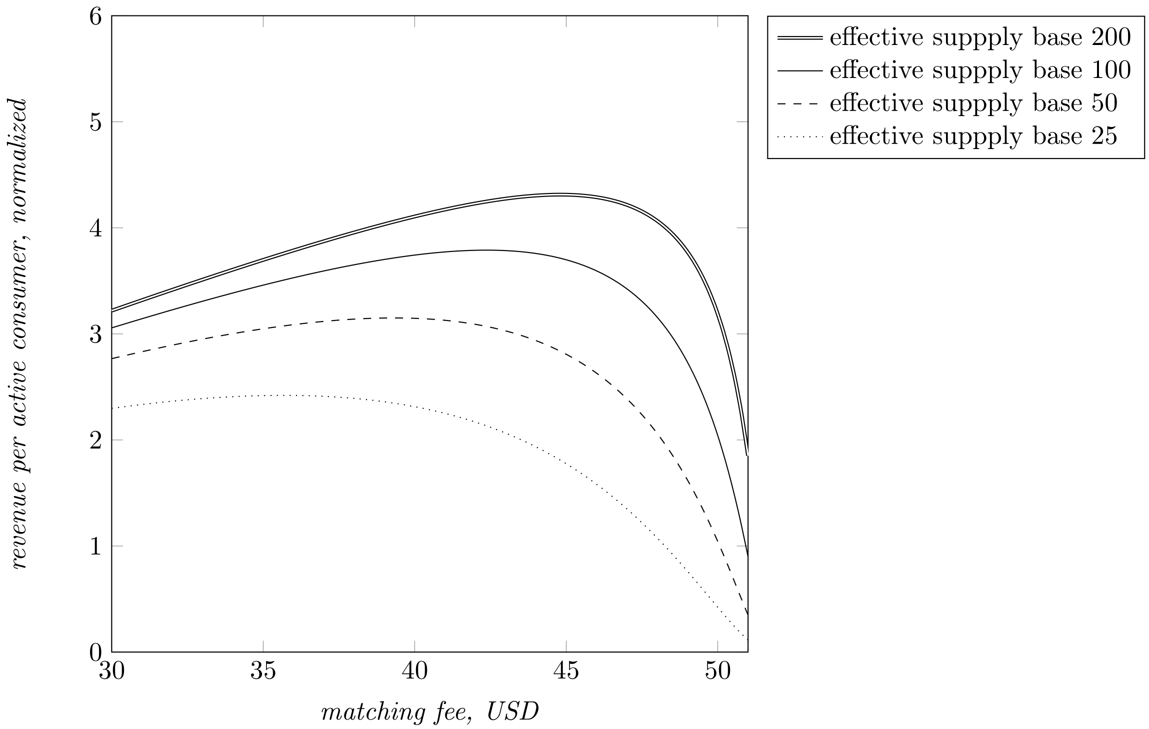


Figure 3: Revenue rate per active consumer as a function of the matching fee

Figure 3 shows that the ability of the marketplace to increase the matching fee depends on the effective supply base. The higher the supply base, the higher the fee that can be charged to optimize short-run revenue. Why didn't the marketplace charge higher matching fees? This is probably due to the difference between short-term and long-term optimization. During our sample period, the marketplace still had to grow, and probably preferred to attract suppliers and consumers at the expense of short-term profitability.

7 Robustness checks

7.1 Weekly estimations

We estimated our model using weekly cross-sectional estimations and averaging the resulting weekly coefficients. For instance, if an estimation for week t yields coefficient β_t, then the our estimate of the coefficient is given by

$$\hat{\beta} = \frac{1}{T} \sum \hat{\beta}_t$$

where $T \approx 400$ is the total number of weeks in which data are available in the dataset. The standard error of the coefficient is computed as the standard deviation of the coefficient in the sample divided by the squared root of total number of coefficients.

$$s.e.\left(\hat{\beta}\right) = \frac{s.d.\left(\left\{\hat{\beta}_t\right\}\right)}{\sqrt{T}}$$

The significance is then verified via the standard t-test:

$$t = \frac{\hat{\beta}}{s.e.\left(\hat{\beta}\right)} \underset{H_0}{\sim} N\left(0,1\right)$$

Since the marketplace is growing over time the pooled estimation utilizes more information from the later periods due to higher number of observations. Therefore, quantitatively, the results of the pooled and robust methods may be slightly different. While this method is less efficient than the pooled estimation, it is less sensitive to specific assumptions and therefore presents a robust alternative. The outcomes of the weekly estimations are presented in exhibits 11-17.

The results strongly reinforce the pooled estimation results. The coefficients of the variables of interest are strongly significant and are quantitatively similar to the ones estimated via pooling. For instance, the discovery rate estimated using weekly estimation is slightly higher than the pooled one (0.04 vs 0.03), the matching quality threshold is very similar (0.27 vs 0.29) and the asymptotic success probability is slightly lower for the weekly estimation (0.53 vs 0.58), which highlights the fact that the weekly estimation treats all time periods equally. The expected impatience parameter of consumers is again very close (0.022 vs 0.0023 decrease in success probability due to one day of matching delay).

At the same time, there are some differences between the weekly and pooled estimations. First, the pricing policy coefficients in the weekly estimations are not significant, but this is due to the small size and thus noisy estimates of the first two years of marketplace operations. For the years 2008-2014 and 2010-2014, the supply-side coefficients become significant and for the 2010-2014 sample, they are very close to the pooled estimates (0.0057 vs 0.0058 for local supply and 0.0013 vs 0.0014 for remote supply). Second, the weekly estimates of consumer adoption assign a higher weight to value-based adoption than the pooled estimation, and makes the non-value-based adoption coefficient insignificant. This suggest that it is the pure *interaction* between the liquidity and the viral effects that drives adoption. In other words, without network effects there would be no viral effect. Third, the effect of remote suppliers in the weekly estimations is insignificant, however the effect is numerically small in the pooled estimation, so this probably results from lower efficiency of the weekly estimations.

One of the largest quantitative differences between the weekly and pooled estimations is the 1.5 higher response probability of remote suppliers. This may be related to the decline in remote match value over time observed in Exhibit 14. This decline also results in a higher estimate for the value of a remote match in the weekly estimations (0.057 vs 0.031). As we noted above, while the pooled estimations put more weight on later periods, which have more observations, weekly estimations give each week the same weight.

Overall, the signs, significance and magnitudes of our weekly estimates are similar to those obtained in our pooled estimation, which reinforces our confidence in our model and results.

7.2 Vertical-by-vertical estimations

Exhibits 18 - 23 show the results of pooled estimations for each service vertical separately. With a small number of exceptions the results are consistent across as well as with the full sample results, which further reinforces our confidence in the model. The marketplace fee strongly depends on the remote supply base and more weakly on the local supply base. The reply sensitivity coefficients are of the same order of magnitude as the full-sample estimations and are strongly significant (only one out of 18 sensitivity coefficients is not highly significant, partially due to the lower sample size). Secondly, discovery rates, average request expiration times, and asymptotic success probabilities are close to each other and full-sample estimates. The probability of an assignment following a local reply varies across verticals suggesting that the liquidity effect varies by type of service.

The key difference of demand adoption equation coefficients from the estimation that pools all verticals together is as follows. Vertical-by-vertical estimation suggests that if the conversion probability is low enough, then the system may exhibit negative virality. However, a higher conversion strengthens the viral coefficient (7 out of 9 coefficients are significant at 0.001 level) which again speaks in favor of the *interaction* of viral and network effects hypothesis. The supply side adoption also support this hypothesis. for 6 out o 9 verticals there is a highly significant viral-network effects interaction. In addition, small value does not cause negative virality on the supply side. What causes it is the remote supply which tends to act as a signal for barrier to entry, which increases with the value that an average supplier can generate. We also observe that local users have a much higher (two orders of magnitude) effect on adoption than remote users.

Overall, estimation of the model for each vertical separately confirms the validity of the results obtained by pooling all verticals together and further reinforces our belief in the model and obtained results.

8 Concluding remarks

This paper developed a framework for analyzing online marketplaces and applied it to a specific service marketplace. Our approach was based on modeling the operational features of the matching process which are often treated as a "black box" characterized by two-sided network effects. Opening the "black box" allowed us to examine the relationship between the consumer behavior, supplier response, the marketplace pricing policy, and the performance of the marketplace. It also allowed us to delineate different streams of suppliers (local vs. remote) and to quantify their relative value to consumers. While local service provision is strongly preferred by both consumers and suppliers, there are substantially more remote than local suppliers available to satisfy each service request. As a result, in 20% of the cases service is provided by a local supplier, and in 80% of the cases it is provided by remote suppliers.

The "local" vs. "remote" dichotomy may take other forms in other marketplaces. For example, the "distance" metric may apply not only to geographic distance but also to other attributes that reflect the quality of a marketplace match. Our analysis clarifies the meaning of viral and network effects in the marketplace context. Viral effects bring more users to the marketplace and are interdependent with network effects, which reflect the increasing probability of conversion with more suppliers. Viral effects first help spread the word about the platform. Once prospective customers learn about the platform, they evaluate it in light of the network effects that influence their adoption decisions. Both of these effects strongly depend on geography. As a result, the supply base that really matters to a consumer requesting service is the effective supply base which discounts remote suppliers. The average success probability is a concave function of the effective supply base with an asymptotic value of 58%, which is outside the (direct) control of the marketplace. While adding a supplier is extremely useful in early years, when the marketplace is of limited size, the higher the effective supply base, the lower the effect of each additional supplier. For example, in the latter half of our sample period (2010-2014), all neighborhoods had a conversion probability above 30%, i.e., the marketplace passed its "critical mass" of supply in each neighborhood. It is not necessary for the effective supply base to be extremely high to achieve effective conversions. The critical mass is achieved with 6 effective suppliers per neighborhood and with an average effective supply base of 30, the conversion probability is within 20% of its upper bound. Finally, the network effects along with the associated critical mass depend not only on the installed supply base, but also the pricing policy of the marketplace, and we studied the interaction among these factors.

Exhibits

Exhibit 1: Summary statistics for key variables

(a) Aggregate variables

Variable	Mean	Std. Dev.	Min.	Max.	N
D_{xvt}, active consumers in x	0.937	6.055	0	363	6263406
D_{-xvt}, active consumers outside of x	1592.881	2189.653	0	15717	6263406
S_{xvt}, active suppliers in x	1.12	7.764	0	741	6263406
S_{-xvt}, active suppliers outside of	1874.105	2046.631	0	10256	6263406
D^{reg}_{xvt}, consumers' requests in x	0.039	0.337	0	39	6263406
D^{reg}_{xvt}, consumers' requests outside of x	66.993	104.748	0	1261	6263406
S^{reg}_{xvt}, suppliers' registrations in x	0.015	0.204	0	36	6263406
S^{reg}_{xvt}, suppliers' registrations outside of x	23.904	24.521	0	178	6263406
M_{vt}, web marketing, USD	10.011	78.066	0	1674.62	6263406

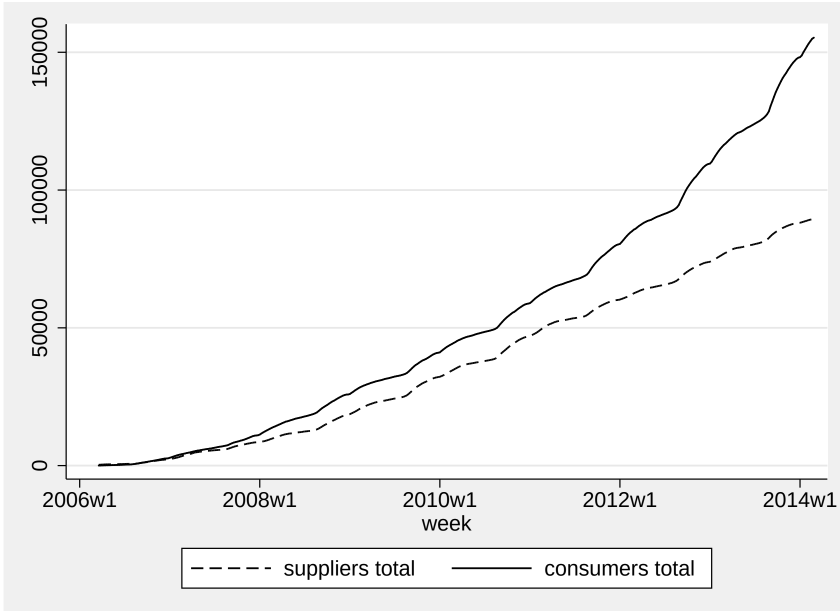
Sample period: March 2006 to February 2014.
Unit of observation: location-vertical-week(xvt) and vertical-week(vt).

(b) Request-related variables

Variable	Mean	Std. Dev.	Min.	Max.	N
fee charged to supplier, USD	26.967	15.18	0	1114.522	239371
request matched to supplier	0.953	0.212	0	1	239371
request assigned to local supplier	0.198	0.398	0	1	155188
time until assignment, days	5.994	17.719	0	714.883	228063
request is succcessful	0.636	0.481	0	1	239371

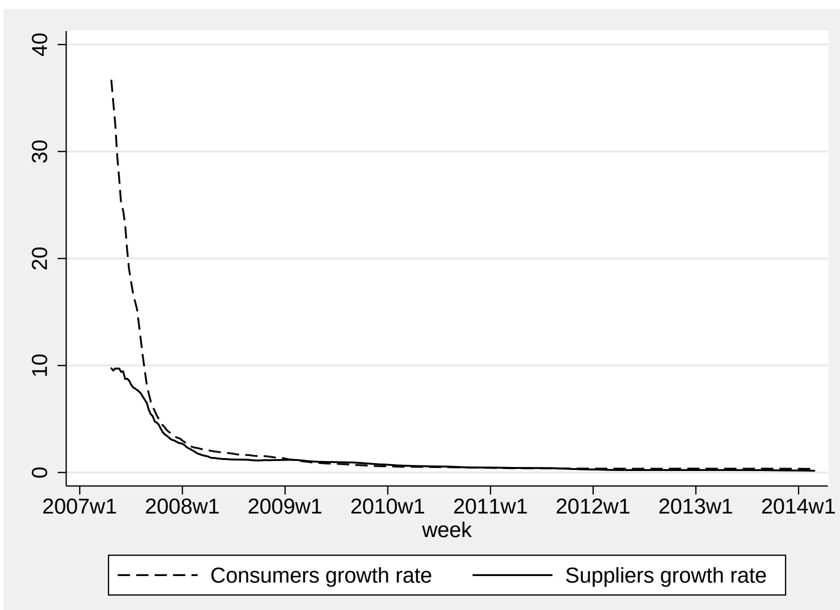
Sample period: March 2006 to February 2014.
Unit of observation: request.

Exhibit 2: Platform growth over time



"Consumers total" denotes the total number of consumers who placed requests on the platform. "Suppliers total" defines the total number of registered suppliers at a particular point in time.

Exhibit 3: Platform growth rate over time



The yearly growth rate of variable x_t, where t denotes time in weeks, is defined to be $\frac{x_t - x_{t-52}}{x_t}$. The numbers on the y-axis are absolute, not percentages. That is, number 10 means 1,000% growth. The mean consumers and suppliers growth rates in 2014 are 34% and 18%.

Exhibit 4: Pricing policy. Pooled estimation

$$f_{ixvt} = \vec{\omega} \cdot \left[S_{xvt},\ S_{-x,vt},\ D^{orders}_{xvt},\ D^{orders}_{-x,vt}\right] + \epsilon_{ixvt}$$

	(1)
	f_{ivt}, USD
S_{xvt}	0.0058***
	(3.63)
S_{-xvt}	0.0014***
	(14.12)
D^{orders}_{xvt}	-0.0858
	(-1.90)
D^{orders}_{-xvt}	-0.0007
	(-1.75)
R^2	0.07
Observations	239371

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: least squares, standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. For variable definitions, see the Glossary.

Exhibit 5: Suppliers' reply sensitivity to the marketplace matching fee. Pooled estimation.

(a) Local replies

$$I\left(replied_{ij}\right)=\left(1-\frac{v_j(0)-A_0}{v_j^0}\right)-\frac{f_{ixvt}}{v_j^0}+\epsilon_{ijxvt}$$

	(1)
	$I\left(replied_{ij}\right)$
f_{ixvt}, USD	-0.0172***
	(-15.98)
$const$	0.775***
	(24.44)
ρ^2	1.44e-06
Observations	246273

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) Remote replies

$$I\left(replied_{ij}\right)=exp\left(-\frac{\beta^S}{v_j^0}\right)\left(1-\frac{v_j(0)-A_0}{v_j^0}\right)-exp\left(-\frac{\beta^S}{v_j^0}\right)\frac{f_{ixvt}}{v_j^0}+\epsilon_{ijxvt}$$

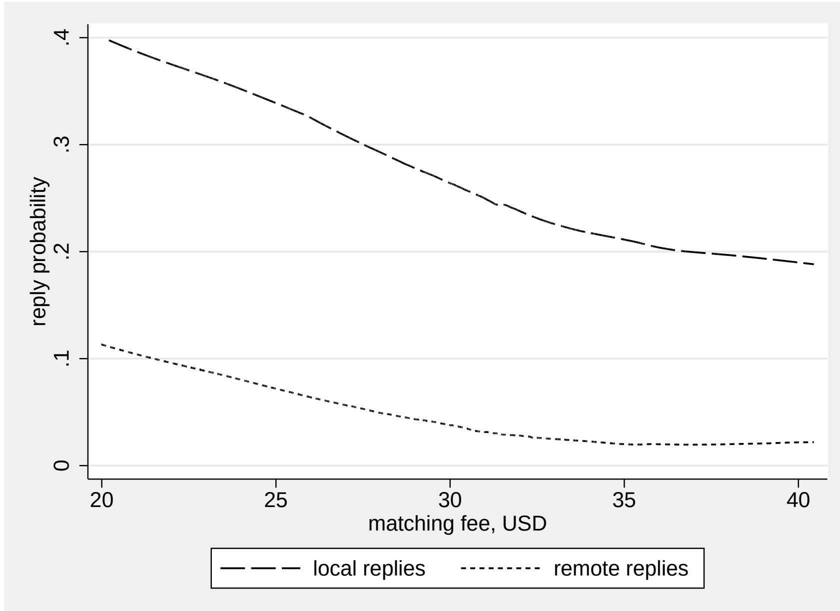
	(1)
	$I\left(replied_{ij}\right)$
f_{ixvt}, USD	-0.00539***
	(-18.85)
$const$	0.203***
	(25.73)
ρ^2	1.61e-04
Observations	25852052

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: two-stage least squares with $z_{xvt} = [S_{xvt}, S_{-xvt}]$ as instruments for marketplace matching fee f_{ixvt}. Standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. ρ^2 is the squared correlation between the predicted values of the "dependent" variable and the actual values. This value reflects goodness of fit. For variable definitions, see the Glossary.

Exhibit 6: Non-parametric smooth fitting for suppliers' reply probability



x-axis contains predicted matching fees from the first stage of IV regression. That is, $\hat{f} = \hat{\beta}_0 + \hat{\beta}_x S_x + \hat{\beta}_{-x} S_{-x}$

Exhibit 7: Network effects geographic structure. Pooled estimation

$$I\left(\text{assigned supplier is local}\right) = \frac{S_{xvt}}{S_{xvt} + S_{-xvt}exp\left(-\left(\frac{\beta^S}{v_j^0} + \frac{\beta^D}{v_i^0}\right)\right)} + \epsilon_{ixvt}$$

	(1)
	f_{ivt}, USD
$exp\left(-\left(\frac{\beta^S}{v_j^0} + \frac{\beta^D}{v_i^0}\right)\right)$	0.0306^{***}
	(8.77)
R^2	0.04
Observations	155177

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: method of moments, standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. For variable definitions, see the Glossary.

Exhibit 8: Request success and expiration. Pooled estimation

$$I\left(success_{ixvt} \middle| \text{one attempt}\right) = \frac{1}{\hat{\pi}_1}\left(\pi_0 - \frac{\pi_0}{\tau_e}\tau^*_{ixvt}\right) + \epsilon_{ixvt}$$

	(1)
	f_{ivt}, USD
π_0	0.5836***
	(333.84)
$1/\tau_e$	0.0213***
	(-48.26)
R^2	0.02
Observations	173003

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: least squares, standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. The equation is estimated on the subsample of request for which matching time was less than 33 days(.99 percentile). $\hat{\pi}_1 = 0.78$, obtained from the data sample, is used to compute estimates of π_0 and $1/\tau_e$. For variable definitions, see the Glossary.

Exhibit 9: Demand adoption. Pooled estimation

$$D^{req}_{xvt} = \frac{\phi_{xvt}-\phi_0}{\phi}\vec{V}^D \cdot [D_{xvt}, D_{-xvt}, 1] + \mu^D M_{vt} + \epsilon_{xvt}$$

	(1) D^{req}_{xvt}
D_{xvt}	1.6e-02**
	(3.02)
$D_{xvt}\phi_{xvt}$	4.8e-02***
	(4.43)
D_{-xvt}	-3.7e-06
	(-1.17)
$D_{-xvt}\phi_{xvt}$	2.9e-06
	(0.50)
ϕ_{xvt}	2.6e-03
	(1.21)
M_{vt}	1.6e-04***
	(9.67)
Constant	3.6e-04
	(1.42)
R^2	0.42
Obs	6154724

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: least squares, standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. The first column contains variables for which the corresponding coefficients' estimates are provided. For variable definitions, see the Glossary.

Exhibit 10: Supply adoption. Pooled estimation

$$S^{reg}_{xvt} = \frac{\theta_{xvt} - \theta_0}{\phi} \vec{V}^S \cdot [S_{xvt}, S_{-xvt}, 1] + \mu^S M_{vt} + \epsilon_{xvt}$$

	(1)
	S^{reg}_{xvt}
S_{xvt}	7.8e-03***
	(9.08)
$S_{xvt}\theta_{xvt}$	1.0e-01***
	(3.64)
S_{-xvt}	-8.9e-07**
	(-2.72)
$S_{xvt}\theta_{-xvt}$	-3.4e-05**
	(-3.09)
θ_{xvt}	2.3e-01***
	(4.28)
M_{vt}	-1.9e-05***
	(-6.01)
Constant	-8.5e-04***
	(-5.05)
R^2	0.18
Obs	6159676

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: least squares, standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. The first column contains variables for which the corresponding coefficients' estimates are provided. For variable definitions, see the Glossary.

Exhibit 11: Pricing policy. Weekly estimation

$f_{ixvt} = \vec{\omega} \cdot \left[S_{xvt},\ S_{-x,vt},\ D^{orders}_{xvt},\ D^{orders}_{-x,vt}\right] + \epsilon_{ixvt}$

	2006-2014	2008-2014	2010-2014
S_{xvt}	0.0338	0.0016	0.0057***
	(0.90)	(1.27)	(6.10)
S_{-xvt}	0.0325	0.0013***	0.0013***
	(1.33)	(7.37)	(10.70)
D^{req}_{xvt}	-3.76	0.0821	-0.0500
	(-1.16)	(0.05)	(-1.02)
D^{req}_{xvt}	-0.0385	-0.0053	-0.0089
	(-0.66)	(-1.12)	(-1.63)
constant	25.24***	21.96***	23.80***
	(7.93)	(34.20)	(110.06)
N	414	321	217

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficients estimates for each week were obtained by using least squares estimation for that week. Means of the coefficients and standard errors of the means are reported. For variable definitions, see the Glossary.

Exhibit 12: Suppliers' reply sensitivity to the marketplace matching fee. Weekly estimation.

(a) Local replies

$$I\left(replied_{ij}\right)=\left(1-\frac{v_j(0)-A_0}{v_j^0}\right)-\frac{f_{ixvt}}{v_j^0}+\epsilon_{ijxvt}$$

	(1)
	$I\left(replied_{ij}\right)$
f_{ixvt}, USD	-0.0122***
	(-3.31)
$const$	0.7024***
	(9.31)
Observations	388

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) Remote replies

$$I\left(replied_{ij}\right)=exp\left(-\frac{\beta^S}{v_j^0}\right)\left(1-\frac{v_j(0)-A_0}{v_j^0}\right)-exp\left(-\frac{\beta^S}{v_j^0}\right)\frac{f_{ixvt}}{v_j^0}+\epsilon_{ijxvt}$$

	(1)
	$I\left(replied_{ij}\right)$
f_{ixvt}, USD	-0.0084***
	(-4.21)
$const$	0.3447***
	(8.21)
Observations	414

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficients wre estimated for each week using two-stage least squares estimation with $z_{xvt} = [S_{xvt}, S_{-xvt}]$ as instruments for marketplace matching fee. Means of the coefficients and standard errors of the means are reported. For variable definitions, see the Glossary.

Exhibit 13: Network effects geographic structure. Weekly estimation

$$I\,(\text{assigned supplier is local}) = \frac{S_{xvt}}{S_{xvt}+S_{-xvt}exp\left(-\left(\frac{\beta^S}{v_j^0}+\frac{\beta^D}{v_i^0}\right)\right)} + \epsilon_{ixvt}$$

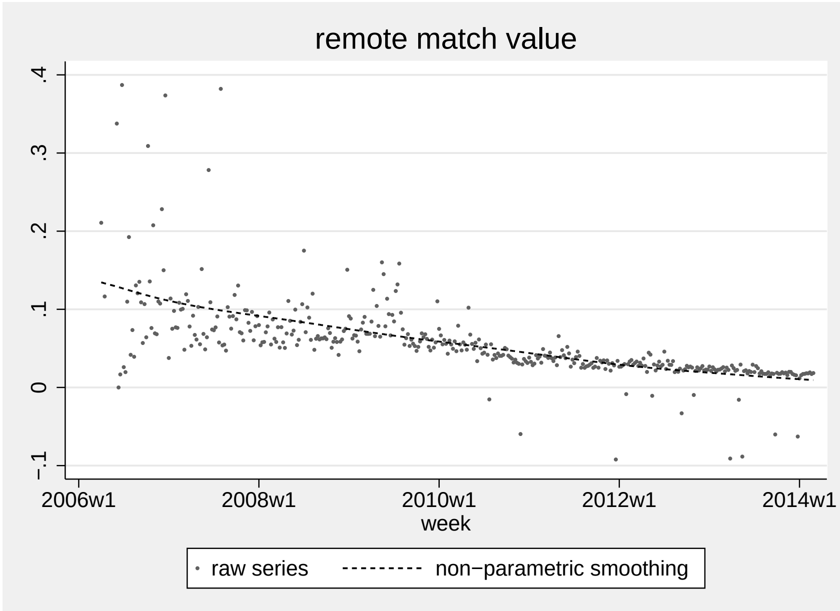
	(1)
	f_{ivt}, USD
$exp\left(-\left(\frac{\beta^S}{v_j^0}+\frac{\beta^D}{v_i^0}\right)\right)$	0.0569***
	(21.67)
Observations	404

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficient were estimated for each week using method of moments for that given week. Means of the coefficients and standard errors of the means are reported. For variable definitions, see the Glossary.

Exhibit 14



Remote match value (relative to local match value) series, obtained via weekly estimations.

Exhibit 15: Request success and expiration. Weekly estimation

$$I\left(success_{ixvt}\middle| \text{one attempt}\right) = \frac{1}{\hat{\pi}_1}\left(\pi_0 - \frac{\pi_0}{\tau_e}\tau^*_{ixvt}\right) + \epsilon_{ixvt}$$

	(1)
	f_{ivt}, USD
π_0	0.56***
	(333.84)
$1/\tau_e$	0.0205***
	(184.66)
Observations	414

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficients were estimated for each week using least squares estimation for that week. The equations were estimated on the subsample of requests for which the matching time was less than 33 days(.99 percentile). Means of the coefficients and standard errors of the means are reported. For variable definitions, see the Glossary.

Exhibit 16: Demand adoption. Weekly estimation

$$D^{req}_{xvt} = \frac{\phi_{xvt} - \phi_0}{\phi} \vec{V}^D \cdot [D_{xvt}, D_{-xvt}, 1] + \epsilon_{xvt}$$

	(1) D^{req}_{xvt}
D_{xvt}	1.2e-03
	(0.27)
$D_{xvt}\phi_{xvt}$	8.1e-02***
	(8.16)
D_{-xvt}	-1.8e-04
	(-1.13)
$D_{-xvt}\phi_{xvt}$	3.3e-04
	(1.13)
ϕ_{xvt}	-2.8e-01
	(-1.09)
Constant	1.5e-01
	(1.11)
Obs	413

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficients were estimated for each week using least squares estimation for that week. Means of the coefficients and standard errors of the means are reported. For variable definitions, see the Glossary.

Exhibit 17: Supply adoption. Weekly estimation

$$S_{xvt}^{reg} = \frac{\theta_{xvt}-\theta_0}{\phi}\vec{V}^S \cdot [S_{xvt}, S_{-xvt}, 1] + \epsilon_{xvt}$$

	(1)
	S_{xvt}^{reg}
S_{xvt}	1.4e-02***
	(17.50)
$S_{xvt}\theta_{xvt}$	9.3e-02***
	(3.08)
S_{-xvt}	-3.8e-05
	(-1.9)
$S_{xvt}\theta_{-xvt}$	3.3e-03
	(1.51)
θ_{xvt}	2.4e-01*
	(2.11)
Constant	2.9e-03***
	(3.92)
R^2	0.18
Obs	413

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficients were estimated for each week are obtained using least squares estimation for that week. Means of the coefficients and standard errors of the means are reported. For variable definitions, see the Glossary.

Exhibit 18: Pricing policy. Pooled estimation vertical-by-vertical

$$f_{ixvt} = \vec{\omega} \cdot \left[S_{xvt},\ S_{-x,vt},\ D^{orders}_{xvt},\ D^{orders}_{-x,vt}\right] + \epsilon_{ixvt}$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
S_{xvt}	0.00512***	0.00292	0.0205***	0.0147***	0.0143***	0.00645	0.0265	0.0202	0.0181
	(4.18)	(1.20)	(7.82)	(3.32)	(5.04)	(1.22)	(1.72)	(1.11)	(0.61)
S_{-xvt}	0.00160***	0.00453***	0.00473***	0.00263***	0.00325***	0.00165***	0.00662***	0.00637***	0.00577***
	(34.21)	(17.36)	(17.22)	(11.49)	(32.01)	(24.09)	(19.12)	(14.88)	(14.07)
D^{orders}_{xvt}	-0.0479	-0.0353	-0.249	-0.307	-0.229**	-0.0401	0.245	0.173	-0.201
	(-1.17)	(-0.21)	(-1.85)	(-1.93)	(-2.70)	(-0.75)	(1.68)	(0.79)	(-1.37)
D^{orders}_{-xvt}	0.00671***	0.0536***	0.0497***	0.0193**	0.0157***	-0.00125*	-0.00663	-0.00268	0.0293***
	(11.35)	(7.73)	(9.61)	(2.97)	(8.81)	(-2.30)	(-1.45)	(-0.41)	(7.69)
Const	15.53***	16.43***	15.90***	15.65***	16.15***	19.38***	16.78***	16.72***	14.28***
	(92.10)	(46.99)	(50.79)	(29.46)	(58.62)	(76.69)	(54.17)	(29.40)	(55.57)
R^2	0.112	0.114	0.110	0.0548	0.123	0.0694	0.0988	0.0566	0.0897
Obs	67316	11748	11860	8164	24379	81320	11320	11481	11783

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: least squares, standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. For variable definitions, see the Glossary.

Exhibit 19: Suppliers' reply sensitivity w.r.t. marketplace matching fee. Pooled estimation vertical-by-vertical.

(a) Local replies

$$I\left(replied_{ij}\right) = \left(1 - \frac{v_j(0) - A_0}{v_j^0}\right) - \frac{f_{ixvt}}{v_j^0} + \epsilon_{ijxvt}$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
f_{ixvt}	-8.7e-03***	-1.5e-02***	-1.1e-02*	-1.8e-02***	-1.2e-02***	-9.3e-03***	-2.0e-02***	-2.1e-02***	-7.5e-03
	(-9.95)	(-5.13)	(-1.98)	(-5.85)	(-8.29)	(-10.97)	(-6.61)	(-6.44)	(-1.02)
Constant	5.3e-01***	7.9e-01***	6.9e-01***	7.8e-01***	6.9e-01***	5.1e-01***	9.0e-01***	9.1e-01***	5.2e-01***
	(17.92)	(8.89)	(4.86)	(11.16)	(15.79)	(19.01)	(11.04)	(10.83)	(3.69)
R^2	-1.0e-01	-2.8e-01	-1.1e-01	-3.0e-01	-1.5e-01	-1.3e-01	-2.6e-01	-3.1e-01	-1.4e-02
Observations	7.2e+04	6.8e+03	6.3e+03	5.2e+03	1.8e+04	1.2e+05	6.6e+03	5.6e+03	4.5e+03

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) Remote replies

$$I\left(replied_{ij}\right) = exp\left(-\frac{\beta^S}{v_j^0}\right)\left(1 - \frac{v_j(0) - A_0}{v_j^0}\right) - exp\left(-\frac{\beta^S}{v_j^0}\right)\frac{f_{ixvt}}{v_j^0} + \epsilon_{ijxvt}$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
f_{ixvt}	-4.1e-03***	-8.0e-03***	-1.5e-02***	-8.5e-03***	-3.9e-03***	-1.6e-03***	-9.4e-03***	-2.3e-03**	-5.0e-03***
	(-16.60)	(-6.82)	(-4.40)	(-5.94)	(-5.69)	(-8.40)	(-9.46)	(-2.80)	(-8.15)
Constant	1.6e-01***	3.3e-01***	5.5e-01***	2.9e-01***	1.8e-01***	7.9e-02***	3.5e-01***	2.0e-01***	1.7e-01***
	(23.83)	(9.94)	(6.10)	(8.36)	(9.99)	(14.00)	(13.91)	(9.42)	(13.43)
Observations	8.1e+06	4.1e+05	3.4e+05	4.2e+05	1.3e+06	1.4e+07	3.5e+05	3.4e+05	3.3e+05

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: two-stage least squares with $z_{xvt} = \left[S_{xvt}, S_{-xvt}\right]$ as instruments for marketplace matching fee f_{ixvt}. Standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. ρ^2 is the squared correlation between the predicted values of the "dependent" variable and the actual values. This value reflects goodness of fit. For variable definitions, see the Glossary.

Exhibit 20: Network effects geographic structure. Pooled estimation vertical-by-vertical

$$I\,(\text{assigned supplier is local}) = \frac{S_{xvt}}{S_{xvt}+S_{-xvt}exp\left(-\left(\frac{\beta^S}{v_j^0}+\frac{\beta^D}{v_i^0}\right)\right)}+\epsilon_{ixvt}$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$exp\left(-\left(\frac{\beta^S}{v_j^0}+\frac{\beta^D}{v_i^0}\right)\right)$	0.0321***	0.0391***	0.0433***	0.0304***	0.0289***	-0.00827***	0.0336***	0.0475***	0.0342***
	(8.09)	(6.74)	(6.34)	(5.88)	(7.15)	(-43827.95)	(10.22)	(7.33)	(6.89)
Observations	44188	7737	8082	4122	15876	54329	7560	7546	5737

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: method of moments, standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. For variable definitions, see the Glossary.

Exhibit 21: Request success and expiration. Pooled estimation vertical-by-vertical

$$I\left(success_{ixvt}\middle|\text{one attempt}\right) = \frac{1}{\hat{\pi}_1}\left(\pi_0 - \frac{\pi_0}{\tau_e}\tau^*_{ixvt}\right) + \epsilon_{ixvt}$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
τ^*_{ixvt}	-0.0166***	-0.0149***	-0.0154***	-0.0160***	-0.0178***	-0.0144***	-0.0185***	-0.0139***	-0.0163***
	(-20.31)	(-8.98)	(-7.60)	(-9.19)	(-15.67)	(-27.89)	(-12.83)	(-8.26)	(-15.56)
Const	0.782***	0.787***	0.796***	0.707***	0.760***	0.683***	0.721***	0.711***	0.701***
	(321.23)	(148.42)	(150.27)	(99.85)	(170.19)	(244.75)	(127.55)	(121.90)	(103.05)
R^2	0.0134	0.0139	0.0101	0.0180	0.0185	0.0142	0.0215	0.0119	0.0280
Observations	51229	9075	9275	5590	17726	55321	8173	8119	8495

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: least squares, standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. The equation was estimated on the subsample of request for which the matching time was less than 33 days(.99 percentile). For variable definitions, see the Glossary.

Exhibit 22: Demand adoption. Pooled estimation vertical-by-vertical

$$D_{xvt}^{req} = \frac{\phi_{xvt} - \phi_0}{\phi} \vec{v}^D \cdot [D_{xvt}, D_{-xvt}, 1] + \mu^D M_{vt} + \epsilon_{xvt}$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
D_{xvt}	3.0e-02*	-4.5e-02***	-4.9e-02***	-1.5e-01***	6.2e-03	3.7e-02***	-3.4e-02***	-1.6e-02*	-3.5e-02
	(2.15)	(-4.37)	(-4.99)	(-3.32)	(0.87)	(5.68)	(-3.70)	(-2.07)	(-1.89)
$D_{xvt}\phi_{xvt}$	1.8e-02	1.3e-01***	1.4e-01***	4.1e-01***	6.7e-02***	2.7e-03	1.4e-01***	1.0e-01***	1.5e-01***
	(0.70)	(7.73)	(8.11)	(4.17)	(4.54)	(0.20)	(7.50)	(6.14)	(3.82)
D_{-xvt}	-1.3e-04***	-4.7e-05*	-7.2e-05***	-3.7e-05	-1.9e-05**	-2.0e-05	-3.8e-05***	-3.1e-05***	-1.2e-05
	(-7.52)	(-2.29)	(-5.37)	(-1.92)	(-2.62)	(-0.55)	(-4.97)	(-3.44)	(-0.67)
$D_{-xvt}\phi_{xvt}$	2.1e-04***	6.6e-05*	1.0e-04***	6.4e-05	2.3e-05	2.2e-05	5.6e-05***	4.4e-05**	7.9e-06
	(7.31)	(2.07)	(4.99)	(1.57)	(1.87)	(0.39)	(4.17)	(2.85)	(0.25)
ϕ_{xvt}	1.3e-02	-7.0e-03***	7.7e-03***	-5.6e-03*	8.1e-03**	-5.0e-01***	5.4e-03***	8.3e-03***	4.5e-03
	(1.34)	(-4.01)	(4.86)	(-2.40)	(2.70)	(-5.80)	(3.66)	(4.08)	(0.97)
M_{vt}	1.5e-04***	1.8e-04***	3.1e-04***	1.1e-03***	2.1e-04***	1.0e-04***	1.9e-04***	1.0e-04**	1.1e-03***
	(8.76)	(5.46)	(5.88)	(4.71)	(6.42)	(6.30)	(6.24)	(2.64)	(3.74)
Constant	4.0e-02***	5.8e-03***	1.7e-03***	3.7e-03**	2.0e-03**	3.8e-01***	2.8e-03***	1.1e-03***	6.3e-04**
	(7.87)	(4.14)	(5.56)	(2.77)	(2.87)	(5.67)	(6.48)	(3.58)	(2.94)
R^2	4.3e-01	1.9e-01	1.9e-01	2.2e-01	2.9e-01	5.3e-01	2.0e-01	2.0e-01	2.3e-01
Observations	6.9e+05	6.9e+05	6.9e+05	6.7e+05	6.9e+05	3.4e+05	6.7e+05	6.9e+05	6.7e+05

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: least squares, standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. For variable definitions, see the Glossary.

Exhibit 23: Supply adoption. Pooled estimation vertical-by-vertical

$$S^{reg}_{xvt} = \frac{\theta_{xvt}-\theta_0}{\phi}\vec{V}^S \cdot [S_{xvt}, S_{-xvt}, 1] + \mu^D M_{vt} + \epsilon_{xvt}$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
S_{xvt}	7.9e-03***	9.3e-03***	7.4e-03***	8.5e-03***	6.6e-03***	9.7e-03***	5.2e-03***	4.8e-03***	3.0e-03***
	(10.03)	(8.61)	(6.56)	(8.35)	(7.81)	(11.09)	(4.99)	(3.86)	(3.56)
$S_{xvt}\theta_{xvt}$	1.1e-01**	-2.8e-02	4.7e-02***	1.9e-01*	3.2e-01***	1.0e-02	9.8e-02**	3.2e-01***	1.8e-02
	(3.01)	(-0.40)	(4.72)	(2.06)	(11.22)	(0.56)	(3.01)	(4.59)	(1.08)
S_{-xvt}	-1.6e-07	-1.8e-06***	-2.4e-06***	-6.2e-07	1.3e-06	-5.2e-06***	-5.6e-07*	6.4e-07	-1.3e-08
	(-0.56)	(-4.10)	(-4.46)	(-1.30)	(1.38)	(-6.07)	(-2.37)	(1.77)	(-0.05)
$S_{-xvt}\theta_{xvt}$	-2.0e-04***	-3.3e-04**	-5.6e-05**	-9.9e-04***	-1.0e-03**	-8.6e-04***	-2.6e-04***	-1.2e-03***	-1.3e-04***
	(-5.37)	(-3.02)	(-3.01)	(-4.35)	(-2.68)	(-7.92)	(-3.57)	(-3.89)	(-3.43)
θ_{xvt}	8.4e-01***	2.3e-01***	3.6e-02**	1.5e+00***	1.3e+00**	6.2e+00***	2.0e-01**	5.4e-01*	1.0e-01***
	(4.10)	(4.27)	(2.91)	(5.12)	(2.86)	(8.18)	(3.11)	(2.36)	(4.16)
Const	-3.4e-03*	1.4e-03	1.2e-03***	1.1e-03*	-9.2e-04	3.0e-02***	4.0e-04	9.3e-05	-2.1e-04
	(-2.10)	(1.35)	(3.52)	(2.02)	(-0.53)	(6.04)	(1.07)	(0.12)	(-1.17)
R^2	2.2e-01	1.8e-01	1.6e-01	1.1e-01	1.3e-01	2.4e-01	6.1e-02	6.7e-02	2.5e-02
Observations	6.9e+05	6.9e+05	6.9e+05	6.7e+05	6.9e+05	4.2e+05	6.7e+05	6.9e+05	6.7e+05

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimation method: least squares, standard errors are clustered (Wooldridge (2002), sec. 13.8.2) by location. The first column contains variables for which the corresponding coefficients' estimates are provided. For variable definitions, see the Glossary.

Exhibit 24: Other parameters. Vertical-by-vertical

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
π_0	0.63^{***}	0.65^{***}	0.65^{***}	0.52^{***}	0.59^{***}	0.51^{***}	0.56^{***}	0.53^{***}	0.58^{***}
τ_e	47.2^{***}	52.7^{***}	51.6^{***}	44.2^{***}	42.7^{***}	47.5^{***}	38.9^{***}	51.3^{***}	42.91^{***}
$exp\left(-\frac{u_0}{v_i^0}\right)$	0.19^{***}	0.46^{***}	0.45^{***}	0.29^{***}	0.31^{***}	0.22^{***}	0.30^{***}	0.31^{***}	0.64^{***}
δ	0.036^{***}	0.032^{***}	0.032^{***}	0.026^{***}	0.029^{***}	0.028^{***}	0.034^{***}	0.040^{***}	0.029^{***}

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

$\hat{\pi_1}$, obtained from the data sample, is used to compute estimates of π_0 and $1/\tau_e$. Estimates of δ and $exp\left(-\frac{u_0}{v_i^0}\right)$ are computed w.r.t. estimators described in empirical specification section.

Glossary

The list of variables used in the estimations:

- D_{xvt}: active consumers in location x, vertical v, time period t
- D_{-xvt}: active consumers in vertical v, time period t, and locations other than x
- S_{xvt}: active suppliers in location x, vertical v, time period t
- S_{-xvt}: active suppliers in vertical v, time period t, and locations other than x
- D^{req}_{xvt}: number of consumers' requests in location x, vertical v, time t
- D^{req}_{-xvt}: number of consumers' requests in vertical v, time t, and locations other than x
- S^{reg}_{xvt}: number of suppliers' registrations in location x, vertical v, time t
- f_{ixvt}: fee per successful match for request i in location x, vertical v, time period t
- $I\left(replied_{ij}\right) \in \{0,1\}$: indicator of whether supplier j replied to request i
- ξ_{ixvt}: inter-arrival time of suppliers' replies to request i
- ρ_{xvt}, ρ_{-xvt}: share of replies to evaluations of local and remote requests by suppliers in location x, vertical v, time period t
- τ_{ixvt}: time from request initiation to supplier assignment
- $I\left(success_{ixvt}\right) \in \{0,1\}$: indicator of whether request i was successful
- γ_t: seasonality factor computed according to the methodology described in subsection 8
- D^{req}_{xvt}: number of consumers' requests in location x, vertical v, time t
- M_{vt}: web marketing expenditure per vertical v in time period t in USD

Appendix

Lemma. If $N \sim Pois\left(\lambda T\right)$ and $X_1 \ldots X_N$ are iid with cdf $F\left(x\right)$ then

$$Pr\left(max_1^N X_i \leq x\right) = exp\left(-\lambda T\left(1 - F(x)\right)\right)$$

□

Let

$$X_{max}\left(N\right) = max_1^N\left(X_i\right)$$

and its cdf

$$F_N\left(x\right) = Pr\left(X_{max}\left(N\right) \leq x\right) = Pr\left(X_1 \leq x, \ldots, X_N \leq x\right)$$

$F_N\left(x\right)$ is the probability that none of events in the Poisson stream N over time T has magnitude X_i higher than x. Events with such property arrive at rate $\lambda\left(1 - F\left(x\right)\right)$. The probability that none of these events happen over time T is

$$exp\left(-\lambda T\left(1 - F\left(x\right)\right)\right)$$

and thus

$$F_N\left(x\right) = exp\left(-\lambda T\left(1 - F\left(x\right)\right)\right)$$

■

Computing the cdf of a maximal value function $u_{ixvt}\left(t\right)$

□

Omitting argument t for clarity of exposition from random variables $u_{ixvt}\left(t\right)$, $u_{ixvt}^{x}\left(t\right)$ and $u_{ixvt}^{-x}\left(t\right)$ we get:

$$\begin{aligned}
& Pr\left(u_{ixvt} \geq u\right) \\
= \; & Pr\left(max\left(u_{ixvt}^{x}, u_{ixvt}^{-x}\right) \geq u\right) \\
= \; & 1 - Pr\left(max\left(u_{ixvt}^{x}, u_{ixvt}^{-x}\right) < u\right) \\
= \; & 1 - Pr\left(u_{ixvt}^{x} < u_0, u_{ixvt}^{-x} < u\right) \\
= \; & 1 - Pr\left(u_{ixvt}^{x} < u_0\right) Pr\left(u_{ixvt}^{-x} < u\right)
\end{aligned}$$

In order to proceed we need to obtain cdfs of u_{ixvt}^{x} and u_{ixvt}^{-x}.

Since, the reply count process is Poisson-distributed, the Lemma above helps us obtain

$$Pr\left(u_{ixvt}^{x}\left(t\right)\leq y\right) = exp\left(-S_{xvt}\delta texp\left(-\frac{f_{xvt}+v_{j}\left(0\right)}{v_{j}^{0}}\right)exp\left(-\frac{y}{v_{i}^{0}}\right)\right) \tag{15}$$

$$Pr\left(u_{ixvt}^{-x}\left(t\right)\leq y\right) = exp\left(-S_{-xvt}\delta texp\left(-\frac{f_{xvt}+v_{j}\left(0\right)+\beta^{S}}{v_{j}^{0}}\right)exp\left(-\frac{y+\beta^{D}}{v_{i}^{0}}\right)\right) \tag{16}$$

Thus

$$Pr\left(u_{ixvt}\left(t\right)\leq u\right)=exp\left(-\delta texp\left(-\frac{f_{xvt}+v_{j}\left(0\right)}{v_{j}^{0}}-\frac{u_{0}}{v_{i}^{0}}\right)\left[S_{xvt}+S_{-xvt}exp\left(-\frac{\beta^{S}}{v_{j}^{0}}-\frac{\beta^{D}}{v_{i}^{0}}\right)\right]\right)$$

Computing the request success probability

□

Let the request expiration time $t_{e}\sim Exp\left(1/\tau_{e}\right)$, arrival time of a suitable local reply $t_{x}\sim Exp\left(1/\tau_{x}\right)$ and arrival time of a suitable remote reply $t_{-x}\sim Exp\left(1/\tau_{-x}\right)$. Then the probability that the request is successful is given by

$$\phi_{x}=\pi_{0}Pr\left(min\left(t_{x},t_{-x}\right)<t_{e}\right)$$

Since $t_{x}$and t_{-x} are exponentially distributed, then $min\left(t_{x},t_{-x}\right)\sim Exp\left(1/\left(\frac{1}{\tau_{x}}+\frac{1}{\tau_{-x}}\right)\right)$ and thus

$$\begin{aligned}\phi_{x}&=\pi_{0}\frac{1/\tau_{x}+1/\tau_{-x}}{1/\tau_{x}+1/\tau_{-x}+1/\tau_{e}}\\&=\pi_{0}\frac{1}{1+\frac{1}{\tau_{e}}/\left(1/\tau_{x}+1/\tau_{-x}\right)}\\&=\left(1+\left(\mu\tau_{e}\right)^{-1}\right)^{-1},\end{aligned}$$

where $\mu=\frac{1}{\tau_{x}}+\frac{1}{\tau_{-x}}=\delta exp\left(-\frac{f_{x}}{v_{j}^{0}}-\frac{u_{0}}{v_{i}^{0}}\right)\left[S_{x}+S_{-x}exp\left(-\frac{\beta^{S}}{v_{j}^{0}}-\frac{\beta^{D}}{v_{i}^{0}}\right)\right]$ is the arrival rate of a suitable match. Indices i,v and t are omitted for clarity.

■

Computing the probability that the assigned match is local

□

Let the request expiration time $t_{e}\sim Exp\left(1/\tau_{e}\right)$, arrival time of a suitable local reply $t_{x}\sim Exp\left(1/\tau_{x}\right)$ and arrival time of a suitable remote reply $t_{-x}\sim Exp\left(1/\tau_{-x}\right)$. Then the probability that the match is assigned and it is local, is given by

$$\begin{aligned}\phi_x^x &= Pr\left(t_x < t_{-x}, min\left(t_x, t_{-x}\right) < t_e\right)\\ &= Pr\left(t_x = min\left(t_x, t_{-x}, t_e\right)\right)\\ &= Pr\left(t_x = min\left(t_x, min\left(t_{-x}, t_e\right)\right)\right)\end{aligned}$$

Since τ_{-x} and τ_e are exponentially distributed, then $min\left(\tau_{-x}, \tau_e\right)$ is also exponentially distributed, therefore

$$\begin{aligned}\phi_x^x &= \frac{1/\tau_x}{1/\tau_x + 1/\tau_{-x} + 1/\tau_A}\\ &= \frac{1/\tau_x + 1/\tau_{-x}}{1/\tau_x + 1/\tau_{-x} + 1/\tau_A}\frac{1/\tau_x}{1/\tau_x + 1/\tau_{-x}}\\ &= Pr\left(\text{match is assigned}\right)\frac{1/\tau_x}{1/\tau_x + 1/\tau_{-x}}\end{aligned}$$

Probability of the assigned match being local is given by

$$\begin{aligned}Pr\left(\text{match is local}\middle|\text{match is assigned}\right) &=\\ \frac{Pr\left(\text{match is local, match is assigned}\right)}{Pr\left(\text{match is assigned}\right)} &=\\ \frac{\phi_x^x}{Pr\left(\text{match is assigned}\right)} &=\\ \frac{1/\tau_x}{1/\tau_x + 1/\tau_{-x}} &=\\ \frac{S_x}{S_x + S_{-x}exp\left(-\left(\frac{\beta^S}{v_j^0} + \frac{\beta^D}{v_i^0}\right)\right)}\end{aligned}$$

Indices v and t are omitted for clarity.

■

Computing expected value of the request for supplier

□

Let $\theta_i = A_i - f_i - v_j\left(0\right) - \beta^S I\left(remote\right)$, then the expected surplus from a local request is given by

$$
\begin{aligned}
E\left[\left(\theta_i+\epsilon_j\left(ij\right)\right)I\left(\theta_i+\epsilon_j\left(ij\right)>0\right)\right] &= \\
E\left[\theta_i+\epsilon_j\left(ij\right)\middle|\theta_i+\epsilon_j\left(ij\right)>0\right]Pr\left(\theta_i+\epsilon_j\left(ij\right)>0\right) &= \\
\left(\theta_i+E\left[\epsilon_j\left(ij\right)\middle|\epsilon_j\left(ij\right)>-\theta_i\right]\right)exp\left(\frac{\theta_i}{v_j^0}\right) &= \\
\left(\theta_i-\theta_i+v_j^0\right)exp\left(\frac{\theta_i}{v_j^0}\right) &= \\
v_j^0 exp\left(\frac{\theta_i}{v_j^0}\right) &
\end{aligned}
$$

In addition, the value of a remote request relative to a local request is thus

$$
\frac{v_j^0 exp\left(\frac{A_i-f_i-v_j(0)-\beta^S}{v_j^0}\right)}{v_j^0 exp\left(\frac{A_i-f_i-v_j(0)}{v_j^0}\right)} = exp\left(-\frac{\beta^S}{v_j^0}\right)
$$

■

Revenue optimization

□

$$
f\left(\frac{1}{1+kexp\left(\alpha f\right)}\right)^2 \to \max_f
$$

FOC

$$
\left(\left(\frac{1}{1+kexp\left(\alpha f\right)}\right)^2\right) - \frac{2f}{1+kexp\left(\alpha f\right)}\frac{k\alpha exp\left(\alpha f\right)}{\left(1+kexp\left(\alpha f\right)\right)^2} = 0
$$

implies

$$
k\left(2\alpha f-1\right)e^{\alpha f} = 1
$$

Let $y = 2\alpha f - 1$, then

$$
\frac{y}{2}e^{\frac{y}{2}} = \frac{1}{2k\sqrt{e}}
$$

Apply the principal branch of Lambert function $W_0\left(\cdot\right)$ to each side of the equation to yield

$$
W_0\left(\frac{y}{2}e^{\frac{y}{2}}\right) \equiv \frac{y}{2} = W_0\left(\frac{1}{2k\sqrt{e}}\right)
$$

Thus

$$
f = \frac{y+1}{2\alpha} = \frac{1}{\alpha}\left(\frac{1}{2}+W_0\left(\frac{1}{2k\sqrt{e}}\right)\right)
$$

■

Revenue optimization. Linear approximation

□

Given the linear approximation, the objective is

$$f\left(\frac{1}{1+\frac{1}{k_1(\alpha_0-\alpha f)}}\right)^2 \to \max_f$$

FOC corresponding to the objective

$$\left(\frac{1}{1+k_1\left(\alpha_0-\alpha f\right)}\right)^2\left(1-\frac{2\alpha k_1 f}{1+\frac{1}{k_1(\alpha_0-\alpha f)}}\frac{1}{\left(k_1\left(\alpha_0-\alpha f\right)\right)^2}\right)=0$$

implies the following quadratic equation

$$\left(k_1\left(\alpha_0-\alpha f\right)\right)^2+k_1\left(\alpha_0-\alpha f\right)-2\alpha k_1 f=0$$

Equation w.r.t. f has two roots

$$f_{1,2}=\frac{\alpha\left(2k_1\alpha_0+3\right)\pm\sqrt{\alpha^2\left(2k_1\alpha_0+3\right)^2-4\alpha_0\left(k_1\alpha_0+1\right)k_1\alpha^2}}{2k_1\alpha^2}$$

Only the smaller one is of interest since it defines the maximum on the feasible fee range $f\in\left[0,\frac{\alpha_0}{\alpha}\right]$.

Hence

$$
\begin{aligned}
f^* &= \frac{\alpha\left(2k_1\alpha_0+3\right)-\sqrt{\alpha^2\left(2k_1\alpha_0+3\right)^2-4\alpha_0\left(k_1\alpha_0+1\right)k_1\alpha^2}}{2k_1\alpha^2} \\
&= \frac{2k_1\alpha_0+3}{2k_1\alpha}\left(1-\sqrt{1-\frac{4\alpha_0 k_1\left(\alpha_0 k_1+1\right)}{\left(2k_1\alpha_0+3\right)^2}}\right) \\
&= \frac{\alpha_0}{\alpha}\left(1+\frac{3}{2k_1}\right)\left(1-\sqrt{1-\frac{1+\frac{1}{\alpha_0 k_1}}{\left(1+\frac{3}{2\alpha_0 k_1}\right)^2}}\right) \\
&= \frac{\alpha_0}{\alpha}\left(1+\frac{3}{2k_1}\right)\left(1-\sqrt{1-\frac{1+\frac{1}{\alpha_0 k_1}}{1+\frac{3}{\alpha_0 k_1}+o\left(\frac{1}{(\alpha_0 k_1)^2}\right)}}\right) \\
&= \frac{\alpha_0}{\alpha}\left(1+\frac{3}{2k_1}\right)\left(1-\sqrt{1-1+\frac{1}{\alpha_0 k_1}\left(1-\frac{3}{\alpha_0 k_1}+o\left(\frac{1}{\left(\alpha_0 k_1\right)^2}\right)\right)}\right) \\
&= \frac{\alpha_0}{\alpha}\left(1+\frac{3}{2k_1}\right)\left(1-\sqrt{1-\left(1+\frac{1}{\alpha_0 k_1}\right)\left(1-\frac{3}{\alpha_0 k_1}+o\left(\frac{1}{\left(\alpha_0 k_1\right)^2}\right)\right)}\right) \\
&= \frac{\alpha_0}{\alpha}\left(1+\frac{3}{2k_1}\right)\left(1-\sqrt{1-\left(1-\frac{2}{\alpha_0 k_1}+o\left(\frac{1}{\left(\alpha_0 k_1\right)^2}\right)\right)}\right) \\
&= \frac{\alpha_0}{\alpha}\left(1+\frac{3}{2k_1}\right)\left(1-\sqrt{\frac{2}{\alpha_0 k_1}+o\left(\frac{1}{\left(\alpha_0 k_1\right)^2}\right)}\right) \\
&\simeq \frac{\alpha_0}{\alpha}\left(1+\frac{3}{2k_1}\right)\left(1-\sqrt{\frac{2}{\alpha_0 k_1}}\right)
\end{aligned}
$$

Linear approximation of the supply sensitivity is a good approximation according to non-parametric fitting – see Figure 6).

■

Correction for supply variables, taking into account suppliers with missing location

Let in a vertical v, in time period t and location x be S_{xvt} suppliers for whom location data was available, and $S_{?vt}$ suppliers in a vertical v and time period t for whom location data was not available. Then the supplier with missing location were distributed among all locations in proportion to their count, i.e. the location-corrected supply variable for location x is as follows:

$$\tilde{S}_{xvt} = \left(S_{?vt} + \sum S_{yvt}\right)\frac{S_{xvt}}{\sum S_{yvt}}$$

De-seasonalization methodology

In pooled estimations for adoption process the dependent variables are de-seasonalized in the following way. First, a centered moving average for the variable is taken over the period $[t-26, t+26]$:

$$x_t^{cma} = \frac{\frac{1}{2}x_{t-26} + \frac{1}{2}x_{t+26} + \sum_{k=-25}^{25} x_{t+k}}{52}$$

Then the seasonality factor is computed as follow:

$$\gamma_t = \frac{1}{n_{years}} \sum_y \frac{x_t}{x_{yt}^{cma}}$$

where the sum is taken across all years y where a week with number t is present.

References

Acemoglu D, Robert Shimer (1999) Holdups and Efficiency with Search Frictions. *International Economic Review* 40(4).

Akbarpour M, Li S, Gharan SO (2014) Dynamic Matching Market Design. *arXiv preprint arXiv:1402.3643* (February), URL http://arxiv.org/abs/1402.3643.

Amihud Y, Mendelson H, Pedersen LH (2006) *Liquidity and Asset Prices* (Now Publishers Inc.).

Aral S, Muchnik L, Sundararajan A (2009) Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences of the United States of America* 106(51):21544–21549, ISSN 0027-8424.

Aral S, Walker D (2011) Creating Social Contagion Through Viral Product Design: A Randomized Trial of Peer Influence in Networks. *Management Science* 57(9):1623–1639, ISSN 0025-1909.

Aral S, Walker D (2014) Tie Strength, Embeddedness, and Social Influence: A Large-Scale Networked Experiment. *Management Science* 60(6):1352–1370, ISSN 0025-1909, URL http://pubsonline.informs.org/doi/abs/10.1287/mnsc.2014.1936.

Armstrong M (2006) Competition in Two-Sided Markets. *RAND Journal of Economics* 37(3):668–691, ISSN 07416261.

Arnosti N (2015) Centralized clearinghouse design: a quantity-quality tradeoff.

Ashlagi I, Jaillet P, Manshadi VH (2013) Kidney Exchange in Dynamic Sparse Heterogenous Pools. *Arxiv preprint arXiv:1206.3522* 1–32.

Bakshy E, Rosenn I, Marlow C, Adamic L (2012) The Role of Social Networks in Information Diffusion. *arXiv* 1201.4145:10, ISSN 1450312292, URL http://arxiv.org/abs/1201.4145.

Bass FM (1969) A newproduct growth for model consumer durables. *Management Science* 15(5):215–227, ISSN 0025-1909.

Becker GS (1973) A Theory of Marriage: Part I. *The Journal of Political Economy* 81(4).

Boudreau K, Hagiu A (2008) Platform rules: Multi-sided platforms as regulators. *Available at SSRN 1269966* 163–191, URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1269966.

Butters GR (1977) Equilibrium distributions of sales and adveristing prices. *The Review of Economic Studies* 44(3):465–491, ISSN 0034-6527, URL http://www.jstor.org/stable/10.2307/2296902\npapers3://publication/uuid/9861BA2A-25C7-4654-859E-BBF2E5

Caillaud B, Jullien B (2003) Chicken & Egg: Competition among Intermediation Service Providers. *The RAND Journal of Economics* 34(2):309–328, ISSN 07416261, URL http://doi.wiley.com/10.2307/1593720.

Cullen Z, Farronato C (2016) Outsourcing Tasks Online : Matching Supply and Demand on Peer-to-Peer Internet Platforms .

Eeckhout J, Kircher P (2010) Sorting versus screening: Search frictions and competing mechanisms. *Journal of Economic Theory* 145(4):1354–1385, ISSN 00220531, URL http://dx.doi.org/10.1016/j.jet.2010.01.005.

Eisenmann T, Parker G, Alstyne MWV (2006) Strategies for Two- Sided Markets. *Harvard Business Review* ISSN 00178012, URL www.hbr.org.

Evans D, Schmalensee R (2013) The Antitrust Analysis of Multi-Sided Platform Businesses, URL http://www.nber.org/papers/w18783.

Fradkin A (2015) Search Frictions and the Design of Online Marketplaces .

Fraiberger SP, Sundararajan A (2015) Peer-to-Peer Rental Markets in the Sharing Economy. *NYU Stern School of Business Research Paper* ISSN 1556-5068, URL http://papers.ssrn.com/abstract=2574337.

Gale D, Shapley L (1962) College admissions and the stability of marriage. *The American Mathematical Monthly* 69(1):9–15.

Gawer A (2011) *Platforms, markets and innovation* (Edward Elgar Publishing).

Godes D, Mayzlin D (2004) Using Online Conversations to Study Word-of-Mouth Communication. *Marketing Science* 23(4):545–560, ISSN 0732-2399.

Goldenberg J, Libai B, Muller E (2001) Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth. *Marketing Letters* 12(3):211–223.

Granovetter MS (1978) Threshold Models of Collective Behavior. *American Journal of Sociology* 83(6):1420–1443, ISSN 00029602, URL http://www.journals.uchicago.edu/doi/abs/10.1086/226707.

Hagiu A (2014) Strategic Decisions for Multisided Platforms. *MIT Sloan Management Review* 55(2):71–80, ISSN 15329194, URL http://sloanreview.mit.edu/article/strategic-decisions-for-multisided-platforms/.

Harris M, Raviv A (1981) Allocation Mechanisms and the Design of Auctions. *Econometrica* 49(6):1477–1499.

Jullien B (2005) Two-sided markets and electronic intermediaries. *CESifo Economic Studies* 51(2-3):233–260, ISSN 1610241X.

Katona Z, Zubcsek PPaL, Sarvary M (2011) Network Effects and Personal Influences : The Diffusion of an Online Social Network. *Journal of Marketing Research* XLVIII(June):425–443, ISSN 00222437, URL http://www.journals.marketingpower.com/doi/pdf/10.1509/jmkr.48.3.425.

Katz ML, Shapiro C (1994) Systems Competition and Network Effects.

Kenneth Burdett, Shouyong Shi, Randall Wright (2001) Pricing and Matching with Frictions. *Journal of Political Economy* 109(5):1060–1085, ISSN 00223808, URL http://www.jstor.org/stable/10.1086/322835.

Koopmans T, Beckmann M (1957) Assignment problems and the location of economic activities. URL http://www.jstor.org/stable/1907742\nhttp://www.jstor.org/stable/10.2307/1907742.

Leskovec J, Adamic L, Huberman B (2007) The Dynamics of Viral Marketing 1(1), ISSN 15591131, URL http://eprints.pascal-network.org/archive/00005344/.

Leskovec J, Backstrom L, Kumar R, Tomkins A (2008) Microscopic Evolution of Social Networks. *KDD'08* 462–470, URL http://dx.doi.org/10.1145/1401890.1401948.

Mahajan V, Muller E, Bass FM (1990) New Product Diffusion Models in Marketing: A Review and Directions for Research. *The Journal of Marketing* 54(1):1–26.

McAfee RP (2008) Mechanism Design by Competing Sellers. *Econometrica* 61(1993):1281–1312.

McBride S (2013) Greylock makes $100 million venture bet on marketplaces. URL http://www.reuters.com/article/2013/08/29/venture-gre

McCall J (1970) Economics of Information and Job Search. *The Quarterly Journal of Economics* 84(1):113–126.

McCall JJ (1965) The Economics of Information and Optimal Stopping Rules. *The Journal of Business* 38(3):300–317.

Mendelson H (1982) Market behavior in a clearing house. *Econometrica: Journal of the Econometric Society* 50(6):1505–1524, URL http://www.jstor.org/stable/1913393.

Mendelson H, Tunca TI (1999) Strategic Spot Trading in Supply Chains. *Management Science* 53(5):742–759.

Milgrom PR, Weber RJ (1982) A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society* 50(5):1089–1122, ISSN 0012-9682.

Myerson RB (1979) Incentive Compatibility and the Bargaining Problem. *Econometrica* 47(1):61–73.

Myerson RB (1981) Optimal Auction Design. *Mathematics of Operations Research* 6(1):58–73.

Nair H (2007) Intertemporal price discrimination with forward-looking consumers: Application to the US market for console video-games. *Quantitative Marketing and Economics* 5(3):239–292, ISSN 15707156.

Nair H, Chintagunta P, Dubé J (2004) Empirical Analysis of Indirect Network Effects in the Market for Personal Digital Assistants. *Quantitative Marketing and Economics* 23–58, URL `http://www.springerlink.com/index/j8162065m0758532.pdf`.

Nosko C, Tadelis S (2015) The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment. *Working Paper* .

Parker GG, Alstyne MWV (2005) Two-Sided Network Effects: A Theory of Information Product Design. *Management Science* 51(10):1494–1504, ISSN 0025-1909.

Peters M (1984) Bertrand Equilbrium with Capacity Constraints and Restricted Mobility. *Econometrica* 52(5):1117–1127.

Resnick P, Zeckhauser R (2002) *Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system.*

Resnick P, Zeckhauser R, Swanson J, Lockwood K (2006) The value of reputation on eBay: A controlled experiment. *Experimental Economics* 9(2):79–101, ISSN 13864157.

Ribeiro B (2014) Modeling and Predicting the Growth and Death of Membership-based Websites. *Proc. WWW* 653–663.

Riley JG, Samuelson WF (1981) Optimal Auctions. *The American Economic Review* 71(3):381–392.

Rochet JC, Tirole J (2006) Two-Sided Markets: A Progress Report. *The RAND Journal of Economics* 37(3):645–667.

Rogers EM (1962) Diffusion of innovations.

Roth A (1984) The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory. *The Journal of Political Economy* 92(6):991–1016, ISSN 0022-3808.

Roth AE (2008) What have we learned from market design? *Economic Journal* 118(2006):285–310, ISSN 0013-0133.

Roth AE, Sotomayor M (1992) Two-sided matching: a study in game-theoretic modeling and analysis (No. 18).

Rysman M (2009) The Economics of Two-Sided Markets. *The Journal of Economic Perspectives* 23(3):125–143.

Shapley LS, Shubik M (1971) The assignment game I: The core. *International Journal of Game Theory* 1(1):111–130, ISSN 00207276.

Smith A (2016) Shared, Collaborative and On Demand: The New Digital Economy. Technical report, Pew Research Center.

Stigler GJ (1961) The Economic of Information 69(3):213–225.

Taylor HM, Karlin S (2014) *An Introduction to stochastic modeling* (Academic Press).

Tucker C, Zhang J (2010) Growing Two-Sided Networks by Advertising the User Base: A Field Experiment. *Marketing Science* 29(5):805–814, ISSN 0732-2399.

Vickrey W (1961) Counterspeculation, Auctions, and Competitive Sealed Tenders. *The Journal of Finance* 16(1):8–37, ISSN 00221082, URL `http://doi.wiley.com/10.1111/j.1540-6261.1961.tb02789.x`.

Walras L (1874) *Elements of Pure Economics, or the theory of social wealth* (Lausanne, Paris).

Weyl EG (2010) A Price Theory of Multi-Sided Platforms. *American Economic Review* 100(4):1642–1672.

Wilson RB (1967) Competitive Bidding with Asymmetric Information. *Management Science* 13(11):816–820, ISSN 00251909.

Wooldridge JM (2002) *Econometric Analysis of Cross Section and Panel Data* (Cambridge, MA: MIT Press).

Young HP (2009) Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning. *American Economic Review* 99(5).