Management Science

Large-Scale Service Marketplaces: The Role of the Moderating Firm

Gad Allon, Achal Bassamboo, Eren B. Çil,

MANAGEMENT SCIENCE
Vol. 58, No. 10, October 2012, pp. 1854–1872
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.1120.1533

Large-Scale Service Marketplaces: The Role of the Moderating Firm

Gad Allon, Achal Bassamboo
Kellogg School of Management, Northwestern University, Evanston, Illinois 60208
{g-allon@kellogg.northwestern.edu, a-bassamboo@kellogg.northwestern.edu}

Eren B. Çil
Lundquist College of Business, University of Oregon, Eugene, Oregon 97403,
erencil@uoregon.edu

Recently, large-scale, Web-based service marketplaces, where many small service providers compete among themselves in catering to customers with diverse needs, have emerged. Customers who frequent these marketplaces seek quick resolutions and thus are usually willing to trade prices with waiting times. The main goal of this paper is to discuss the role of the moderating firm in facilitating information gathering, operational efficiency, and communication among agents in service marketplaces. Surprisingly, we show that operational efficiency may be detrimental to the overall efficiency of the marketplace. Furthermore, we establish that to reap the "expected" gains of operational efficiency, the moderating firm may need to complement the operational efficiency by enabling communication among its agents. The study emphasizes the scale of such marketplaces and the impact it has on the outcomes.

Key words: service operations; fluid models; asymptotic analysis; large games; noncooperative game theory
History: Received August 5, 2009; accepted January 18, 2012, by Yossi Aviv, operations management. Published online in *Articles in Advance* June 15, 2012.

1. Introduction

Recently, large-scale, Web-based service marketplaces, where many small service providers (agents) compete among themselves in catering to customers with diverse needs, have emerged. Customers who frequent these marketplaces seek quick resolutions for their temporary problems and thus are usually willing to trade prices with waiting times. These marketplaces are typically operated by an independent firm, which we shall refer to as the *moderating firm*. The moderating firm establishes the infrastructure for the interaction between customers and agents. In particular, it provides the customers and the agents with the information required to make their decisions. These moderating firms vary with respect to their involvement in the marketplace. They can introduce operational tools that specify how the customers and the agents are matched together. For instance, whereas some of the moderating firms allow customers to choose a specific service provider directly, others allow customers to post their needs and let service providers apply, postponing the service provider selection decision of the customers until they obtain enough information about agents' availability. Moreover, moderating firms can introduce strategic tools that allow communication and collaboration among the agents themselves. These different involvements result in different economic and operational systems and thus vary in their level of efficiency and the outcomes for both customers and service providers.

A typical example of such a marketplace is oDesk.com, where around 1,000,000 programmers compete to provide software solutions. oDesk.com allows for two types of interaction between customers and service providers. On one hand, customers can go directly to a programmer and ask him to provide the service. The customers are then queued for this specific agent. In this type of interaction, most of the time is spent waiting for the agent to complete his previous jobs (36% of the waiting time is spent from the moment the customer chooses the agent until the agent begins working).[1] On the other hand, oDesk.com also allows customers to post jobs and wait while agents apply for the job. In this type of interaction, a negligible amount of time passes until more than 10 agents apply, leaving the decision at the hands of the customer. Another large-scale, online service marketplace is ServiceLive.com, which is a start-up owned by Sears Holding Company. ServiceLive.com (with the slogan of "your price, your time") caters to time- and price-conscious customers and service providers in the home repair and improvement arena. ServiceLive.com allows its

[1] This is based on data obtained from oDesk.com for about 10,000 randomly chosen transactions.

customers to choose among multiple agents once they describe their projects. This type of interaction between customers and service providers is equivalent to the second one described for oDesk.com. Both oDesk.com and ServiceLive.com receive 10% of the revenue obtained by the providers at service completion. In both marketplaces, the moderating firms allow customers to browse among tens of thousands of agents and communicate with different providers.

Both oDesk.com and ServiceLive.com are part of a growing industry of online service marketplaces. Alok Aggarwal, the chairman of Evalueserve.com, a market research company in Saratoga, California, said, "this market [the market for work outsourcing] is expected to grow 20% to $300 million in sales this year, with transactions between employers and the free-lancers totaling about $1.8 billion" (Flandez 2008). In line with this, Gary Swart, chief executive officer of oDesk.com, said that "the number of free-lancers registered with the firm in America has risen from 28,000 at the end of 2008 to 247,000 at the end of April" (*The Economist* 2010).

Motivated by these online service marketplaces, we aim to study the moderating firm's role in the service marketplace where the objective of the individual players, customers as well as service providers, is to maximize their own utility. We distinguish between three degrees of moderating firms' involvement in such markets: (1) *No-intervention*: the moderating firm restricts its involvement to providing the facility for agents to advertise their services and set their prices and for customers to compare the different agents. (2) *Operational efficiency*: the moderating firm provides additional mechanisms that facilitate efficient matching between customers and service providers. These mechanisms aim at reducing the inefficiency associated with having the right agent with the right capability idle while a customer with similar needs is waiting in line for another agent. As we will discuss, a system in which customers post their needs and wait for agents' applications is an example of such a mechanism. (3) *Enabling communication*: the moderating firm may allow providers to communicate among themselves and exchange information on prices and job requirements.

To study the different configurations possible in such marketplaces, we consider a sequence of related games where the set of possible strategies and the solution concepts vary to reflect the different modes of interaction available in the marketplace, either between the customer and the agents or between the agents themselves. Specifically, we study the following three games:

No-intervention Model: In this game, each agent chooses his price and operates as a single-server queue. Customers then choose agents based on prices and waiting times. We characterize the subgame perfect Nash equilibrium in this game.

Operational Efficiency Model: In this game, the mechanism introduced by the moderating firm efficiently matches customers interested in purchasing the service at a particular price with the available agents charging that or a lower price. This mechanism achieves the desired level of efficiency by virtually grouping all agents charging the same price. In contrast to the no-intervention model, customers do not need to commit to a specific agent upon their arrival.

Communication Enabled Model: In this game, agents can exchange information in a noncommittal, costless manner. As in the model with operational efficiency, all the agents charging the same price are virtually grouped, and customers choose the price/subpool. We would be interested in allowing limited preplay communication among the agents within a noncooperative structure; i.e., the agents are free to discuss their pricing strategies but not allowed to make binding commitments. Ray (1996) claims that the possibility of preplay communication has motivated the notion of strong Nash equilibrium, see Aumann (1959), which requires stability against deviations by every conceivable coalition. Following this idea, we use a refinement of the subgame perfect Nash equilibrium concept that requires the equilibrium to be (limited size) coalition proof.

We next state our key findings along with the contributions of the paper:

1. We appear to be the first to distinguish between tools aimed at increasing the operational efficiency and tools aimed at changing the nature of the strategic interaction by enabling communication. We show these tools have a nontrivial impact on the outcomes for all involved parties.

2. In analyzing a market with operational efficiency, we first show that only the prices in a small neighborhood of the operating cost of agents are sustained as equilibrium outcomes when supply exceeds demand. Further, when demand exceeds supply, we are able to show that operational efficiency leads to multiple equilibria in markets with a sufficiently large number of agents. In many of these equilibria, the emerging prices are lower than those arising in the market with no intervention.

3. We show that to overcome the possible deterioration of the profits discussed above, the moderating firm can allow for communication among the agents, even if done through a nonbinding mechanism. The main contribution of this result is in showing that the operational efficiency needs to be complemented with the ability to communicate to obtain desirable outcomes for the involved parties. These desirable outcomes are only achievable in a marketplace where

demand exceeds supply. Therefore, our contribution is also in highlighting the fact that it is crucial to understand the specific market conditions in terms of the ratio between demand and supply.

2. Literature Review

The previous work related with our paper can be divided into two categories. The first category consists of research that studies the applications of queueing theory in service systems. The second one consists of research focused on developing approximations to analyze complex service systems.

Service systems with customers, who are both price and time sensitive, have attracted the attention of researchers for many years. The analysis of such systems dates back to Naor's seminal work (see Naor 1969), which analyzes customer behavior in a single-server queueing system. Motivated by his work, many researchers study the service systems facing price- and delay-sensitive customers in various settings. We refer the reader to Hassin and Haviv (2003) for an extensive summary of the early attempts in this line of research. More recently, Cachon and Harker (2002) and Allon and Federgruen (2007) study the competition between multiple firms offering substitute but differentiated services by modeling customer behavior implicitly via an exogenously given demand function. An alternative approach is followed in Chen and Wan (2003), where authors examine the customers' choice problem explicitly by embedding it into the firms' pricing problem. Other notable examples focusing on the customers' demand decision in competition models are Ha et al. (2003) and Cachon and Zhang (2007).

The pricing and the capacity planning problem of service systems can easily become analytically intractable when trying to study more complex models, such as a multiserver queueing systems. Recognizing this difficulty, many researchers seek robust and accurate approximations to analyze multiserver queues. Halfin and Whitt (1981) is the first paper that proposes and analyzes a multiserver framework. This framework is aimed at developing approximations, which are asymptotically correct, for multiserver systems. It has been applied by many researchers to study the pricing and service design problem of a monopoly in more realistic and detailed settings. Armony and Maglaras (2004) and Maglaras and Zeevi (2005) are examples of recent work using the asymptotic analysis to tackle complexity of these problems. Furthermore, Garnett et al. (2002), Ward and Glynn (2003), and Zeltyn and Mandelbaum (2005) extend the asymptotic analysis of Markovian queueing systems by considering customer abandonments.

The idea of using approximation methods can also be applied to characterize the equilibrium behavior of the firms in a competitive environment. To our knowledge, Allon and Gurvich (2010) is the first paper to study competition among complex queueing systems by using asymptotic analysis to approximate the queueing dynamics. Another recent paper that studies the equilibrium characterization of a competitive marketplace using asymptotic analysis is by Chen et al. (2008). They consider a marketplace with multiple suppliers competing with each other over their prices and target lead times. There are two main differences between these two papers and our work. First, both of them study a service environment with a fixed number of decision makers (firms), whereas the number of decision makers in our marketplace (agents) is growing. Second, they only consider a competitive environment where the firms behave individually. In contrast, we study the noncooperative case as well as the case where the agents have a limited level of collaboration.

In the field of operations management (OM), the majority of the papers employing game-theoretic foundations study noncooperative settings. For an excellent survey, we refer to Cachon and Netessine (2004). There is also a growing literature that studies the OM problems in the context of cooperative game theory. Nagarajan and Sošić (2008) provide an extensive summary of the applications of cooperative game theory in supply chain management. Notable examples are the formation of coalitions among retailers to share their inventories, suppliers, and marketing powers (see Granot and Sošić 2005, Sošić 2006, Nagarajan and Sošić 2007). This body of research is related to our work, where we look for the limited collaboration among agents.

Our work may also be viewed as related to the literature on labor markets that studies the wage dynamics (see Burdett and Mortensen 1998; Manning 2003, 2004; Michaelides 2010). In both our model and labor economics literature, people or firms with service needs seek an employee or an agent to perform the job they requested. In our model, service seekers trade off time they need to wait until their job starts and cost, the phenomenon generally disregarded in labor economics literature. Further, our focus is on a market for temporary help, which means that the engagement between sides ends upon the service completion. This stands in contrast to the labor economics literature in which the engagement is assumed to be permanent. It is also important to note the difference between interventions studied in our model and the ones in the labor economics literature. Unlike the interventions we studied, which focus on improving operational efficiency, the interventions discussed in labor economics are usually aimed at regulating wages directly. Our paper also differs from the literature on market microstructure. This body of literature studies market makers who can set prices and hold inventories of assets to stabilize markets (see Garman 1976, Amihud and Mendelson 1980, Ho and Stoll 1983, and a comprehensive survey by Biais et al. 2005). However, the moderating firm considered in our paper has no direct price-setting power and cannot respond to customers' service requests. Furthermore, papers that study market microstructure disregard the operational details such as waiting and idleness.

3. Model Formulation

Consider a service marketplace where agents and customers make their decisions to maximize their individual utilities. Customers' need for the service is generated according to a Poisson process with rate Λ. This forms the "potential demand" for the marketplace. A customer decides whether to join the marketplace or not: If she decides not to join the system, her utility is zero. If she joins the system, she decides who would process her job. The customers who join the marketplace form the "effective demand" for the marketplace. The exact nature of this decision depends on the specific structure of the marketplace, decided upfront by the moderating firm. We shall elaborate on the choices of customers in §§4–6. We assume that the service time required to satisfy the requests of a given customer is exponentially distributed with rate μ. Without loss of generality, we let $\mu = 1$. When the service of a customer is successfully completed, she pays the price of the service, earns a reward of R, and incurs a waiting cost of c per unit time until her service commences.[2] Because the customers visiting the marketplace seek temporary help, a customer joining the system may become impatient while waiting for her service to start and abandon. In this case, the abandoning customer does not pay any price or earn any reward, but she does incur a waiting cost for the time she spends in the system. We assume that customers' abandonment times are independent of all other stochastic components and are exponentially distributed with mean m_a. Customers decide whether to request service or not and by whom to be served according to their expected utility. The expected utility of a customer is based on the reward, the price, and the anticipated waiting time.

The above summarizes the demand arriving to the marketplace. Next, we discuss the service provision in a marketplace with k ex ante identical agents.[3] The only decision of an agent is to choose a price for his service; each agent makes this decision independently. Let $(p_1, \ldots, p_k)$ denote the resulting price vector, with p_n being the price chosen by the nth agent. We normalized the operating cost of the agents to zero for notational convenience. The expected revenue of an agent depends on the price he chooses and his demand volume.

We refer to the ratio $\Lambda/(\mu k)$ as the demand-supply ratio of the system and denote it by ρ. The demand–supply ratio is a first-order measure for the mismatch between aggregate demand and the total processing capacity. Marketplaces vary with respect to their demand–supply ratio, ρ, and, as we shall discuss, ρ has a significant impact on the market outcome. We broadly categorize marketplaces into two: a buyer's market where $\rho \leq 1$, and a seller's market where $\rho > 1$.

4. No-Intervention Model

The essential role of the moderating firm in a large-scale marketplace is to set up the infrastructure for the interaction between players. This is crucial because all players have to be equipped with the necessary information, such as prices to make their decisions, yet individual players cannot gather this information on their own. When the moderating firm provides only the required information, it has no impact on the strategic interaction taking place in the marketplace. We thus refer to such a setting as the no-intervention model. We analyze the dynamics of a large-scale marketplace in the no-intervention model not only to derive insights about the behavior of the self-interested and competing players in such a system but also to build a benchmark for the cases in which the moderating firm introduces additional features that change the nature of the marketplace. Therefore, in this section, we study the behavior of a marketplace where the moderating firm confines itself to aggregating and providing information.

We model the strategic interaction between the agents and the customers as a sequential move game. Given the setup of §3, along with the above mentioned role of the moderating firm, the agents first announce their prices. Each arriving customer observes these prices and decides whether to request service or not. Further, if a customer decides to join the system, she also chooses the agent who processes her service request. The service of a customer starts immediately if the agent she chooses is available. Otherwise, she joins the queue in front of the agent and waits for her service to commence. We denote the fraction of customers choosing agent n by D_n. Then, ΛD_n is the demand volume for agent n.

More specifically, each agent's operations can be modeled as an $M/M/1+M$ queueing system[4] where

[2] Our model can also be used to study a setting where customers incur waiting cost also during their service. One can incorporate that by modifying the customer reward from R to $R - c/\mu$.

[3] We will discuss a model with heterogeneous agents in §7.

[4] The notation $+M$ denotes the exponential abandonment times.

the arrival rate of customers depends on the strategies of customers and agents.[5] If the rate of customers who request service from an agent charging price p is λ, the utility of a customer requesting service from this agent is $U(\lambda, p) = (R - p)[1 - \beta(\lambda)] - W(\lambda)c$, where $\beta(\lambda)$, which will be referred to as the abandonment function, is the probability of abandonment, and $W(\lambda)$ is the expected waiting time, in an $M/M/1+M$ system with arrival rate λ, service rate 1, and abandonment rate $1/m_a$. Using queueing theory, the utility of customers can be rewritten as $U(\lambda, p) = (R - p + cm_a) \cdot [1 - \beta(\lambda)] - cm_a$. Similarly, the revenue of that agent is $V(\lambda, p) = p\lambda[1 - \beta(\lambda)]$. It is important to note that $V(\lambda, p)$ is the revenue rate of an agent, but throughout the paper we will refer to it as the revenue for ease of exposition.

As we consider a sequential move game, we are interested in the Subgame Perfect Nash Equilibrium (SPNE) of the game. We begin by characterizing the equilibrium in the second-stage game where customers make their service requests given the agents' pricing decisions. Then, based on the second-stage equilibrium, we derive the equilibrium of the first stage in which only agents make pricing decisions.

Fixing the agents' strategies $(p_n)_{n=1}^k$, an arriving customer observes the agents' prices and chooses the agent who maximizes her utility, anticipating the behavior of all other customers. Therefore, in equilibrium a customer chooses an agent only if the utility she obtains from him (weakly) dominates her utility from any other agent. This is also known as "Nash flow equilibrium" (see Roughgarden 2005) in the congestion games literature. We formally define the Customer Equilibrium as follows:

Definition 1 (Customer Equilibrium). Given $(p_n)_{n=1}^k$, we say that $(D_n)_{n=1}^k$ is a Customer Equilibrium if the following conditions are satisfied:

1. For any n with $D_n > 0$, we have that $U(\Lambda D_n, p_n) \geq U(\Lambda D_m, p_m) \geq 0$, for all $m \leq k$.
2. If $U(\Lambda D_n, p_n) > 0$ for some $n \leq k$, then $\sum_{n=1}^k D_n = 1$.

The first condition of the Customer Equilibrium requires that customers request service from an agent in equilibrium only if that agent is one of their best alternatives. Moreover, the second condition ensures that all customers join the system if it is possible to earn strictly positive utility by requesting service from an agent. Customer Equilibrium exists by the continuity of the utility functions and Rath (1992). In the following proposition, we show that for any given price vector, the second-stage game has a unique equilibrium.

Proposition 1. *Given a price vector $(p_n)_{n=1}^k$, there is a unique Customer Equilibrium.*

Because the Customer Equilibrium is unique for any given price vector, we denote the fraction of customers requesting service from agent n in equilibrium by $D_n^{CE}(p_1, \ldots, p_k)$ when $(p_1, \ldots, p_k)$ are the prices announced by agents. $D_n^{CE}(p_1, \ldots, p_k)$ is well defined in light of Proposition 1.

We can now move to the first-stage game, which is played only among the agents. An equilibrium in this stage requires that none of the agents can improve his revenues by deviating unilaterally while taking the customers' response into account. We formalize this in the following definition:

Definition 2 (Subgame Perfect Nash Equilibrium). Let $(D_n, p_n)_{n=1}^k$ summarize the strategy of all players in the market for all $n = 1, \ldots, k$. Then, $(D_n, p_n)_{n=1}^k$ is an SPNE if the following conditions are satisfied:

1. $D_n = D_n^{CE}(p_1, \ldots, p_k)$ for all $n \leq k$.
2. For any $l \leq k$, we have

$$V(\Lambda D_l, p_l) = \max_{p'} V(\Lambda D_l^{CE}(p_1, \ldots, p_{l-1}, p', p_{l+1}, \ldots, p_k), p').$$

The first condition requires that $(D_n)_{n=1}^k$ arises in equilibrium in the second-stage game. The second condition states that none of the agents has incentive to change his price. Note that agents take into account the impact price changes have on the Customer Equilibrium and thus on demand.

4.1. Characterization of SPNE

In this section, we our restrict attention to symmetric SPNE where all agents charge the same price p in the first stage. This is a natural choice because all agents are identical. We will discuss nonsymmetric equilibria in §7.

A price p emerges in equilibrium in the first stage if a single agent chooses to charge p to maximize his revenues given that all other agents announce p. When all other $k-1$ agents announce p, a generic agent, say agent l, solves the following maximization problem to determine his best response:

$$\max_{p_l \geq 0} \left\{ P p_l \Lambda D_l^{CE}(p, \ldots, p, p_l, p, \ldots, p) \cdot [1 - \beta(\Lambda D_l^{CE}(p, \ldots, p, p_l, p, \ldots, p))] \right\}. \quad (1)$$

In this problem, the objective function is the revenue of agent l when he charges p_l and the remaining agents charge p. Thus, p is a symmetric equilibrium in the first-stage game if it is a solution to the

[5] Note that an agent can process more than one job at the same time in certain settings. In such settings, a processor sharing model will be a more appropriate queueing model, yet these models are known to be significantly more complex than our queueing model. Our model can be viewed as an approximation of such settings.

above problem. We denote the symmetric SPNE by (D^*, p^*), where all agents charge p^* and each agent has a demand of ΛD^*; i.e., $D_n^{CE}(p, \ldots, p) = D^*$ for any $n \leq k$. We characterize the symmetric SPNE in the following theorem:

Theorem 1. *If $\beta(\lambda)$ is concave, then there exists a symmetric SPNE. Furthermore, the symmetric SPNE is characterized as follows*:

1. *If $\Lambda \geq k\lambda^0$, then the symmetric SPNE is* $(D^*, p^*) = (\min\{\lambda^{\mathrm{mon}}, \rho\}/\Lambda, R + cm_a - cm_a/(1-\beta(\min\{\lambda^{\mathrm{mon}}, \rho\})))$.
2. *If $\Lambda \leq k\lambda^0$, then the symmetric SPNE is* $(D^*, p^*) = \big(1/k, R + cm_a - (R + cm_a)(k-1)/(k/(1-\nu(\rho)) - 1)\big)$.

Here λ^{mon} is the unique solution to $1 - \beta(\lambda) - \lambda\beta'(\lambda) = cm_a/(R + cm_a)$, *$\lambda^0$ is the unique solution to* $(R + cm_a) \cdot (k-1) - (cm_a/(1-\beta(\lambda)))(k/(1-\nu(\lambda)) - 1) = 0$, *and* $\nu(\lambda) = \lambda\beta'(\lambda)/(1-\beta(\lambda))$.

Similar to Theorems 1–3 in Chen and Wan (2003), the above result suggests that agents behave as local monopolists and charge their monopoly prices when the arrival rate is sufficiently high. Moreover, in this case, agents may choose not to cover the market completely. However, once the arrival rate becomes less than λ^0, the equilibrium price will be pushed down because the agents are engaged in a cutthroat competition, where intensity of competition can be quantified by the strictly positive utility left for customers in the equilibrium. It is also worth noting that utility of customers in the equilibrium increases as the arrival rate decreases.

Remark 1. Concavity of the abandonment function, $\beta(\lambda)$, is a sufficient condition for the existence of symmetric equilibrium. In Lemma 1 in Allon et al. (2012), we show that $\beta(\lambda)$ is concave when $m_a \leq 1$; i.e., abandonment rate is higher than service rate. Furthermore, conducting a numerical study, we observe that $\beta(\lambda)$ is concave even for $1 \leq m_a \leq 2$. However, for higher values of m_a, the function $\beta(\lambda)$ is not concave in λ. This is not surprising given the complicated structure of queueing systems with impatient customers. For instance, Armony et al. (2009) show the difficulty of proving the convexity of the expected head count in the steady state of a system with customer abandonments. Even though $\beta(\lambda)$ is not concave, there can be a symmetric SPNE, and the above theorem characterizes this symmetric equilibrium. Numerically, we see that the equilibrium candidate characterized above still emerges as the symmetric SPNE when $\beta(\lambda)$ is not concave. In this numerical study, we consider a marketplace where $R = 1$, $c \in \{0.05, 0.06, \ldots, 0.2\}$, and $k = 50$. Then we study five scenarios that differ in the average abandonment time m_a and lead to nonconcave $\beta(\lambda)$. We assume $m_a \in \{5, 6, \ldots, 10\}$. For each of these scenarios, we show that the price proposed as equilibrium price in Theorem 1 is equilibrium by varying the arrival rate Λ on a grid from 10 to 50 with a step size of 1.

5. Operational Efficiency Model

In the previous section, we characterized the market outcome in the absence of any intervention on the part of the moderating firm. We now turn to discuss the impact of different mechanisms used by the moderating firm. As we discussed in the introduction, the moderating firm may provide a mechanism that improves the operational efficiency of the whole system by efficiently matching customers and agents. This mechanism aims at reducing inefficiency from the possibility of having a customer waiting in line for a busy agent while an agent who can serve her is idle. This efficiency improvement is equivalent to virtually grouping the agents charging the same price. For instance, oDesk.com achieves this goal by allowing customers to post their needs and allowing service providers to apply to these postings. When a customer posts a job at oDesk.com, agents that are willing to serve this customer apply to the posting. Among the applicants charging less than what the customer wants to pay, the customer will favor agents based on their immediate availability. The main driver of the operational efficiency in this setting is that customers no longer need to specify an agent upon their arrival because the job posting mechanism allows customers to postpone their service request decisions until they have enough information about the availability of the providers.

In this section, we modify the service marketplace considered in §4 by assuming that the mechanism introduced by the moderating firm ensures that customers do not stay in line when there is an idle agent willing to serve them by charging the price they want to pay or less. This can be modeled as a queueing network where the agents announcing the same price are virtually grouped together. Once each agent announces a price per customer to be served, we can construct a resulting price vector $(p_n)_{n=1}^N$ where $N \leq k$ is the number of different prices announced by the agents. We refer to the agents announcing the price p_n as subpool n and denote the number of agents in the subpool n by y_n. Hence, $(p_n, y_n)_{n=1}^N$ summarizes the strategy of all agents.

Under this mechanism, we model the customer decision making and experience as follows: If there are different prices announced by the agents, i.e., $N > 1$, the customer chooses a subpool from which she requests the service. We refer to the price charged by this subpool as the "preferred price." Each customer who decides to join the system enters the service immediately if there is an available agent either in the subpool she chooses or in any subpool announcing a price less than her preferred price. Moreover, the customer is served by the subpool offering the lowest price among all available subpools. Otherwise, she waits in a queue until an agent, who charges a price less than or equal to her preferred price, becomes available. We denote the fraction of customers requesting service from subpool n by D_n. In this model of customer experience, there are two crucial features: (1) The service of an arriving customer commences immediately when there are available agents charging less than or equal to her preferred price. (2) If they have to wait, customers no longer wait for a specific agent but rather for any available agent.

Because we model the marketplace as a queueing network, the operations of each subpool depend on the operations of the other subpools. For instance, each subpool may handle customers from the other subpools (giving priority to its "own" customers) while some of the other subpools are serving its customers. Therefore, given the strategies of agents, $(p_n, y_n)_{n=1}^N$, and the service decisions of customers, $(D_n)_{n=1}^N$, the expected utility of a customer choosing the subpool l depends on all of these decisions and can be written as

$$\begin{aligned} &U_l(D_1, \ldots, D_N; p_1, \ldots, p_N; y_1, \ldots, y_N) \\ &\quad = PServ_{ll}[(R - p_l + cm_a)(1-\beta_l) - cm_a] \\ &\qquad + \sum_{m \neq l} PServ_{lm}(R - p_m), \end{aligned}$$

where $\beta_l(D_1, \ldots, D_N; p_1, \ldots, p_N; y_1, \ldots, y_N)$ denotes the probability of abandonment in the subpool l, and $PServ_{lm}(D_1, \ldots, D_N; p_1, \ldots, p_N; y_1, \ldots, y_N)$ denotes the probability that a customer choosing the subpool l is served by the subpool m when ΛD_n is the rate of customer arrival to the subpool n for $n = 1, \ldots, N$. We want to note that for any subpool l, $PServ_{lm} = 0$ for any m such that $p_m > p_l$ because customer choosing subpool l cannot be served by a subpool charging more than p_l. Furthermore, the revenue of an agent in subpool l is

$$\begin{aligned} &V_l(D_1, \ldots, D_N; p_1, \ldots, p_N; y_1, \ldots, y_N) \\ &\quad = p_l \sigma_l(D_1, \ldots, D_N; p_1, \ldots, p_N; y_1, \ldots, y_N), \end{aligned}$$

where $\sigma_l(\ldots; \ldots; \ldots)$ is utilization of agents in subpool l when ΛD_n is the rate of customer arrival to the subpool n for $n = 1, \ldots, N$. Here, we assume that a customer choosing the subpool l pays p_m when she is served by subpool m for $m \neq l$.

It is also worth noting that a marketplace operates as an $M/M/k+M$ system when all agents charge the same price. This allows us to employ the well-known limiting behavior of the multiserver systems to characterize the market outcome. Furthermore, in the case where the agents announce different prices, we will show that the interdependency between the subpools announcing different prices diminishes as the market grows. In fact, large-scale marketplaces operate "almost like" the combination of independent multiserver systems.

The strategic interaction between the agents and the customers is modeled, as before, as a sequential move game. However, we use a slightly different second-stage equilibrium than the one in Definition 1 because the customer's decision and utility is changed by the new mechanism. The new customer equilibrium, which we refer to as Market Customer Equilibrium, uses the concept of Nash flow equilibrium with the requirement that customers only care about the prices announced by the sub-pools instead of individual prices.

Definition 3 (Market Customer Equilibrium). Given $(p_n, y_n)_{n=1}^N$, we say that $(D_n)_{n=1}^N$ is a Market Customer Equilibrium (MCE) if the following conditions are satisfied:

1. For any l with $D_l > 0$, we have that

$$\begin{aligned} &U_l(D_1, \ldots, D_N; p_1, \ldots, p_N; y_1, \ldots, y_N) \\ &\quad \geq U_m(D_1, \ldots, D_N; p_1, \ldots, p_N; y_1, \ldots, y_N), \end{aligned}$$

for all $m \leq N$.

2. If $U_l(D_1, \ldots, D_N; p_1, \ldots, p_N; y_1, \ldots, y_N) > 0$ for some $l \leq N$, then $\sum_{n=1}^N D_n = 1$.

Although MCE always exists by the continuity of the utility functions and Rath (1992), its uniqueness cannot be guaranteed. For notational convenience, we shall assume that the best outcome from the customer perspective arises when there are multiple MCE (in fact, it can be shown that the limit of all MCEs is unique as the number of agents in the market grows). Because the outcome is assumed to be unique, we denote the fraction of customers requesting service from subpool n in a Market Customer Equilibrium by $D_n^{\mathrm{MCE}}(p_1, \ldots, p_N; y_1, \ldots, y_N)$ when $(p_n, y_n)_{n=1}^N$ is a tuple of two vectors whose components are the prices and the number of agents announcing them.

Agents make pricing decisions in the first stage of the game. Unlike the no-intervention model, we need to account for two types of unilateral deviation of agents: an agent can either choose to deviate by joining an existing subpool or announce a new price. Therefore, an equilibrium in the first stage should be immune to any of these two deviations. One can show that as the market grows, there exists a profitable unilateral deviation from any price in a buyer's market. In analyzing such markets, we would like to highlight the following two observations: (1) The arising system dynamic is too complex for exact analysis yet amenable to asymptotic analysis. (2) Although a single agent, indeed, may have profitable deviations from every price in a buyer's market, the gains from deviations are small and diminish as the market grows. Thus, following Dixon (1987) and recently

Allon and Gurvich (2010), we study a somewhat weaker notion of equilibrium, which allows us to characterize the market outcome (if one exists) as the market grows even when a Nash equilibrium does not exist. To this end, we consider a sequence of marketplaces indexed by the number of agents; i.e., there are k agents in the kth marketplace. The arrival rate in the kth marketplace is assumed to be $\Lambda^k = \rho k$. This ensures that the demand–supply ratio is constant along the sequence of marketplaces. Then, in each market, we focus on an equilibrium concept, which requires immunity against only deviations that improve the revenue of an agent by at least $\epsilon \geq 0$ as formally stated in Definition 4 (see below). We refer to ϵ as the level of equilibrium approximation. We denote the level of equilibrium approximation in the kth market by ϵ^k, and we assume that $\epsilon^k \to 0$ and $\epsilon^k\sqrt{k} \to \infty$ as $k \to \infty$. We study the behavior of the equilibrium along the sequence of marketplaces we described above to derive the equilibrium in a marketplace with a large number of agents.

Definition 4 (ϵ-Market Equilibrium). Let $(D_n^k, p_n^k, y_n^k)_{n=1}^N$ summarize the strategy of all players in the kth market with $y_n^k > 0$ for all $n = 1, \ldots, N$. Then, $(D_n^k, p_n^k, y_n^k)_{n=1}^N$ is an ϵ-Market Equilibrium if the following conditions are satisfied:

1. $D_n^k = D_n^{\mathrm{MCE}}(p_1^k, \ldots, p_N^k; y_1^k, \ldots, y_N^k)$ for all $n \leq N$.
2. For any $l \leq N$ and $m \leq N$, we have that

$$V_l(D_1^k, \ldots, D_N^k; p_1^k, \ldots, p_N^k; y_1^k, \ldots, y_N^k) \geq V_l(\hat{D}_1^k, \ldots, \hat{D}_N^k; p_1^k, \ldots, p_N^k; \hat{y}_1^k, \ldots, \hat{y}_N^k) - \epsilon^k,$$

where $\hat{y}_n^k = y_n^k - 1$ if $n = l$, $\hat{y}_n^k = y_n^k + 1$ if $n = m$, $\hat{y}_n^k = y_n^k$ otherwise, and $\hat{D}_n^k = D_n^{\mathrm{MCE}}(p_1^k, \ldots, p_N^k; \hat{y}_1^k, \ldots, \hat{y}_N^k)$ for all $n \leq N$.

3. For any $l \leq N$ and $p' \neq p_n^k$ for all $n = 1, \ldots, N$, we have that

$$V_l(D_1^k, \ldots, D_N^k; p_1^k, \ldots, p_N^k; y_1^k, \ldots, y_N^k) \geq V_{N+1}(\hat{D}_1^k, \ldots, \hat{D}_{N+1}^k; p_1^k, \ldots, p_N^k, p'; \hat{y}_1^k, \ldots, \hat{y}_{N+1}^k) - \epsilon^k,$$

where $\hat{y}_n^k = y_n^k - 1$ if $n = l$, $\hat{y}_n^k = 1$ if $n = N+1$, $\hat{y}_n^k = y_n^k$ otherwise, and $\hat{D}_n^k = D_n^{\mathrm{MCE}}(p_1^k, \ldots, p_N^k, p'; \hat{y}_1^k, \ldots, \hat{y}_{N+1}^k)$ for all $n \leq N+1$.

The first condition in the above definition requires that the vector $(D_n^k)_{n=1}^N$ forms an equilibrium among the customers if the agents choose the strategy $(p_n^k, y_n^k)_{n=1}^N$. The second and third conditions characterize the equilibrium in the first-stage game: The second condition states that an agent cannot improve his revenue by more than ϵ^k when he joins an existing subpool, and the third condition states that an agent cannot improve his revenue by more than ϵ^k when he introduces a new subpool. We next turn to characterize the equilibrium in the kth marketplace. Note that if $\epsilon^k \equiv 0$ for all k, then the above definition reduces to that of the Nash equilibrium.

5.1. Characterization of the Market Equilibrium

In this subsection, we study the symmetric equilibrium for the sequence of marketplaces we constructed above. As a first step toward characterizing the symmetric equilibrium, we derive the revenues of agents when they announce the same price in the kth marketplace. As we noted before, such a marketplace operates as an $M/M/k+M$ system with arrival rate $\Lambda^k D_1^{\mathrm{MCE}}(p^k; k)$, service rate 1, and abandonment rate $1/m_a$, where $D_1^{\mathrm{MCE}}(p^k; k)$ is the Market Customer Equilibrium when all k agents charge p^k. Therefore, the revenue of an agent in this case is given by

$$V_1(D_1^{\mathrm{MCE}}(p^k; k); p^k; k) = p\rho D_1^{\mathrm{MCE}}(p^k; k)[1 - \beta^M(\Lambda^k D_1^{\mathrm{MCE}}(p^k; k); k)], \quad (2)$$

where $\beta^M(\lambda; k)$ is probability of abandonment in $M/M/k+M$ system with arrival rate λ, service rate 1, and abandonment rate $1/m_a$.

To characterize an ϵ^k-symmetric Market Equilibrium, we need to verify that a single agent does not have any incentive to deviate to a price other than p^k in the kth marketplace. Recall that if an agent chooses $p' \neq p^k$, this amounts to creating his own subpool, and his revenue is given by $V_2(D_1^{\mathrm{MCE}}(p^k, p'; k-1, 1), D_2^{\mathrm{MCE}}(p^k, p'; k-1, 1); p^k, p'; k-1, 1)$, where $(D_n^{\mathrm{MCE}}(p^k, p'; k-1, 1))_{n=1}^2$ is the Market Customer Equilibrium given that $k-1$ agents charge p^k and one agent charges p'. We then say that a price p^k emerges as the symmetric ϵ^k-Market Equilibrium if

$$V_1(D_1^{\mathrm{MCE}}(p^k; k), p^k, k) \geq \max_{0 \leq p' \leq R} V_2\big(D_1^{\mathrm{MCE}}(p^k, p'; k-1, 1), D_2^{\mathrm{MCE}}(p^k, p'; k-1, 1); p^k, p'; k-1, 1\big) - \epsilon^k, \quad (3)$$

where the left-hand side is the revenues of agents when all agents charge p^k, and the right-hand side is the maximum revenue that a single agent can obtain by deviating from p^k.

To understand the behavior of the market outcome in large markets, we shall first study the left-hand side of (3) along the trajectory of marketplaces in which all k agents charge p^k and $p^k \to p$ as $k \to \infty$. In a buyer's market, we show that all customers join the system in equilibrium as long as $p < R$ because they experience negligible waiting times and obtain approximately the utility of $R - p$ by joining in a marketplace with a large number of agents. Therefore, the revenue of each agent is approximated by $p\rho$ in a buyer's market when $p < R$. In a seller's market, some of the customers leave the market immediately because of the high congestion level even if $p < R$, but the rate of customers requesting service should, in equilibrium, be higher than the processing capacity when $p < R$. Therefore, agents are always "overutilized" in a seller's market, and the revenue of each agent is approximately p when $p < R$. When $p = R$, the rate of customers requesting service depends on the convergence rate of p^k both in a buyer's and a seller's market. Thus, $p \min\{\rho, 1\}$ constitutes an upper bound for the revenue of each agent if $p = R$. The following proposition presents these results formally.

Proposition 2. *Let $D_1^{\text{MCE}}(p^k; k)$ be the Market Customer Equilibrium when all agents charge p^k in the kth marketplace such that $\lim_{k\to\infty} p^k = p$. When $= p < R$, we have that $\lim_{k\to\infty} D_1^{\text{MCE}}(p^k; k) = \min\{1, (R - p + cm_a)/(\rho cm_a)\}$. Furthermore,*

$$\lim_{k\to\infty} V_1(D_1^{\text{MCE}}(p^k; k); p^k; k) = \begin{cases} p\rho & \text{if } \rho \le 1, \\ p & \text{if } \rho > 1. \end{cases}$$

When $p = R$, we have that $\limsup_{k\to\infty} D_1^{\text{MCE}}(p^k; k) \le \min\{1, 1/\rho\}$, and

$$\lim_{k\to\infty} V_1(D_1^{\text{MCE}}(p^k; k); p^k; k) \le \begin{cases} p\rho & \text{if } \rho \le 1, \\ p & \text{if } \rho > 1. \end{cases}$$

After approximating the revenue of the agents when they charge the same price, we now focus on the maximum revenue that an agent can obtain by creating his own subpool. As we did above, we again distinguish between buyer's and seller's markets.

5.1.1. Buyer's Market. When all agents charge the same price p^k in a buyer's market, we next show that a single agent can improve his revenue when he decreases his price. Such a cut will allow a single agent to serve not only his own customers but also the customers choosing the price p^k. In fact, his revenue can be arbitrarily close to p^k following a small price cut as long as the rate of customers requesting service is bounded away from zero when all agents charge p^k; i.e., $\lim_{k\to\infty} D_1^{\text{MCE}}(p^k; k) > 0$. The following proposition proves this observation formally.

Proposition 3. *Let*

$$V'(p^k; k) = \max_{0\le p' < p^k} V_2\big(D_1^{\text{MCE}}(p^k, p'; k-1, 1), D_2^{\text{MCE}}(p^k, p'; k-1, 1); p^k, p'; k-1, 1\big)$$

for any sequence of p^k such that $\lim_{k\to\infty} p^k = p$. Then we have that $\liminf_{k\to\infty} V_k'(p^k; k) > 0$ when $p > 0$. Furthermore, when $\lim_{k\to\infty} D_1^{\text{MCE}}(p^k; k) > 0$, we have that $\lim_{k\to\infty} V'(p^k; k) = p$.

As we established in Proposition 2, the revenue of an agent when all agents charge the same price p^k can be bounded from above by $p^k \rho$ in large marketplaces. Then Proposition 3 implies that any p^k satisfying $\lim_{k\to\infty} p^k = p > \epsilon^k/(1-\rho)$ cannot emerge as the equilibrium price of a symmetric ϵ^k-Market Equilibrium for large k. Thus, as $\lim_{k\to\infty} \epsilon^k = 0$, we obtain that any sequence of prices except the ones converging to zero cannot be sustained as the equilibrium price of a symmetric ϵ^k-Market Equilibrium along the trajectory of marketplaces. Note that we do not need to analyze the revenue of an agent after a price increase because it is sufficient to demonstrate the existence of one profitable deviation to show that a given price cannot be an equilibrium outcome. We formalize these observations in the following theorem.

Theorem 2. *In a buyer's market with $\rho < 1$:*

1. *Let p_{EQ}^k be a price emerging as the equilibrium price of a symmetric ϵ^k-Market Equilibrium in the kth marketplace. Then for any $\xi > 0$, there exists a K such that $p_{EQ}^k < \xi$ for all $k > K$.*
2. *There exists a K such that zero is an equilibrium price of a symmetric ϵ^k-Market Equilibrium in the kth marketplace for all $k > K$.*
3. *Let Π_{OE}^k and Π_{NI}^k be the total revenue in the kth marketplace with and without operational efficiency, respectively. Then for any $\xi > 0$, there exists a K such that $\Pi_{OE}^k/\Pi_{NI}^k < \xi$ for all $k > K$.*

The above theorem states that if a moderating firm provides efficient matching in a buyer's market, the equilibrium outcome of the marketplace will converge to zero. Because the profit of the firm is the share of the revenue generated in the marketplace, providing efficient matching deteriorates the profit of the firm, compared to the no-intervention case, as well as the revenue of the agents. In fact, we show that the ratio between the total revenue generated in a marketplace under operational efficiency and under the no-intervention converges to zero. We also establish that zero can emerge as the equilibrium price in large marketplaces. In §7, we discuss the extension of the above theorem, which is based on showing that the revenues of agents converge to zero even in a nonsymmetric equilibrium.

5.1.2. Seller's Market. After discussing the impact of providing efficient matching in a buyer's market, we now focus on a seller's market. Unlike in a buyer's market, a single agent cannot improve his revenue after a price cut because it does not improve his utilization significantly. Note that agents are already "overutilized" and earning a revenue of p^k while they are charging the same price p^k in a seller's market. Therefore, in a seller's market, the only possible profitable deviation for a single agent is to increase his price in large enough marketplaces.

In such a deviation, a single agent loses some of his customers because of his high price, and he also loses the benefits of efficient matching because he becomes an individual provider. Both of these factors will limit his ability to make higher profit. In fact, the following proposition establishes an upper bound on the asymptotic revenue that a single agent can generate by increasing his price.

Proposition 4. *Let*

$$V'(p^k;k) = \max_{p^k \le p' \le R} V_1\big(D_1^{\text{MCE}}(p', p^k; 1, k-1), D_2^{\text{MCE}}(p', p^k; 1, k-1); p', p^k; 1, k-1\big)$$

for any given sequence of prices p^k such that $\lim_{k\to\infty} p^k = p$. When $p < R$ in a seller's market ($\rho > 1$), we have that

$$\limsup_{k\to\infty} V'(p^k;k) \le (R + cm_a)\lambda^\Delta(p;R)[1 - \beta(\lambda^\Delta(p;R))] - \lambda^\Delta(p;R)(\Delta(p;R) + cm_a),$$

where $\Delta(p;R) = \max\{0, (R - p + cm_a)/\rho - cm_a\}$, and $\lambda^\Delta(p;R)$ is the unique solution to $1 - \beta(\lambda) - \lambda\beta'(\lambda) = (\Delta(p;R) + cm_a)/(R + cm_a)$.

When a single provider increases his price, we show that the demand for agents who do not change their prices is almost the same as their original demand before deviation. Hence, the utility of customers choosing the subpool consisting of $k-1$ agents is $\Delta(p;R)$, which is the utility that the customers obtain in the Market Customer Equilibrium in a large marketplace when all agents charge p^k. Then to approximate the maximum post-deviation revenue, one can treat the deviating agent as a monopoly whose customers have an outside option with the value of $\Delta(p;R)$. In fact, the above proposition shows that this approximation constitutes an upper bound on the agent's post-deviation revenue. A monopoly always makes sure that the utility of customers is exactly equal to their outside option by setting the price to $R + cm_a - (\Delta(p;R) + cm_a)/(1 - \beta(\lambda))$ for any given target of demand rate λ. He then picks λ, maximizing his revenue, and sets his price accordingly. We refer the reader to the proof of Proposition 4 for a more detailed discussion on the revenue maximization problem of a monopoly.

Combining the two observations above, it is clear that in a large marketplace, a price p^k emerges as the symmetric ϵ^k-Market Equilibrium outcome if p^k is greater than the profit of a monopoly serving customers with outside option $\Delta(p;R)$. We state this result in the following theorem.

Theorem 3. *In a seller's market ($\rho > 1$), let*

$$p^* \in \mathcal{P}(\rho;R) \equiv \{p\colon p > (R + cm_a)\lambda^\Delta(p;R)[1 - \beta(\lambda^\Delta(p;R))] - \lambda^\Delta(p;R)(\Delta(p;R) + cm_a),\ 0 \le p < R\},$$

where $\Delta(p;R)$ and $\lambda^\Delta(p;R)$ are defined as in Proposition 4. Then for any given sequence of prices p^{*k} that converges to p^* as $k \to \infty$, there exists a K such that p^{*k} emerges as the equilibrium price of a symmetric ϵ^k-Market Equilibrium in the kth marketplace for all $k > K$. Furthermore, for any $\rho_1 > \rho_2$, we have that $\mathcal{P}(\rho_1;R) \subseteq \mathcal{P}(\rho_2;R)$.

The above theorem characterizes the set of symmetric ϵ^k-Market Equilibria for large marketplaces. The theorem does not guarantee the uniqueness of such an equilibrium; i.e., $\mathcal{P}(\rho;R)$ may not be a singleton. In fact, $\mathcal{P}(\rho;R)$ may consist of uncountably many prices. Furthermore, we show that $\mathcal{P}(\rho;R)$ shrinks as ρ increases. As the demand–supply ratio increases, customers experience significant waiting times even if they are served by a price-generated pool. Therefore, the level of customer surplus that a deviating agent has to forego declines as ρ rises. As a result of this, a single agent has more room to deviate and improve his revenue when demand is high. It is also worth highlighting that a single agent has such a profitable deviation opportunity even though the number of agents grows to infinity.

Characterizing the set of symmetric equilibria, $\mathcal{P}(\rho;R)$, is difficult in general. For illustrative purposes, we consider the case where the abandonment rate is equal to the service rate. We show that a similar structure holds for the settings when $\mu \ne m_a$ using a numerical study (see Allon et al. 2012). The next corollary characterizes the correspondence $\mathcal{P}(\rho;R)$ as well as the asymptotic behavior of the unique equilibrium price under the no-intervention model.

Corollary 1. *Suppose the abandonment rate is equal to the service rate. Then we have that*

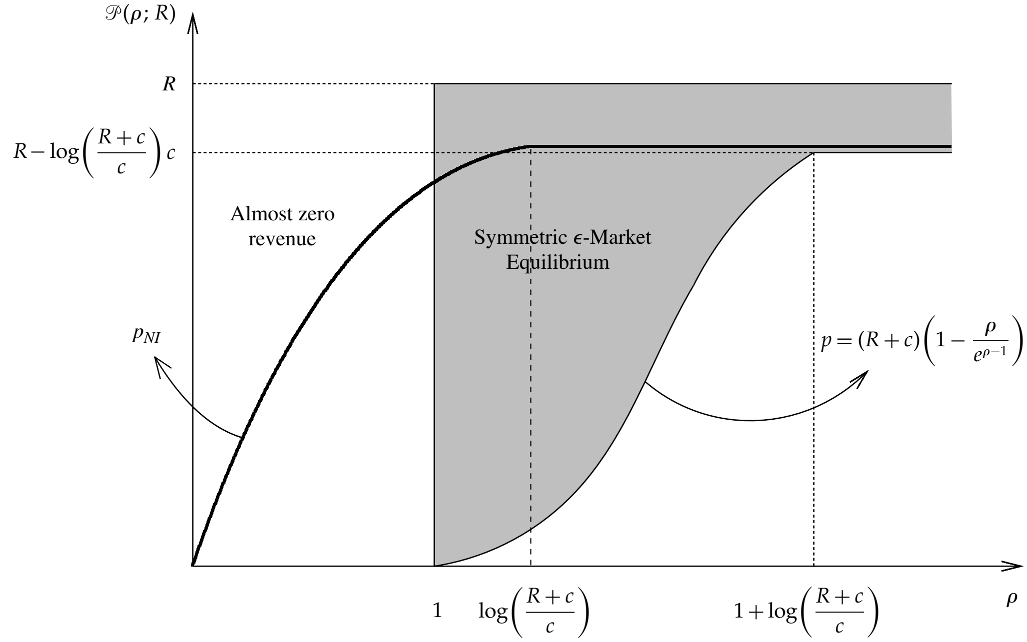
1. *$\lambda^\Delta(p;R) = \log((R+c)/(\Delta(p;R)+c))$ where $\Delta(p;R)$ is defined as in Proposition 4. Furthermore, the correspondence $\mathcal{P}(\rho;R)$ defined in Theorem 3 can be expressed as*

$$\mathcal{P}(\rho;R) = \left\{p\colon p > \left[R + c - \left(1 + \log\left(\frac{R+c}{\Delta(p;R)+c}\right)\right)\cdot[\Delta(p;R)+c]\right],\ 0 \le p < R,\right\}.$$

2. *$\lim_{k\to\infty} p_{NI}^k = p_{NI} \equiv (R + c)\min\{1 - \rho/(e^\rho - 1), 1 - (c/R)\log((R+c)/c)\}$, where p_{NI}^k is the unique equilibrium price under no-intervention setting in the kth marketplace.*

Figure 1 displays the correspondence $\mathcal{P}(\rho;R)$ and the limit p_{NI}. More specifically, the gray area

Figure 1 **Prices That Form a Symmetric Market Equilibrium as a Function of the Demand–Supply Mismatch** (ρ)



Note. The service rates and abandonment rates are assumed to be one.

represents the prices that can emerge as the equilibrium price of a symmetric equilibrium in a large marketplace and the bold curve depicts p_{NI}. We observe that for all $\rho > 1$, the set $\mathcal{P}(\rho; R)$ is not a singleton. In fact, we have a wide range of prices that can form an equilibrium. Furthermore, many of the possible equilibrium prices in $\mathcal{P}(\rho; R)$ are lower than p_{NI}. The intuition behind this result is the following: In a marketplace where the moderating firm efficiently matches customers and agents, a single agent who deviates by increasing his price loses benefits of efficient matching and thus cannot sustain the same quality of service (in terms of waiting times) as his "original" pool. It turns out that the deviating agent cannot improve his "original" revenue by decreasing his price either. Thus, in a seller's market, the price-generated pool serves as a deterrent against single agent deviations even if prices are unappealing from a system point of view. It is also important to note that such lower prices lead to loss in total revenue for the marketplace compared to the no-intervention setting. Although one may expect operational efficiency tools to be a leverage for higher revenues in the market, it is surprising to see that reducing the unnecessary waiting and idleness present in a system with no intervention may deteriorate the revenues.

Myerson (1991) argues that the question of which equilibrium would emerge as the outcome of a game with multiple equilibria can be answered with the focal-point effect phenomenon.[6] Our goal in this paper is not to conclude that only low prices can be a focal point. In fact, when comparing the equilibrium prices in a market with and without operational efficiency, one should also observe that operational efficiency does not only serve as a deterrent for deviations from low prices but also prevents deviations from high prices for any level of demand-supply ratio. Moreover, when the aggregate demand is sufficiently high, efficient matching always leads to higher profits, although the equilibrium prices under operational efficiency may be slightly lower than the unique equilibrium in a market without operational efficiency.

Our analysis in this section assumes that a customer with a preferred price p_l pays p_m when she is served by subpool m for any $m \neq l$. In real marketplaces such as oDesk.com, customers may end up paying a price between their preferred price and the prices asked by the providers when these prices are different. To account for that, one may envision an extension of our model in which a customer choosing subpool l pays $\phi p_l + (1-\phi)p_m$, where $\phi \in (0, 1)$, when an agent from subpool m with $m \neq l$ serves her. In such a model, our key findings—equilibrium prices are close to zero in a buyer's market and some of the equilibrium outcomes may lead to profit loss in a seller's market—would continue to hold.

In this section, we study a specific mechanism that the moderating firm uses to achieve operational efficiency. There are other mechanisms, such as providing real-time congestion information, that may be used by the moderating firm. When customers are able to

[6] Focal-point effects are any psychological or cultural norms that tends to focus players' attention on one equilibrium.

obtain real-time congestion information, Allon et al. (2012) show that our analytical results for a buyer's market continue to hold, and a simulation experiment demonstrates that multiple, possibly harmful equilibria also exist in a seller's market.

6. Communication Enabled Model

In this section, we continue to study the impact of different mechanisms used by the moderating firm. As we mentioned in the introduction, the moderating firm may complement its operational tool discussed in the previous section with a strategic tool that changes the nature of the interaction among agents. In a marketplace such as oDesk.com, service providers are offered discussion boards in which they are allowed to exchange information. Moreover, the market supports the creation of affiliation groups, which are self-enforcing entities. We will thus focus on the impact of enabling communication among agents on the market outcome.

The economics literature suggests that when the players have the opportunity to perform nonbinding preplay communication among themselves, the stability of an outcome can be threatened by potential deviations formed by coalitions, even in noncooperative games. Following this idea, the well-known notion of Strong Nash Equilibrium (SNE) requires stability against deviations formed by any conceivable coalitions (see Aumann 1959). The main drawback of SNE is that many of the games do not have any SNE.

In this section, we modify the marketplace we study in the previous section by assuming that agents have opportunities to make nonbinding communication prior to making their decisions so that they can try to self-coordinate their actions in a mutually beneficial way, despite each agent selfishly maximizing his own utility.

Echoing the ideas in the economics literature, allowing communication among agents changes the equilibrium concept we use to characterize the outcome in the marketplace. We model this by proposing a new equilibrium concept that allows several agents to deviate together. More specifically, the new concept requires that a strategy of agents should be immune to any coalitions. Because a marketplace tends to be large, e.g., there are hundreds of thousands of agents in oDesk.com, one has to restrict the possible size of a coalition. We denote the largest fraction of agents that is allowed to deviate together by $\delta \in (1/k, 1]$. As in §5, we focus on the deviations that improve the revenues of agents at least by $\epsilon \geq 0$. Furthermore, we again study the behavior of the equilibrium along the sequence of marketplaces we described in §5. Recall that there are k agents, the arrival rate is $\Lambda^k = \rho k$, and the level of equilibrium approximation is ϵ^k, with the same asymptotic properties as in §5, in the kth marketplace. We let δ^k be the largest fraction of agents that is allowed to deviate together in the kth marketplace. We assume that $\delta^k k \to \infty$ as $k \to \infty$. This condition states that the number of agents allowed to deviate increases without bound as the market size increases. We refer to our new equilibrium concept as (δ, ϵ)-Market Equilibrium, which is defined as follows:

Definition 5 ((δ, ϵ)-Market Equilibrium). Let $(D_n^k, p_n^k, y_n^k)_{n=1}^N$ summarize the strategy of all players in the kth market with $y_n^k > 0$ for all $n = 1, \ldots, N$. Then, $(D_n^k, p_n^k, y_n^k)_{n=1}^N$ is a (δ^k, ϵ^k)-Market Equilibrium if the following conditions are satisfied:

1. $D_n^k = D_n^{\text{MCE}}(p_1^k, \ldots, p_N^k; y_1^k, \ldots, y_N^k)$ for all $n \leq N$.
2. For any $l \leq N$, $m \leq N$, and $0 < d \leq \min\{y_l^k, \lfloor \delta^k k \rfloor\}$, we have that

$$V_l(D_1^k, \ldots, D_N^k; p_1^k, \ldots, p_N^k; y_1^k, \ldots, y_N^k) \geq V_l(\hat{D}_1^k, \ldots, \hat{D}_N^k; p_1^k, \ldots, p_N^k; \hat{y}_1^k, \ldots, \hat{y}_N^k) - \epsilon^k,$$

where $\hat{y}_n^k = y_n^k - d$ if $n = l$, $\hat{y}_n^k = y_n^k + d$ if $n = m$, $\hat{y}_n^k = y_n^k$ otherwise, and $\hat{D}_n^k = D_n^{\text{MCE}}(p_1^k, \ldots, p_N^k; \hat{y}_1^k, \ldots, \hat{y}_N^k)$ for all $n \leq N$.

3. For any $l \leq N$, $0 < d \leq \min\{y_l^k, \lfloor \delta^k k \rfloor\}$, and $p' \neq p_n$ for all $n = 1, \ldots, N$, we have that

$$V_l(D_1^k, \ldots, D_N^k; p_1^k, \ldots, p_N^k; y_1^k, \ldots, y_N^k) \geq V_{N+1}(\hat{D}_1^k, \ldots, \hat{D}_{N+1}^k; p_1^k, \ldots, p_N^k, p'; \hat{y}_1^k, \ldots, \hat{y}_{N+1}^k) - \epsilon^k,$$

where $\hat{y}_n^k = y_n^k - d$ if $n = l$, $\hat{y}_n^k = d$ if $n = N+1$, $\hat{y}_n^k = y_n^k$ otherwise, and $\hat{D}_n^k = D_n^{\text{MCE}}(p_1^k, \ldots, p_N^k, p'; \hat{y}_1^k, \ldots, \hat{y}_{N+1}^k)$ for all $n \leq N+1$.

The above definition is closely related to the definition of ϵ-Market Equilibrium in §5. The key difference between these two equilibrium definitions is that (δ, ϵ)-Market Equilibrium allows a group of agents to deviate by either forming a new subpool or joining an existing one. In fact, our new equilibrium concept is a refinement of the ϵ-Market Equilibrium. Therefore, any (δ, ϵ)-Market Equilibrium is also a ϵ-Market Equilibrium. Employing the (δ, ϵ)-Market Equilibrium concept, we expect that the set of prices that can be sustained as a ϵ-Market Equilibrium will shrink because (δ, ϵ)-Market Equilibrium is more restrictive. Kalai (2004) and Gradwohl and Reingold (2008) study large games and show that all Nash equilibria of certain large games are resilient to deviations by coalitions. Such a phenomena does not exist in our model.[7]

[7] According to the definition in Gradwohl and Reingold (2008), a Nash equilibrium is resilient to coalitions if players cannot improve their revenues "too much" even after a coordinated deviation. In our setting, "too much" has to be almost as much as the customer reward, R, in order to apply their results to our game. Clearly, this makes the definition of resilience vacuous because none of the agents can increase his revenue by more than R.

6.1. Characterization of the (δ, ϵ)-Market Equilibrium

Similar to §5, we focus on the symmetric (δ, ϵ)-Market Equilibrium where all agents charge the same price. The revenue of an agent when all agents charge the same price p^k is the same as in (2), and thus Proposition 2 establishes its asymptotic behavior.

In a buyer's market with $\rho < 1$, we showed that only the prices in a small neighborhood of zero can emerge as a symmetric ϵ-Market Equilibrium in large marketplaces. Because the (δ, ϵ)-Market Equilibrium is a refinement of the ϵ-Market Equilibrium, this implies any sequence of prices that emerge as symmetric (δ, ϵ)-Market Equilibrium converges to zero as the market size grows. Furthermore, we show that $p = 0$ can emerge as the equilibrium price in large marketplaces.

Theorem 4. *Let p^k_{EQ} be a price emerging as a symmetric (δ^k, ϵ^k)-Market Equilibrium in the kth marketplace where $\rho < 1$. Then, for any $\xi > 0$, there exists a K such that $p^k_{EQ} < \xi$ for all $k > K$. Furthermore, when $\lim_{k\to\infty} \delta^k = 0$, there exists a K such that zero is an equilibrium price of a symmetric (δ^k, ϵ^k)-Market Equilibrium in the kth marketplace for all $k > K$.*

In a seller's market, Proposition 2 shows that the rate of customers requesting service will exceed the processing capacity of agents when all agents charge a price lower than R. Therefore, customers experience significant waiting times and not only pay the price of the service but also incur a strictly positive waiting cost. Then we show that a small group of agents, because customers pay an extra cost, can increase their prices while ensuring that they are still "overutilized" after the price increase. Because this small group of agents increases their prices without hurting their utilization, this deviation clearly improves their revenues (this is in contrast to the setting in §5 where the utilization of a single agent does drop after a price increase). Thus, in a seller's market, only the prices, which are very close to R, can emerge as the equilibrium price of a symmetric (δ, ϵ)-Market Equilibrium in large marketplaces. To contrast this result with the result in Theorem 3, it is worth noting that a single agent has only a limited opportunity to improve his revenue by increasing his price because in most cases the revenue improvement from the price increase is overcome by the drop in utilization. Therefore, without the communication opportunity, it was possible to observe low prices as the market outcome even though demand exceeds supply.

Theorem 5. *Let p^k_{EQ} be a price emerging as a symmetric (δ^k, ϵ^k)-Market Equilibrium in the kth marketplace where $\rho > 1$. Then for any $\xi > 0$, there exists a K such that $p^k_{EQ} > R - \xi$ and $D_1^{\mathrm{MCE}}(p^k_{EQ}; k) > 1/\rho - \xi$ for all $k > K$. Furthermore, there exist a sequence p^{*k} and a K such that p^{*k} forms a symmetric (δ^k, ϵ^k)-Market Equilibrium in the kth marketplace, for all $k > K$.*

The above result shows that agents can sustain a price, which extracts all of the customer surplus, as the equilibrium outcome in a seller's market. Moreover, it also implies that the marketplace cannot be congested in the equilibrium even in a seller's market because any level of congestion can be capitalized by agents through a price increase.

Theorem 5 characterizes the unique limit of symmetric (δ, ϵ)-Market Equilibrium, but this result can be extended by showing that R is indeed the unique limit of all possible (δ, ϵ)-Market Equilibria as discussed in §7. Furthermore, Allon et al. (2012) show that the ability to communicate leads to high equilibrium prices when the moderating firm provides real-time queue information to reduce the mismatch between customers and agents as long as the largest fraction of agents that is allowed to deviate together is close to one.

7. A Marketplace with Nonidentical Agents

In §3, we introduced a model where all of the agents in the marketplace are a priori identical. However, it is natural to imagine that large service marketplaces attract service providers with different skill sets, which provides their customers different values for the service. Thus, we explore the robustness of the conclusions of the previous sections to the heterogeneity among providers.

To this end, we consider a marketplace where agents provide the same service but in different quality levels, low and high. We assume that customers value the service with respect to its quality. Particularly, customers earn a reward of R_H and R_L when they are served by a high-quality and a low-quality agent, respectively. Without loss of generality, we assume $R_L \le R_H$. The model setup is the same as in §3, and we use a similar mode of analysis as in §§4–6.

The behavior of the marketplace when the moderating firm confines itself to setting up the necessary infrastructure is very similar to the equilibrium in Theorem 1: Agents may behave as local monopolists when the arrival rate is sufficiently high. Furthermore, once the arrival rate is less than a certain threshold, customers observe lower prices, which allow them to earn strictly positive utility, because of the intensified competition. However, we also encounter new results when we allow for heterogeneous agents. First, unlike the identical agent model, we observe that the main driver of equilibrium outcomes for certain parameters is not only the competition between providers but also the agents' different quality of service. For instance, when the demand rate is in a certain range, high-quality agents charge a low price and forego a significant customer surplus both because of the low demand and because they want to keep the low-quality agents out of the marketplace. We also show that it is possible to have a continuum of symmetric equilibria, whereas we always have a unique symmetric equilibrium with the identical agents.

The impact of improving the operational efficiency in a marketplace with nonidentical agents is also similar to our findings in the identical agents model: When demand is sufficiently low in a buyer's market, the revenues of agents are always in a small neighborhood of zero in large marketplaces. In a seller's market, there are multiple equilibria, which may lead to profit loss for the firm compared to the no-intervention model. Unlike the identical agents model, we show that there may be multiple equilibria even in a buyer's market as long as demand exceeds the total capacity of high-quality agents. However, most of these equilibrium prices may be very low compared to the equilibrium outcome in the no-intervention model. Thus, providing tools to improve operational efficiency may still deteriorate the moderating firm's profit.

Finally, we explore the impact of enabling communication among agents in a market with nonidentical agents. As in §6, we establish that preplay communication helps agents sustain the profit maximizing one among the multiple equilibria arising from providing operational efficiency.

Our results in the nonidentical agents model also provide insights about the nonsymmetric equilibrium outcomes in the identical agents model. In particular, our model with nonidentical agents helps us to prove that the nonsymmetric equilibrium may exist only for a small range of demand–supply ratio ρ in the no-intervention model with identical agents, and this range becomes negligible as the number of agents grows. Furthermore, using our results in this section, we show that in the operational efficiency model with identical agents, the revenues of all agents in any nonsymmetric equilibrium (if it exists) should be in a small neighborhood of zero in a buyer's market. We also show that in the communication model with identical agents, even if there are any nonsymmetric equilibria in a seller's market, the revenue of each agent in equilibrium, as well as the price they charge, should converge to R.

We refer the reader to Allon et al. (2012) for a detailed discussion of our findings in this section.

8. Conclusion

In this paper, we study a marketplace in which many small service providers compete with each other in providing service to self-interested customers looking for temporary help. The main focus of the paper is on the role of the moderating firm, which sets up the marketplace and creates the infrastructure where agents and customers interact. To this end, we explore the impact of different strategies employed by the moderating firm by considering three market models.

We characterize the market outcomes in each of these models. We observe that outcomes critically depend on the moderating firm's involvement and market conditions, i.e., whether it is a buyer's or a seller's market. Because different types of involvement of the moderating firm result in different equilibrium prices and customer demand, the moderating firm aims to intervene in the marketplace to make sure that the "right" prices and customer demand emerge in equilibrium. Specifically, the moderating firm tries to maximize the revenues of agents because its profit is a share of the agents' revenues.

We show that when the firm ensures efficient operational matching and enables agent communication in a seller's market, the natural upper bound on the revenue generated in a marketplace[8] is asymptotically achievable; thus, using these two tools together dominates any other strategy from the moderating firm's perspective in a seller's market. We also show that efficient operational matching in a buyer's market leads to arbitrarily small total marketplace revenue compared to the total revenue under the no-intervention model. Hence, using the matching mechanism we discuss in this paper is not advisable in a buyer's market despite its reduction of the mismatch between demand and supply. This result is somewhat counterintuitive because the efficiency improvement from better matching is not necessarily translated into additional profits. It seems other tools aimed at improving operational efficiency, such as providing real-time queue information, will have a similar impact on the moderating firm's profit in a buyer's market.

Both oDesk.com and ServiceLive.com are currently in their growth stage and have not achieved their full potential in terms of demand for their services. However, both firms can and should project the "mature" market conditions and decide on their appropriate measures to adopt. Given the moderate level of congestion in oDesk.com, one may infer that the marketplace can be identified as a seller's market. Following the discussion before, oDesk.com's decision to offer operational tools complemented with strategic tools is well justified.

[8] In a given marketplace, the total revenues of the agents cannot exceed $\min\{\Lambda, k\}R$ because they cannot charge more than R, and their effective demand is the minimum of their processing capacity and the aggregate demand.

There are also other possible ways for a moderating firm to be involved in the marketplace, including contracting with agents or providing a suggested price. Particularly, the setting in which the firm provides a suggested price can be viewed as preplay communication and will indeed shrink the set of equilibria. However, these types of interactions between the moderating firm and agents are outside the scope of this paper because these settings are not a market per se anymore. In such environments, the firm would decide on prices as well as the allocation of agents to customers.

Modeling operational efficiency, we assume that agents give priority to their own customers. One may consider an extension of our model in which agents are allowed to choose both priority and prices simultaneously. The equilibria that arise in our model with fixed priority rule would still be sustained in such an extended game. Hence, the main spirit of our findings, namely, that providing operational efficiency may lead to profit loss, would not change. Additional equilibria would be possible in the extended model only when demand exceeds supply.

Acknowledgments

The authors are grateful to Martin Lariviere, Jan Van Mieghem, and Laurens Debo for their helpful suggestions. The authors also thank the review team for their valuable comments, which substantially improved this paper.

Appendix A. Proofs in §4[9]

A.1. Proof of Proposition 1

Suppose there are two Customer Equilibria, say, $(D_n)_{n=1}^k$ and $(D'_n)_{n=1}^k$; given $(p_n)_{n=1}^k$ with $p_n < R$ for some n (when $p_n = R$ for all $1 \le n \le k$, the unique equilibrium is clearly $D_n = 0$ for all $1 \le n \le k$). Let $S = \{n \le k \colon D_n > 0\}$, and $S' = \{n \le k \colon D'_n > 0\}$.

As Lemma 2 in Allon et al. (2012) shows, we have that $S = S'$. Then let $U(\Lambda D_n, p_n) = u$ for any $n \in S$ and $U(\Lambda D'_n, p_n) = u'$ for any $n \in S'$. Because $S = S'$ and $D_n \ne D'_n$ for some $n \in S$, we have that $u \ne u'$. Without loss of generality, assume $u > u'$. This implies that $\sum_{n=1}^k D_n < \sum_{n=1}^k D'_n \le 1$. However, because $\sum_{n=1}^k D_n < 1$, we have that $u' < u = 0$, which is a contradiction.

A.2. Proof of Theorem 1

Existence and Uniqueness of λ^{mon} and λ^0. After a birth–death chain analysis of an $M/M/1+M$ system with arrival rate λ, service rate 1, and abandonment rate $1/m_a$, we have that $\beta(\lambda) = 1 - g(\lambda)/(\lambda(1+g(\lambda)))$, and $W(\lambda) = m_a\beta(\lambda)$ where $a_0 = 1$, $a_n = 1/(\prod_{i=0}^{n-1}(1+i/m_a)) = m_a^n/(\prod_{i=0}^{n-1}(m_a+i))$ for any $n \ge 1$, and $g(\lambda) = \sum_{n=1}^{\infty} a_n\lambda^n$.

Observe that $1 - \beta(\lambda) - \lambda\beta'(\lambda)$ is strictly decreasing in λ because $\lambda[1-\beta(\lambda)]$ is strictly concave by Lemma 1.4 in Allon et al. (2012). Moreover, $\lim_{\lambda\to 0}[1-\beta(\lambda)-\lambda\beta'(\lambda)] = \lim_{\lambda\to 0}(g'(\lambda)/[1+g(\lambda)]^2) = 1$ because $\lim_{\lambda\to 0} g(\lambda) = 0$, and $\lim_{\lambda\to 0} g'(\lambda) = 1$. It is also true that $\lim_{\lambda\to\infty}[1-\beta(\lambda)-\lambda\beta'(\lambda)] \le 0$ because $\lim_{\lambda\to\infty}\beta(\lambda) = 1$. Therefore, it is clear that λ^{mon} exists and it is unique.

Let $z(\lambda) = (R+cm_a)(k-1) - [cm_a/(1-\beta(\lambda))][k/(1-\nu(\lambda)) - 1]$. Then, $z(\lambda)$ is strictly decreasing in λ because $\nu(\lambda)$ and $\beta(\lambda)$ are strictly increasing in λ by Lemma 1. Moreover, $z(\lambda^{\text{mon}}) = cm_a/(1-\beta(\lambda^{\text{mon}})) - (R+cm_a) < 0$. Therefore, it is clear that λ^0 exists, it is unique, and $\lambda^0 < k\lambda^{\text{mon}}$.

Necessary Conditions for the Symmetric Equilibrium. The best response problem of agent l in (1) can be rewritten as follows:

$$\begin{aligned}
\max_{p_l\ge 0,\, D_l\ge 0,\, D_{-l}\ge 0}\quad & p_l\Lambda D_l[1-\beta(\Lambda D_l)]\\
\text{s.t.}\quad & (R-p_l+cm_a)[1-\beta(\Lambda D_l)] - cm_a \ge 0,\\
& (R-p_l+cm_a)[1-\beta(\Lambda D_l)]\\
&\quad = (R-p+cm_a)[1-\beta(\Lambda D_{-l})],\\
& D_l + (k-1)D_{-l} \le 1.
\end{aligned}$$

In this new problem, we state the conditions of the Customer Equilibrium as the constraints of the problem. In other words, for any (D_l, D_{-l}) satisfying the constraints, we have that $D_l = D_l^{CE}(p,\ldots,p,p_l,p,\ldots,p)$ and $D_{-l} = D_n^{CE}(p,\ldots,p,p_l,p,\ldots,p)$ for any $n \ne l$. We denote the solution to the above problem by $(D_l(p), D_{-l}(p), p_l(p))$ for a given p.

After some algebra and using the first-order-conditions of the above problem, any symmetric SPNE (D,p) should satisfy the following conditions:

$$\left.\begin{aligned}
D &= (\min\{\lambda^{\text{mon}}, \Lambda/k\})/\Lambda\\
p &= R + cm_a - \frac{cm_a}{1-\beta(\min\{\lambda^{\text{mon}}, \Lambda/k\})}
\end{aligned}\right\} \Leftrightarrow \Lambda \ge \lambda^0,$$

$$\left. D = 1/k,\quad p = R + c - \frac{(R+c)(k-1)}{\dfrac{k}{1-(\Lambda/k\beta'(\Lambda/k))/(1-\beta(\Lambda/k))} - 1}\right\} \Leftrightarrow \Lambda < \lambda^0.$$

Sufficient Conditions for the Symmetric Equilibrium. Lemma 3 in Allon et al. (2012) establishes the existence of the symmetric SPNE when $\beta(\lambda)$ is concave.

Appendix B. Proofs in §5

B.1. Proof of Proposition 2

We prove Proposition 2 through a case-by-case analysis focusing two cases: (1) $\mathbf{p} < \mathbf{R}$ and (2) $\mathbf{p} = \mathbf{R}$.

Case 1 ($\mathbf{p} < \mathbf{R}$). To prove our claim in this case, we first argue that $\liminf_{k\to\infty} D_1^{\text{MCE}}(p^k;k) \ge \min\{1, (R-p+cm_a)/(\rho cm_a)\}$. Suppose on the contrary that the result does not hold. Then there exists a convergent subsequence of $D_1^{\text{MCE}}(p^k;k)$, say $D_1^{\text{MCE}}(p^k;k)$ (we do not use a new notation for the subsequence for notational convenience) such that

$$\lim_{k\to\infty} D_1^{\text{MCE}}(p^k;k) = \liminf_{k\to\infty} D_1^{\text{MCE}}(p^k;k) < \min\left\{1, \frac{R-p+cm_a}{\rho cm_a}\right\},$$

[9] Allon et al. (2012) provide the formal presentations and the proofs of the supplementary lemmas used in the appendices.

because $D_1^{\text{MCE}}(p^k;k) \in [0,1]$ for any $k = 1, 2, \ldots$. Let $\tilde{D} = \lim_{k\to\infty} D_1^{\text{MCE}}(p^k;k)$. Then because the system behaves as a multiserver queue when all the agents charge the same price, we have that

$$\begin{aligned}\lim_{k\to\infty} U_1(D_1^{\text{MCE}}(p^k;k);p^k;k) &= \lim_{k\to\infty}(R-p^k+cm_a)(1-\beta^M(\Lambda^k D_1^{\text{MCE}}(p^k;k);k))-cm_a\\ &= \frac{R-p+cm_a}{\max\{\tilde{D}\rho,1\}}-cm_a > cm_a-cm_a=0,\end{aligned}$$

where the equality holds by Lemma 4 in Allon et al. (2012) and the last inequality holds because $\tilde{D}(p) < \min\{1, (R-p+cm_a)/(\rho cm_a)\}$ and $p < R$. Therefore, there exists a K^* such that for any $k > K^*$, we have $U_1(D_1^{\text{MCE}}(p^k;k);p^k;k) > 0$, whereas $D_1^{\text{MCE}}(p^k;k) < 1$. However, this contradicts the definition of Market Customer Equilibrium.

We now argue that $\limsup_{k\to\infty} D_1^{\text{MCE}}(p^k;k) \le \min\{1, (R-p+cm_a)/(\rho cm_a)\}$. To do this it is sufficient to show $\limsup_{k\to\infty} D_1^{\text{MCE}}(p^k;k) \le (R-p+cm_a)/(\rho cm_a)$ because $D_1^{\text{MCE}}(p^k;k) \le 1$ for any k. Suppose on the contrary that the result does not hold. Then there exists a convergent subsequence of $D_1^{\text{MCE}}(p^k;k)$, say $D_1^{\text{MCE}}(p^k;k)$, such that

$$\lim_{k\to\infty} D_1^{\text{MCE}}(p^k;k) = \limsup_{k\to\infty} D_1^{\text{MCE}}(p^k;k) > \frac{R-p+cm_a}{\rho cm_a},$$

because $D_1^{\text{MCE}}(p;k) \in [0,1]$ for any $k = 1, 2, \ldots$. Let $\tilde{D} = \lim_{k\to\infty} D_1^{\text{MCE}}(p^k;k)$. Then observe that

$$\begin{aligned}\lim_{k\to\infty} U_1(D_1^{\text{MCE}}(p^k;k);p^k;k) &= \lim_{k\to\infty}(R-p^k+cm_a)(1-\beta^M(\Lambda^k D_1^{\text{MCE}}(p^k;k),k))-cm_a\\ &= \frac{R-p+cm_a}{\max\{\tilde{D}\rho,1\}}-cm_a < cm_a-cm_a=0,\end{aligned}$$

where the equality holds by Lemma 4 and the last inequality holds because $\rho\tilde{D}(p) > (R-p+cm_a)/(cm_a) \ge 1$. Therefore, there exists a K^* such that for any $k > K^*$, we have $U_1(D_1^{\text{MCE}}(p^k;k);p^k;k) < 0$. However, this contradicts with the definition of Market Customer Equilibrium because $D_1^{\text{MCE}}(p^k;k) > 0$ for large k.

Once we establish that

$$\lim_{k\to\infty} D_1^{\text{MCE}}(p^k;k) = \min\{1, (R-p+cm_a)/(\rho cm_a)\},$$

we have that

$$\lim_{k\to\infty}\left[1-\beta^M(\Lambda^k D_1^{\text{MCE}}(p^k;k);k)\right] = \frac{1}{\max\{\rho\min\{1,(R-p+cm_a)/(\rho cm_a)\},1\}}$$

by Lemma 4. Finally, combining these two, we have that $\lim_{k\to\infty} V_1(D_1^{\text{MCE}}(p^k;k);p^k;k) = p\min\{\rho,1\}$.

Case 2 ($\mathbf{p}=\mathbf{R}$). Note that we do not use the condition $p < R$ to argue that $\limsup_{k\to\infty} D_1^{\text{MCE}}(p^k;k) \le \min\{1,(R-p+cm_a)/(\rho cm_a)\}$. Therefore, $\min\{1,(R-p+cm_a)/(\rho cm_a)\}$ is also the upper bound for the fraction of customers requesting service in the limit when $p=R$. As a direct implication of that, the upper bound for the revenues of the agents in the limit is $p\min\{\rho,1\}$.

B.2. Proof of Proposition 3

We start proving the proposition by considering the case $p < R$. Note that $\lim_{k\to\infty} D_1^{\text{MCE}}(p^k;k) > 0$ when $p < R$ by Proposition 2. Thus, for the case where $p < R$, we need to show that $\lim_{k\to\infty} V'(p^k;k) = p$. Note that this statement is trivially true when $p = 0$. To prove Proposition 3 for $p > 0$, we consider a deviation by a single agent where he decreases his price by an arbitrary small amount $\varepsilon > 0$.

Let $D_{\text{pool}}(k) = D_1^{\text{MCE}}(p^k, p-\varepsilon; k-1, 1)$ and $D_{\text{one}}(k) = D_2^{\text{MCE}}(p^k, p-\varepsilon; k-1, 1)$. We first argue that

$$\liminf_{k\to\infty} D_{\text{pool}}(k) + D_{\text{one}}(k) \ge \min\{1, 1/\rho\}$$

for any $p < R$. We prove this claim by contradiction, so we suppose $\liminf_{k\to\infty} D_{\text{pool}}(k) + D_{\text{one}}(k) < \min\{1, 1/\rho\}$. Then there should exist convergent subsequences of $D_{\text{pool}}(k)$ and $D_{\text{one}}(k)$ such that $\lim_{k\to\infty} D_{\text{pool}}(k) + D_{\text{one}}(k) < \min\{1, 1/\rho\}$. Using this observation, and letting $P_{\text{one}}(k) = PServ_{12}(D_{\text{pool}}(k), D_{\text{one}}(k); p^k, p-\varepsilon; k-1, 1)$ for notational convenience, we have that

$$\begin{aligned}&\lim_{k\to\infty} U_1(D_{\text{pool}}(k), D_{\text{one}}(k); p^k, p-\varepsilon; k-1, 1)\\ &\quad= \Big(1-\lim_{k\to\infty} P_{\text{one}}(k)\Big)\Big[(R-p+cm_a)\\ &\qquad\cdot\Big[1-\lim_{k\to\infty}\beta_1(D_{\text{pool}}(k), D_{\text{one}}(k); p^k, p-\varepsilon; k-1, 1)\Big]-cm_a\Big]\\ &\qquad+(R-p+\varepsilon)\lim_{k\to\infty} P_{\text{one}}(k) \ge (R-p+cm_a)\\ &\qquad\cdot\Big[1-\lim_{k\to\infty}\beta_1(D_{\text{pool}}(k), D_{\text{one}}(k); p^k, p-\varepsilon; k-1, 1)\Big]-cm_a\\ &\quad\ge (R-p+cm_a)\Big[1-\lim_{k\to\infty}\beta^M(\Lambda^k D_{\text{pool}}(k); k-1)\Big]-cm_a\\ &\quad= R-p > 0,\end{aligned}$$

where the second inequality holds because some customers choosing subpool 1 may be served by subpool 2, and the last equality holds because $\lim_{k\to\infty}((\Lambda^k D_{\text{pool}}(k))/(k-1)) < \min\{1,\rho\} \le 1$. However, this contradicts with the definition of the customer equilibrium because we suppose $\lim_{k\to\infty} D_{\text{pool}}(k) + D_{\text{one}}(k) < 1$, i.e., some customers choose not to request service for sufficiently large k. Hence, we should have that $\liminf_{k\to\infty} D_{\text{pool}}(k) + D_{\text{one}}(k) \ge \min\{1, 1/\rho\}$.

Then using the fact that $\liminf_{k\to\infty} D_{\text{pool}}(k) + D_{\text{one}}(k) \ge \min\{1, 1/\rho\}$, we have that

$$\begin{aligned}&\liminf_{k\to\infty} V_2(D_{\text{pool}}(k), D_{\text{one}}(k); p^k, p-\varepsilon; k-1, 1)\\ &\quad= (p-\varepsilon)\liminf_{k\to\infty}\sigma_2(D_{\text{pool}}(k), D_{\text{one}}(k); p^k, p-\varepsilon; k-1, 1)\\ &\quad\ge (p-\varepsilon)\lim_{k\to\infty}\frac{\Lambda^k(D_{\text{pool}}(k)+D_{\text{one}}(k))}{1+\Lambda^k(D_{\text{pool}}(k)+D_{\text{one}}(k))} = p-\varepsilon,\end{aligned}$$

where the inequality holds by Lemma 5 in Allon et al. (2012). Note that the revenue of a single agent after the deviation we propose is less than the optimal deviation $V'(p^k;k)$; thus we have that

$$\begin{aligned}&\liminf_{k\to\infty} V'(p^k;k)\\ &\quad\ge \liminf_{k\to\infty} V_2(D_{\text{pool}}(k), D_{\text{one}}(k); p^k, p-\varepsilon; k-1, 1) \ge p-\varepsilon.\end{aligned}$$

Finally, our claim holds because ε can be arbitrarily small and $V'(p^k;k) \le p^k$ by construction.

Now, we consider the case where

$$p = R \quad \text{and} \quad \liminf_{k\to\infty} D_1^{\text{MCE}}(p^k;k) = \tilde{D} > 0$$

(note that $\rho\tilde{D} \le 1$ by Proposition 2). This time, we will show that $\liminf_{k\to\infty} D_{\text{pool}}(k) + D_{\text{one}}(k) \ge \tilde{D}$. As above, we assume the contrary. Then there exists a subsequence of $D_{\text{pool}}(k)$ such that $\liminf_{k\to\infty} D_{\text{pool}}(k) = \tilde{D}_{\text{pool}} < \tilde{D}$. Then we have that

$$\begin{aligned} &U_1(D_{\text{pool}}(k), D_{\text{one}}(k); p^k, p^k - \varepsilon; k-1, 1) \\ &\quad \ge (R - p^k + cm_a)[1 - \beta^M(\Lambda^k D_{\text{pool}}(k); k-1)] - cm_a, \\ &\quad > (R - p^k + cm_a)[1 - \beta^M(\Lambda^k D_1^{\text{MCE}}(p^k;k);k)] - cm_a \ge 0, \end{aligned}$$

for large k, where the strict inequality holds because

$$\beta^M(\Lambda^k D_{\text{pool}}(k); k-1) \simeq \zeta_1(\rho\tilde{D}_{\text{pool}} e^{1-\rho\tilde{D}_{\text{pool}}})^k \quad \text{and}$$
$$\beta^M(\Lambda^k D_1^{\text{MCE}}(p^k;k);k) \simeq \zeta_2(\rho\tilde{D}e^{1-\rho\tilde{D}})^k$$

for some constants ζ_1 and ζ_2 by Theorem 5 in Zeltyn and Mandelbaum (2005) and because $\rho\tilde{D} \le 1$. Note that $(\rho\tilde{D}_{\text{pool}} e^{1-\rho\tilde{D}_{\text{pool}}})^k/(\rho\tilde{D}e^{1-\rho\tilde{D}})^k \to 0$ as $k \to 0$ because $\tilde{D}_{\text{pool}} < \tilde{D}$ and xe^{1-x} is strictly increasing in x for any $x < 1$. However, this contradicts with the definition of the customer equilibrium because we suppose $\lim_{k\to\infty} D_{\text{pool}}(k) + D_{\text{one}}(k) < 1$; i.e., some customers choose not to request service for sufficiently large k. Hence, we should have that $\liminf_{k\to\infty} D_{\text{pool}}(k) + D_{\text{one}}(k) \ge \tilde{D}$. Using this result, we can again show that the utilization of the deviating agent will converge to one, and thus his revenue will be $R - \varepsilon$.

Finally, we need to show that $\lim_{k\to\infty} V'(p^k;k) > 0$ when $p = R$ and $\liminf_{k\to\infty} D_1^{\text{MCE}}(p^k;k) = 0$. Consider a deviation where a single agent cuts his price and charges $R/2$. Let, $\hat{\lambda}$ solves $(R/2 + cm_a)[1 - \beta(\lambda)] = cm_a$. There exists such $\hat{\lambda}$ because $\beta(\lambda)$ is increasing in λ and $\lim_{\lambda\to\infty} \beta(\lambda) = 1$. Then by construction $\Lambda^k(D_{\text{pool}}(k) + D_{\text{one}}(k)) \ge \hat{\lambda}$ because otherwise the customer choosing the deviating agent would earn a strictly positive utility while $D_{\text{pool}}(k) + D_{\text{one}}(k) < 1$ for large k, and that would be a contradiction. Therefore, using Lemma 5, we have that $\lim_{k\to\infty} V'(p^k;k) \ge R/2(\hat{\lambda}/(1+\hat{\lambda})) > 0$.

B.3. Proof of Theorem 2

1. To prove our claim, it is sufficient to show that $\limsup_{k\to\infty} p_{EQ}^k = 0$ because for any $\xi > \limsup_{k\to\infty} p_{EQ}^k$, there is a K such that $p_{EQ}^k < \xi$ for all $k > K$ by Theorem 3.17 in Rudin (1976). We prove that $\limsup_{k\to\infty} p_{EQ}^k = 0$ by contradiction. Thus, we suppose that $\limsup_{k\to\infty} p_{EQ}^k > 0$. Then there should exist a convergent subsequence of p_{EQ}^k such that $\lim_{k\to\infty} p_{EQ}^k = \tilde{p} > 0$ because the equilibrium prices p_{EQ}^k are bounded from above by R. Let

$$\begin{aligned} V'(p_{EQ}^k;k) = \max_{0\le p'\le p_{EQ}^k} V_2\big(&D_1^{\text{MCE}}(p_{EQ}^k, p'; k-1, 1), \\ &D_2^{\text{MCE}}(p_{EQ}^k, p'; k-1, 1); p_{EQ}^k, p'; k-1, 1\big). \end{aligned}$$

When $\rho < 1$ and $\lim_{k\to\infty} D_1^{\text{MCE}}(p_{EQ}^k;k) > 0$, we have that $\liminf_{k\to\infty} V'(p_{EQ}^k;k) = \tilde{p} > \rho\tilde{p} \ge \lim_{k\to\infty} V_1(D_1^{\text{MCE}}(p_{EQ}^k;k); p_{EQ}^k;k) + \lim_{k\to\infty} \epsilon^k$, by Proposition 3 and by the definition of ϵ^k. Then for sufficiently large k, we should have that $V'(p_{EQ}^k;k) > V_1(D_1^{\text{MCE}}(p_{EQ}^k;k); p_{EQ}^k;k) + \epsilon^k$, which implies that p_{EQ}^k cannot emerge as the equilibrium price of a symmetric ϵ-Market Equilibrium for large k.

Similarly, when $\lim_{k\to\infty} D_1^{\text{MCE}}(p_{EQ}^k;k) = 0$ (and thus, $\lim_{k\to\infty} V_1(D_1^{\text{MCE}}(p_{EQ}^k;k); p_{EQ}^k;k) = 0$), we have that

$$\liminf_{k\to\infty} V'(p_{EQ}^k;k) > 0 = \lim_{k\to\infty} V_1(D_1^{\text{MCE}}(p_{EQ}^k;k); p_{EQ}^k;k) + \lim_{k\to\infty} \epsilon^k,$$

by again Proposition 3 and because $\tilde{p} > 0$. Thus, for sufficiently large k, we should again have that $V'(p_{EQ}^k;k) > V_1(D_1^{\text{MCE}}(p_{EQ}^k;k); p_{EQ}^k;k) + \epsilon^k$, which implies that p_{EQ}^k cannot emerge as the equilibrium price of a symmetric ϵ-Market Equilibrium for large k.

2. To prove this claim, we suppose, on the contrary, that for any K, there exists a $k > K$ such that zero cannot be a symmetric equilibrium in the kth market. Thus, there should be a sequence $\hat{p}^k$ such that a single agent can improve his revenue by increasing his price to $\hat{p}^k$ in the kth marketplace. Let $U_{\text{pool}}(k)$ and $U_{\text{dev}}(k)$ be the utility of customers choosing price zero and $\hat{p}^k$, respectively. Because we suppose that the deviating agent improves his revenue, a strictly positive fraction of customers should pick him, and thus we should have that $U_{\text{dev}}(k) \ge U_{\text{pool}}(k)$ for any k. Using this observation we have that

$$\begin{aligned} &(R - \hat{p}^k)[1 - P_{12}(k)] + RP_{12}(k) \\ &\quad \ge U_{\text{dev}}(k) \ge U_{\text{pool}}(k) \ge (R - cm_a)(1 - \beta^M(\rho k; k-1)) - cm_a, \end{aligned}$$

where $P_{12}(k)$ is the probability that a customer picking $\hat{p}^k$ is served by the agents charging zero in the kth marketplace. The first inequality above holds because customers, who pick $\hat{p}^k$ and are served by the deviating agents, may abandon, and the last inequality holds because agents charging zero may not serve all customers, and they give priority to their own customers. Because $\rho < 1$ and using Theorem 5.1 of Zeltyn and Mandelbaum (2005), we have that $U_{\text{pool}}(k)$ converges to R with an exponential speed; i.e., there exists a constant ζ such that $U_{\text{pool}}(k) = R - e^{-\zeta k}$. Then the above inequality implies that $\hat{p}^k[1 - P_{12}(k)] \le e^{-\zeta k}$.

Note that the revenue of the agent deviating, say $V_{\text{dev}}(k)$, in the kth marketplace is less than $\rho k p_k'[1 - P_{12}(k)]$ because the rate of customers he can serve cannot be greater than $\rho k[1 - P_{12}(k)]$. As a result of this observation, we have that $V_{\text{dev}}(k) \le \rho k e^{-\zeta k} \Rightarrow V_{\text{dev}}(k)/\epsilon^k \le \rho k e^{-\zeta k}/\epsilon^k \to 0$ as $k \to \infty$, where the convergence holds because $\epsilon^k\sqrt{k} \to \infty$ and $e^{\zeta k}k^{-3/2} \to \infty$ as $k \to \infty$. Because $V_{\text{dev}}(k)/\epsilon^k$ converges to zero, we should have that $V_{\text{dev}}(k) < \epsilon^k$ for large k, which contradicts that $\hat{p}^k$ is a profitable deviation. Hence, there should be a K such that zero emerges as the equilibrium price of a symmetric ϵ^k-Market Equilibrium for all $k > K$.

3. Let λ_k^0 be the constant λ^0 defined in the kth marketplace in the no-intervention model. Note that $\lim_{k\to\infty} \lambda_k^0/k = \lambda^{\text{mon}}$. Using this observation, we have that

$$\lim_{k\to\infty} p_{NI}^k = \begin{cases} (R+c)\left[1 - \dfrac{1 - \beta(\rho) - \rho\beta'(\rho)}{1 - \beta(\rho)}\right] & \textit{if } \rho \le \lambda^{\text{mon}} \\[2ex] (R+c)\left[1 - \dfrac{c}{(R+c)(1 - \beta(\lambda^{\text{mon}}))}\right] & \textit{if } \rho > \lambda^{\text{mon}}, \end{cases}$$

where p^k_{NI} is the unique equilibrium price under no-intervention setting. Furthermore, we have that the utilization of a single agent in the no-intervention model converges to $\rho[1-\beta(\rho)]$ when $\rho \le \lambda^{\text{mon}}$ and $\lambda^{\text{mon}}[1-\beta(\lambda^{\text{mon}})]$ otherwise. Thus, it is clear that the revenue of an agent converges to a strictly positive limit, say $\bar{v}$, under the no-intervention model. Thus, there is a K_1 such that $\Pi^k_{NI} > \bar{v}/2$ for all $k > K_1$. Furthermore, by part 1, there is a K_2 such that revenue of an agent under the operational efficiency model is less than $\xi\bar{v}/2$ for all $k > K_2$. Thus, we have that $\Pi^k_{OE}/\Pi^k_{NI} < (k\xi\bar{v}/2)/(k\bar{v}/2) < \xi$ for all $k > \max\{K_1, K_2\}$.

B.4. Proof of Proposition 4

Consider the following problem:

$$\Pi(p) = \max_{p\le p'\le R,\, \lambda\ge 0} \; p'\lambda[1-\beta(\lambda)]$$
$$\text{s.t.} \;\; (R-p'+cm_a)[1-\beta(\lambda)] \ge \Delta(p;R)+cm_a.$$

One can easily show that $\Pi(p) = \max_\lambda (R+cm_a)\lambda[1-\beta(\lambda)] - \lambda[\Delta(p;R)+cm_a]$. Then by using the first-order condition, we have that $\Pi(p) = (R+cm_a)\lambda^\Delta(p;R)[1-\beta(\lambda^\Delta(p;R))] - \lambda^\Delta(p;R)(\Delta(p;R)+cm_a)$, where $\lambda^\Delta(p;R)$ solves $1-\beta(\lambda)-\lambda\beta'(\lambda) = (\Delta(p;R)+cm_a)/(R+cm_a)$.

Now, we are ready to prove Proposition 4. For notational convenience, let $\hat{p}(k) = \arg\max_{p^k\le p'\le R} V_1(D_1^{\text{MCE}}(p',p^k;1,k-1), D_2^{\text{MCE}}(p',p^k;1,k-1); p',p^k;1,k-1)$, $\lambda(k) = \Lambda^k D_1^{\text{MCE}}(\hat{p}(k),p^k;1,k-1)$, and $D_{\text{pool}}(k) = D_2^{\text{MCE}}(\hat{p}(k),p^k;1,k-1)$. Observe that for any $k = 1,2,\ldots$, $\hat{p}(k)\in[0,R]$, $\lambda(k)\in[0,\bar{\lambda}]$, $D_{\text{pool}}(k)\in[0,1]$, $V'(p^k;k)\in[0,R]$, where $\bar{\lambda}$ is defined as in Lemma 6 in Allon et al. (2012). Then there exists a convergent subsequence of $V'(p^k;k)$, which is denoted by $V'(p^k;k)$ (for notational convenience), such that $\lim_{k\to\infty} V'(p^k;k) = \limsup_{k\to\infty} V'(p^k;k)$. Moreover, $\hat{p}(k)$, $\lambda(k)$, $D_{\text{pool}}(k)$ have convergent subsequences, and we let $\tilde{p}^{\text{dev}} = \lim_{k\to\infty}\hat{p}(k)$, $\tilde{\lambda} = \lim_{k\to\infty}\lambda(k)$, and $\tilde{D}_{\text{pool}} = \lim_{k\to\infty} D_{\text{pool}}(k)$.

Then by the continuity of $\beta(\lambda)$ and definition of Market Customer Equilibrium, we have that

$$\begin{aligned}
(R-\tilde{p}^{\text{dev}}+cm_a)[1-\beta(\tilde{\lambda})]
&= \lim_{r\to\infty}(R-\hat{p}(k)+cm_a)[1-\beta(\lambda(k))]\\
&= \lim_{k\to\infty} U_1(\lambda(k)/\Lambda^k, D_{\text{pool}}(k); \hat{p}(k), p^k; 1, k-1) + cm_a\\
&= \lim_{k\to\infty} U_2(\lambda(k)/\Lambda^k, D_{\text{pool}}(k); \hat{p}(k), p^k; 1, k-1) + cm_a\\
&= \frac{(R-p+cm_a)}{\rho\tilde{D}_{\text{pool}}} = \Delta(p;R)+cm_a,
\end{aligned}$$

where the second equality follows by Lemmas 6.2 and 6.3, and the last two equalities hold by Lemmas 6.5 and 6.6. Therefore, $(\tilde{p}^{\text{dev}}, \tilde{\lambda})$ satisfy the constraint in the limit problem. And this implies that $\limsup_{k\to\infty} V'(p^k;k) = \tilde{p}^{\text{dev}}\tilde{\lambda}(1-\beta(\tilde{\lambda})) \le \Pi(p)$.

B.5. Proof of Theorem 3

We first show that when $\rho > 1$, a single provider who cuts his price cannot improve his revenue by more than ϵ^k for large enough k. Thus, the only possible profitable deviation for a single agent is to increase his price for large enough, yet finite, k. To argue that, let $V^{\text{cut}}(p^k;k) = \max_{0\le p'\le p^k} V_2(D_1^{\text{MCE}}(p^k,p';k-1,1), D_2^{\text{MCE}}(p^k,p';k-1,1); p^k,p';k-1,1)$ for any given sequence p^k with limit $p < R$. Note that $\rho\lim_{k\to\infty} D_1^{\text{MCE}}(p;k) > 1$ by Proposition 2. Therefore, using Theorem 6.1 in Zeltyn and Mandelbaum (2005), we have that the revenue of an agent converges to p exponentially as $k\to\infty$ when all agents charge p^k in a seller's market; i.e., there exists a constant ζ such that $V_1(D_1^{\text{MCE}}(p^k;k);p^k;k) = p^k(1-e^{-\zeta k})$. Using this observation, for large enough k, we have that $V_1(D_1^{\text{MCE}}(p^k;k);p^k;k) = p^k(1-e^{-\zeta k}) \ge p^k - Re^{-\zeta k} > p^k - \epsilon^k \ge V^{\text{cut}}(p;k) - \epsilon^k$. The second inequality holds for large enough k because $\lim_{k\to\infty} e^{-\zeta k}/\epsilon^k = 0$ by our assumption of $\lim_{k\to\infty}\epsilon^k\sqrt{k} = \infty$ and because $\lim_{k\to\infty} e^{\zeta k}/\sqrt{k} = \infty$. This implies that a single agent cannot have a profitable deviation by decreasing his price in large marketplaces. Hence, to verify that any sequence of prices p^{*k} with limit $p^* \in \mathcal{P}(\rho;R)$ can emerge as an equilibrium outcome of a symmetric ϵ^k-Market Equilibrium, it is sufficient to check any single agent deviation where the agent increases his price.

Let

$$V'(p^{*k};k) = \max_{p^{*k}\le p'\le R} V_1\big(D_1^{\text{MCE}}(p',p^{*k};1,k-1), D_2^{\text{MCE}}(p',p^{*k};1,k-1); p',p^{*k};1,k-1\big).$$

Because $p^* \in \mathcal{P}(\rho;R)$, we have that

$$\begin{aligned}
&\lim_{k\to\infty} V_1(D_1^{\text{MCE}}(p^{*k};k);p^{*k};k)\\
&\quad = p^* > (R+cm_a)\lambda^\Delta(p^*;R)[1-\beta(\lambda^\Delta(p^*;R))]\\
&\qquad - \lambda^\Delta(p^*;R)(\Delta(p^*;R)+cm_a)\\
&\quad \ge \limsup_{k\to\infty} V'(p^{*k};k),
\end{aligned}$$

where the last inequality holds by Proposition 4. Then there should exist a K such that $V'(p^{*k};k) < V_1(D_1^{\text{MCE}}(p^{*k};k);p^{*k};k)$ for all $k > K$, which implies that p^{*k} can emerge as the equilibrium price of a symmetric ϵ^k-Market Equilibrium for all $k > K$.

Monotonicity of $\mathcal{P}(\rho;R)$. Let $\Pi(p,\rho) = \max_\lambda (R+cm_a)\cdot\lambda[1-\beta(\lambda)] - \lambda[\Delta(p;R)+cm_a]$. We first want to note that $\Pi(p,\rho)$ is increasing in ρ for all p because $\Delta(p;R)$ is decreasing in ρ. Now, suppose $\rho_1 > \rho_2$ and $p\in\mathcal{P}(\rho_1;R)$. Then we have that $p\in\mathcal{P}(\rho_1;R) \Rightarrow p > \Pi(p,\rho_1) \ge \Pi(p,\rho_2) \Rightarrow p \in \mathcal{P}(\rho_2;R)$, where the inequality holds because $\Pi(p,\rho)$ is increasing in ρ. Hence, we have that $\mathcal{P}(\rho_1;R)\subseteq\mathcal{P}(\rho_2;R)$.

Appendix C. Proofs in §6

C.1. Proof of Theorem 4

We showed, in Theorem 2, that $p^k < \xi$ for large k for any $\xi > 0$ even we allowed for only single agent deviations. Thus, it is only necessary to argue that $p = 0$ is an equilibrium price. In fact, the proof of such a claim is the same as the proof of Theorem 2.2. The details can be seen in Allon et al. (2012). It is worth noting that we could only show the existence of equilibrium for $\rho < 1-\tilde{\delta}$ if $\lim_{k\to\infty}\delta^k = \tilde{\delta}$ for some $\tilde{\delta} > 0$.

C.2. Proof of Theorem 5

Before proving the theorem, we first state the following proposition. The proof of this proposition can be seen in Allon et al. (2012). This proposition simply proves that a group of agents can improve its revenues by increasing prices in a seller's market when the agents are allowed to deviate together.

Proposition 5. *In a seller's market* ($\rho > 1$), *we have that* $\liminf_{k\to\infty} D_1^{\mathrm{MCE}}(k)/\delta^k \geq 1/\rho$, *where* $D_n^{\mathrm{MCE}}(k) = D_n^{\mathrm{MCE}}(\hat{p}^k, p^k; \lfloor\delta^k k\rfloor, k - \lfloor\delta^k k\rfloor)$, $\lim_{k\to\infty} p^k = p < R$, $\lim_{k\to\infty} \hat{p}^k = p'$, *and* $p < p' < \min\{R, p + (1-(1/\rho))(R - p + cm_a)\}$. *Furthermore, we have that* $\lim_{k\to\infty} V_1(D_1^{\mathrm{MCE}}(k), D_2^{\mathrm{MCE}}(k); \hat{p}^k, p^k; \lfloor\delta^k k\rfloor, k - \lfloor\delta^k k\rfloor) = p'$.

To prove our claim on the equilibrium prices, it is sufficient to show that $\liminf_{k\to\infty} p_{EQ}^k = R$. We prove this by supposing on the contrary that the result does not hold. Then we can find a convergent subsequence of equilibrium prices p_{EQ}^k such that $\lim_{k\to\infty} p_{EQ}^k = p < R$. Using Proposition 5, agents can improve their utility by increasing their price to p' in a marketplace with sufficiently large number of agents. However, this contradicts that p_{EQ}^k is a (δ^k, ϵ^k)-Market Equilibrium. Hence, for any given $\xi > 0$, there should exist a large K such that $p_{EQ}^k > R - \xi$ for any $k > K$.

The proof of our claim on the customer equilibrium and the existence of the equilibrium sequence can be seen in Allon et al. (2012).

References

Allon G, Federgruen A (2007) Competition in service industries. *Oper. Res.* 55(1):37–55.

Allon G, Gurvich I (2010) Pricing and dimensioning competing large-scale service providers. *Manufacturing Service Oper. Management* 12(3):449–469.

Allon G, Bassamboo A, Çil E (2012) Supplementary material to "Large-scale service marketplaces: The role of the moderating firm." http://www.kellogg.northwestern.edu/research/operations/workingpapers.htm.

Amihud Y, Mendelson H (1980) Dealership market: Market-making with inventory. *J. Financial Econom.* 8(1):31–53.

Armony M, Maglaras C (2004) On customer contact centers with a call-back option: Customer decisions, routing rules and system design. *Oper. Res.* 52(2):271–292.

Armony M, Plambeck E, Seshadri S (2009) Sensitivity of optimal capacity to customer impatience in an unobservable M/M/S queue (Why you shouldn't shout at the DMV). *Manufacturing Service Oper. Management* 11(1):19–32.

Aumann RJ (1959) Acceptable points in general cooperative *n*-person games. Tucker AW, Luce RD, eds. *Contributions to the Theory of Games* (Princeton University Press, Princeton, NJ), 287–324.

Biais B, Glosten L, Spattt C (2005) Market microstructure: A survey of microfoundations, empirical results, and policy implications. *J. Financial Markets* 8(2):217–264.

Burdett K, Mortensen D (1998) Wage differentials, employer size, and unemployment. *Internat. Econom. Rev.* 39(2):257–273.

Cachon G, Harker PT (2002) Competition and outsourcing with scale economies. *Management Sci.* 48(10):1314–1333.

Cachon G, Netessine S (2004) Game theory in supply chain analysis. Simchi-Levi D, Wu SD, Shen ZM, eds. *Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era* (Kluwer Academic, Boston), 13–66.

Cachon G, Zhang F (2007) Obtaining fast service in a queueing system via performance-based allocation of demand. *Management Sci.* 53(3):408–420.

Chen H, Wan Y (2003) Price competition of make-to-order firms. *IIE Trans.* 35(9):817–832.

Chen Y, Maglaras C, Vulcano G (2008) Design of an aggregated marketplace under congestion effects: Asymptotic analysis and equilibrium characterization. Working paper, University of California, Berkeley, Berkeley. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1919161##.

Dixon H (1987) Approximate bertrand equilibria in a replicated industry. *Rev. Econom. Stud.* 54(1):47–62.

Economist, The (2010) Work in the digital age: A clouded future. *The Economist* (May 13), http://www.economist.com/node/16116919.

Flandez R (2008) Help wanted and found. *Wall Street Journal* (October 13), http://online.wsj.com/article/SB122347721312915407.html.

Garman M (1976) Market microstructure. *J. Financial Econom.* 3(3):257–275.

Garnett O, Mandelbaum A, Reiman M (2002) Desining a call center with impatient customers. *Manufacturing Service Oper. Management* 4(3):208–227.

Gradwohl R, Reingold O (2008) Fault tolerance in large games. *Proc. Ninth ACM Conf. Electronic Commerce* (ACM, New York), 274–283.

Granot D, Sošić G (2005) Formation of alliances in Internet-based supply exchanges. *Management Sci.* 51(1):92–105.

Ha AY, Li L, Ng S (2003) Price and delivery logistics competition in a supply chain. *Management Sci.* 49(9):1139–1153.

Halfin S, Whitt W (1981) Heavy-traffic limits for queues with many exponential servers. *Oper. Res.* 29(3):567–588.

Hassin R, Haviv M (2003) *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems* (Kluwer Academic, Boston).

Ho T, Stoll H (1983) The dynamics of dealer markets under competition. *J. Finance* 38(4):1053–1074.

Kalai E (2004) Large robust games. *Econometrica* 72(6):1631–1665.

Maglaras C, Zeevi A (2005) Pricing and design of differentiated services: Approximate analysis and structural insights. *Oper. Res.* 53(2):242–262.

Manning A (2003) The real thin theory: Monopsony in modern labour markets. *Labour Econom.* 10(2):105–131.

Manning A (2004) Monopsony and the efficiency of labour market interventions. *Labour Econom.* 11(2):145–163.

Michaelides M (2010) Labour market oligopsonistic competition: The effect of worker immobility on wages. *Labour Econom.* 17(1):230–239.

Myerson RB (1991) *Game Theory: Analysis of Conflict* (Harvard University Press, Cambridge, MA).

Nagarajan M, Sošić G (2007) Stable farsighted coalitions in competetive markets. *Management Sci.* 53(1):29–45.

Nagarajan M, Sošić G (2008) Game-theoretic analysis of cooperation among supply chain agents: Review and extensions. *Eur. J. Oper. Res.* 187(3):719–745.

Naor P (1969) The regulation of queue size by levying tolls. *Econometrica* 37(1):15–24.

Rath KP (1992) A direct proof of the existence of pure strategy equilibria in games with a continuum of players. *Econom. Theory* 2(3):427–433.

Ray I (1996) Coalition-proof correlated equilibirum: A definition. *Games Econom. Behav.* 17:56–79.

Roughgarden T (2005) *Selfish Routing and the Price of Anarchy* (MIT Press, Cambridge, MA).

Rudin W (1976) *Principles of Mathematical Analysis*, 3rd ed. (McGraw-Hill, New York).

Sošić G (2006) Transhipment of inventories among retailers: Myopic vs. farsighted staility. *Management Sci.* 52(10):1493–1508.

Ward AR, Glynn PW (2003) Diffusion approximation for a Markovian queue with reneging. *Queueing Systems* 43(1–2):103–128.

Zeltyn S, Mandelbaum A (2005) Call centers with impatient customers: Many-server asymptotics of the M/M/*n*+G queues. *Queueing Systems* 51(3–4):361–402.