European Journal of Operational Research 245 (2015) 470–479

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Production, Manufacturing and Logistics

Capacity reservation for time-sensitive service providers: An application in seaport management

L. Jeff Hong[a], Xiaowei Xu[b], Sheng Hao Zhang[c,*]

[a] Department of Economics and Finance and Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong, China

[b] Department of Supply Chain Management and Marketing Sciences, Rutgers, The State University of New Jersey, Newark, NJ 07102, USA

[c] Department of Industrial Engineering, School of Management, The Key Laboratory of the Ministry of Education for Process Control and Efficiency Engineering, Xi'an Jiaotong University, Xi'an 710049, China

A R T I C L E I N F O

Article history:
Received 26 May 2013
Accepted 20 March 2015
Available online 28 March 2015

Keywords:
Capacity management
Pooling
Seaport management
Time-sensitive demands
Frequency competition

A B S T R A C T

This paper analyzes a capacity management problem in which two service providers utilize a common facility to serve two separate markets with time-sensitive demands. The facility provider has a fixed capacity and all parties maximize demand rates. When the service providers share the facility, they play a frequency competition game with a unique Nash equilibrium. When the service providers have dedicated facilities, the facility provider leads two separate Stackelberg games. A centralized system with the first-best outcome is also examined. Based on closed-form solutions under all three scenarios, we find that facility capacity competition is a prerequisite condition for not pooling the service providers. Moreover, we establish the rankings of preferred strategies for all parties with respect to the ratio of the service providers' demand loss rates, which are proportional to the time sensitivity of demand and the potential market size. Interestingly a triple-agreement situation for the pooling strategy exists if the rates are close, and the facility provider permits a request for dedicated facilities only if the service provider has an overwhelming dominance at the demand loss rate. We connect these managerial insights with strategic seaport capacity management.

1. Introduction

Many service systems involve with multiple parties and increasing the service capacity of one service provider may not help improving the overall service system performance. For instance, a maritime system includes carriers and port authorities. Since a seaport has a limited capacity in processing vessels, a carrier may not be able to shorten the cargo delivery time when pushing the vessel frequency close to the port's handling capacity. Moreover, the bottleneck of public logistic facilities becomes severe when multiple service providers compete on a fixed amount of facility capacity. It is well known that a user of a public resource often ignores the negative externality that she/he imposes on other users (Hardin, 1968). This ignorance can cause congestion and massive losses in many logistic systems. For example, Ball, Barnhart, Dresner, Hansen, Neels, et al. (2010) estimate that the total cost of US domestic air traffic delays is around $31.2 billion for calendar year 2007. One way to solve this issue is to use incentive-compatible pricing schemes (see Ha, 1998; Mendelson & Whang, 1990), which have been widely adopted by public transportation authorities. Another way is to allocate dedicated facilities to certain types of users, which is commonly practiced by port authorities and is the focus of this paper.

A strategic problem for a port authority is to decide whether to pool all carriers together to share the port facilities or to allocate dedicated facilities to individual carriers. When a port pools the vessels from all carriers together and fully utilizes its facilities, this pooling effect generally leads to more efficient usage of the facilities. However, the pooling strategy is not perfect for the port. When carriers are put together, they may compete for the port facilities by increasing the vessel frequency in order to provide better service for their customers. This competition effect may result in congestion and offset the benefit of the pooling effect. Since using dedicated facilities separates the operations of different carriers, this reservation strategy eliminates the competition effect as well as the pooling effect. From carriers' perspective, a busy port may cause long and unpredictable time delays, which often cause a loss to carriers as their customers are usually sensitive to the time spent on the transportation route. To reduce the time delay and avoid competition with other carriers for port facilities, carriers are inclined to having dedicated port facilities. Hence, it is important for both a port authority and carriers to

* Corresponding author. Tel.: +86 15829361566.
E-mail addresses: jeffhong@cityu.edu.hk (L. Jeff Hong), xiaoweix@andromeda.rutgers.edu (X. Xu), shhzhang@mail.xjtu.edu.cn (S.H. Zhang).

http://dx.doi.org/10.1016/j.ejor.2015.03.027
0377-2217/

understand the tradeoff between pooling and reservation strategies and the interactions among their capacity decisions.

This paper studies a three-tier model, where a facility provider (a destination seaport) offers its facility to two service providers (carriers), who ship customers' cargos from two different origin ports to the same destination port. We assume that customer demand rate on each route decreases linearly in the total transportation time spent at the origin and destination ports and all parties maximize their cargo volumes. Three scenarios are considered. In the first scenario, the facility provider adopts the pooling strategy. The service providers determine their service capacity levels and compete for facility usage. We find a unique Nash equilibrium for this scenario. In the second scenario, the facility provider allocates facility capacity to each service provider, who determines the service capacity given its dedicated facility capacity. Finally, in the third scenario, we study the first-best outcome of a centralized system, in which a central planner jointly chooses the facility capacity management strategy and the service capacity levels. Based on closed-form solutions under all three scenarios, our study identifies conditions under which the facility provider and service providers should adopt the pooling or reservation strategy.

Our work contributes to the literature on capacity pooling and reservation strategies, which will be reviewed in the next section, in the following aspects.

First, we assume that both the facility provider and service providers maximize their demand rates, as cargo volume is one of the most important performance measures in the maritime industry (see Stopford, 2009; World Bank, 2007, p. 85). A port authority run by a local government weights much more on the economic contribution of cargo traffic to the local economy than its own profitability. A long-distance oversea shipper needs to defense its market share when its clients have an expensive alternative of air shipping. This distinguishes our model from many works on time-sensitive demands, in which pricing is often the central concern. Moreover, we pay attention to the case where two service providers face separate markets. Hence our model avoids the complexity introduced by market competition between two service providers, which is often the main theme of literature on time-based competition. By focusing on capacity management from an operational perspective, we find that pooling is always optimal under the centralized system, which suggests that facility capacity competition is a prerequisite condition for not pooling the service providers.

Second, our three-tier model allows time-sensitive customer demands depending on the transportation time spent at both the origin and destination ports. Notice that increasing the shipping frequency on a route decreases the time that cargo spends at the origin port but increases the time that vessels spend at the destination port of the route. When sharing the common facility capacity, self-interested service providers ignore the negative externality of their frequency decisions on others and cause facility over-utilization. Essentially our model under the pooling strategy examines a frequency competition between two service providers on the common facility capacity and hence contributes to the literature on frequency competition.

Third, we find that the facility provider's optimal choice between the pooling and reservation strategies critically depends on the ratio of the demand loss rates of two service providers. The demand loss rate is proportional to the potential cargo volume and the time sensitivity of demand on a route. Our result complements observations in the queueing literature that pooling is not optimal if customer characteristics, such as service time distributions and time sensitivity, are significantly different. Furthermore, we show that dedicated facilities are not always preferred by service providers and their optimal choices are also determined by the ratio of the demand loss rates of two service providers. This view is missing in the queueing literature, which only concern about the optimal choice of the facility provider.

Finally, in reality, the allocation of the facility provider's capacity is often done with service providers through tough negotiation processes, which may involve many other economy factors, for example, port charges, long-term relationship, etc. No matter how complex these processes are, all players have to understand the tradeoff between pooling and reservation strategies from an operational perspective, which is exactly the focus of our work. The managerial insights developed in this paper, e.g. the rankings of their preferred strategies and the existence of the triple-agreement situation, help all players to understand the interactions among their capacity decisions and lay down a sound foundation upon which to incorporate other factors in the tradeoff between pooling and reservation strategies.

The rest of the paper is organized as follows. In Section 2, we provide a brief review of related literature. The model is introduced in Section 3. Then, we study the pooling strategy, the reservation strategy and the centralized system in Section 4. We make comparisons between the pooling and reservation strategies in Section 5 and draw conclusions in Section 6. All proofs are relegated to the online supplements of the Appendix.

2. Literature review

The tradeoffs between capacity pooling and reservation strategies have been studied from many different perspectives. We briefly review the literature from four aspects below.

2.1. Queueing systems

It is well know that combining separate subsystems into one system may improve the overall system efficiency, since the combination reduces the chance of idleness of subsystems and generates economies of scale. However, if customers have heterogenous characteristics, then merging queues may be counterproductive. Smith and Whitt (1981) and Whitt (1999) show that if customers fall into classes with different service time distributions, then keeping different types of customers into separate queues may be optimal. Yu, Benjaafar, and Gerchak (2015) study a capacity sharing problem among a set of independent queues. They find that capacity pooling may not be optimal if the workloads of queues are significantly different. Rothkopf and Rech (1987) provide other reasons of not merging queues. van Dijk and van der Sluis (2009) propose rules to further reduce average waiting time under both pooled and unpooled scenarios for two customer groups with different service time distributions.

The preferred choice between pooling and reservation strategies highly depends on the congestion caused by negative externalities that a user imposes on other users in queueing systems. Haviv and Ritov (1998) derive measures of such negative externalities under different queue disciplines. Osorio and Bierlaire (2009) explain the propagation of congestion. Mendelson and Whang (1990) and Ha (1998) develop incentive-compatible pricing schemes to regulate the negative externality effects. Our model demonstrates under what market conditions the pooling benefit dominates (is dominated by) the negative impact of facility capacity competition for both the facility provider and service providers.

2.2. Time-sensitive demands

When customer utility or demand is time sensitive, capacity pooling and reservation strategies can serve as market segmentation tools. For instance, Pangburn and Stavrulaki (2008) study a joint pricing and capacity management problem and find that capacity pooling is suboptimal if customers are heterogenous in their time sensitivity. Our model reveals that another customer characteristic, the potential market size, also affects the pooling decision.

However, most studies consider profit-maximizing problems with pooled service capacity under various settings. Since we focus on capacity management from an operational perspective, we only review a few studies and refer to them for a more comprehensive review.

Boyaci and Ray (2003) study a product differentiation problem in which a firm determines the prices of a regular and an express product and the delivery time of the express product. They examine the relationship among capacity cost, time differentiation and price differentiation by assuming a linear demand model on price and guaranteed delivery time. Ray and Jewkes (2004) consider a linear demand model on price and lead time. They derive the profit maximizing optimal policy and present the conditions under which overlooking price and lead time dependence will lead to a sub-optimal decision. Sinha, Rangaraj, and Hemachandra (2010) study a surplus capacity pricing problem with two classes of customers, where secondary class customers's demand rate depends linearly on unit price and service level offered. They optimize unit admission price and quality of service (QoS) offered to secondary class customers while maintaining a prespecified QoS to primary class customers. Our model also assumes a linear demand function on the transportation time but maximizes customer demand rates instead.

2.3. Time-based competition

When managers' attention shifts from internal operations to outside competition, speed is also an important weapon for firms to gain market share (Stalk & Hout, 1990). Again most studies do not consider the option of splitting capacity. For example, Kalai, Kamien, and Rubinovitch (1992) consider a duopoly game in which two firms compete for market shares by choosing individual server capacity levels in a queueing system. Armony and Haviv (2003) investigate a game with two firms and two classes of customers, where firms decide their service capacity and price and customers' utility depends on price charged and expected queueing delays. They characterize properties of the equilibrium. So (2000) studies a price and delivery time competitive game in which firms satisfy customer demand within a guaranteed delivery period at a prefixed probability level. He finds that high capacity firms offer better time guarantees than do low capacity firms and an increase of time sensitivity in customer demand strengthens this differentiation. Hassin and Haviv (2003) present a detailed review on time-based competition models and Allon and Federgruen (2008) provide more updated references.

There are only a few papers on capacity pooling and splitting strategies. Motivated by Internet access service, Mandjes and Timmer (2007) consider a game similar to Armony and Haviv (2003). But they allow firms to split their service capacity and customers' utility depends on utilization of the service resources, rather than queueing delay. They find that it may pay off for both firms to split their service capacities, even if the firms can only choose between offering either a single network or two networks of equal capacity. Unlike the previous literature, we study a three-tier model, where two service providers compete on the fixed facility capacity and customer demands are sensitive to the total transportation time. We find that the pooling decision of the facility provider depends on not only customer characteristics but also whether or not the service providers compete for facility usage. We also derive the preferred choice between two strategies from the service providers' perspective.

2.4. Frequency competition

In transportation industries, expanding capacity by providing more frequency on a route not only shortens the waiting time but also offers more departure options for customers. The importance of service frequency has long been recognized by airline companies in gaining market share. Adler (2001) adopts a logit model to quantify the impact of frequencies on market share and studies airline competition on fares, frequencies, and aircraft sizes. He derives equilibrium results for a network comprising four airports and two airlines. Vaze and Barnhart (2012) model the connection between the market share and the frequency share by the so-called *S*-curve or sigmoidal relationship and propose a game-theoretic model for airline frequency competition under slot constraints. They demonstrate that a small reduction in the total number of allocated slots translates into a substantial reduction in flight and passenger delays and also a considerable improvement in airlines' profits. Surprisingly few researchers consider how to optimize the service frequency in maritime industry. Meng and Wang (2011) comment that *"With regard to the service frequency, researchers either consider no requirement on the frequency, or impose a minimum number of services within a planning horizon, or require a fixed weekly service frequency. In other words, they have not sought to optimize the service frequency."* Meng and Wang (2011) optimize service frequency, containership fleet deployment plan, and sailing speed for a long-haul liner service route. Apparently, frequency competition is beyond the scope of their study.

Because the focus of our paper is on the tradeoff between capacity pooling and reservation strategies, we do not model service providers' competition in the transportation market but study the impact of their competition on utilizing a common facility. A feature of our model is that providing more frequency by a service provider will shorten the cargo waiting time at the origin port but lengthen the vessel waiting and docking time at the destination port. We investigate the impact of both service and public facility capacity levels on the total transportation time and specify the preference over the pooling and reservation strategies for both service providers and the facility authority.

3. The model

We consider a facility provider that offers its facilities to service providers. For instance, a port provides berths to carriers, who ship cargos to and from the port. The facility provider has a total facility capacity of K, which is measured as the number of vessels that the facility can handle per unit time and is mainly determined by the facility infrastructure (e.g. the number of berths). Throughout the paper, we assume that the total facility capacity is fixed, since it is often very time-consuming and/or expensive to change the facility infrastructure.

For simplicity, we consider only two service providers. This is sufficient to demonstrate the tradeoff between the capacity pooling and reservation strategies. We assume that the service providers face separate markets that do not interact with each other. For instance, carriers may serve different routes, which have different origins but share the same destination port (see Fig. 1). In this case, the demands for the carrier's services rarely affect each other. We denote the capacity level of service provider i as μ_i, where $i = 1, 2$, which is measured as the frequency of vessels traveling on the route. We let λ_i be the demand rate for service provider i, which is measured as the cargo volume per unit time (or equivalently, the number of vessels needed to ship the cargo per unit time), and assume that the demand rate strictly decreases in the total transportation time of service i, where $i = 1, 2$.

Since the capacity allocation strategy of the facility provider is a long-term decision, we assume the sequence of events as follows. The facility provider first announces its choice between pooling two service providers together to share the total facility capacity K and reserving dedicated facility capacity K_i to service provider i, where $i = 1, 2$. In the latter option, $K_1 + K_2 = K$ holds obviously. Next, service providers determine their capacity levels.

The total transportation time t_i of cargos on route i can be decomposed into three phases: the dwell time $t_{d,\,i}$ that the cargos spend at the origin port, the shipping time $t_{s,\,i}$ on the ocean, and the facility time $t_{f,\,i}$ that a vessel spends on waiting for and using port facilities at the destination port. Notice that the shipping phase is independent of the service and facility capacity levels.

In the dwell phase, only when both the facility and vessels of service providers are available at the origin port, cargos can be loaded

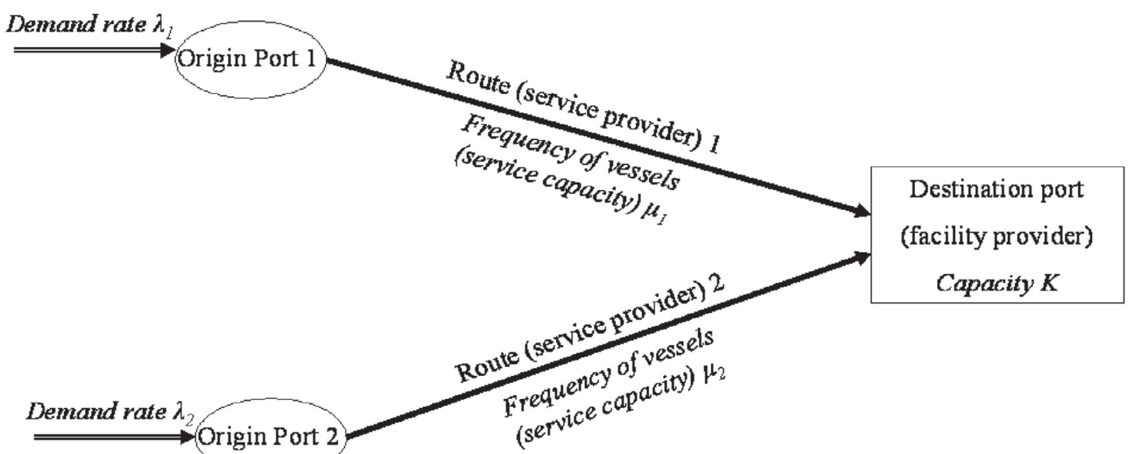

Fig. 1. The maritime system.

and shipped out. Therefore, the major part of the dwell time is the time that cargos spend on waiting for the next available vessels. To focus on the interaction between the destination port and carriers' capacity decisions, we assume that the vessel and cargo processing time at the origin port is negligible compared to the cargo waiting time. This is likely to occur if the origin port is not a busy one. In maritime industry, arrivals and departures of vessels seldom follow the shipping schedule punctually. Our investigation on the vessel arrival and departure pattern at Kwai Tsing Container Terminals of Hong Kong (please refer to the Appendix for details) indicates that vessel arrivals should be modeled via a stochastic process. Hence we model the cargo waiting process as an M/M/1 queue and the dwell time as $t_{d,i} = \frac{1}{\mu_i - \lambda_i}$ on route i.

In the facility phase, the vessel first waits on the sea out of the port and then is permitted to dock and handle cargos until a berth is available in the port. For a busy port, waiting and docking typically take much longer than handling cargos at the berth. For instance, the average containership waiting time at the Port of Cartagena in Colombia is about 2 hours when the berth occupancy rate is 50 percent, but the time jumps to 10 days when the berth occupancy rate increases to 90 percent (World Bank, 2007, p. 2). Hence we assume that the cargo processing time is negligible at the destination port and model the vessel waiting and docking process as an M/M/1 queue. If service providers share the facility, the facility time is $t_{f,1} = t_{f,2} = \frac{1}{K-\mu_1-\mu_2}$. If service provider i has dedicated facility capacity K_i, the facility time is $t_{f,i} = \frac{1}{K_i-\mu_i}$. Our investigation at Kwai Tsing Container Terminals of Hong Kong also provides some empirical evidence to support such formulation.

We assume a linear time-dependent demand function $\lambda_i(t_i) = a_i - \eta_i t_i$, where $i = 1, 2$. The linear demand form is widely adopted in the literature (see Kalai et al., 1992; So, 2000). Since $t_i = t_{d,i} + t_{f,i} + t_{s,i}$, we can rewrite $\lambda_i(t_i) = A_i[1 - \theta_i(t_{d,i} + t_{f,i})]$, where $A_i = a_i - \eta_i t_{s,i}$ and $\theta_i = \frac{\eta_i}{a_i - \eta_i t_{s,i}}$. Because $t_{s,i}$ is independent of the service and facility capacity levels, A_i and θ_i are constant. We interpret A_i as the potential market size (after adjusting the shipping time) and θ_i as the time sensitivity of demand.

Due to the large capital investment needed to build ports, the total cargo volume is a primary concern of ports. Moreover, to a port authority, the overall contribution of the port operations to the local economy, which is correlated with the cargo volume, is often more important than its own profitability. Hence, we assume that the facility provider maximizes its demand rate. To skip the carrier's marketing decisions but concentrate on its operations side, we assume that the objective of each service provider is also maximizing cargo volume, which is equivalent to maximizing its profit with a constant margin. Cost-plus pricing is a common pricing procedure and has widely adopted in many industries (Nagle, Hogan, & Zale, 2011).

With slight abuse of notation, we let $K_i = K - \mu_{3-i}$ if the facility provider adopts the pooling strategy. This enables us to write the facility time as $t_{f,i} = \frac{1}{K_i-\mu_i}$ and unify many derivations under both strategies. The demand rate λ_i, as a function of K_i and μ_i, is

$$\lambda_i(K_i, \mu_i) = \frac{1}{2}\left[A_i - \frac{A_i\theta_i}{K_i - \mu_i} + \mu_i - \sqrt{\delta_i^2 + 4A_i\theta_i}\right], \tag{1}$$

where $\delta_i = A_i - \frac{A_i\theta_i}{K_i-\mu_i} - \mu_i$, if $\frac{1}{2}(K_i - \sqrt{K_i^2 - 4K_i\theta_i}) < \mu_i < \frac{1}{2}(K_i + \sqrt{K_i^2 - 4K_i\theta_i})$ and $K_i > 4\theta_i$; otherwise, $\lambda_i(K_i, \mu_i) = 0$. (See derivation details in the Appendix.)

We solve the service provider's demand maximization problem as shown below.

Proposition 1. *Assume that $K_i > 4\theta_i$. (1) The optimal service capacity is $\mu_i^*(K_i) = \frac{3K_i+A_i}{4} - \frac{1}{4}\sqrt{(K_i - A_i)^2 + 16A_i\theta_i}$. (2) The optimal demand rate is $\lambda_i^*(K_i) = \lambda_i(K_i, \mu_i^*(K_i)) = \frac{K_i+A_i}{2} - \frac{1}{2}\sqrt{(K_i - A_i)^2 + 16A_i\theta_i}$. (3) $\mu_i^*(K_i) > \lambda_i^*(K_i) > 0$. (4) The optimal service capacity $\mu_i^*(K_i)$ and the demand rate $\lambda_i^*(K_i)$ are increasing in K_i and A_i, but decreasing in θ_i.*

By Proposition 1, as the facility capacity K_i and/or the market size A_i increase, service provider i increases its service capacity to attract customer demand. But as customers become more time sensitive (i.e. an increase of θ_i), the demand for the service drops, which reduces the required service capacity. The condition of $K_i > 4\theta_i$ in Proposition 1 implies that customers are not extremely time sensitive and/or the facility capacity is moderately large. An invalidation of this condition shuts down the operations of service provider i (i.e. $\mu_i^*(K_i) = \lambda_i^*(K_i) = 0$). Since this is a trivial case, we will avoid it in the rest of this paper.

4. Three scenarios

In this section, we solve the optimal capacity decisions of involved parties under three scenarios: the pooling strategy, the reservation strategy and the centralized system.

4.1. The pooling strategy

When the facility provider adopts the pooling strategy, the service providers determine their individual service capacity levels and compete on the facility capacity. Since $K_i = K - \mu_{3-i}$ depends on the service capacity of the other service provider (i.e. μ_{3-i}), each service provider's capacity decision affects the other's capacity decision. We model this as a simultaneous game and study the pure strategy Nash equilibrium of the game. Notice that $\mu_1 + \mu_2 < K$ at a Nash equilibrium since overloading the facility implies infinity processing time and kills the demand for each service provider.

Let $\mu_i^{\text{POOL}*}$ and $\lambda_i^{\text{POOL}*}$ denote the equilibrium service capacity level and demand rate of service provider i, respectively, and $\Lambda^{\text{POOL}*} = \lambda_1^{\text{POOL}*} + \lambda_2^{\text{POOL}*}$ denote the total demand rate of the facility provider under the equilibrium. To ensure that both service providers use the facility, we make the following assumption.

Assumption 1. $K \geq \max(A_1, A_2) + 8\max(\theta_1, \theta_2)$.

By Claim 1 of Proposition 1 and Assumption 1,

$$K - \mu_{3-i}^*(K) = \frac{K - A_{3-i}}{4} + \frac{1}{4}\sqrt{(K - A_{3-i})^2 + 16A_{3-i}\theta_{3-i}} > (K - A_{3-i})/2 > 4\theta_i$$

for $i = 1, 2$. By Claim 4 of Proposition 1, $\mu_i^*(K_i) \leq \mu_i^*(K)$ for any $K_i \leq K$, where $K_i = K - \mu_{3-i}$. Therefore, we have $K - \mu_{3-i}^*(K_{3-i}) > 4\theta_i$ for any $K_{3-i} \leq K$ and the condition of Proposition 1 always holds at any equilibrium. By Claim 3 of Proposition 1, $\mu_i^{\text{POOL}*} > \lambda_i^{\text{POOL}*} > 0$ for $i = 1, 2$. Hence, Assumption 1 implies that both service providers use the facility under the pooling strategy.

By Claims 1 and 2 of Proposition 1, the equilibrium service capacity $\mu_i^{\text{POOL}*}$ and the demand rate $\lambda_i^{\text{POOL}*}$ must satisfy the following equations:

$$\mu_i^{\text{POOL}*} = \frac{3\left(K - \mu_{3-i}^{\text{POOL}*}\right) + A_i}{4} - \frac{1}{4}\sqrt{(K - \mu_{3-i}^{\text{POOL}*} - A_i)^2 + 16A_i\theta_i},$$

$$\lambda_i^{\text{POOL}*} = \frac{(K - \mu_{3-i}^{\text{POOL}*}) + A_i}{2} - \frac{1}{2}\sqrt{(K - \mu_{3-i}^{\text{POOL}*} - A_i)^2 + 16A_i\theta_i},$$

for both $i = 1, 2$. By solving the equations through various transformations, we obtain a unique pure strategy Nash equilibrium in Theorem 1.

Theorem 1. *There exists a unique pure strategy Nash equilibrium such that both service providers use the facility. At the equilibrium, the service capacity and demand rate of service provider i are*

$$\mu_i^{\text{POOL}*} = A_i - \frac{A_i\theta_i}{2(A_1\theta_1 + A_2\theta_2)}\left(\sqrt{M^2 + \Xi_{\text{POOL}}^2} + M\right) + \frac{1}{6}\left(\sqrt{M^2 + \Xi_{\text{POOL}}^2} - M\right) > 0,$$

$$\lambda_i^{\text{POOL}*} = A_i - \frac{A_i\theta_i}{2(A_1\theta_1 + A_2\theta_2)}\left(\sqrt{M^2 + \Xi_{\text{POOL}}^2} + M\right) > 0,$$

respectively, and the total demand rate of the facility provider is

$$\Lambda^{\text{POOL}*} = \frac{1}{2}\left[A_1 + A_2 + K - \sqrt{M^2 + \Xi_{\text{POOL}}^2}\,\right],$$

where $M = A_1 + A_2 - K$ *and* $\Xi_{\text{POOL}} = \sqrt{24(A_1\theta_1 + A_2\theta_2)}$.

With the closed-form equilibrium in Theorem 1, we derive the comparative statics of service capacity levels and demand rates as the market conditions and facility capacity change.

Proposition 2. *(1) Service provider i's equilibrium service capacity* $\mu_i^{\text{POOL}*}$ *and demand rate* $\lambda_i^{\text{POOL}*}$ *are increasing in K,* A_i *and* θ_{3-i}*, but decreasing in* A_{3-i} *and* θ_i*. (2) The facility provider's equilibrium total demand rate* $\Lambda^{\text{POOL}*}$ *is increasing in K,* A_1 *and* A_2*, but decreasing in* θ_1 *and* θ_2.

Proposition 2 shows that an increase of facility capacity allows the service providers to increase their service capacity levels. As a result, the demand rate of each service provider and the total demand rate increase. A service provider increases its service capacity to attract customers as its potential market size increases. This causes a drop in the facility capacity index for the other service provider and hence a reduction in its capacity and demand rate. However, the total demand rate increases. An increase of customers' time sensitivity drives down the customer demand rate and hence the required service capacity of service i. This increases the facility capacity index for the other service provider, who then increases its service capacity to attract more customers.

4.2. The reservation strategy

Applying reservation strategy eliminates not only the gaming behavior of the service providers but also the pooling effect. By Proposition 1, if $K_i > 4\theta_i$, then service provider i chooses service capacity $\mu_i^*(K_i)$, which brings in demand $\lambda_i^*(K_i)$ for the facility provider; if $K_i \leq 4\theta_i$, which implies that the dedicated facility capacity for service provider i is too small, then service provider i chooses to drop out of its market and not to use the facility. In the latter case, the dedicated facility generates no demand and is wasted. Hence, it is not optimal for the facility provider to allocate any dedicated capacity $K_i \in (0, 4\theta_i]$. Let Λ^{RES} denote the total demand rate for the facility under this scenario, we have

$$\Lambda^{\text{RES}}(K_1, K_2) = \begin{cases} \displaystyle\sum_{i=1}^{2} \frac{1}{2}(K_i + A_i) - \frac{1}{2}\sqrt{(K_i - A_i)^2 + 16A_i\theta_i}, \\ \text{if } K_i > 4\theta_i \text{ and } K_1 + K_2 = K; \\ \frac{1}{2}(K + A_i) - \frac{1}{2}\sqrt{(K - A_i)^2 + 16A_i\theta_i}, \\ \text{if } K_i = K \text{ and } K_{3-i} = 0. \end{cases}$$

Let $K_i^{\text{RES}*}$ denote the optimal dedicated facility capacity for service provider i. With the dedicated capacity $K_i^{\text{RES}*}$, we let $\mu_i^{\text{RES}*}$ and $\lambda_i^{\text{RES}*}$ denote the optimal service capacity and demand rate of service provider i, respectively. The optimal total demand rate of the facility provider is $\Lambda^{\text{RES}*} = \Lambda^{\text{RES}}(K_1^{\text{RES}*}, K_2^{\text{RES}*})$. We make the following assumption to avoid the trivial case that the facility provider causes one service provider to drop out.

Assumption 2. $A_i \geq 16\theta_i$, where $i = 1, 2$.

This assumption implies that customers are not extremely time sensitive and/or the potential market size is large.

Theorem 2. *The optimal dedicated facility capacity, the service capacity and the demand rate of service provider i are*

$$K_i^{\text{RES}*} = A_i - \frac{\sqrt{A_i\theta_i}}{\sqrt{A_1\theta_1} + \sqrt{A_2\theta_2}}(A_1 + A_2 - K) > 0,$$

$$\mu_i^{\text{RES}*} = A_i - \frac{\sqrt{A_i\theta_i}}{4(\sqrt{A_1\theta_1} + \sqrt{A_2\theta_2})}\left(\sqrt{M^2 + \Xi_{\text{RES}}^2} + 3M\right) > 0,$$

$$\lambda_i^{\text{RES}*} = A_i - \frac{\sqrt{A_i\theta_i}}{2(\sqrt{A_1\theta_1} + \sqrt{A_2\theta_2})}\left(\sqrt{M^2 + \Xi_{\text{RES}}^2} + M\right) > 0,$$

respectively, and the total demand rate of the facility provider is

$$\Lambda^{\text{RES}*} = \frac{1}{2}[A_1 + A_2 + K - \sqrt{M^2 + \Xi_{\text{RES}}^2}],$$

where $\Xi_{\text{RES}} = 4(\sqrt{A_1\theta_1} + \sqrt{A_2\theta_2})$.

With the closed-form solutions in Theorem 2, we derive the comparative statics of the service capacity levels and demand rates as the market conditions and facility capacity change.

Proposition 3. *(1) The optimal dedicated facility capacity* $K_i^{\text{RES}*}$ *for service provider i, its service capacity* $\mu_i^{\text{RES}*}$ *and demand rate* $\lambda_i^{\text{RES}*}$ *are increasing in K and* A_i*, but decreasing in* A_{3-i}*. (2) If* $A_1 + A_2 \geq K$*, then the optimal dedicated facility capacity* $K_i^{\text{RES}*}$ *for service provider i, its* *service capacity $\mu_i^{\mathrm{RES}*}$ and demand rate $\lambda_i^{\mathrm{RES}*}$ are increasing in θ_{3-i}, but decreasing in θ_i. (3) If $A_1 + A_2 < K$, then the optimal service capacity $\mu_i^{\mathrm{RES}*}$ of service provider i and its demand rate $\lambda_i^{\mathrm{RES}*}$ are decreasing in θ_i and θ_{3-i}, but the optimal dedicated facility capacity $K_i^{\mathrm{RES}*}$ is increasing (decreasing) in θ_i (θ_{3-i}). (4) The facility provider's total demand rate $\Lambda^{\mathrm{RES}*}$ is increasing in K, A_1 and A_2, but decreasing in θ_1 and θ_2.*

Most comparative statics in Proposition 3 are parallel to the ones in Proposition 2, except those with respect to the time sensitivity parameters. Notice that $A_1 + A_2$ represents the overall potential market size, and $A_1 + A_2 > K$ ($A_1 + A_2 < K$) indicates that the total facility capacity is (not) tight. When the facility capacity is tight, as the time sensitivity θ_{3-i} increases, service provider i's market becomes relatively more attractive. Hence, the facility provider should allocate more capacity to service provider i and decrease the capacity for the other one. As more facility capacity becomes available, service provider i increases its service capacity to attract more customer demand. In contrast, the other one cuts its service capacity and loses demand, as its dedicated facility capacity is reduced.

When the facility capacity is not tight, as the time sensitivity θ_{3-i} increases, the facility provider adopts a very different strategy. That is, she moves facility capacity from service provider i to the other and tries to keep the time-sensitive customers of the latter one, even though this causes service provider i to cut its service capacity and lose demand.

4.3. The centralized system

We consider a centralized system in which a central planner determines not only the facility operations strategy but also the capacity levels of two service providers to maximize the total demand rate. The central planner can operate the facility with two strategies: (1) By the pooling strategy, the central planner allows both service providers to share the entire facility. (2) By the reservation strategy, the central planner allocates facility capacity K_i to serve only service provider i, where $K_1 + K_2 = K$.

Because the central planner determines the service capacity levels, this eliminates the competition effect and the negative externality of the service providers' gaming behavior demonstrated in Subsection 4.1. Hence, the pooling strategy is expected to dominate the reservation strategy. We establish this result in the following theorem.

Theorem 3. *It is always optimal to adopt the pooling strategy under the centralized system.*

By Theorem 3, we focus on the pooling strategy. Similarly, let $\mu_i^{\mathrm{CEN}*}$ and $\lambda_i^{\mathrm{CEN}*}$ denote the optimal service capacity level and demand rate of service provider i, respectively, and $\Lambda^{\mathrm{CEN}*} = \lambda_1^{\mathrm{CEN}*} + \lambda_2^{\mathrm{CEN}*}$ denote the optimal total demand rate under the centralized system.

Theorem 4. *The optimal service capacity and demand rate for service provider i under the centralized system are*

$$\mu_i^{\mathrm{CEN}*} = A_i - \frac{1}{\Xi_{\mathrm{CEN}}}\left[\frac{A_i\theta_i}{\sqrt{A_1\theta_1 + A_2\theta_2}}\left(M + \sqrt{M^2 + \Xi_{\mathrm{CEN}}^2}\right) + 2\sqrt{A_i\theta_i}\,M\right] > 0,$$

$$\lambda_i^{\mathrm{CEN}*} = A_i - \frac{1}{\Xi_{\mathrm{CEN}}}\left(\frac{A_i\theta_i}{\sqrt{A_1\theta_1 + A_2\theta_2}} + \sqrt{A_i\theta_i}\right)\left(M + \sqrt{M^2 + \Xi_{\mathrm{CEN}}^2}\right) > 0,$$

respectively, and the optimal total demand rate of the facility provider is

$$\Lambda^{\mathrm{CEN}*} = \frac{1}{2}\left[A_1 + A_2 + K - \sqrt{M^2 + \Xi_{\mathrm{CEN}}^2}\right],$$

where $\Xi_{\mathrm{CEN}} = 2(\sqrt{A_1\theta_1 + A_2\theta_2} + \sqrt{A_1\theta_1} + \sqrt{A_2\theta_2})$.

With the closed-form solutions in Theorem 4, we derive the comparative statics of the service capacity levels and demand rates as the market conditions and facility capacity change.

Proposition 4. *(1) Service provider i's service capacity $\mu_i^{\mathrm{CEN}*}$ and demand rate $\lambda_i^{\mathrm{CEN}*}$ are increasing in K. (2) The facility provider's total demand rate $\Lambda^{\mathrm{CEN}*}$ is increasing in K, A_1 and A_2, but decreasing in θ_1 and θ_2.*

By Proposition 4, an increase in the facility capacity allows the central planner to increase the service capacity levels and demand rates. Although the comparative statics of the total demand rate $\Lambda^{\mathrm{CEN}*}$ with respect to the market conditions (A_i and θ_i) are parallel to the ones in Propositions 2 and 3, the optimal service capacity $\mu_i^{\mathrm{CEN}*}$ and individual demand rate $\lambda_i^{\mathrm{CEN}*}$ may not be monotonic as the market conditions change. This is demonstrated in Example 1.

Example 1. We let $K = 650$, $A_1 = A_2 = 300$ and $\theta_1 = 5$ and vary $\theta_2 \in (0, 5]$. Fig. 2 shows the optimal service capacity $\mu_i^{\mathrm{CEN}*}$, the individual demand rate $\lambda_i^{\mathrm{CEN}*}$ and the total demand rate $\Lambda^{\mathrm{CEN}*}$.

As illustrated in Fig. 2, when customers on the route 2 become more time-sensitive, the central planner initially cuts the dwell time at the origin port by increasing service capacity $\mu_2^{\mathrm{CEN}*}$, and shortens the facility time at the destination port by decreasing service capacity $\mu_1^{\mathrm{CEN}*}$, which reduces the demand on the route 1 slightly. But the eventual demand decrease on the route 2 lowers required service capacity $\mu_2^{\mathrm{CEN}*}$, and drives up service capacity $\mu_1^{\mathrm{CEN}*}$ and demand rate $\lambda_1^{\mathrm{CEN}*}$. The total demand rate $\Lambda^{\mathrm{CEN}*}$ is decreasing as the market conditions of service 2 deteriorate, which is consistent with Proposition 4.

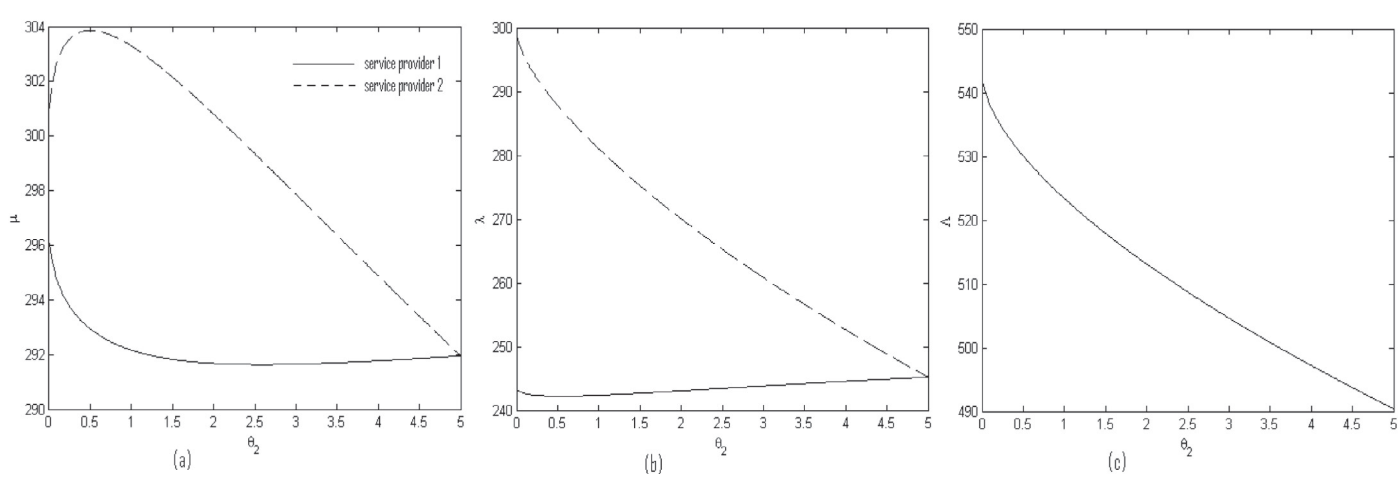

Fig. 2. (a)Optimal service capacity levels $\mu_1^{\mathrm{CEN}*}$ and $\mu_2^{\mathrm{CEN}*}$; (b) optimal demand rates $\lambda_1^{\mathrm{CEN}*}$ and $\lambda_2^{\mathrm{CEN}*}$; (c) optimal total demand rate $\Lambda^{\mathrm{CEN}*}$.

5. Comparisons and managerial insights

In this section, we compare the three scenarios studied in Section 4. We let $\beta_i = A_i\theta_i$, which represents the demand loss of service provider i if the total transportation time of service i increases by one unit. In the example of port operations, a carrier with a large market size often has a large value of β_i. Furthermore, we define $\gamma_i = \beta_i/\beta_{3-i}$ as the ratio of the two demand loss rates. Notice that $\gamma_1 = 1/\gamma_2$.

5.1. The facility provider's preference

First, we study the facility provider's preference for the pooling and reservation strategies.

Theorem 5. *There are two thresholds for γ_1 such that $\Lambda^{\text{POOL}*} > \Lambda^{\text{RES}*}$ if $7-4\sqrt{3} < \gamma_1 < 7+4\sqrt{3}$ and $\Lambda^{\text{POOL}*} < \Lambda^{\text{RES}*}$ if $\gamma_1 < 7-4\sqrt{3}$ or $\gamma_1 > 7+4\sqrt{3}$.*

From the facility provider's viewpoint, using capacity reservation eliminates the competition effect between the service providers, but it also erases the pooling effect. If the negative externality of the service providers' gaming behavior offsets the pooling benefit, the facility provider should adopt the reservation strategy. Theorem 5 gives a clear criterion on this tradeoff, which solely depends on the demand loss ratio γ_1. It is better for the facility provider to adopt the reservation strategy if and only if the demand loss rates of the two service providers are significantly different, i.e. $\gamma_1 < 7-4\sqrt{3} \approx 1/14$ or $\gamma_1 > 7+4\sqrt{3} \approx 14$.

Notice that the demand loss rate is proportional to the potential cargo volume and the time sensitivity of demand on a route. When the potential market sizes on two routes are close, the reservation strategy outperforms the pooling strategy if the service providers have significantly different time sensitivities of demand. As mentioned in Section 2, many works on when to pool separate subsystems together have found that pooling is not optimal if customer characteristics, e.g. service time distributions and time sensitivity, are significantly different. This observation is confirmed in our model.

When the time sensitivities of demand on two routes are similar, one service provider has to be dominant so that the facility provider adopts the reservation strategy. This explains the real practice in port operations: as reported by the Port Reform Toolkit (World Bank, 2007), ports dominated by one carrier often provide dedicated facilities to the dominant carrier, but ports that serve equally sized carriers are unwilling to provide dedicated facilities. For instance, the Maersk Line accounts for almost 80–90 percent of the traffic at the Port of Salalah, and it is much larger than other carriers using the port. Hence, it is not surprising that the port provides the Maersk Line with dedicated facilities. However, at the ports of Shanghai, Singapore and Shenzhen, which are the three largest ports in the world now (Hong Kong Marine Department, 2014), none of the carriers has such a clear size dominance over other carriers. Therefore, these ports do not provide dedicated facilities (World Bank, 2007, p. 86). Our model reveals a new customer characteristic, i.e. potential market size, to pay attention to in the tradeoff between capacity pooling and reservation strategies.

To generate more managerial insights, we compare the pooling and reservation strategies with the centralized system below.

Theorem 6. *(1) $\Lambda^{\text{CEN}*} > \Lambda^{\text{POOL}*}$ and $\Lambda^{\text{CEN}*} > \Lambda^{\text{RES}*}$. (2) $\mu^{\text{POOL}*} > \mu^{\text{CEN}*} > \mu^{\text{RES}*}$, where $\mu^{\text{POOL}*} = \mu_1^{\text{POOL}*} + \mu_2^{\text{POOL}*}$, $\mu^{\text{RES}*} = \mu_1^{\text{RES}*} + \mu_2^{\text{RES}*}$ and $\mu^{\text{CEN}*} = \mu_1^{\text{CEN}*} + \mu_2^{\text{CEN}*}$.*

Claim 1 of Theorem 6 confirms that the centralized system indeed achieves the first-best outcome for the overall cargo volume as it takes advantage of the pooling effect while eliminates the competition effect between the service providers. Recall that pooling is always optimal under the centralized system. Hence, facility capacity competition is a prerequisite condition for not pooling the service providers together.

Claim 2 of Theorem 6 suggests that the facility utilization rate is lowest under the reservation strategy. Pooling the service providers together leads to a more efficient usage of the facility capacity, which stimulates the total traffic amount, i.e. $\mu^{\text{POOL}*} > \mu^{\text{RES}*}$ and $\mu^{\text{CEN}*} > \mu^{\text{RES}*}$. However, when there is no authority which regulates the negative externality of the service providers' gaming behavior, pooling self-interested service providers also introduces facility capacity competition, which ends with an overused facility, i.e. $\mu^{\text{POOL}*} > \mu^{\text{CEN}*}$. In other words, the pooling effect can be measured by performance differences between the centralized system and the reservation strategy, i.e. $\Lambda^{\text{CEN}*} - \Lambda^{\text{RES}*}$ and $\mu^{\text{CEN}*} - \mu^{\text{RES}*}$; while the competition effect can be measured by performance differences between the centralized system and the pooling strategy, i.e. $\Lambda^{\text{CEN}*} - \Lambda^{\text{POOL}*}$ and $\mu^{\text{POOL}*} > \mu^{\text{CEN}*}$.

5.2. The service providers' preference

Next we examine the service providers' preference for the pooling and reservation strategies.

Theorem 7. *There is a threshold $\overline{\gamma}_i \in (1, 4)$ such that $\lambda_i^{\text{POOL}*} > \lambda_i^{\text{RES}*}$ if $\gamma_i < \overline{\gamma}_i$ and $\lambda_i^{\text{POOL}*} < \lambda_i^{\text{RES}*}$ if $\gamma_i > \overline{\gamma}_i$.*

Theorem 7 points out that the demand loss ratio is also the sole determinant of the service providers' preferred choice between two strategies. From the service providers' perspective, sharing the facility capacity with the other one gains the benefit of accessing the whole facility but suffers the loss of an overused facility. For a service provider with a large demand loss rate, it is vulnerable to a long facility time and usually allocated a large share of the facility capacity under the reservation strategy. Hence choosing the pooling strategy tends to bring in less benefits due to the pooling effect but more losses due to the competition effect, and becomes an inferior choice if and only if its relative loss index is beyond a certain threshold, i.e. $\gamma_i > \overline{\gamma}_i$. This explains why large carriers often ask for dedicated facilities.

We summarize in Fig. 3 the strategy preference with respect to $\ln(\gamma_1)$ for the facility provider and two service providers, respectively. For example, if the condition $\ln(\gamma_1) \in (7+4\sqrt{3}, \infty)$ holds, both the facility provider and service provider 1 choose the reservation strategy, while service provider 2 prefers the pooling strategy.

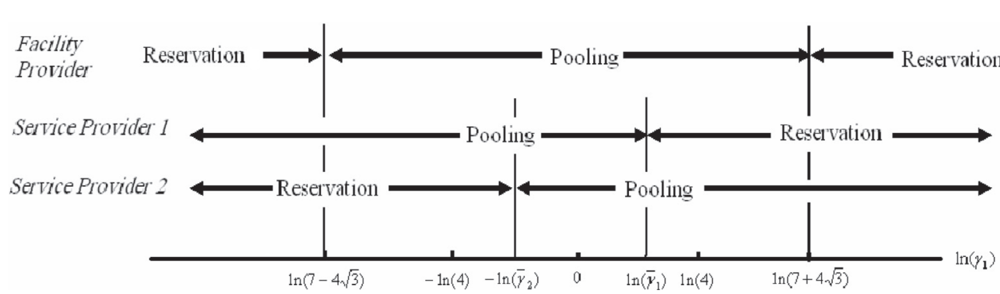

Fig. 3. The preferences for the capacity reservation and pooling strategies.

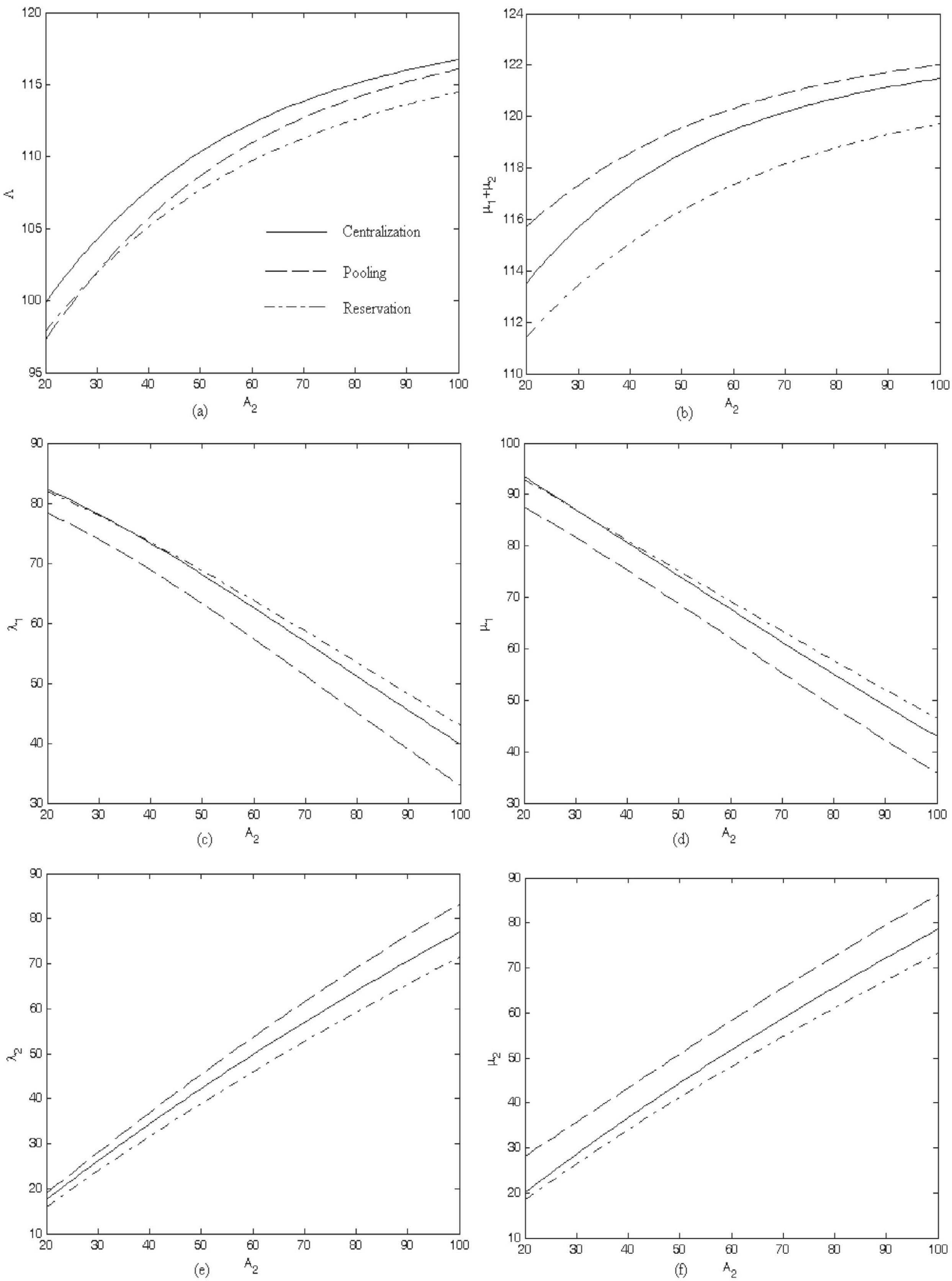

Fig. 4. (a) Optimal total demand rate $\Lambda^{\mathtt{I}*}$; (b)optimal total service capacity $\mu^{\mathtt{I}*}$; (c) optimal demand rate of service provider 1 $\lambda_1^{\mathtt{I}*}$; (d) optimal service capacity of service provider 1 $\mu_1^{\mathtt{I}*}$; (e)optimal demand rate of service provider 2 $\lambda_2^{\mathtt{I}*}$; (f)optimal service capacity of service provider 2 $\mu_2^{\mathtt{I}*}$, where $\mathtt{I} \in \{\mathtt{POOL}, \mathtt{RES}, \mathtt{CEN}\}$.

Obviously, the optimal choice is dependent on the view of the concerned party.

Interestingly, in the region where the condition $\ln(\gamma_1) \in (-\ln(\overline{\gamma}_2), \ln(\overline{\gamma}_1))$ is satisfied, a triple-agreement situation exists. That is, when the facility provider chooses pooling two carriers together to maximize the total cargo volume, both service providers maximize their individual demand rates as well. The condition for the triple-agreement situation means that the demand loss rates of the two service providers are similar. When their demand loss rates become different, i.e. $\ln(\gamma_1) \in (-\infty, -\ln(\overline{\gamma}_2))$ or $\ln(\gamma_1) \in (\ln(\overline{\gamma}_1), \infty)$, the service provider with the larger demand loss rate appeals to the facility provider for dedicated facilities. However, its request may be denied by the facility authority if its demand loss rate is not sufficiently dominant, i.e. $\ln(\gamma_1) \in (7 - 4\sqrt{3}, -\ln(\overline{\gamma}_2))$ or $\ln(\gamma_1) \in (\ln(\overline{\gamma}_1), 7 + 4\sqrt{3})$.

5.3. Numerical example

Finally, we consider a numerical example to have a better visual representation of the interactions among the facility provider and two service providers' capacity decisions.

Example 2. We let $K = 125$, $A_1 = 100$, $\theta_1 = 1$ and $\theta_2 = 0.25$ and vary $A_2 \in [20, 100]$. Fig. 4 shows the optimal demand rates $\Lambda^{\mathtt{I}*}$, $\lambda_1^{\mathtt{I}*}$ and $\lambda_2^{\mathtt{I}*}$, and service capacity levels $\mu^{\mathtt{I}*}$, $\mu_1^{\mathtt{I}*}$ and $\mu_2^{\mathtt{I}*}$, where $\mathtt{I} \in \{\mathtt{POOL}, \mathtt{RES}, \mathtt{CEN}\}$.

As shown in Fig. 4(a), if the potential market size of service provider 2 is very small (i.e. $A_2 < 30$), the optimal total demand rate $\Lambda^{\mathtt{RES}*}$ is higher than $\Lambda^{\mathtt{POOL}*}$, which implies that the facility provider prefers the reservation strategy. Otherwise, $\Lambda^{\mathtt{RES}*} < \Lambda^{\mathtt{POOL}*}$, which implies that the facility provider prefers the pooling strategy. This is consistent with Theorem 5.

Fig. 4 (a) also demonstrates that the total demand rate $\Lambda^{\mathtt{CEN}*}$ under the centralized system is the highest among the three scenarios (i.e. pooling, reservation and centralization). Moreover, Fig. 4(b) shows that the total traffic amount $\mu^{\mathtt{POOL}*}$ under the facility competition is the highest among the three scenarios. This is consistent with Theorem 6.

Notice that service provider 1 has a larger market size and is more time-sensitive than service provider 2, which implies that service provider 1 has a larger demand loss rate than service provider 2. As shown in Fig. 4(c) and (e), service provider 1 prefers the reservation strategy since $\lambda_1^{\mathtt{RES}*} > \lambda_1^{\mathtt{POOL}*}$, but service provider 2 prefers the pooling strategy since $\lambda_2^{\mathtt{RES}*} < \lambda_2^{\mathtt{POOL}*}$. This is consistent with Theorem 7.

Since service provider 2 is less time-sensitive than service provider 1, the facility competition and its congestion consequence has less negative effect on service provider 2 than on service provider 1. Hence, service provider 2 behaves more aggressively under the facility competition than service provider 1. This causes that service provider 2's optimal service capacity $\mu_2^{\mathtt{POOL}*}$ is the highest among the three scenarios as shown in Fig. 4(f), but service provider 1's optimal service capacity $\mu_1^{\mathtt{POOL}*}$ is the lowest among the three scenarios as shown in Fig. 4(d). Hence, the monotonic ranking of the total traffic rate in Theorem 6 may be reversed at the individual service provider level.

Finally, as shown in Fig. 4, the optimal demand rates and service capacity levels are monotonic in A_2, which is consistent with Propositions 2–4.

6. Concluding remarks

In this paper, we consider the interactions among a facility provider and two service providers' capacity decisions. The facility provider has a fixed amount of capacity and offers its facilities to two service providers, who determine their vessel frequencies to serve two separate transportation markets. We assume that the demand rate in each market is linearly decreasing in the total transportation time and all parties maximize demand rates. When the service providers share the facility capacity, they play a simultaneous game to compete for facility usage. We prove that a unique Nash equilibrium exists. When the service providers have their dedicated facilities, the facility provider leads two separate Stackelberg games with the service providers. We also examine a centralized system, where a central planner makes all capacity decisions to achieve the first-best outcome for the overall system performance. By proving that pooling is always optimal under the centralized system, we find that facility capacity competition is a prerequisite condition for not pooling the service providers.

We prove that the choice between the pooling and reservation strategies critically depends on the service providers' demand loss rates, which are proportional to two customer characteristics: the time sensitivity of demand and the potential market size. When the demand loss rates are close, we identify a triple-agreement situation in which the pooling benefit offsets the negative impact of facility capacity competition for both the facility provider and two service providers. When the demand loss rates are different, the competition effect may dominate the pooling effect for both the facility provider and the service provider with the larger demand loss rate. However, because the threshold of the facility provider is much larger than that of the service provider, the former permits a request for dedicated facilities only if the latter's demand loss rate is sufficiently dominant.

Our research provides important guidelines for strategic seaport capacity management. Our results highlight that facility capacity competition is a prerequisite condition for not pooling the service providers and the potential market size is a new customer characteristic to pay attention to in the tradeoff between capacity pooling and reservation strategies. We provide quantitative criteria on how the facility provider and service providers should determine their capacity decisions given the market conditions. Our finding, that the facility provider allocates dedicated facilities only to the service provider with an overwhelming dominance at the demand loss rate, is consistent with the observations in practice. It also provides an insight into when a service provider appeals to the facility provider for dedicated facilities and why this is often denied by such seaports as Shanghai, Singapore and Shenzhen. This as well as the existence of the triple-agreement situation are important qualitative insights for practitioners.

There are several directions to extend this research. First, we focus on capacity management from an operational perspective. From a pricing perspective, the facility provider can price the facility capacity usage to penalize a profit-maximizing service provider who overuses the facility system. It is an interesting research question to design a joint optimal pricing policy and a capacity management strategy for a port authority who deals with profit-maximizing carriers. Second, we consider two service providers to demonstrate the tradeoff between the capacity reservation and pooling strategies. But, in practice, a port often serves multiple carriers, who may benefit from forming alliances to share dedicated port facilities. This can be studied under a cooperative game framework. Finally, seaports face heavy competition from local competitors. Capacity reservation can be used as a strategic weapon to attract carriers. A game-theoretic model may help understand competition among seaports.

Acknowledgements

The authors would like to thank the editor and two anonymous referees for their helpful comments and suggestions that greatly improved the presentation of this paper. L. Jeff Hong is supported in part by the Hong Kong Research Grants Council Theme-Based Research Project No. T32-620/11. Sheng Hao Zhang is supported in part by National Natural Science Foundation of China under Grant 71390333, 71271167 and 70901062, and Program for New Century Excellent Talents in University under Grant NCET-13-0465.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.ejor.2015.03.027

References

Adler, N. (2001). Competition in a deregulated air transportation market. *European Journal of Operational Research, 129*(2), 337–345.
Allon, G., & Federgruen, A. (2008). Service competition with general queueing facilities. *Operations Research, 56*(4), 827–849.
Armony, M., & Haviv, M. (2003). Price and delay competition between two service providers. *European Journal of Operational Research, 147*(1), 32–50.
Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., et al. (2010). *Total delay impact study: An assessment of the costs and impacts of flight delay in the United States: Technical report*. Washington, DC: National Center of Excellence for Aviation Operations Research.
Boyaci, T., & Ray, S. (2003). Product differentiation and capacity cost interaction in time and price sensitive markets. *Manufacturing & Service Operations Management, 5*(1), 18–36.
Ha, A. (1998). Incentive-compatible pricing for a service facility with joint production and congestion externalities. *Management Science, 44*(12), 1623–1636.
Hardin, G. (1968). The tragedy of the commons. *Science, 162*(3859), 1243–1248.
Hassin, R., & Haviv, M. (2003). *To queue or not to queue: Equilibrium behavior in queueing systems*. Boston: Kluwer.
Haviv, M., & Ritov, Y. (1998). Externalities, tangible externalities, and queue disciplines. *Management Science, 44*(6), 850–858.
Hong Kong Marine Department. (2014). Ranking of container ports of the world. http://www.mardep.gov.hk/en/publication/pdf/portstat_2_y_b5.pdf.
Kalai, E., Kamien, M. E., & Rubinovitch, M. (1992). Optimal service speeds in a competitive environment. *Management Science, 38*(8), 1154–1163.
Mandjes, M., & Timmer, J. (2007). A duopoly model with heterogeneous congestion-sensitive customers. *European Journal of Operational Research, 176*(1), 445–467.
Mendelson, H., & Whang, S. (1990). Optimal incentive-compatible priority pricing for the M/M/1 queue. *Operations Research, 38*(5), 870–883.
Meng, Q., & Wang, S. (2011). Optimal operating strategy for a long-haul liner service route. *European Journal of Operational Research, 205*(1), 105–114.
Nagle, T. T., Hogan, J. E., & Zale, J. (2011). *The strategy and tactics of pricing*. Upper Saddle River, New Jersey: Prentice Hall.
Osorio, C., & Bierlaire, M. (2009). An analytic finite capacity queueing network model capturing the propagation of congestion and blocking. *European Journal of Operational Research, 196*(3), 996–1007.
Pangburn, M. S., & Stavrulaki, E. (2008). Capacity and price setting for dispersed, time-sensitive customer segments. *European Journal of Operational Research, 184*, 1100–1121.
Ray, S., & Jewkes, E. M. (2004). Customer lead time management when both demand and price are lead time sensitive. *European Journal of Operational Research, 153*(3), 769–781.
Rothkopf, M. H., & Rech, P. (1987). Perspectives on queues: Combining queues is not always beneficial. *Operations Research, 35*(6), 906–908.
Sinha, S. K., Rangaraj, N., & Hemachandra, N. (2010). Pricing surplus server capacity for mean waiting time sensitive customers. *European Journal of Operational Research, 205*(1), 159–171.
Smith, D. R., & Whitt, W. (1981). Resource sharing for efficiency in traffic systems. *The Bell System Technical Journal, 60*, 39–55.
So, K. C. (2000). Price and time competition for service delivery. *Manufacturing & Service Operations Management, 2*(4), 392–409.
Stalk, G., & Hout, T. M. (1990). *Competing against time: How time-based competition is reshaping global markets*. New York: The Free Press.
Stopford, M. (2009). *Maritime economics* (3rd ed.). New York, NY: Routledge.
van Dijk, N., & van der Sluis, E. (2009). Pooling is not the answer. *European Journal of Operational Research, 197*(1), 415–421.
Vaze, V., & Barnhart, C. (2012). Modeling airline frequency competition for airport congestion mitigation. *Transportation Science, 46*(4), 512–535.
Whitt, W. (1999). Partitioning customers into service groups. *Management Science, 45*(11), 1579–1592.
World Bank. (2007). *The port reform toolkit* (2nd ed.). http://www.ppiaf.org/sites/ppiaf.org/files/documents/toolkits/Portoolkit/toolkit.html.
Yu, Y., Benjaafar, S., & Gerchak, Y. (2015). Capacity sharing and cost allocation among independent firms with congestion. *Production and Operations Management*, forthcoming.