Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Product Differentiation and Capacity Cost Interaction in Time and Price Sensitive Markets

Tamer Boyaci, Saibal Ray,

Product Differentiation and Capacity Cost Interaction in Time and Price Sensitive Markets

Tamer Boyaci • Saibal Ray
Faculty of Management, McGill University, Montreal, Quebec, Canada H3A 1G5
tamer.boyaci@mcgill.ca • *saibal.ray@mcgill.ca*

In this paper, we study a profit-maximizing firm selling two substitutable products in a price and time sensitive market. The products differ only in their prices and delivery times. We assume that there are dedicated capacities for each product and that there is a standard industry delivery time for the regular (slower) product. The objective of the firm is to determine the delivery time of the express (faster) product and appropriately price the two products, taking into consideration the impact of delivery time reduction on capacity requirements and costs. We develop a model that integrates pricing and delivery time decisions with capacity requirements and costs, and study scenarios where the firm is constrained in capacity for none, one, or both product(s). We show how product differentiation decisions are influenced by capacity costs, and how the firm should adapt its differentiation strategy in response to a change in its operating dynamics. We first identify a market characteristic that governs the optimal pricing structure. We then show that the degree of product differentiation depends on both the absolute, as well as the relative values of the capacity costs. Provided that the capacity cost differential remains the same, higher capacity costs induce less time differentiation and less price differentiation. An increase in capacity cost differential increases price differentiation, but decreases time differentiation. The optimal prices depend, in addition to the above, on the market characteristic. We find that prices can actually decrease when the firm incurs capacity-related costs. We also explore the impact of substitutability on product differentiation, and illustrate our results in a numerical study.
(*Product Differentiation; Pricing; Delivery-Time Guarantees; Time-Based Competition; Capacity Management; Substitution*)

1. Introduction

Speed has emerged, in addition to price, as the dominant factor in determining competitive advantage. The academic and popular literature on time-based competition presents ample evidence as to how firms can use speed (or equivalently, time) as a strategic weapon to gain competitive advantage (Stalk and Hout 1990, Blackburn et al. 1992, Hum and Sim 1996, Suri 1998). A recent survey of 110 firms in five major manufacturing sectors conducted by Performance Management Group (a subsidiary of high-tech management consultants PRTM) substantiates these claims. They report that the best in class performers focus their operations on achieving breakthroughs not only in cost, but also in speed (Geary and Zonnenberg 2000). It is also well known in the business logistics literature that speed and consistency of delivery time are the two most important elements of customer service in addition to price (Sterling and Lambert 1989, Ballou 1998, chap. 4).

Successful implementations of time-based strategies run the whole gamut—transportation and warehousing services, manufacturing (especially make-

1523-4614/03/0501/0018$05.00
1526-5498 electronic ISSN

to-order industries), information and data services, web hosting, online content distribution, and online shopping, among others. Such strategies, naturally, have particularly strong demand-enhancing effects in industries where delivery time is a key concern. It is therefore essential for firms to adopt time-based strategies to maintain a high degree of responsiveness, while sustaining a competitive cost structure (Hum and Sim 1996). This brings capacity-related issues to the forefront, because both the speed of product delivery and the cost at which this can be achieved, hinges on the firm's capacity. The motivation of our study stems partly from this salient fact.

An effective way to maintain customer responsiveness and to enhance demand is through time-based product differentiation and segment pricing. If firms develop a thorough understanding of customers' preferences to time and their willingness to pay, different customer segments can be served profitably, through careful time *and* price-based market segmentation. In many such cases, the firm offers *the same product or service, differentiated only in terms of the delivery times offered and the prices charged.* Examples of such market segmentation are abundant. Photo development stores offer one-hour express service and a cheaper three-day regular service. In the transportation sector, companies offer different modes of transportation; for example, air delivery and ground delivery for many origin-destination pairs and types of physical goods. Obviously, the air-delivery mode is faster, but is also more expensive, at least for sufficiently large distances. Going a step further, transportation companies create segments within a particular delivery mode. For example, within ground transportation, there is the "time-definite" fast delivery option, as well as a more cost-effective (and less reliable) regular delivery option.

It is presumable that time-based differentiation would have a direct influence on the operational system that produces and delivers these products. A natural question that comes to mind then is whether firms also differentiate the systems that produce these products or not. In certain instances, the distinction is natural (e.g., different modes of transportation). But a close examination of practice reveals that both cases, shared capacity for both segments or dedicated capacity for each niche, are possible and there are advantages and disadvantages of both. Production and delivery through shared capacity exploits the benefits of pooling; however, focusing too much on it leads to the phenomena of "averaging," for example, all customers are served at an average speed and charged an average cost. The result is that firms get "stuck in the middle"—neither enough time sensitive, nor enough cost-effective (Fuller et al. 1993, O'Conor 2002). Hence, when the "needs" of the customer groups are very different and each such segment is broad enough, it might be prudent to serve these segments through dedicated channels. Academicians and practitioners have advocated this strategy for quite some time in the name of tailored logistics or tailored supply chain management (Chopra and Meindl 2001, Ballou 1998, Fuller et al. 1993). The operational setting analyzed in this paper is the latter one, where each product is produced and delivered by *a tailored supply system* with a *dedicated capacity*.

There are many practical settings in which firms establish dedicated facilities for each customer niche. In the *transportation and logistics* industry, recent years have witnessed a dramatic rise in the demand for "time-definite" premium transportation services, which used to be a small niche service exclusively for handling emergencies. Realizing that shipping via truck could be just as quick as airfreight for relatively short-haul shipments, many carriers have started moving time sensitive freight via ground using 100% truck delivery. In order to accommodate this shift, courier companies had to arrange for dedicated ground transportation fleet for time-definite shipments (Farris 2002). As a result, these carriers now have two totally separate capacities (trucks, information system). A well-known example is FedEx, which has two dedicated systems for ground transportation—*FedEx Ground* for business-to-business, cost-effective small package delivery, and *FedEx Custom Critical* (*Surface Expedite*), previously known as Roberts Express, for time sensitive premium transportation. Similar philosophy is also being utilized by a number of other logistics companies like Roadway Express and Consolidated Freightways (Foster 1999). Fuller et al. (1993) also provides an example of a tailored supply system from the transportation industry. National Freight Consortium of the United Kingdom created four distinct customer segments, based partly on their time sensitivity, and decided to serve the logistical needs of each segment through dedicated assets. Likewise, in *warehousing*, it is common to have dedicated warehouses for delivery-committed "express" orders, which are processed individually by a pick-and-pack system. The "regular" orders, on the other hand, are processed separately by a less costly (automated) picking system (Bowman 2001).

Time-based differentiation with dedicated capacities is also common in *information and data services*. There are many data archival companies that provide electronic data storage for customers' legacy/original data as well as "processed" data. Data with immediate access requirements is stored "online" within the service company's artifact database for web access, and is stored "offline" on physical media for less time sensitive applications (see ⟨www.ixtsolutions.com⟩ for an example). These archival services even enable customers to select the access time requirements for the emergency data. Obviously, the prices charged are different for the two applications. In a similar vein, *web hosting and content delivery* firms maintain separate servers for customers whose content for online delivery (e.g., news) is time sensitive. This configuration enables the firm to do content updates in real time. The regular customers, whose content do not change very frequently, are pooled in a separate system, where the updating is done during offline hours (see ⟨www.cacheflow.com⟩ for an example).

In *manufacturing*, time-based differentiation with dedicated capacities is most evident in make-to-order industries such as printing. For example, a Southeast Asia based printing company has separate facilities for time sensitive magazines like the Asian editions of *The Economist*, *Time*, *Newsweek*, etc., and for mass-produced books, even though most of the processing steps are similar for both types of products. The company even uses a dedicated delivery system for time sensitive materials to deliver them to the distribution points (or even end customers) through their distribution networks ⟨www.tpl.com.sg⟩.

There are many other practical examples that fit into the above-discussed operational setting. We refer the reader to Smith et al. (2000) for an interesting example related to *online shopping*. More examples can be found in Cachon and Harker (2002), Chopra and Meindl (2001) and Fuller et al. (1993).

As noted earlier, demand-enhancing efforts of time-based differentiation cannot be decoupled from capacity-related decisions. The success of any time-based strategy adapted by the firm is intrinsically related to the supply capabilities of the firm. An example of such time-based strategies is *a uniform delivery time guarantee for customers*.[1] The whole idea of premium transportation (e.g., guaranteed next day delivery) or dedicated web hosting facility (e.g., guaranteed response time of less than 10 seconds for online trading companies) is based on the issue of time guarantees offered to customers. A shorter guaranteed delivery time may generate more demand, but this in return puts pressure on the supply resources and can lead to higher capacity-induced costs. Furthermore, there is a risk that demand may exceed the firm's capacity to respond, which can lead to a penalty cost for the seller and/or a decrease in repeat business. For example, Ameritrade advertises on its website that for certain types of online trading, if the service takes more than 10 seconds, it will be totally commission-free ⟨www.Ameritrade.com⟩. Similar guarantees and associated penalties are present in other online trading companies, as well as in the transportation industry (for more examples, refer to So and Song 1998, and Rao et al. 2000). Therefore, it is necessary to have some internal mechanism in place to ensure that the promised lead times are feasible and reliably met, and managers of such firms have to decide *simultaneously on the prices, delivery times, and capacity levels*.

It is also important to recognize that when a firm offers an "express" product, it would generate new customers, but there will also be some degree of cannibalization of the "regular" product. If the difference

[1] There are also other popular time-based strategies like serving customers as fast as possible, or encouraging potential customers to get a delivery time quote prior to ordering (see Hum and Sim 1996, So and Song 1998).

in prices is not significant, while the time differential is, customers will gravitate towards the express product and vice versa. It is therefore imperative for managers to take into account the *degree of substitutability* when making product differentiation decisions.

In this paper, we study a single firm that is providing a service, or selling a make-to-order product to customers, and is using the strategy of uniform delivery time guarantee to attract customers. The firm may be a monopolist, or may be one of several competing firms that offer similar products. However, we do not explicitly model competition in our study. Our focus is on the optimal decisions of a single firm in a capacitated environment. We assume that there is a well-established "lengthy" industry standard delivery time, L_2, for the regular product, which the firm cannot alter. For example, in most of the online retail web hosting services, any updating, if not done in express fashion, must be done within one day. The aim of the firm is to offer a shorter delivery time ($L_1 < L_2$) variant in order to attract time sensitive customers and obtain a price premium.[2] For the express product (which we will refer to as P1), the firm provides a guaranteed delivery time of L_1 to customers and charges a price of p_1, while the regular product (which we will refer to as P2) has a guaranteed delivery time of L_2 and is charged a price p_2. Note that the "products" are differentiated only in terms of prices and delivery times and are served by two dedicated capacities. The objective of the firm is to maximize its profit by suitable selection of the prices for both products and the length of the delivery time for the express product. In doing so, the firm has to take into account the prevalent industry characteristics discussed before: The products are substitutable and therefore customer demand for each product depends on the price and delivery time of both products; reducing delivery time by increasing capacity requires some cost; and the firm must be able to satisfy the guaranteed delivery time according to a prespecified reliability level. We look into different capacity (capacity cost for none/one/both product(s)) structures that the firm might face. Some of the issues that we address are:

- What will be the optimal prices and the delivery times for the firm under each capacity structure?
- What are the effects of capacity structures on the firm's optimal price and delivery time decisions—both on the absolute values and on the extent of differentiation?
- What is the effect of substitutability on the firm's optimal decisions?

The analytical models that we present in this paper determine the optimal prices and delivery times for the firm under different capacity structures. Comparing these models, we gain critical managerial insights as to how capacity costs *and* cost differentials both interact with optimal product positioning and differentiation, and how the firm should react to a change in its operating environment, which can be influenced by a shift in the firm's make-or-buy strategy. Specifically, higher marginal capacity costs prompt the firm to differentiate less, both in price and time, when the cost differential is constant, whereas higher cost differentials induce the firm to opt for less time differentiation but more price differentiation. In addition, we show the significance of understanding the sensitivity of the customer market to price and speed, as the firm's optimal pricing strategy is governed by this characteristic. We also show that ignoring product substitutability can result in higher or lower prices, delivery time, and differentiation, depending on the capacity costs and the market characteristic.

In §2, we present our modeling framework and a summary of the relevant literature. In §3, we develop and analyze our "benchmark" model without any capacity-related costs. Section 4 analyzes and compares different capacitated models. Section 5 deals with the effect of substitutability on the firm's optimal decisions. Section 6 concludes with the summary of insights, remarks on some model assumptions, and possible extensions for future research.

[2] Presumably, the firm could have opted for an even longer delivery time and reduced prices to attract more customers. However, our setting is where the industry standard is sufficiently long that the only option is to go for a shorter delivery time. We believe this is a more realistic representation of the industry trend.

2. Model Framework and Related Literature

We assume that customers arrive to take delivery of the products at two separate facilities (one for the regular product and another for the express) according to a Poisson process. The mean arrival rate for each of the products depends on the price and guaranteed delivery time of both of the products, and the price and delivery time differentials between each of the products. The service time for each product is exponentially distributed and the customers are served on a first-come first-served basis. We also assume that the marginal cost for increasing capacity, through, for example, hiring extra workers or acquiring improved equipment, will cost the same, which implies that the capacity cost function is increasing and linear. The firm has a predetermined internal delivery time reliability target, α ($0 < \alpha < 1$)—the probability that a random customer will have an actual waiting time less than the guaranteed delivery time. We assume for simplicity that the reliability level given is the same for both products, and close to one. Since α is not a decision variable, this assumption does not change the nature of our results.

Our demand model is a linear model with substitution: (i) each product's mean demand is decreasing in its own price and delivery time offered, and (ii) each product's price hike can only increase the other product's mean demand, and each product's delivery time reduction can only decrease the other product's mean demand. The exact form of our demand model is inspired by Tsay and Agrawal (2000). Before we describe the details, we define the following notation ($j = 1, 2$ refers to P1 and P2, respectively):[3]

$2a$—potential market size for the product (this is the total demand if the price and delivery time for both products are zero),

λ_j—mean demand for product j (units/time),

μ_j—capacity for product/facility j (per time),

m—unit operating cost (excluding capacity induced costs) for each product,

β_p—price sensitivity of demand,

β_L—delivery time sensitivity of demand,

[3] Later, we will introduce a second subscript for the different models studied.

θ_p—sensitivity of switchovers toward price difference,

θ_L—sensitivity of switchovers toward guaranteed delivery time difference.

As indicated before, the guaranteed delivery time and price for the faster product (P1) are denoted by L_1 and p_1, respectively, while those for the regular product (P2) are denoted by L_2 and p_2, respectively. Based on the above notation, the mean demand for each product is given by:

$$\lambda_1 = a - \beta_p p_1 + \theta_p(p_2 - p_1) - \beta_L L_1 + \theta_L(L_2 - L_1), \quad (1)$$

$$\lambda_2 = a - \beta_p p_2 + \theta_p(p_1 - p_2) - \beta_L L_2 + \theta_L(L_1 - L_2). \quad (2)$$

Note that $\lambda_1 + \lambda_2 = 2a - \beta_p p_1 - \beta_L L_1 - \beta_p p_2 - \beta_L L_2$, which means that all else being equal, a change in the θs will not alter the total demand for the two products. The firm can attract "new" customers only by lowering the prices for its products (at a rate β_p), or by offering a shorter delivery time (at a rate β_L). However, the switching of the customers is governed by the difference between the prices and delivery times. Higher p_1 induces customers to switch from P1 to P2 (at a rate θ_p), whereas lower L_1 induces customers to switch to P1 from P2 (at a rate θ_L).

To complete our modeling framework, we define the following:

W—steady state actual waiting time in the facility,

S—actual delivery time reliability level, $P(W < L)$,

$M(\mu_j)$—capacity cost per unit time for the product/facility j ($=A_j\mu_j$, where A_j is the marginal cost of capacity for product/facility j),

π—profit per unit time for the firm.

Then the firm's problem in its general form can be formulated as:

$$\begin{aligned} \text{Maximize} \quad & \pi(p_1, p_2, L_1, \mu_1, \mu_2) \\ & = (p_1 - m)\lambda_1 + (p_2 - m)\lambda_2 \\ & \quad - M(\mu_1) - M(\mu_2), \end{aligned} \quad (3)$$

subject to:

$$S_1 = P(W < L_1) = 1 - e^{-(\mu_1 - \lambda_1)L_1} \geq \alpha,$$

$$S_2 = P(W < L_2) = 1 - e^{-(\mu_2 - \lambda_2)L_2} \geq \alpha,$$

$$\mu_1 > \lambda_1, \mu_2 > \lambda_2, p_1 > m, p_2 > m, L_1 < L_2, \lambda_1, \lambda_2 > 0,$$

where λ_1 and λ_2 are given by Equations (1) and (2), respectively.

Two remarks are in order here. First, observe that the last set of constraints is necessary to define a realistic problem setting and therefore has to hold for all the models considered in this study. Hence, we will not repeat them for the rest of the paper. Obviously, we impose restrictions on our model parameter values so that these constraints are satisfied. Second, note that under our Poisson arrivals and exponential service time assumptions, the form of the delivery reliability constraint is exact. Furthermore, as shown in the literature (see So and Song 1998, and references therein), for high service levels the tail of the waiting time distribution is well approximated by the exponential distribution even for a $G/G/s$ queue. Hence, our model is approximately valid for more general demand and service time characteristics.

It is well known (see Palaka et al. 1998, So and Song 1998) that at optimality, delivery reliability constraints must be binding, which implies that the capacity requirements μ_j of the products will be:

$$\mu_j(p_1, p_2, L_1) = \frac{-\ln(1-\alpha)}{L_j} + \lambda_j(p_1, p_2, L_1), \quad j = 1, 2. \tag{4}$$

As a result, the firm's problem reduces to maximizing Equation (3) with μ_j's given as above. Note that Equation (4) enables us to explicitly model the impact of prices and delivery times (i.e., demand) on capacity requirements and costs. Observe also that the system stability constraint, for example, $\mu_j > \lambda_j$, is automatically satisfied by Equation (4). As a final remark, notice that since the reliability constraints are binding and α is not a decision variable, the capacity costs A_js can be assumed to also include a penalty cost per unit (So 2000).

There are three streams of research that are related to our work: (i) marketing/business logistics literature indicating the importance of delivery time as an element of customer service and its effect on demand; (ii) models that deal with substitution and/or differentiation in the presence of price and service sensitive demand, but does not explicitly model the effect of demand on capacity; and (iii) models that take into account the effect of demand on capacity requirements and costs, but does not consider differentiation of products or substitution effects.

According to Baker et al. (2001), less than 10% of end-consumers and less than 30% of corporate customers base their purchase decisions on an item's selling price only; the rest care about both price and other customer service attributes. Recent marketing/business logistics literature suggests that delivery time length and reliability consistently rank as the two most important customer service elements, irrespective of the type of industry. Although the effect of price on demand has been well known for a long time, there is now definitive evidence that customer service, in general, and delivery time, in particular, has a considerable positive impact on demand (Jackson et al. 1986, Sterling and Lambert 1989, Heskett 1994). A summary of these customer service studies and other related studies can be found in Ballou (1998, chap. 4).

There is extensive literature (especially within economics and marketing) in the second stream of related research, which assumes demand to be both price and customer service sensitive. In these works, customer demand is a deterministic function of price and other attributes like quality, location, advertising, delivery time, etc. Most of these studies analyze a competitive setting. Furthermore, supply- and capacity-related issues are not usually modeled in detail. We refer the reader to excellent books by Anderson et al. (1992) and Tirole (1988) for an extensive review of this literature. Our model fundamentally differs from this approach by considering a noncompetitive setting with a rich (essentially) deterministic demand framework and an integrated capacity cost model. Nevertheless, there are two existing models that are structurally comparable to ours.

Birge et al. (1998) study a pricing and capacity setting problem for a firm that produces two substitutable products in a newsvendor framework, both in centralized and decentralized settings. There is a cost of adding capacity and the firm has to decide on the price and/or capacity levels of one or both of the products. However, delivery time is not a part of the model and the interaction between capacity and demand is not modeled. Tsay and Agrawal (2000) analyze the competition between two retailers, based on both price and customer service when the products are substitutable. The focus is on investigating the impact of competition between retailers on prices, service levels, and profits, and identifying coordination mechanisms with a common manufacturer. Our demand framework is similar to Tsay and Agrawal, and their centralized scenario is philosophically similar to our model. However, we model the internal dynamics of the operating system in more detail, and explore in-depth the interaction between capacity, prices, and delivery times for a noncompeting firm.

The most salient feature of the third line of research is the recognition that longer delivery times might have a negative impact on customer demand. Utilizing queuing theory to estimate delivery times and delays, the majority of these models are developed under the assumption that the firm knows the cost of waiting, from a customer's viewpoint as well as the facility's capacity costs. It is further assumed that any arriving customer's decision to join the system depends on the "full" price, which is equal to the sum of the explicit price charged and the cost of waiting. A number of papers in this stream investigate issues that arise due to competition between firms in terms of price and/or speed. For example, Kalai et al. (1992) study the competition between firms based on mean processing time, whereas Li (1992) explores the role of inventory when firms compete on speed of delivery and the impact of this competition on the firm's optimal choice between make-to-order and make-to-stock strategies. Lederer and Li (1997), on the other hand, study a setting in which firms compete for price and time sensitive customers, based on price and allocation of a fixed capacity. They show the effects of delay costs on pricing and production planning. More recently, Cachon and Harker (2002) model competition between two firms with price and time sensitive demand. They show how economies of scale can provide motivation for outsourcing in this competitive setting. Our model differs from this stream in a number of ways. First, we do not assume delay costs which are, in practice, hard to quantify, especially when customers are external. In our model, customer delay is managed by maintaining a delivery time guarantee for each customer type. Second, these models, unlike ours, do not take into account the issue of substitution between customer niches. A third distinction is that our analysis is based on a noncompetitive setting.

However, there is a second group of queueing-based research, that studies a noncompetitive setting and addresses pricing and/or capacity selection issues for one or multiple sets of customers (Dewan and Mendelson 1990, Mendelson and Whang 1990, Stidham 1992). Since these models assume a single processing facility, customer segmentation necessitates developing pricing schemes for different customer segments and utilizing priority rules in serving customers (e.g., Mendelson and Whang 1990). Our model is significantly different from these papers because we assume separate capacities for each customer type with the possibility of substitution among the products. Also note that these noncompetitive models also use the assumption that the firm knows the delay costs for the users, which is fundamentally different from our approach of using delivery time guarantee.

Under a guaranteed delivery time strategy, firms advertise a uniform delivery lead time within which they promise to satisfy "most" of the customer orders. The length of the guaranteed delivery time is a decision variable that directly affects demand. This concept has been dealt with in detail by several recent papers including So and Song (1998), Palaka et al. (1998), So (2000), Rao et al. (2000), and Ray and Jewkes (2002). The firm is still modeled as a queueing system, where the mean customer demand is dependent on the price and/or length of the guaranteed delivery time. The firm incurs linear capacity costs and a reliability constraint is used to ensure a satisfactory service level. Although we model delivery reliability and capacity costs in a similar manner, our model differs significantly from these works. The existing models either assume a single product noncompetitive setting, or analyze price and guaranteed delivery time selection in a competitive setting (only So 2000). None of these models take into account the product differentiation and/or substitution effects.

The main contribution of our study is in providing managerial insights into the effect of capacity cost on product differentiation. By combining substitutability, differentiation, and price and time sensitive demand with a detailed capacity cost framework, our integrated operations marketing model captures the interaction between pricing, delivery time guarantees, and capacity requirements and costs. We start by analyzing the case with no capacity costs.

3. Uncapacitated Benchmark Model (Model 0)

In this section we analyze a "pure marketing" model where there is no capacity constraint involved with either product, and hence, there is no need for delivery reliability constraints or capacity costs in the firm's profit maximization problem (i.e., $A_1 = A_2 = 0$). Though in reality, firms will always have some capacity cost associated with at least one product, we present this model to act as a "benchmark" for the firm (since this is the best that the firm can do), and also to get a better idea about the effect of capacity on the optimal decisions when capacity induced costs are introduced in the latter models.

We first assume that L_1 is fixed and determine the optimal prices of P1 and P2 (see Appendix A for details). The optimal prices are given by:[4]

$$p_{10}^*(L_1) = \frac{m}{2} + \frac{a}{2\beta_p} + L_2 \frac{\beta_p\theta_L - \theta_p\beta_L}{2\beta_p(\beta_p + 2\theta_p)} - L_1 \frac{\beta_p\beta_L + \beta_p\theta_L + \theta_p\beta_L}{2\beta_p(\beta_p + 2\theta_p)}, \tag{5}$$

$$p_{20}^*(L_1) = \frac{m}{2} + \frac{a}{2\beta_p} - L_2 \frac{\beta_p\beta_L + \beta_p\theta_L + \theta_p\beta_L}{2\beta_p(\beta_p + 2\theta_p)} + L_1 \frac{\beta_p\theta_L - \theta_p\beta_L}{2\beta_p(\beta_p + 2\theta_p)}, \tag{6}$$

and

$$p_{10}^*(L_1) - p_{20}^*(L_1) = (L_2 - L_1)\frac{\beta_L + 2\theta_L}{2\beta_p(\beta_p + 2\theta_p)}. \tag{7}$$

Analyses of Equations (5) and (6) show that behavior of the optimal prices depends not only on the guaranteed delivery times, but also depends crucially on a market characteristic. Specifically, whether "new" customers are attracted more by lower prices or lower delivery times, and whether switching of the customers is influenced more by the disparity between the prices or delivery times, governs the optimal pricing strategy.

Proposition 1. *An increase in the guaranteed delivery time L_1 for* P1 *results in*:

(i) *A linear decrease in optimal* P1 *price.*

(ii) *A linear increase in optimal* P2 *price if $\beta_p\theta_L > \theta_p\beta_L$ and a linear decrease in* P2 *optimal price if $\beta_p\theta_L < \theta_p\beta_L$.*

Proof. Immediate. □

Part (i) of Proposition 1 is intuitive because when the firm provides shorter delivery time for P1, for example, lower L_1, it will extract a price premium from those customers. Part (ii) shows the impact of market characteristics on the optimal pricing strategy. As we will see afterwards, this particular characteristic influences many of our results, so it deserves deeper investigation.

When $\beta_p\theta_L > \theta_p\beta_L$, the market is more sensitive to price, but switchovers are more governed by difference in delivery times, which will henceforth be referred to as a "price and time difference sensitive" (PTD) market. For $\beta_p\theta_L < \theta_p\beta_L$, the market is more sensitive to delivery time, but switchovers are more governed by price difference, which will be referred to as a "time and price difference sensitive" (TPD) market. Note that when $\theta_L \approx \theta_p$, PTD implies a price sensitive market, while TPD implies a lead-time sensitive market. A closer examination of the market condition yields further insights on the PTD/TPD distinction of the market. A unit decrease in delivery time of P1 generates customers at a rate of $(\beta_L + \theta_L)$ for P1, out of which β_L are "new" customers and θ_L are cannibalized customers, substituting P1 for P2. Similarly, a unit price decrease for P1 (resp., P2) generates customers at a rate $(\beta_p + \theta_p)$, out of which β_p are "new" customers and θ_p are cannibalized from P2 (resp., P1). Notice that TPD condition $\beta_L\theta_p > \theta_L\beta_p$ is equivalent to $[\theta_p/(\beta_p + \theta_p)] > [\theta_L/(\beta_L + \theta_L)]$, which means that the fraction of "cannibalized" demand (with respect to the total demand generated) is higher for a price reduction rather than a reduction in delivery time. Clearly, in a TPD market, the cannibalization effect is stronger for price. On the other hand, PTD condition

[4] Note that we introduced a second subscript to refer to the model number.

$\beta_p\theta_L > \theta_p\beta_L$ is equivalent to $[\theta_p/(\beta_p+\theta_p)] < [\theta_L/(\beta_L+\theta_L)]$, which means that the cannibalization effect is stronger for delivery time. Our results indicate the importance for managers to make this clear distinction. This can be achieved through an empirical study, which is beyond the scope of this paper.[5]

The above market segregation makes result (ii) quite intuitive. For a PTD market, as L_1 is decreased, there is a relatively small gain of new customers but more customers buy P1. By decreasing the P2 price, the firm can attract more new customers, some of which will opt for P1, increasing further the profit of the firm. For a TPD market, a decrease in L_1 attracts a significant set of new customers, but relatively less switchover to P1. Because the switchovers are more sensitive to price differential now, the firm can afford to increase the price of P2 without considerably decreasing new demand (β_p is relatively low), while still encouraging more customers to opt for P1 and pay a price premium.

Comparing Equations (5) and (6) we can also show that for $L_1 < L_2$, the optimal P1 price is always higher and more sensitive toward L_1 than the optimal P2 price. From Equation (7) we conclude that differentiation in price and time go hand-in-hand—higher differentiation in terms of delivery time will lead to higher price differentiation.

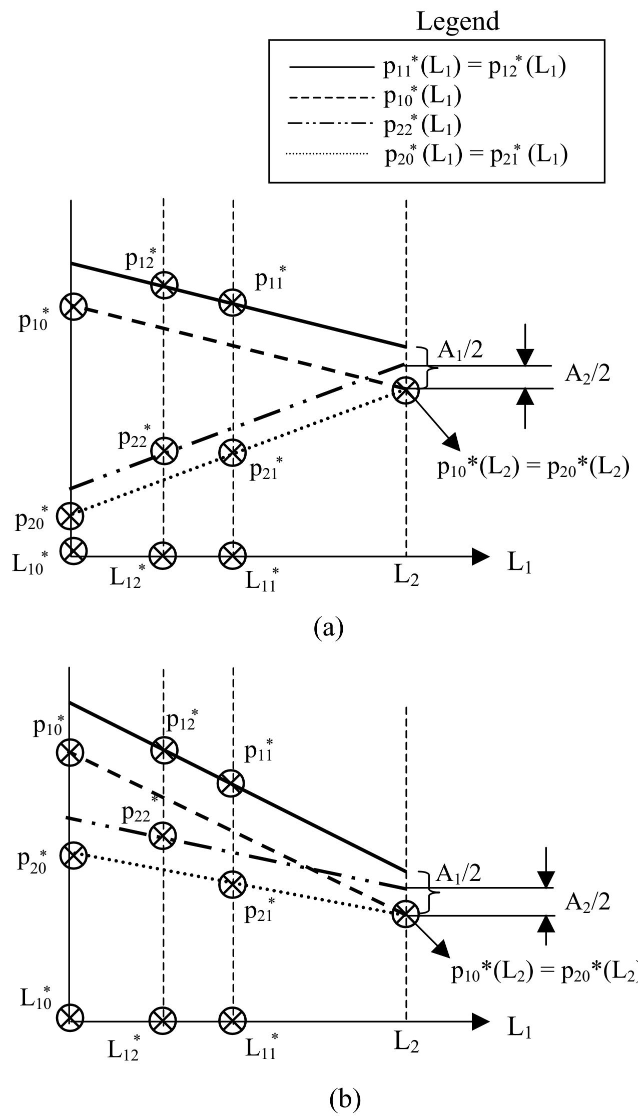
Proposition 2. *The optimal* P1 *demand is linear decreasing in* L_1, *the optimal* P2 *demand is linear increasing in* L_1, *while the optimal total firm demand is linear decreasing in* L_1.

Proof. See Appendix B. □

Proposition 2 is intuitive. Because we are, in principle, focusing on the time sensitive market, we would expect both P1 demand and total customer demand to decrease with an increase in delivery time. However, note that the demand for P2 individually increases with delivery time due to substitution. If there were no product differentiation (i.e., $L_1 = L_2$), the firm would have charged a uniform price $p = p^*_{10}(L_2) = p^*_{20}(L_2)$ (see Figure 1). The decision to differentiate in terms of delivery time to attract more customers leads to differentiation in terms of price. The firm then increases the price of one product and decreases/increases the price of the other, depending on the market characteristic (PTD or TPD).

Until this point in our analysis, we assumed L_1 to be fixed and concentrated on providing insights about the optimal prices and customer demand. Now let us see what the value of the optimal delivery time will be.

Figure 1 **(a) Optimal Decision Variable Values for PTD Market (b) Optimal Decision Variable Values for TPD Market**



[5] Interestingly, we have not been able to locate such a study within the existing academic literature.

PROPOSITION 3. *When the firm faces no capacity constraints, its profit is decreasing convex in L_1 on $[0, L_2]$ implying that $L_{10}^* = 0$.*

PROOF. See Appendix B. □

Because there is no cost associated with reducing delivery time, it is expected that the *optimal strategy for the firm will be maximum differentiation in terms of delivery time* so as to capture maximum demand and price premium. The optimal prices for this model (p_{10}^* and p_{20}^*) can be obtained by substituting $L_{10}^* = 0$ in Equations (5) and (6). The optimal prices and delivery time for this model are shown in Figure 1. However, as we noted before, in almost all practical scenarios there will be some capacity costs related with one or both of the products. We investigate such cases in the next section.

4. Capacitated Models

In this section, we introduce models with integrated capacity requirements and costs, and explore the impact of capacity-induced costs on the optimal prices and guaranteed delivery times, as well as the price and time differentiations.

4.1. Capacity Cost for Shorter Delivery Time Product P1 Only (Model 1)

The basic setting for this model remains the same as in Model 0, except for the fact that there is a capacity cost associated with processing the express product. It might correspond to a scenario in which the firm outsources only the slower service. Under such an outsourcing arrangement, it is quite credible that the firm would pay on a per-unit-of-service basis (m), subject to the subcontractor satisfying some service-level constraints, for example, L_2 and α. Hence, the firm does not have to incur capacity costs for the regular product. There are also other cases in which this model might be applicable. For example, many transportation companies rent "expensive," relatively small warehouses in the airport premises for express deliveries only, while the regular items are stored in a central large facility where the rental cost is much cheaper and capacity is not a major concern. Comparing this model with Model 0, we can gain insight into how different capacity environments can shape optimal product positioning and differentiation.

The problem for the firm in this case is:

$$\begin{aligned}\text{Maximize}\quad & \pi(p_1, p_2, L_1, \mu_1)\\ &= (p_1 - m)\lambda_1 - (p_2 - m)\lambda_2 - A_1\mu_1, \quad (8)\end{aligned}$$

where $\mu_1 = (-\ln(1-\alpha)/L_1) + \lambda_1$ as in Equation (4). As before, we first determine the optimal prices for a fixed L_1 and then determine the optimal L_1 (see Appendix A). Notice that $p_{21}^*(L_1)$ will remain the same as $p_{20}^*(L_1)$ in Equation (6), but $p_{11}^*(L_1)$ will now be of the form:

$$\begin{aligned}p_{11}^*(L_1) = \frac{A_1+m}{2} + \frac{a}{2\beta_p} + L_2\frac{\beta_p\theta_L - \theta_p\beta_L}{2\beta_p(\beta_p+2\theta_p)}\\ - L_1\frac{\beta_p\beta_L + \beta_p\theta_L + \theta_p\beta_L}{2\beta_p(\beta_p+2\theta_p)}. \quad (9)\end{aligned}$$

Comparing Equation (9) with Equation (5), it is evident that the capacity cost associated with P1 results in an upward shift in its optimal price by a factor of $(A_1/2)$ for any L_1 (see Figure 1).

Substituting $p_{11}^*(L_1)$ and $p_{21}^*(L_1)$ in Equation (8), we obtain the profit function in terms of L_1 only. A mild restriction on the parameter values can guarantee a unique interior maximizer, L_{11}^* (see Appendix A). Clearly, unlike the uncapacitated case, it would be prohibitively expensive for a firm to maximally differentiate products on the basis of time. Meaning that $L_{11}^* > 0$. A more interesting issue is the effect of the marginal capacity cost on L_{11}^*.

PROPOSITION 4. *The optimal delivery time L_{11}^* for* P1 *is increasing in A_1.*

PROOF. See Appendix B. □

Proposition 4 implies that *higher capacity costs induce less time-based differentiation*. This has an immediate impact on optimal prices and price differentiation.

PROPOSITION 5. *In comparison with the uncapacitated case*:

(i) *The optimal* P2 *price is higher for a PTD market and lower for a TPD market for any A_1, and the difference increases with A_1. Hence, the optimal* P2 *price is increasing in A_1 for a PTD market and decreasing for a TPD market.*

(ii) *There is a threshold value of* A_1 *below which the optimal* P1 *price is lower, and above which the optimal* P1 *price is higher. Implying that the optimal* P1 *price is not monotone in* A_1.

(iii) *There is a threshold value of* A_1 *below which price differentiation is less, and above which price differentiation is greater.*

(iv) *For any* A_1, *the optimal* P1 *demand is lower, the optimal* P2 *demand is higher, and the optimal total firm demand is lower.*

Proof. See Appendix B. □

The implication of Proposition 5 is that the price impact of capacity-induced costs for the express product is critically linked to both the marginal capacity cost and the market profile. We already know from Proposition 4 that a capacity cost A_1 results in less time differentiation, $L_{11}^* > L_{10}^*$. The effect of this on P2 is evident from the fact that $p_{21}^*(L_1) = p_{20}^*(L_1)$ which is given by Equation (6) and Proposition 1. Clearly, the optimal P2 price depends on the market characteristics, as stated in (i) (see also Figure 1). What is more striking and counterintuitive is the impact of capacity costs on the express product P1. Proposition 4(ii) shows that the price of P1 can actually decrease even when the associated marginal capacity cost increases. Likewise, (iii) shows that a low marginal capacity cost will lead to less price differentiation. Only when the marginal capacity cost is sufficiently high, should the firm increase the price for its express product and the price differentiation. Note the contrast with the uncapacitated case where price and time differentiation go hand-in-hand. Capacity cost and optimal prices have a predictable effect on customer demand as shown in (iv). Higher costs for the express product P1 reduces its optimal demand and induces more switching to P2, whereas the total demand decreases.

In summary, the managerial insights of this model is that *when offering a shorter delivery time product with an associated capacity cost, the firm should reduce time differentiation to maintain control over capacity induced costs. At the same time, it should decrease price differentiation if the marginal cost of capacity is low, and increase price differentiation if the cost is high. The optimal prices, on the other hand, will be governed by both the marginal cost of capacity and the market characteristic.*

4.2. Capacity Cost for Both Products (Model 2)

In reality, many firms would incur some capacity-related costs for both of its products (e.g., for Fed Ex there is a capacity-induced cost for both *FedEx Ground* and *FedEx Custom Critical*). Analysis of this more general case expands our understanding of the impact of capacity costs and structures on optimal product differentiation.

In this model, the firm incurs two capacity costs—a linear capacity cost A_1 for P1 and another linear investment cost A_2 for the regular product P2. We assume that $A_2 \le A_1$, because even when there is a marginal capacity cost for the regular product, we would expect it to be less than that for the express product (e.g., most time-definite delivery companies require more sophisticated information technology for coordination). The problem for the firm is:

$$\text{Maximize} \quad \pi(p_1, p_2, L_1, \mu_1, \mu_2) = (p_1 - m)\lambda_1 - (p_2 - m)\lambda_2 - A_1\mu_1 - A_2\mu_2, \tag{10}$$

where $\mu_1 = (-\ln(1-\alpha)/L_1) + \lambda_1$ and $\mu_2 = (-\ln(1-\alpha)/L_2) + \lambda_2$ as shown in Equation (4). Note that Model 1 can be interpreted as a special case of this model when $A_2 = 0$. Again, we first determine the optimal prices for a fixed L_1 and then determine the optimal L_1 (see Appendix A). Notice that $p_{12}^*(L_1)$ will remain the same as $p_{11}^*(L_1)$ in Equation (9), but $p_{22}^*(L_1)$ will now be:

$$p_{22}^*(L_1) = \frac{A_2 + m}{2} + \frac{a}{2\beta_p} - L_2 \frac{\beta_p\beta_L + \beta_p\theta_L + \theta_p\beta_L}{2\beta_p(\beta_p + 2\theta_p)} + L_1 \frac{\beta_p\theta_L - \theta_p\beta_L}{2\beta_p(\beta_p + 2\theta_p)}. \tag{11}$$

Comparing Equation (11) with Equation (6), it is evident that the capacity cost associated with P2 results in an upward shift in its optimal price by a factor of $(A_2/2)$ for any L_1 (see Figure 1). Substituting $p_{12}^*(L_1)$ and $p_{22}^*(L_1)$ in Equation (10), we can again guarantee an interior optimal L_{12}^* under some mild restrictions (see Appendix A). However, the optimal delivery time $L_{12}*$ will be different from that of Model 1, as indicated in the following proposition.

PROPOSITION 6. *When $A_1, A_2 > 0$,*

(i) *For any A_1 and A_2, the optimal* P1 *delivery time L^*_{12} is less than L^*_{11}. Furthermore, L^*_{12} is increasing in A_1, and decreasing in A_2.*

(ii) *As the difference between marginal capacity costs $(A_1 - A_2)$ decreases, the difference between the optimal delivery times $(L^*_{11} - L^*_{12})$ increases.*

(iii) *Assuming that the marginal capacity cost differential $(A_1 - A_2)$ remains constant, an increase in A_1 increases the optimal delivery time L^*_{12}. Implying that A_1 has a stronger effect on L^*_{12} than A_2.*

PROOF. See Appendix B. □

We would normally expect that an additional capacity related cost for P2 would increase the optimal delivery time for P1. Interestingly, (i) indicates the contrary. This has an intuitive explanation. Recall that in the presence of added capacity costs, the optimal P2 price would increase (see Equation 11). If the delivery time for P1 stayed the same or increased, this would decrease the total firm demand, and due to less time differentiation, customers would switch to P2, leading to even more capacity cost for the firm. By decreasing L_1, the firm not only generates more demand and induces more customers to buy P1, but also charges a premium to these customers. As indicated in (ii), the extent of increase in time differentiation is governed by the difference in the marginal capacity costs. For $A_2 = 0$, $L^*_{12} = L^*_{11}$, while for $A_2 = A_1$, the difference will be maximum. Note from (iii) that higher capacity costs induce less time-based differentiation for a constant capacity cost differential.

Proposition 6 clearly links the optimal time differentiation to absolute and relative costs of capacity for Model 2. The next proposition shows the impact of capacity costs on optimal prices and optimal demands for Model 2 in comparison with Model 1.

PROPOSITION 7. *In comparison with the case where there is no associated capacity costs for* P2 (*i.e.*, *Model* 1), *if there are capacity-induced costs for both products*, *then*, *for any $0 < A_2 \leq A_1$:*

(i) *The optimal* P1 *price will be higher and the difference will increase as $(A_1 - A_2)$ decreases. Hence, for a fixed A_1, the optimal* P1 *price for Model 2 is increasing in A_2.*

(ii) *The optimal* P2 *price will be higher for a TPD market and the difference will increase as $(A_1 - A_2)$ decreases. Implying that for a fixed A_1, the optimal* P2 *price for Model 2 is increasing in A_2. For a PTD market, the comparison will depend on the relative values of A_1 and A_2.*

(iii) *The optimal* P1 *demand will be higher and the optimal* P2 *demand will be lower, but the comparison of the optimal total firm demand will depend on the relative values of A_1 and A_2.*

PROOF. See Appendix B. □

From Proposition 6 we know that when there is also a capacity cost associated with the regular product, the firm will guarantee a shorter delivery time for P1. The effect of this on the optimal price for P1 is evident from (i) (Figure 1). Since the difference in optimal delivery times increases as $(A_1 - A_2)$ decreases, the price premium is the largest when $A_2 = A_1$. The effect on optimal P2 price (ii) is less clear because $p^*_{22}(L_1) = p^*_{21}(L_1) + A_2/2$ as in Equation (11) and the monotonicity of $p^*_{22}(L_1)$ depends on the market characteristics. For a TPD market, $p^*_{22}(L_1)$ is decreasing in L_1, therefore the same argument for P1 applies. However, for a PTD market, $p^*_{22}(L_1)$ is increasing in L_1, and we cannot ascertain any conclusion analytically. Nevertheless, our extensive numerical studies indicate that the same intuition also applies for the PTD case (i.e., optimal P2 price also increases). An immediate question then is what happens to the optimal price differentiation between the two products. Our numerical studies indicate that price differentiation decreases, despite the fact that the prices of both products increase. The effect on optimal demands (iii) is more predictable. An added capacity cost for the regular product decreases its optimal demand, but increases the demand of the express product. Our numerical experiments suggest that the optimal firm demand also decreases, although we cannot analytically establish this unequivocally.

The above analysis sheds light on the significant changes in the optimal pricing and delivery time strategies when firms face the possibility of incurring capacity-induced costs not only for the express product, but also for the regular one. *The new strategy should be to focus more on time-based differentiation by guaranteeing a shorter delivery time for the express product. At the same time, the firm should increase prices of both products, while making sure that price differentiation decreases.*

While Propositions 6 and 7 focused on the comparison between Model 1 and Model 2, in the following proposition we compare Model 2 ($0 < A_2 \leq A_1$) with Model 0 ($A_1 = A_2 = 0$). Clearly, $L_{12}^* > L_{10}^* = 0$, which implies less time-based differentiation.

PROPOSITION 8. *In comparison with the uncapacitated case, in Model* 2 ($0 < A_2 \leq A_1$):

(i) *There is a threshold value of* A_2 *(say* A_2'*): If* $A_2 > A_2'$*, then the optimal* P1 *price is higher for any* A_1*. If* $A_2 < A_2'$*, there is a threshold of* A_1*, below which the optimal* P1 *price is lower and above which the optimal* P1 *price is higher.*

(ii) *The optimal* P2 *price is higher for a PTD market and the difference is increasing in* A_2 *(for a fixed* A_1*). In the case of a TPD market, there are two threshold values of* A_2 *(say* A_2'' *and* $A_2''' > A_2''$*): If* $A_2 > A_2'''$*, then the optimal* P2 *price is higher for any* A_1*. If* $A_2 < A_2''$*, the optimal* P2 *price is lower for any* A_1*. If* $A_2'' \geq A_2 \geq A_2'''$*, there is a threshold of* A_1*, below which the optimal* P2 *price is higher and above which the optimal* P2 *price is lower.*

(iii) *Assuming that the marginal capacity cost differential* ($A_1 - A_2$) *remains constant, an increase in* A_1 *decreases the optimal price differentiation.*

(iv) *For any fixed* A_2*, there is a threshold* A_1 *below which the optimal price differentiation is less and above which the optimal price differentiation is greater.*

(v) *The optimal total firm demand is smaller.*

PROOF. See Appendix B. □

Proposition 8*, in essence, generalizes the results of Proposition* 5. It substantiates our earlier claim that a capacity cost does not necessarily increase the prices of the product or the price differentiation. Only when the sum of the marginal capacity costs is sufficiently high, the optimal prices of both products increase. Price differentiation, however, depends on both the absolute and the relative marginal capacity costs. If the marginal capacity cost differential remains constant, an increase in capacity costs decreases price differentiation as well (iii). As far as the effect of capacity cost differential is concerned, if the marginal capacity costs are high but equal, then price differentiation is lower. On the other hand, if the marginal capacity costs are low, a high capacity cost differential can still make the price differentiation higher (iv). Demand implication of capacity costs (v), on the other hand, is rather intuitive.

We note that our findings in this paper regarding the properties of optimal product differentiation are quite robust even if the analysis is done with an increasing convex capacity cost (e.g., a quadratic capacity cost structure for each product/facility of the form $A_j\mu_j^2$). The primary difference is that under more expensive capacity cost structures (i.e., convex), the absolute capacity costs become more dominant over the cost differentials. Also, as the capacity investment cost becomes more expensive, it becomes more important for managers to explicitly relate capacity to demand.

4.3. Numerical Illustrations

We have performed extensive numerical studies for our models. In this subsection we present a small sample of these studies to illustrate some of our results (refer to Table 1). We focused on three different markets: (1) Strictly PTD ($\beta_p\theta_L/\beta_L\theta_p = 5$; $\beta_p = 50$, $\theta_L = 25$, $\beta_L = 25$, $\theta_p = 10$), (2) Neither PTD nor TPD ($\beta_p\theta_L/\beta_L\theta_p = 1$; $\beta_p = 50$, $\theta_L = 10$, $\beta_L = 50$, $\theta_p = 10$), and (3) Strictly TPD ($\beta_p\theta_L/\beta_L\theta_p = 0.2$; $\beta_p = 25$, $\theta_L = 10$, $\beta_L = 50$, $\theta_p = 25$). In the table we present the optimal decision variable values for each market, as well as the optimal price differential ($\Delta^* = p_1^* - p_2^*$). We kept the values of the following parameters at a constant

Table 1 Numerical Results

	Strictly PTD Market				Neither PTD nor TPD Market				Strictly TPD Market			
A_1, A_2	L_1^*	p_1^*	p_2^*	Δ^*	L_1^*	p_1^*	p_2^*	Δ^*	L_1^*	p_1^*	p_2^*	Δ^*
0, 0	0	11.93	10.32	1.61	0	11.50	10.00	1.50	0	20.70	19.30	1.40
7, 0	0.66	15.17	10.42	4.75	0.38	14.81	10.00	4.81	0.22	24.04	19.24	4.80
7, 3.5	0.52	15.22	12.14	3.08	0.37	14.82	11.75	3.06	0.21	24.04	20.99	3.05
7, 7	0.44	15.25	13.88	1.37	0.36	14.83	13.50	1.32	0.20	24.04	22.74	1.30
5, 1.5	0.39	14.27	11.13	3.14	0.28	13.86	10.75	3.11	0.17	23.07	20.00	3.07

level: $a = 1000$, $m = 3$, $\alpha = 0.99$, and $L_2 = 3$ (hence, the optimal delivery time differential $= 3 - L_1^*$).

Comparing Row 1 and Row 2 in the table, it is clear that a marginal capacity cost associated with P1 increases its optimal delivery time (Proposition 4). Since this marginal cost is quite high, the optimal P1 price also increases. However, while optimal P2 price increases for "strictly PTD" and "neither PTD nor TPD" markets, it actually decreases for a "strictly TPD" market (Proposition 5). Comparing Row 3 and Row 5, we note that the optimal delivery time for P1 increases with its capacity cost. Meaning that time differentiation decreases (Proposition 6) and so does the optimal price differentiation (Proposition 8) as long as the capacity cost differential remains constant. On the other hand, from Rows 2, 3, and 4, we observe that decrease in capacity cost differentials increases time differentiation (Proposition 6); however, the optimal prices of the products increases and the optimal price differentiation decreases, independent of the market characteristics (Proposition 7). Also note that our assertion in Proposition 8, that the price differentiation for a capacitated firm will be less than that for an uncapacitated one when the marginal capacity costs are equal remains true (Row 1 and Row 4).

5. The Effect of Substitution

One of the distinct features of our model is that we take into account substitutability of the products, in terms of both price and delivery time. In order to address the effect of substitution, we focus on Model 2 (since Model 1 is a special case of Model 2), and set the substitution rates $\theta_p = \theta_L = 0$. This analysis can serve two managerial purposes. It can show *the consequence of neglecting the substitution effect when it is present*, and also *the impact of presence of substitutable products* on a firm's differentiation strategy.[6]

The optimal prices of P1 and P2, assuming L_1 is fixed, will be obtained by substituting θ_p, $\theta_L = 0$ in Equations (9) and (5), respectively. Note that the P2 (regular) price will now be independent of L_1. Though the price of P1 will still be decreasing in L_1, the slope will be different from Model 2. Comparing the price functions with Model 2, we can have the following proposition.

Proposition 9. *For a particular value of L_1, not taking into account the substitution effect, will lead to a lower optimal* P1 *price and a higher optimal* P2 *price* (*i.e., less price differentiation*) *for a PTD market and a higher optimal* P1 *price and a lower optimal* P2 *price* (*i.e., more price differentiation*) *for a TPD market.*

Proof. Immediate. □

Obviously, not taking into account product substitutability means ignoring a crucial feature of the market in which the firm operates, and would result in suboptimal delivery times and prices. In general, the degree of suboptimality would depend on the capacity costs A_1, A_2, the differential $(A_1 - A_2)$, as well as the market characteristics. The next proposition shows that indeed the effect of substitution on optimal product differentiation can be in either direction.

Proposition 10. *When $A_1 = A_2$, not taking into account substitution would result in:*

(i) *Longer optimal* P1 *delivery time (i.e., less time differentiation), lower optimal* P1 *price, and higher optimal* P2 *price (i.e., less price differentiation) for a PTD market.*

(ii) *Shorter optimal* P1 *delivery time (i.e., more time differentiation), higher optimal* P1 *price, and lower optimal* P2 *price (i.e., more price differentiation) for a strongly TPD market.*

(iii) *Longer optimal* P1 *delivery time (i.e., less time differentiation) and lower optimal* P2 *price for a mildly TPD market. However, the comparison for optimal* P1 *price and price differentiation in this case will depend on the values of A_1 and A_2.*

Proof. See Appendix B. □

Without θs, the PTD market is simply a price sensitive market ($\beta_p > \beta_L$). Taking substitutability into account reduces the effective price sensitivity of such a market. Hence, firms can increase time differentiation by guaranteeing a shorter delivery time for the express product and extract a price premium. In order to maintain customer demand, the firm reduces the price of P2. Similarly, for a TPD market, taking note of cannibalization reduces the effective time sensitivity of the market.

[6] Note that in an empirical study, if the possibility of substitution is ignored (i.e., if θs are assumed to be zero), the estimates of βs might also change. We do not account for this possibility in our analysis.

6. Conclusions and Future Research

In this paper, we studied a single profit-maximizing firm selling two variants of a product in a price and time sensitive market. Our primary objective has been to explore the effect of capacity-related costs on the optimal prices and delivery time decisions, which generate insights on the optimal positioning of the two alternatives. We developed and analyzed a model that can determine the firm's optimal product positioning and also explicitly relates prices and uniform delivery time guarantees to capacity requirements and costs under high service levels.

A more in-depth comparison of the uncapacitated and capacitated models leads to important managerial lessons. First of all, it shows that managers need to understand their market's relative sensitivity towards price and speed, which governs optimal product positioning and product differentiation strategies. We categorized markets as "price and time difference sensitive" (PTD) and "time and price difference sensitive" (TPD). Naturally, the optimal prices are different for PTD and TPD markets. We found that the underlying market type has the most profound impact on the optimal price of the regular product P2. These effects are summarized in Table 2.

With respect to capacity costs, we show that managers have to consider both the absolute as well as the relative capacity costs in positioning their products. As capacity becomes more expensive, firms would move away from differentiation based on price and speed, provided relative costs remain the same. If the difference in costs increase, the price-based differentiation increases, but time-based differentiation decreases. We note that under heavier capacity structures, managers have to relate product differentiation and the resulting demands more closely with the capacity requirements because the cost impact would be more significant. As a result, absolute capacity costs become more important than relative costs in optimal product positioning.

The comparison of different capacitated models (capacity cost for no/one/both product(s)) that we present in this paper yields interesting insights on the product differentiation strategies of firms facing different capacity environments and how firms should respond to a change in its internal operating dynamics. As we have indicated, if a firm that is outsourcing all its products to a facility with ample supply capacity decides to internalize one product and its associated capacity costs, it does not necessarily have to increase the product's price or the price differentiation between the two products. When the internal capacity cost is high enough, such an action is justifiable. When the firm decides to also internalize the other product, it should increase both products' prices and differentiate more in time. A managerially significant implication is *firms that are more tightly constrained in supply options have a larger preference for a time-based strategy*, provided there is a capacity cost involved with at least one product. Alternatively, when a firm producing internally decides to outsource part of its operation to a facility with abundant supply capability and maintain only the high price-premium product in-house, it should differentiate more in terms of price and less in terms of time.

There are a number of model assumptions that are worth revisiting. Relaxation of some of these assump-

Table 2 Effects of Market Characteristics on Optimal Differentiation Strategy

	PTD Market	TPD Market
General Observation	Optimal P2 price is increasing in L_1	Optimal P2 price is decreasing in L_1
Comparison between Model 0 and Model 1	$p^*_{21} > p^*_{20}$; p^*_{21} and $(p^*_{21} - p^*_{20})$ increasing in A_1	$p^*_{21} < p^*_{20}$; p^*_{21} is decreasing and $(p^*_{20} - p^*_{21})$ increasing in A_1
Comparison between Model 1 and Model 2	$p^*_{22} > p^*_{21}$; p^*_{22} and $(p^*_{22} - p^*_{21})$ increasing in A_2 (for a fixed A_1)	$p^*_{22} > p^*_{21}$; p^*_{22} and $(p^*_{22} - p^*_{21})$ increasing in A_2 (for a fixed A_1)
Comparison between Model 0 and Model 2	$p^*_{22} > p^*_{20}$	$p^*_{22} < p^*_{20}$ below A''_2; $p^*_{22} > p^*_{20}$ above A'''_2; $p^*_{22} \gtreqless p^*_{20}$ for $A''_2 \geq A_2 \geq A'''_2$
Effect of not taking θs into account, $A_1 = A_2$	Longer L^*_1, lower p^*_1 and higher p^*_2	Shorter L^*_1, higher p^*_1 and lower p^*_2

tions can pave the way for interesting extensions which are left for future research. One critical assumption in our model is that the two products have dedicated facilities. Although practical examples of such systems are abundant, a shared capacity framework can shed further light on the interaction between capacity and product differentiation. In particular, it would be possible to study the impact of substitution on the capacity allocation. Such a model would require a different methodology (i.e., priority queueing framework as in Mendelson and Whang 1990), but can lead to additional managerial insights. Another set of assumptions that can be relaxed is the fixedness of the delivery time for the regular product (L_2) and the delivery reliability level (α). Presumably, both of these elements can be decision variables affecting the demand curve. Incorporating L_2 as a decision variable would definitely yield a more general model. However, our preliminary analysis suggests that most of the insights generated from our study would remain valid as long as L_2 is constrained to be greater than L_1. Incorporating α into our demand model might be more interesting because this would bring forth a new trade-off between offering shorter delivery times and higher service levels. This would be particularly interesting in a shared capacity framework. In a similar vein, our model can be extended to accommodate more general demand functions (e.g., Cobb-Douglas form). Though we have not studied this extension, we would expect most of our results to remain true. However, some of the proofs would surely run into the problem of analytical tractability. Another direction of research would be to analyze a competitive scenario. Since the first step of any competitive analysis is to determine the optimal reactions of each firm in the market, our current model will serve as an integral part of such an extension.

Acknowledgments

The authors thank two anonymous referees, the Senior Editor, and the Deputy Editor for their valuable comments. This research is supported by grants from the Natural Sciences and Engineering Research Council of Canada.

Appendix A. Optimization of the Profit Function for Linear Capacity Cost

We assume that the parameter values are restricted so that $\lambda_j > 0$ and $p_j > m$. We also assume that $L_1 < L_2$. Differentiating π obtained by replacing Equation (4) in Equation (3) (which implies $\mu_j > \lambda_j$) with respect to (wrt) p_1 we obtain, $\partial^2\pi/\partial p_1^2 = -2\beta_p - 2\theta_p < 0 \Rightarrow \pi$ is concave in p_1. Substituting $p_1^*(p_2, L_1)$, obtained by solving $\partial\pi/\partial p_1 = 0$, into π and differentiating twice wrt p_2 we have $-(2\beta_p(\beta_p + 2\theta_p))/(\beta_p + \theta_p) \Rightarrow \pi(p_2, L_1)$ is concave in p_2. Hence, we can determine $p_2^*(L_1)$ by solving $\partial\pi(p_2, L_1)/\partial p_1 = 0$ and $p_1^*(L_1)$ from $p_1^*(p_2^*(L_1), L_1)$, the expressions for which are given by Equations (9) and (11), respectively. Substituting the optimal prices into the profit function and differentiating wrt L_1 we have:

$$\frac{\partial\pi(L_1)}{\partial L_1} = (-\beta_L - \theta_L)(p_1^*(L_1) - m - A_1) + \theta_L(p_2^*(L_1) - m - A_2) - \frac{A_1\ln(1-\alpha)}{L_1^2}, \tag{A1}$$

and

$$\frac{\partial^2\pi(L_1)}{\partial L_1^2} = -(\beta_L + \theta_L)\frac{\partial p_1^*(L_1)}{\partial L_1} + \theta_L\frac{\partial p_2^*(L_1)}{\partial L_1} + \frac{2A_1\ln(1-\alpha)}{L_1^3}. \tag{A2}$$

As $L_1 \to 0$, the profit function is increasing concave in L_1 for any $A_1 > 0 \Rightarrow$ for any $A_1 > 0$, the optimal delivery time for P1 will be positive. Since (A2) $= 0$ has a unique solution (it starts from $-\propto$ and is increasing), say L_U, the profit function is concave until L_U and then convex. Hence, $\partial\pi/\partial L_1 < 0$ for $L_1 = L_2$ is sufficient to guarantee an interior profit-maximizing solution. The condition simplifies to:

$$\frac{-a\beta_L + (A_1 + m)\beta_L\beta_p + (A_1 + A_2)\theta_L\beta_p}{2\beta_p} + \frac{\beta_L^2}{2\beta_p}L_2 - \frac{A_1\ln(1-\alpha)}{L_2^2} < 0. \tag{A3}$$

Note that a necessary condition for Equation (A3) to hold is a to be high. A sufficiently high value of a also guarantees $\lambda_j > 0$ and $p_j > m$. We will refer to the expression (A1) $= 0$ as the first order condition (FOC). Model 0 and Model 1 will be special cases of the above with $A_2 = 0$ and $A_1 = A_2 = 0$, respectively. Note that for Model 0 the profit function will be decreasing convex in L_1.

Appendix B.[7]

Proof of Proposition 2. Substituting the values of $p_1^*(L_1)$ and $p_2^*(L_1)$ (for any A_1, A_2) in λ_1 and λ_2 of Equations (1) and (2), respectively, the demand functions can be expressed in terms of the single decision variable L_1. Differentiating λ_1 wrt L_1 we have $(\partial\lambda_1(L_1))/\partial L_1 = ((-\beta_L - \theta_L)/2) < 0$. Similar straightforward calculation proves the claims about λ_2 and $(\lambda_1 + \lambda_2)$. □

[7] Note that we have tried to keep the proofs clear and concise for expository purposes. Some expressions in the proofs are extremely lengthy which prohibits explicit presentation. In such circumstances, rather than providing the detailed expressions, we present their general structure and behavior.

Proof of Proposition 3. Substituting A_1, $A_2 = 0$ in Equations (A1) and (A2), we note that Equation (A2) > 0, implying π is convex in L_1. Also from Equation (A1) we can see that for $p^*_{10}(L_1) > p^*_{20}(L_1)$, π is decreasing in L_1. From Equation (7) it is clear that for $L_1 < L_2$, $p^*_{10}(L_1) > p^*_{20}(L_1)$. Hence, the profit for the firm is decreasing convex in L_1 on $[0, L_2]$, which implies that $L^*_{10} = 0$. □

Proof of Proposition 4. Total differentiation of the FOC for Model 1 wrt A_1 yields:

$$\frac{\partial L^*_{11}}{\partial A_1} = -\left(\frac{\partial^2\pi/\partial L_{11}\partial A_1}{\partial^2\pi/\partial L^2_{11}}\right)\bigg|_{L_1=L^*_{11}} \quad \text{where}$$

$$\frac{\partial^2\pi}{\partial L_1\partial A_1}\bigg|_{L_1=L^*_{11}} = \frac{-\ln(1-\alpha)}{L^{*2}_{11}} + \frac{\beta_L}{2} > 0. \tag{A4}$$

Since we know that

$$\frac{\partial^2\pi}{\partial L^2_{11}}\bigg|_{L_1=L^*_{11}} < 0, \quad \frac{\partial L^*_{11}}{\partial A_1} > 0. \quad \square$$

Proof of Proposition 5. (i) The result follows because $p^*_2(L_1)$ is increasing in L_1 for a PTD market and decreasing in L_1 for a TPD market and $L^*_{11} > L^*_{10} = 0$. The second part follows since L^*_{11} is increasing in A_1, and hence we can conclude that p^*_{21} is increasing in A_1 for a PTD market and decreasing for a TPD market (because p^*_{20} is independent of A_1).

(ii) Comparing the optimal prices for Model 0 and Model 1 it is clear that if $L^*_{11} > (A_1/2)\left|(\partial p^*_1(L_1)/\partial L_1)\right|$ (def. $L_C(A_1)$), $p^*_{11} < p^*_{10}$. If $L^*_{11} < L_C(A_1)$, $p^*_{11} > p^*_{10}$. Substituting $L_1 = L_C(A_1)$ in the FOC for Model 1 we can show that the expression will be of the form[8] $f(A_1) = E_1 + E_2A_1 + (E_3/A_1)$ where $E_3 > 0$. Therefore, f is convex in A_1. Also as $A_1 \to 0$, f will be positive and decreasing. Hence, for low values of A_1, $L^*_{11} > L_C(A_1)$. Hence, $p^*_{11} < p^*_{10}$. Recall that $(\partial\pi(L_1)/\partial L_1) < 0$ for $L_1 = L_2$. Hence, for high values of A_1, $L^*_{11} < L_C(A_1)$. Hence, $p^*_{11} > p^*_{10}$. Since f is convex, initially positive, and is negative for A_1 that makes $L_C(A_1) = L_2$, it is clear that there is a unique threshold A_1 below which $p^*_{11} < p^*_{10}$ and above which $p^*_{11} > p^*_{10}$.

(iii) The proof for this part follows from similar arguments in (ii) except that now

$$L_C(A_1) = \frac{A_1/2}{|(\partial p^*_1(L_1)/\partial L_1) - (\partial p^*_2(L_1)/\partial L_1)|}.$$

(iv) Optimal demand of P2 for Model 1 − Optimal demand of P2 for Model 0

$$= \frac{\theta_p A_1}{2} + L^*_{11}\left[-(\beta_p+\theta_p)\frac{\partial p^*_2(L_1)}{\partial L_1} - \theta_p\frac{\partial p^*_1(L_1)}{\partial L_1} + \theta_L\right],$$

which can be shown to be positive. Optimal total demand for Model 1 − Optimal total demand for Model 0

$$= -\frac{\beta_p A_1}{2} - \frac{\beta_L}{2}L^*_{11},$$

which is negative. Because the optimal demand for P2 increases while the total demand decreases, it implies that the optimal demand for P1 must decrease. □

Proof of Proposition 6. (i) $[\partial\pi(L_1)/\partial L_1]$ for Model 1 − $[\partial\pi(L_1)/\partial L_1]$ for Model 2 = $(A_2\theta_L)/2 > 0$. So, FOC for Model 2 at L^*_{11} is negative, which implies $L^*_{12} < L^*_{11}$. $[\partial^2\pi/\partial L_1\partial A_1]$ at $(L_1 = L^*_{12}) = [(\beta_L+\theta_L)/2 - (\ln(1-\alpha)/L^2_1)] > 0 \Rightarrow$ for a fixed A_2, L^*_{12} is increasing in A_1. Similarly, $[\partial^2\pi/\partial L_1\partial A_2]$ at $(L_1 = L^*_{12}) = -\theta_L/2 < 0 \Rightarrow$ for a fixed A_1, L^*_{12} is decreasing in A_2.

(ii) Let the difference between the marginal capacity costs be d $(0 \le d \le A_1)$. Replacing $A_2 = A_1 - d$ in Model 2 and differentiating the FOC wrt d at $L_1 = L^*_{12}$ gives: $[\partial^2\pi/\partial L_1\partial d]$ at $(L_1 = L^*_{12}) = \theta_L/2 > 0$. Then following Proposition 4, we can show that $\partial L^*_{12}/\partial d > 0$. It is also clear that L^*_{11} is independent of d. Hence, $(L^*_{11} - L^*_{12})$ is decreasing in d.

(iii) Differentiating the FOC for Model 2 wrt A_1 at $L_1 = L^*_{12}$ (with $A_2 = A_1 - d$ and constant d) we have

$$\frac{\partial^2\pi}{\partial L_1\partial A_1}\bigg|_{L_1=L^*_{12}} = \frac{-\ln(1-\alpha)}{L^{*2}_{12}} + \frac{\beta_L}{2} > 0.$$

Now following Proposition 4 we can show that $\partial L^*_{12}/\partial A_1 > 0$. □

Proof of Proposition 7. (i) $p^*_{12} - p^*_{11} = (L^*_{12} - L^*_{11})(\partial p^*_1(L_1)/\partial L_1) > 0$, (based on Proposition 1 and Proposition 6), which proves the first part. As the difference in marginal capacity costs decreases, the difference in the optimal delivery times, $L^*_{11} - L^*_{12}$, increases (Proposition 6), which proves the second part. Since, for a fixed A_1, p^*_{11} is a constant and the decrease of $(A_1 - A_2)$ is synonymous with increase of A_2, we have the third result.

(ii) $p^*_{22} - p^*_{21} = (A_2/2) + (L^*_{12} - L^*_{11})(\partial p^*_2(L_1)/\partial L_1)$. Since p^*_2 is decreasing in L_1 for a TPD market, we can conclude that $p^*_{22} > p^*_{21}$ for a TPD market. Also based on Proposition 1 and Proposition 6 we can establish $(p^*_{22} - p^*_{21})$ is increasing in A_2 for a fixed A_1 which also implies that p^*_{22} is increasing in A_2 for a fixed A_1. However, for a PTD market, p^*_2 is increasing in L_1. In that case, the comparison will depend both on A_1 and A_2. Note that the difference in the optimal price differentiation for the two models,

$$(p^*_{12} - p^*_{22}) - (p^*_{11} - p^*_{21})$$
$$= (-A_2/2) + (L^*_{12} - L^*_{11})\left(\frac{\partial p^*_1(L_1)}{\partial L_1} - \frac{\partial p^*_2(L_1)}{\partial L_1}\right)$$

the sign of which will also depend on both A_1 and A_2.

(iii) Demand for P1 for Model 2 − Demand for P1 for Model 1 = $(A_2\theta_p/2) + (L^*_{11} - L^*_{12})[(\theta_L + \beta_L)/2] > 0$, since $L^*_{11} - L^*_{12} > 0$. Demand for P2 for Model 2 − Demand for P2 for Model 1 = $(-A_2[\beta_p + \theta_p]/2) + (L^*_{12} - L^*_{11})(\theta_L/2) < 0$. However, the sign for difference in the total demands = $((A_2\beta_p/2) - (L^*_{11} - L^*_{12})(\beta_L/2))$ will depend on both A_1 and A_2. □

Proof of Proposition 8. (i) Comparing the optimal prices of P1 for Model 0 and Model 2 it is clear that if $L^*_{12} > L_C(A_1) = ((A_1/2)/|\partial p^*_1(L_1)/\partial L_1|)$, $p^*_{12} < p^*_{10}$. If $L^*_{12} < L_C(A_1)$, $p^*_{12} > p^*_{10}$. Substituting $L_1 = L_C(A_1)$ in FOC for Model 2 and assuming that $A_1 = A_2 = A$, we can show that the expression will be of the form $g(A) = E_1 + E_2A + (E_3/A)$ where $E_3 > 0$. Then, following the Proof of Proposition 5(ii), we can conclude that there is a threshold value A' of A

[8] Note that when we use the terms E_i in different sections of the proof, it signifies different values, for example, E_1 in the Proof of Proposition 5 is different from E_1 in the Proof of Proposition 6(i), which is not the same as E_1 in the Proof of Proposition 6(ii).

below which $p^*_{12} < p^*_{10}$ and above A', $p^*_{12} > p^*_{10}$. Repeating the same process as above for a fixed A_2 results in an expression $h(A_1) = E_4 + E_5 A_1 + (E_6/A_1)$ where $E_6 > 0$. Note that when $A_1 = A_2 = A$, $h(A_1 = A) = g(A)$. The general behavior of the expressions g and h are the same; however, for $h(A_1)$ we are only interested in $A_1 \geq A_2$. The value of the fixed A_2 will determine whether $h(A_1)$ will have any threshold A_1 (like A') or not. If $A_2 < A'$, then $h(A_1)$ will be positive at that $A_1 = A_2$ (since g will be positive for $A_1 = A_2 = A < A'$). Based again on the arguments of Proof of Proposition 5(ii), there will be a unique $A_1 > A_2$, below which $L^*_{12} > L_C(A_1) \Rightarrow p^*_{12} < p^*_{10}$, and above which $L^*_{12} < L_C(A_1) \Rightarrow p^*_{12} > p^*_{10}$. If $A_2 > A'$, then at $A_1 = A_2$, $h(A_1)$ will be negative (since g will be negative at $A_1 = A_2 = A > A'$) and will remain so for the relevant $A_1 > A_2$. Hence, $p^*_{12} > p^*_{10}$ for any $A_1 > A_2$.

(ii) $p^*_{22} - p^*_{20} = (A_2/2) + L^*_{12}(\partial p^*_2(L_1)/\partial L_1)$. For a PTD market, $(\partial p^*_2(L_1)/\partial L_1) > 0$; hence, $p^*_{22} > p^*_{20}$. Because p^*_{20} is independent of A_2 while p^*_{22} is increasing in A_2, $(p^*_{22} - p^*_{20})$ is increasing in A_2. For TPD market, note that $(p^*_{22} - p^*_{20})$ is decreasing in A_1 (from Propositions 1 and 6). Let

$$L_C(A_2) = \frac{(A_2/2)}{|(\partial p^*_2(L_1)/\partial L_1)|}.$$

Comparing the optimal prices of P2 for Model 0 and Model 2 it is clear that if $L^*_{12} > L_C(A_2)$, $p^*_{22} < p^*_{20}$. If $L^*_{12} < L_C(A_2)$, $p^*_{22} > p^*_{20}$. Following a similar argument as in (i), we can obtain an expression $g(A)$ for $A_1 = A_2 = A$, and hence a threshold A'' below which $p^*_{22} < p^*_{20}$ and above A'', $p^*_{22} > p^*_{20}$. Now, repeating the process for a fixed A_2 as in (i) results in $h(A_1)$ and we are interested in $A_1 \geq A_2$. Note that $h(A_1)$ is linear increasing in A_1. If $A_2 < A''$, then $h(A_1)$ will be positive at $A_1 = A_2 \Rightarrow p^*_{22} < p^*_{20}$, and hence it will remain so for any relevant $A_1 \geq A_2$. If $A_2 > A''$, then at $A_1 = A_2$, $h(A_1)$ will be negative. If A_2 is large ($> A'''$), then a starting value of $h(A_1)$ will be highly negative, and hence even though $h(A_1)$ is linear increasing in A_1, it will never become positive $\Rightarrow p^*_{22} > p^*_{20}$. If $A''_2 \geq A_2 \geq A'''_2$, $h(A_1)$ will be initially negative and then positive which proves the proposition.

(iii)

$$(p^*_{12} - p^*_{22}) - (p^*_{10} - p^*_{20}) = [(A_1 - A_2)/2] + L^*_{12}\left(\frac{\partial p^*_1(L_1)}{\partial L_1} - \frac{\partial p^*_2(L_1)}{\partial L_1}\right).$$

For a fixed $A_1 - A_2 = d$, the difference is decreasing in A_1, since L^*_{12} is increasing in A_1 and the other term (difference in the slopes of price functions) is a negative constant.

(iv) We have the expression for the difference in optimal price differentiations in (iii). For this part,

$$L_C(A_1) = \frac{(A_1 - A_2)/2}{|(\partial p^*_1(L_1)/\partial L_1) - (\partial p^*_2(L_1)/\partial L_1)|}.$$

It is clear that when $A_2 = A_1$, $(p^*_{12} - p^*_{22}) > (p^*_{10} - p^*_{20})$. Substituting $L_1 = L_C(A_1)$ in FOC for Model 2 we can show that the resultant expression (def. $k(A_1)$) will be convex and will be positive and decreasing as $A_1 \to A_2$. Then following Proofs of Propositions 5(ii) and 8(i) we can prove that there will be a threshold $A_1 > A_2$, below which $k(A_1)$ will be positive (i.e., $L^*_{12} > L_C$). Implying $(p^*_{12} - p^*_{22}) < (p^*_{10} - p^*_{20})$ and above which $k(A_1)$ will be negative (i.e., $L^*_{12} < L_C$), implying $(p^*_{12} - p^*_{22}) > (p^*_{10} - p^*_{20})$.

(v) Difference in total demand

$$(\text{Model } 2 - \text{Model } 0) = [-(A_1 + A_2)(\beta_p/2)] - L^*_{12}(\beta_L/2) < 0. \quad \square$$

Proof of Proposition 10. Substituting θ_p, $\theta_L = 0$ in Equation (A1) we obtain the expression of $\partial\pi/\partial L_1$ for the model without θs and comparing that with Equation (A1) we have: $(\partial\pi/\partial L_1)$ for Model $2 - (\partial\pi/\partial L_1)$ for the model without

$$\theta\text{s} = (L_1 - L_2)\frac{\beta_L(\beta_p\theta_L - \beta_L\theta_p) + \beta_p\theta_L(\beta_L + 2\theta_L)}{2\beta_p(\beta_p + 2\theta_p)} + \frac{(A_1 - A_2)\theta_L}{2}.$$

For $A_1 = A_2$, the expression will be negative for a PTD market. Hence, the optimal delivery time will be longer if we ignore the substitution effect. Even when $\beta_p\theta_L < \beta_L\theta_p$ but the difference is not large, the same result will hold. However, when $\beta_p\theta_L \ll \beta_L\theta_p$, the expression will be positive, and then the optimal delivery time will be shorter if we ignore the substitution effect.

For a PTD market, assuming that the optimal L_1 for both cases (i.e., with and without θs) is equal to the L^*_1 without θs, we know from Proposition 9 that ignoring θ will lead to a lower optimal P1 price, higher optimal P2 price, and less price differentiation. Because for the model with θs optimal delivery time is shorter, the optimal P1 price is decreasing in L_1 and optimal P2 price is increasing in L_1, we obtain (i). Using the same argument we can attain (ii) for a strongly TPD market. However, for a mildly TPD market, the optimal prices and price differentiation will behave similar to the result for TPD market in Proposition 9, whereas the optimal delivery time for P1 will behave like a PTD market. Then, using the same line of reasoning as before, we can show that the optimal P2 price will decrease. However, note that the comparison of the optimal P1 price will depend on the difference in the optimal delivery times of the two models and the difference in the slope of the optimal P1 functions. Because we cannot determine the effect on the optimal P1 price, we cannot decide on the effect on optimal price differentiation. $\square$

References

Anderson, S. P., A. De Palma, J. F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge, MA.

Baker, W., M. Marn, C. Zawada. 2001. Price smarter on the net. *Harvard Bus. Rev.* February 2–7.

Ballou, R. H. 1998. *Business Logistics Management*. Prentice Hall, Upper Saddle River, NJ.

Birge, J. R., J. Drogosz, I. Duenyas. 1998. Setting single-period optimal capacity levels and prices for substitutable products. *The Internat. J. Flexible Manufacturing Systems* **10** 407–430.

Blackburn, J. D., T. Elrod, W. B. Lindsley, A. J. Zahorik. 1992. The strategic value of response time and product variety. C. A. Voss, ed. *Manufacturing Strategy—Process and Content*. Chapman and Hall, London, U.K.

Bowman, R. J. 2001. Bearing in on a modern supply chain for Europe. *Global Logistics and Supply Chain Strategies*, April, ⟨*www.SupplyChainBrain.com*⟩.

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

Chopra, S., P. Meindl. 2001. *Supply Chain Management: Strategy, Planning and Operation*. Prentice Hall, Upper Saddle River, NJ.

Dewan, S., H. Mendelson. 1990. User delay costs and internal pricing for a service facility. *Management Sci.* **36** 1502–1507.

Farris, M. T. 2002. The rise of premium transportation through technology. Working paper, Center for Logistics Education and Research, University of North Texas, Denton, TX.

Foster, T. 1999. Expedited explodes. *Logistics Management and Distribution Report*, September.

Fuller, J. B., J. O'Conor, R. Rawlinson. 1993. Tailored logistics: The next advantage. *Harvard Business Rev.* (May–June) 87–98.

Geary, S., J. P. Zonnenberg. 2000. What it means to be best in class. *Supply Chain Management Rev.* **4** 42–49.

Heskett, J. L. 1994. Controlling customer logistics service. *Internat. J. Phys. Distribution and Logist. Management* **24** 4–10.

Hum, S-H., H-H. Sim. 1996. Time-based competition: Literature review and implications for modelling. *Internat. J. Oper. Prod. Management* **16** 75–90.

Jackson, D. W., J. E. Keith, R. K. Burdick. 1986. Examining the relative importance of physical distribution service elements. *J. Bus. Logist.* **7** 14–32.

Kalai, E., M. E. Kamien, M. Rubinovitch. 1992. Optimal service speeds in a competitive environment. *Management Sci.* **38** 1154–1163.

Lederer, P. J., L. Li. 1997. Pricing, production scheduling and delivery-time competition. *Oper. Res.* **45** 407–420.

Li, L. 1992. The role of inventory in delivery-time competition. *Management Sci.* **38** 182–197.

Mendelson, H., S. Whang. 1990. Optimal incentive-compatible priority pricing for the $M/M/1$ queue. *Oper. Res.* **38** 870–883.

O'Conor, J. 2002. *Private Correspondences*.

Palaka, K., S. Erlebacher, D. H. Kropp. 1998. Lead time setting, capacity utilization, and pricing decisions under lead time dependent demand. *IIE Trans.* **30** 151–163.

Rao, U. S., J. M. Swaminathan, J. Zhang. 2000. Integrated demand and production management in a periodic, make-to-order setting with uniform guaranteed lead time and outsourcing. Working paper, GSIA, Carnegie Mellon University, Pittsburgh, PA.

Ray, S., E. M. Jewkes. 2002. Customer lead time management when both price and demand are lead time dependent. *Eur. J. Oper. Res.* Forthcoming.

Smith, M., J. Bailey, E. Brynjolfsson. 2000. Understanding digital markets: Review and assessment. *Understanding the Digital Economy*. E. Brynjolfsson, B. Kahin, eds. MIT Press, Boston, MA.

So, K. C. 2000. Price and time competition for service delivery. *Manufacturing & Service Oper. Management* **2** 392–409.

——, J-S. Song. 1998. Price, delivery time guarantees and capacity selection. *Eur. J. Oper. Res.* **111** 28–49.

Stalk, G., T. M. Hout. 1990. *Competing Against Time: How Time-Based Competition is Reshaping Global Markets*. The Free Press, New York.

Sterling, J. U., D. M. Lambert. 1989. Customer service research: Past deficiencies, present solutions and future opportunities. *Internat. J. Phys. Dist. Materials Management* **19** 1–23.

Stidham, S. 1992. Pricing and capacity decisions for a service facility: Stability and multiple local optima. *Management Sci.* **38** 1121–1139.

Suri, R. 1998. *Quick Response Manufacturing: A Companywide Approach to Reducing Lead Times*. Productivity Press, Portland, OR.

Tirole, J. 1988. *The Theory of Industrial Organization*. The MIT Press, Cambridge, MA.

Tsay, A. A. N. Agrawal. 2000. Channel dynamics under price and service competition. *Manufacturing & Service Oper. Management* **2** 372–391.

The consulting Senior Editor for this manuscript was Paul Kleindorfer. This manuscript was received on May 8, 2002, and was with the authors 56 days for 3 revisions. The average review cycle time was 68 days.