Price and leadtime differentiation and operations strategy in a competitive market

Sachin Jayaswal[a,*]

[a] *Indian Institute of Management, Vastrapur, Ahmedabad, Gujarat 380 015, India. Ph: +91-79-6632-4877, Fax: +91-79-6632-6896*

Abstract

We study a duopoly market in which customers are heterogeneous in their sensitivity to price and leadtime, and can be segmented as price sensitive or time sensitive. Each firm tailors (differentiates) its products/services for the two customer classes solely based on price and the corresponding guaranteed leadtime. Our objective is to understand how competition affects price and leadtime differentiation of the firms since the extant literature reports very contradicting results. Our results suggest that when firms use dedicated resources to serve the two market segments, pure price competition always tends to decrease individual prices as well as price differentiation, irrespective of the market behavior. Further, the effect of competition is more pronounced when customers are allowed to self-select, thereby introducing substitutability between the two product options. On the other hand, when firms compete in time, in addition to price, the effect of competition on product differentiation depends crucially on the behavior of the market. We further use our model to study the effects of asymmetry between the competing firms on their product differentiation. Our results suggest that the firm with a larger market base should always maintain a larger price and leadtime differentiation between the two market segments. Similarly, the firm with a capacity cost advantage should also maintain a larger leadtime differentiation.

Keywords: Operations strategy, competition, price, leadtime, product differentiation

1. Introduction

The importance of time as a competitive priority is now well established, especially in make-to-order and service industries. Shorter leadtime guarantee can have a major impact on both demand as well as price. Blackburn (1991) and Maltz and Maltz (1998), besides others, have empirically shown the impact of leadtime on customer demand. In fact, many firms today use leadtime guarantee in their promotion campaigns. For example, Cat Logistics, a subsidiary of Caterpillar, promises to ship service parts within 24 hours to its clients (Schmidt and Aschkenase, 2004). Some firms even exploit customers' sensitivity to time to extract price premium for the same product by promising them a shorter leadtime. Amazon.com, for example, charges more than double the shipping costs to guarantee a delivery in two days against its normal delivery time of around a week (Ray and Jewkes, 2004). Such firms exploit customers' heterogeneity in their preferences for price and time to charge different prices from different market segments (and promise

*Corresponding author
Email address: sachin@iimahd.ernet.in (Sachin Jayaswal)

them different leadtimes) in an attempt to capture a greater fraction of the surplus that uniform pricing leaves with the consumers (Tirole, 1998).

Different firms in an industry may compete with each other by offering better deals, either in the form of lower prices, better leadtime guarantees or both to their customers. Attractive price and leadtime can generate demand. Failure to meet the guaranteed leadtimes may, however, result in penalties, either in the form of a discount or partial refund, denting a firm's margin. A striking example in this context is the case of seven online retailers, including Macys.com, Toysrus.com and CDNOW, that paid fines to the tune of $1.5 million to settle a Federal Trade Commission lawsuit over late deliveries made in 1999 (Pekgun, 2007). Firms, therefore, target to meet their guaranteed leadtimes with at least a given level of reliability. In a make-to-order or service industries, this usually translates into a better (server) capacity management, an operations related decision. Marketing and operations managers of such firms, therefore, need to jointly decide on the optimal prices and leadtimes for the different market segments, and the required capacity levels.

Ours is not the first work to take into account for such a linkage between a firm's marketing decision of price and leadtime differentiation and its operation's capacity related decision. Boyaci and Ray (2003), Boyaci and Ray (2006), Zhao et al. (2008) and Jayaswal et al. (2011) have done a related study, but in a monopolistic setting. Businesses in real world, however, rarely operate as monopolists. To the best of our knowledge, ours is the first attempt to model the linkage between a firm's product differentiation decision and its capacity decision in a *competitive setting*. The problem of product differentiation has been studied extensively in Industrial Organization literature (see Tirole (1998) for an extensive review of the literature). However, the literature in Industrial Organization does not model the linkage between product differentiation and capacity related decisions, which so crucially exists.

Keeping the above discussion in mind, we study a market with two competing firms. Each firm sells a menu of products/services, differentiated only in their prices and leadtime guarantees, to exploit customers' heterogeneity in their preferences for time and price. The main issue that a firm faces in such a market is how to optimally differentiate its products, based on their prices and guaranteed leadtimes, for customers with different preferences, and accordingly decide its optimal capacity levels. In a competitive market, it also needs to take into account the reaction from other firms, and its impact on its own demand.

The objective of this paper is to study the effect of competition on firms' price and leadtime differentiation decisions. While it is well known that competition, in general, drives prices down, its effect on price differentiation (price discrimination, as is popularly termed in Economics literature) is not clear, all the more so when the price discrimination is based on some endogenous category such as the leadtime guarantee. The traditional theory on price discrimination, which predicts that market competition decreases a firms ability to use price discrimination, has often been challenged by very contrasting empirical results (Gerardi and Shapiro, 2007). Further, different empirical studies on the effects of competition on price discrimination have themselves so far produced very conflicting results. For example, the theoretical results of Gale (1993) suggest that price discrimination increases with competition. So does the empirical research by Borenstein and Rose (1994) for the airline industry, which is in sharp contrast to the observations made by Gerardi and Shapiro (2007). Further, a more complex analysis is necessary when firms must price discriminate on the basis of some endogenous category such as the leadtime preference (Varian, 1989). Moreover, the effect of competition on leadtime differentiation between different customer segments itself has not been studied. Through our rich model, which takes into account the important link between marketing and operations, we try to shed more light into the effect of market competition on price and leadtime differentiation. Further, a competing firm may either use separate (dedicated) resources (capacities) for the the different market segments or may share them across the market segments. For example, FedEx uses separate facilities for its express and ground services.[1] UPS, in contrast, delivers express and ground services using one integrated network.[2] By modelling both the dedicated as well shared capacity strategy used by competing firms, we try to understand the role played by a firm's capacity strategy on price and leadtime differentiation strategy in a competitive market. Specifically, we try to address the following issues: (1) How does competition affect the price and leadtime differentiation decisions of a firm relative to a monopolistic setting? Traditional theory suggests competition should result in lower prices; should it result in lower price discrimination as well? How does the result change if firms compete in time, in addition to price? (2) How does the operations strategy, specifically the capacity strategy (dedicated versus shared capacities), used by competing firms affect the equilibrium price and leadtime decisions? (3) How does asymmetry in firms' operating conditions (in terms of capacity cost or market penetration) affect the equilibrium price and leadtime differentiation decisions?

In the first part of the paper, we study a setting where both the competing firms use dedicated capacity strategy. Such a problem setting leads to analytical results. We show that a Nash Equilibrium solution always exists for such a setting, and devise an iterative procedure to compute it. In the later part of the paper, we study (numerically) different settings where each of the competing firms may operate in a dedicated or a shared capacity setting. Some of the important managerial insights we draw from our analysis are: (1) Price competition not only tends to decrease individual prices, but it also decreases price differentiation between the different market segments. Further, the effect of competition is more pronounced in presence of product substitution (second degree price discrimination). (2) When firms compete in time, in addition to price, the precise effect on price and leadtime differentiation depends crucially on the market behavior. (3) Price differentiation at equilibrium is always larger for the firm that shares its capacities compared to the other firm that dedicates its capacities for different market segments. However, sharing capacities results in larger leadtime differentiation only when capacity cost is small. (4) When asymmetry exists between competing firms, the firm with a larger market base should always maintain a larger price and leadtime differentiation between the different market segments. Similarly, the firm with a capacity cost advantage should also maintain a larger leadtime differentiation. However, whether it should maintain a larger or a smaller price differentiation depends on whether the firms use dedicated or shared capacities and also on their capacity cost level.

Rest of the paper is organized as follows. In Section 2, we provide a review of the related literature. We present our mathematical model and the underlying assumptions in Section 3. In Section 4, we describe the best response of a firm, both for dedicated and shared capacity settings, given its competitor's price and leadtime decisions. Section 5 discusses the Equilibrium solution for the duopoly problem, followed by analysis of important results. The paper concludes with a summary of our main results and directions

[1]http://www.fedex.com/us/about/express/pressreleases/pressrelease011900.html?link=4

[2]http://sec.edgar-online.com/1999/10/20/11/0000940180-99-001230/Section2.asp

for future research in Section 6.

2. Literature Review

The work in this paper relates closely to price discrimination, which has been studied extensively in the Economics (Industrial Organization) literature, and to price and/or time competition in Operations Management (OM) field. We briefly review each of these two areas of research separately.

The traditional theory in Industrial Organization suggests that uniform pricing leaves some surplus with the consumers. Firms, therefore, use discriminatory pricing for the same product/service in an attempt to capture a greater fraction of consumers' surplus when customers are heterogeneous in their preferences (Tirole, 1998). Earlier works on price discrimination (e.g., Varian (1985)) have focused on its effect on social welfare. Some of the more recent works have started to focus on the effect of competition on price discrimination. However, there still is no agreement in the literature on whether competition increases or decreases price discrimination.

The textbook theory argues that competitive firms cannot price discriminate since they are price takers, while monopolists can price discriminate to the extent that there exists both homogeneity in consumers' demand elasticities and a useful sorting mechanism to distinguish between consumer types (Gerardi and Shapiro, 2007). The textbook theory, therefore, predicts that competition should decrease price discrimination. This is further corroborated by the theoretical model of Rochet and Stole (2002) on second degree price discrimination. However, the theoretical models of Gale (1993) and Stole (1995) produce exactly the opposite results. As there is no overarching theoretical model, the relation between competition and price discrimination becomes an empirical question. However, different empirical studies have again produced very contrasting results. Borenstein (1989) and Borenstein and Rose (1994) found evidence of increasing price dispersion with competition in airline industry, thereby suggesting that competition increases price discrimination. However, a more detailed empirical study by Gerardi and Shapiro (2007) found a negative relation between market competition and price dispersion, thereby suggesting that competition decreases price discrimination.

Literature in OM on price and/or time competition model a firm's operations in detail. These papers can mostly be classified into (Cachon and Harker, 2002) (i) papers on inventory games, and (ii) those on queueing games. Papers on inventory games are relevant in a make-to-stock setting where firms use inventory as their strategic tool to compete in the market. Papers on queueing games are pertinent to make-to-order or service industries, where firms use better (server) capacity/queue management to adjust their price and leadtime decisions, and thus compete in the market. Since our model is relevant to make-to-order or service industries, our focus is on the latter category.

Literature on queueing games can further be categorized into (i) papers that aggregate price and waiting time into a single measure called "full price", and (ii) those that model price and delivery/waiting time as independent explanatory variables. Levhari and Luski (1978), Loch (1991), Lederer and Li (1997), Chen and Wen (2003) and Armony and Haviv (2003) belong to the former category. All these papers assume that customers associate a specific cost rate with their waiting time, and that they make their selection of a firm based only on their "full price", which is the sum of the actual price charged and the expected delay cost, disregarding any other service attributes. Further, they assume that all customers are in a position to assess the equilibrium steady state waiting times they will experience.

The second category of papers, which model price and leadtime as independent explanatory variables, include So (2000), Pekgun et al. (2006), Allon and Federgruen (2007), Allon and Federgruen (2008). These papers model customers' aggregate demand for a firm's service as a function of its price, leadtime and/or other service attributes, each of which is explicitly advertised by the firm. We position our work in this category since we treat price and leadtimes as independent variables announced by a firm. Although our demand model bears some similarity with those used by Pekgun et al. (2006) and Allon and Federgruen (2007), it is fundamentally different from them in that these models consider only a single customer class, and thus there is no market segmentation. To our knowledge, Loch (1991), Lederer and Li (1997), Armony and Haviv (2003) and Allon and Federgruen (2008) are the only works to have addressed the phenomenon of market segmentation. As noted earlier, these papers, except for Allon and Federgruen (2008), assume that customers aggregate the price and leadtime into a full price, and that they select the service provider on the basis of this full price only. In doing so, they assume that all customers are in a position to assess the equilibrium steady state waiting times they will experience, while in our model, waiting time standards are advertised to the different classes. Moreover, they consider the firms' capacity levels as exogenously given, and not a decision variable.

Thus, Allon and Federgruen (2008) appears to be the closest to our model, although effect of competition on price discrimination is not their focus. However, they study completely segmented markets, which means that each customer is strictly assigned to a specific class, and she cannot switch between different classes. This is tantamount to saying that the specific service package (price and leadtime combination) offered to a given customer class is not available to any other class, and hence is non-substitutable. In marketing/economics parlance, their work is pertinent to third degree price discrimination (see Talluri and Ryzin (2004) for a definition). Our demand model is more generalized, which also captures product substitution. Our model, therefore, allows us to study both second and third degree price discrimination. Moreover, Allon and Federgruen (2008) use a service level that is based on expected leadtimes. In other words, they assume that firms select their capacity levels so that customers from each segment are served within their promised leadtimes on average. This does not provide any bound on instances of unusually long leadtimes. It is quite possible then that a large portion of the demands are actually not served within their promised leadtimes, even if the promised leadtimes are met on average. We, therefore, assume that firms select their capacity levels so as to fulfill their promised leadtimes with a high level of reliability (generally 99%). This makes the leadtime guarantees more attractive, although it makes the problem a lot more challenging to solve.

3. Decision Model

We consider a service or a make-to-order manufacturing industry with two firms, indexed by $i \in \{1, 2\}$ and $j = 3-i$, competing in a market that is segmented into 2 customer classes, indexed by $c \in \{h, l\}$. Class h customers are high priority/express customers who are more time sensitive and are willing to pay a price premium for a shorter leadtime. Class l customers are low priority/regular customers who are more price sensitive and are willing to accept a longer leadtime for a price discount. Firm i competes for its market share by selecting its prices p_c^i and guaranteed leadtimes L_c^i offered to market segment $c \in \{h, l\}$. Firm i faces a demand from class c, generated according to a Poisson process with rate $\lambda_c^i\,(p_c^i, L_c^i, c \in \{h,l\}, i \in \{1,2\})$, which depends on the decisions of both firms in the following way: (i) each firm's expected demand from a given market segment is (i) decreasing in its price and leadtime offered to that segment, (ii) increasing in its price and leadtime offered to the other market segment, and (iii) increasing in the price and leadtime offered offered by the other firm to the same segment. We model this using the following system of linear equations:

$$\lambda_h^i = a^i - \beta_p^h p_h^i + \theta_p(p_l^i - p_h^i) - \beta_L^h L_h^i + \theta_L(L_l^i - L_h^i) + \gamma_p(p_h^j - p_h^i) + \gamma_L(L_h^j - L_h^i) \quad (1)$$

$$\lambda_l^i = a^i - \beta_p^l p_l^i + \theta_p(p_h^i - p_l^i) - \beta_L^l L_l^i + \theta_L(L_h^i - L_l^i) + \gamma_p(p_l^j - p_l^i) + \gamma_L(L_l^j - L_l^i) \quad (2)$$

where,

$2a^i$	:	market base of firm i
β_p^c	:	sensitivity of class c demand to its own price
β_L^c	:	sensitivity of class c demand to its own guaranteed leadtime
θ_p	:	sensitivity of demand to inter-class price difference
θ_L	:	sensitivity of demand to inter-class leadtime difference
γ_p	:	sensitivity of demand to inter-firm price difference
γ_L	:	sensitivity of demand to inter-firm leadtime difference

$2a^i$ parameterizes firm $i's$ market base. Mathematically, it is the demand faced by firm i when price and leadtime offered by each firm to each customer class is zero. It captures the aggregate effect of all the factors such as a firm's brand image, service quality, etc other than price and leadtime on demand. Hence, the firm offering the lowest price and the shortest leadtime to a market segment need not capture its entire demand. The relative values of a^i and a^j can be loosely used to describe comparative advantage in terms of a firm's market penetration. This may reflect the underlying preferences of customers for one firm over the other, which may be due to customers' appeal for a brand.

The above demand model generalizes the demand model used by Jayaswal et al. (2011) to a competitive setting. It also generalizes the demand models used by Tsay and Agrawal (2000) and Pekgun et al. (2006) to segmented markets. Further, it generalizes the demand model used by Allon and Federgruen (2008) by taking into account product substitution (θ_p and θ_L). The total market size $(\sum_{i\in\{1,2\}}\sum_{c\in\{h,l\}}\lambda_c^i)$ in our model is invariant to any changes in inter-firm or inter-class sensitivities, which only affects the distribution of the total demand among the firms and the customer classes. However, the pricing and leadtime decisions of the two firms affect both the total market size as well as the resulting demand for each firm and from each market segment. By definition, $\beta_p^c > 0$, $\beta_L^c > 0$, $\theta_p \geq 0$, $\theta_L \geq 0$, $\gamma_p \geq 0$, $\gamma_L \geq 0$, $\beta_p^h < \beta_p^l$ and $\beta_L^h > \beta_L^l$. $\theta_p > 0$, $\theta_L > 0$ signifies product substitution, while $\gamma_p > 0$, $\gamma_L > 0$ signify the presence of price competition and leadtime competition in the market. $\gamma_p = \gamma_L = 0$ makes the demand of two firms independent, and hence decouples their decision making, resulting in a monopolistic setting.

We use a reliability level α^i with which firm i tries to meet its leadtime guarantee. In our model, we restrict our discussions only to cases where the service reliability for each firm is the same, i.e., $\alpha^i = \alpha$. This is applicable in situations where there exists some industry standard and published reports (like those published by the Aviation Consumer Protection Division of the U.S. Department of Transportation, and Expedia for airline industry) such that the delivery performance of each firm is readily available to customers. In this way, firms are discouraged from performing below the standard such that the market share is then mainly affected by their promised times and prices, as depicted by our demand model. Of course, our model and analysis also allows for different service reliabilities for different firms.

We assume the time firm i takes to serve a customer from class c is exponentially distributed with a rate μ_c^i. Firm i, therefore, behaves like an M/M/$\cdot$ queuing system. Customers within each class are served on a first-come-first-serve (FCFS) basis. The firm can invest in its installed capacity to increase its processing rate μ_c^i. We assume there are no economies of scale in investing in capacity. So a unit increment in μ_c^i always costs $\$A^i$[3]. We also assume that firm i incurs the same operating cost of $\$m^i$ in serving a customer of either class.

The industry is assumed to have established a standard leadtime for regular customers, and hence $L_l^i = L_l^j = L_l$[4]. Firm i selects and announces its express leadtime and the two prices (L_h^i, p_h^i, p_l^i) so as to maximize its profit. Firm i does so taking into account the leadtime and prices (L_h^j, p_h^j, p_l^j) selected by firm $j = 3 - i$ since they have an impact on firm $i's$ demands, and hence on its profit. It also needs to simultaneously select its optimal service rates (i.e., installed capacities) μ_c^i in order to meet the guaranteed leadtimes with at least a minimum level of reliability α^i.

Notations

i, j	:	indices for firm; $i \in \{1, 2\}$, *and* $j = 3 - i$
c	:	index for customer class; $c \in \{h, l\}$
λ_c^i	:	mean demand rate for firm i from customer class c (units/unit time)
μ_c^i	:	mean processing rate of firm i for customer class c (units/unit time)
p_c^i	:	price charged by firm i to customer class c (\$/unit)
L_c^i	:	leadtime quoted by firm i to customer class c (time units)
W_c^i	:	steady state actual sojourn (waiting + service) time of customer class c at firm i (time units)
α	:	target service level set by either firm (no unit).
$S_c^i(L_c^i)$	:	actual service level achieved by firm i for quoted leadtime L_c^i, i.e., $\mathrm{P}(W_c^i \leq L_c^i)$
m^i	:	unit operating cost of firm i (\$/unit)
A^i	:	marginal capacity cost of firm i (\$/unit)

In the next section, we describe the best response of a firm, given its competitor's price, leadtime and capacity decisions.

4. A Firm's Best Response

Given the price, leadtime and capacity decisions $(p_h^j, p_l^j, L_h^j, \mu_h^j, \mu_l^j)$ of firm $j = 3 - i$, firm $i \in \{1, 2\}$ tries to selects its own corresponding decisions $(p_h^i, p_l^i, L_h^i, \mu_h^i, \mu_l^i)$ that maximize its profit and also ensure that its leadtime commitments are met reliably. As clear from the demand model ((1) - (2)), the demands for firm $i \in \{1, 2\}$, and its decisions

[3] A^i may be different for different customer classes if they are served by different service capacities. Using the same marginal capacity cost for the two customer classes, however, allows a meaningful comparison between the dedicated and the shared capacity settings.

[4] Although we make this assumption mainly for the tractability of the model, this is still a reasonably realistic representation of certain business settings. For example, in most of the online retail web hosting services, any updating of content, if not done in express fashion, must be done within one day (Boyaci and Ray, 2003).

in turn, depend on the price and leadtime decisions made by firm $j = 3 - i$. Firm $i's$ demand and its decisions, however, do not depend on the capacity level (μ_h^j, μ_l^j) selected by firm j. While competing with the other firm, each firm, therefore, possesses only two types of essential strategic instruments: prices and the leadtimes. Firm $i's$ strategy can be defined as a vector of its strategic decision variables $s^i := (p_h^i, p_l^i, L_h^i)$, which it uses to compete against the other firm j. The best response of firm i to firm $j's$ strategy, $s^j := (p_h^j, p_l^j, L_h^j)$, is thus a strategy $s^{i*} := (p_h^{i*}, p_l^{i*}, L_h^{i*})$ such that $\pi^i(s^{i*}, s^j) = \max_{s^i} \pi^i(s^i, s^j)$, $i \in \{1,2\}$ and $j = 3 - i$. Firm $i's$ best response is the solution to the following optimization problem (called the pricing and leadtime decision problem $[PLDP^i]$):

$[\mathbf{PLDP^i}]$:

$$\max_{p_h^i, p_l^i, L_h^i, \mu_h^i, \mu_l^i} \quad \pi^i = (p_h^i - m^i)\lambda_h^i + (p_l^i - m^i)\lambda_l^i - A^i(\mu_h^i + \mu_l^i) \tag{3}$$

$$\text{s.t.} \quad L_h^i < L_l^i \tag{4}$$

$$p_h^i, p_l^i, \lambda_h^i, \lambda_l^i, \mu_h^i, \mu_l^i \geq 0 \tag{5}$$

$$\textit{Stability condition} \tag{6}$$

$$S_h^i(L_h^i) = P(W_h^i \leq L_h^i) \geq \alpha \tag{7}$$

$$S_l^i(L_l^i) = P(W_l^i \leq L_l^i) \geq \alpha \tag{8}$$

where λ_h^i and λ_l^i are given by (1) and (2) respectively. The objective function (3) is the profit per unit time. Constraint (4) requires that the guaranteed leadtime for high priority customers be shorter than that for the other class. Constraint set (5) is needed to define a realistic problem setting. Constraint (6) is the stability condition for the queuing system, which models the service facility at the firm. Constraints (7) and (8) are leadtime reliability constraints (also called service level constraints), which require that the steady state actual leadtime W_h^i (resp., W_l^i) of a customer should not exceed its guaranteed leadtime L_h^i (resp., L_l^i) with a probability of at least α.

Note that a firm's best response problem has a form very much similar to a firm's optimal decision in a monopolistic setting, discussed by Jayaswal et al. (2011). Difference still arises between the two due to the strategic interaction between the competing firms, which is captured in the demand model (1) - (2). Therefore, the best response of a firm can also be solved by adapting the solution method developed by Jayaswal et al. (2011) for the monopolistic setting. In what follows, we adapt $[PLDP^i]$ for a firm using dedicated or shared capacity strategy by specifying the form of constraints (6)-(8).

4.1. Dedicated Capacity Setting

For a dedicated capacity setting, where each customer class is served by a separate M/M/1 server, the sojourn time distribution for either class of customers is known to be exponential (Ross, 2010). In this case, there is a separate stability condition for each of the queues. Hence, constraints (6), (7) and (8) can be expressed as:

$$\lambda_c^i - \mu_c^i < 0, \; c \in \{h, l\} \tag{6^{DC}}$$

$$S_h^i(L_h^i) = P(W_h^i \leq L_h^i) = 1 - e^{(\lambda_h^i - \mu_h^i)L_h^i} \geq \alpha \tag{7^{DC}}$$

$$S_l^i(L_l^i) = P(W_l^i \leq L_l^i) = 1 - e^{(\lambda_l^i - \mu_l^i)L_l^i} \geq \alpha \tag{8^{DC}}$$

Proposition 1. *For a given strategy $s^j := (p_h^j, p_l^j, L_h^j)$ by firm $j = 3-i$, the best response express leadtime L_h^{i*} of firm $i \in \{1,2\}$ using dedicated capacities is given by the unique root of (9) in the interval [0, L_l^i)*

$$\frac{\partial \pi^i(L_h^i)}{\partial L_h^i} = -\left(\beta_L^h + \theta_L + \gamma_L\right)\left(p_h^{i*}(L_h^i) - m^i - A^i\right) + \theta_L\left(p_l^{i*}(L_h^i) - m^i - A^i\right) - \frac{A\ln(1-\alpha)}{(L_h^i)^2} \tag{9}$$

where $p_h^{i*}(L_h^i)$ and $p_l^{i*}(L_h^i)$ are the optimal prices for a fixed express leadtime of firm i, given by:

$$\begin{aligned}p_h^{i*}(L_h^i) =& \frac{(\beta_p^l + 2\theta_p + \gamma_p)a^i - \{\beta_p^l(\theta_L + \gamma_L) + \beta_L^h(\beta_p^l + \theta_p + \gamma_p) + \theta_p\gamma_L + \gamma_L\gamma_p + \theta_L\gamma_p\}L_h^i}{2D}\\ & \frac{+\{(\beta_p^l + \gamma_p)\theta_L - (\beta_L^l + \gamma_L)\theta_p\}L_l^i + (\beta_p^l\gamma_p + \gamma_p\theta_p + \gamma_p^2)p_h^j + (\theta_p\gamma_p)p_l^j}{2D}\\ & \frac{+(\beta_p^l\gamma_L + \gamma_L\gamma_p + \theta_p\gamma_L)L_h^j + (\theta_p\gamma_L)L_l^j}{2D} + \frac{A^i + m^i}{2}\end{aligned} \tag{10}$$

$$\begin{aligned}p_l^{i*}(L_h^i) =& \frac{(\beta_p^h + 2\theta_p + \gamma_p)a^i - \{\beta_p^h(\theta_L + \gamma_L) + \beta_L^l(\beta_p^h + \theta_p + \gamma_p) + \theta_p\gamma_L + \gamma_L\gamma_p + \theta_L\gamma_p\}L_l^i}{2D}\\ & \frac{+\{(\beta_p^h + \gamma_p)\theta_L - (\beta_L^h + \gamma_L)\theta_p\}L_h^i + (\theta_p\gamma_p)p_h^j + (\beta_p^h\gamma_p + \gamma_p\theta_p + \gamma_p^2)p_l^j}{2D}\\ & \frac{+(\theta_p\gamma_L)L_h^j + (\beta_p^h\gamma_L + \gamma_L\gamma_p + \theta_p\gamma_L)L_l^j}{2D} + \frac{A^i + m^i}{2}\end{aligned} \tag{11}$$

and $D = \beta_p^h\beta_p^l + \beta_p^h\theta_p + \beta_p^l\theta_p + \beta_p^h\gamma_p + \beta_p^l\gamma_p + 2\theta_p\gamma_p + \gamma_p^2$.

The corresponding optimal price differentiation is then:

$$\begin{aligned}p_h^{i*}(L_h^i) - p_l^{i*}(L_h^i) =& \frac{\{(\beta_p^l - \beta_p^h)a^i + (\beta_p^h + \beta_p^l)\theta_L + (\gamma_L + 2\theta_L)\gamma_p\}(L_l^i - L_h^i)}{2D}\\ & \frac{-(\beta_p^l\beta_L^h + \beta_p^l\gamma_L + \beta_L^h\gamma_p)L_h^i + (\beta_p^h\beta_L^l + \beta_p^h\gamma_L + \beta_L^l\gamma_p)L_l^i + (\beta_p^l + \gamma_p)\gamma_p p_h^j}{2D}\\ & \frac{-(\beta_p^h + \gamma_p)\gamma_p p_l^j + (\beta_p^l + \gamma_p)\gamma_L L_h^j - (\beta_p^h + \gamma_p)\gamma_L L_l^j}{2D}\end{aligned} \tag{12}$$

Proof. See Appendix A. □

(10) and (11) suggest that in pricing its product for a given customer segment, a firm should take into account the price quoted by the other firm not only to the same customer segment but also to the other customer segment. This, at first thought, sounds surprising. This is because our demand functions (1) and (2) suggest that a firm's demand from a given segment is not influenced by what is offered to the other segment by the other firm. To make things clear, our demand function (1), for example, suggests that the demand faced by firm 1 from the express segment depends on the price charged by firm 2 to the express segment, but is not influenced by what firm 2 charges to the regular customers.

However, (10) suggests that in pricing its product for express customers, firm 1 should keep in mind not only the price charged by firm 2 to the express customers but also the price charged by firm 2 to the regular customers. This is because firm $2's$ price to regular customers influences its demand from express customers as well. So, in pricing its product for express customers, firm 1 should take into account the other factors that influence its express customers' decision, which includes the price charged by firm 1 to its regular customers.

Proposition 2. *Given a strategy $s^j := (p_h^j, p_l^j, L_h^j)$ of firm $j = 3 - i$, a decrease in the express leadtime L_h^i by firm $i \in \{1, 2\}$ using dedicated capacities results in: (a) an increase in p_h^{i*} (b) a decrease in p_l^{i*} if $\theta_L/(\beta_L^h + \gamma_L) > \theta_p/(\beta_p^h + \gamma_p)$; an increase in p_l^{i*} if $\theta_L/(\beta_L^h + \gamma_L) < \theta_p/(\beta_p^h + \gamma_p)$.*

Proof. Follows directly from Proposition 1. □

Proposition 2 suggests that given the decisions of the other firm, a firm's best response express price always increases with a decrease in its own express leadtime. The effect of any change in its express leadtime on the regular price, however, depends on the behavior of the market. Specifically, the regular price decreases with any decrease in its express leadtime if the relative sensitivity of customers to the inter-class leadtime difference (with respect to their own leadtime and inter-firm leadtime difference) is greater than their relative sensitivity to the inter-class price difference (with respect to their own price and inter-firm price difference), such that $\theta_L/(\beta_L^c + \gamma_L) > \theta_p/(\beta_p^c + \gamma_p)$, $c \in \{h, l\}$. We call such a market as Time Difference Sensitive (TDS). On the other hand, a decrease in its express leadtime increases its regular prices if the relative sensitivity of customers to the inter-class leadtime difference (with respect to their own price and inter-firm leadtime difference) is smaller than their relative sensitivity to the inter-class price difference (with respect to their own price and inter-firm price difference), such that $\theta_L/(\beta_L^c + \gamma_L) < \theta_p/(\beta_p^c + \gamma_p)$, $c \in \{h, l\}$. We call such a market as Price Difference Sensitive (PDS).

4.2. Shared Capacity Setting

The firm's choice of shared capacity is modelled using a single server, which serves both customer classes employing a simple fixed priority scheme that always gives priority to time-sensitive customers. Customers within each class are served on a first-come-first-served (FCFS) basis. In this paper, we use a preemptive priority scheme, but the analysis can be extended to a non-preemptive priority discipline.

For a shared capacity setting, the sojourn time distribution $S_h^i(\cdot)$ for high priority customers in a preemptive priority queue is known to be exponential (Chang, 1965). Hence, the leadtime reliability constraint (7) has an analytical closed-form representation, similar to that for the dedicated capacity setting. We assume the single server serves customers of either class at the same rate $\mu_h^i = \mu_l^i = \mu^i$. Constraints (6) and (7) in a shared capacity setting can then be expressed as:

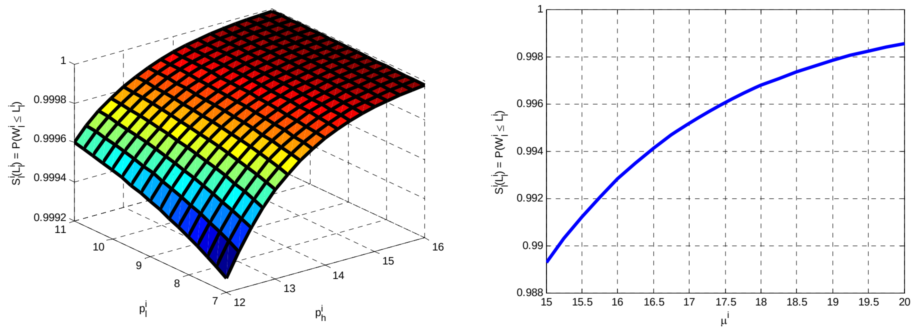
$$\lambda_h^i + \lambda_l^i - \mu^i < 0 \tag{6^{SC}}$$

$$S_h^i(L_h^i) = P(W_h^i \leq L_h^i) = 1 - e^{(\lambda_h^i - \mu^i)L_h^i} \geq \alpha \tag{7^{SC}}$$

However, a closed form expression for the sojourn time distribution $S_l^i(\cdot)$ for low priority customers, appearing in constraint (8), is not known (see Jayaswal et al. (2011) for a detailed discussion). Further, even an analytical characterization of the sojourn time distribution or a good approximation will not produce an analytical solution similar to that for the dedicated capacity setting since it cannot be guaranteed at the outset which of the constraints will be binding at optimality. So, the model in a shared capacity setting does not lend itself to an easy solution using conventional optimization methods. We resolve this difficulty by solving it in two stages. We first solve the problem for a fixed L_h^i (we term it as pricing decision problem $[PDP_{SC}^i]$) numerically using the *matrix geometric method* in a *cutting plane* framework (see Appendices B, C and D). Solution to $[PDP_{SC}^i]$ is then used to solve the pricing and leadtime decision problem $[PLDP_{SC}^i]$ using the *golden section search* method. In the following, we discuss the solution method very briefly, and refer our readers to Jayaswal et al. (2011) for the details.

4.2.1. The Pricing Decision Problem $[PDP_{SC}^i]$

On substituting (1) and (2) into (3), the objective function for $[PDP]$ is quadratic. All constraints are linear, except for (8), which does not have a closed form expression. Although the exact form of $S_l^i(\cdot)$ in constraint (8) is unknown, we exploit its special structure, determined numerically using the matrix geometric method (see Appendix A). Plots of $S_l^i(\cdot)$ vs. $(p_h^i,\, p_l^i)$, and $S_l^i(\cdot)$ vs. μ^i are shown in Figure 1. These plots suggest that $S_l^i(\cdot)$ is concave in $(p_h^i,\, p_l^i)$ and separately in μ^i. However, this does not necessarily show the joint concavity of $S_l^i(\cdot)$ in $(p_h^i,\, p_l^i,\, \mu^i)$. We will, therefore, integrate into our solution method a mechanism to ensure that the concavity assumption is not violated.



(a) S_l^i vs. (p_h^i, p_l^i) (b) S_l^i vs. μ^i

Figure 1: Service level vs. prices and capacity

Assuming $S_l^i(\cdot)$ is concave, it can be approximated by a set of tangent hyperplanes at various points $(p_h^{ik},\, p_l^{ik},\, \mu^{ik})$, $\forall\, k \in K$, that is:

$$S_l^i(\cdot) = \min_{k\in K}\left\{ S_l^{ik}(\cdot) + (p_h^i - p_h^{ik})\left(\frac{\partial S_l^{ik}(\cdot)}{\partial p_h^i}\right) + (p_l^i - p_l^{ik})\left(\frac{\partial S_l^{ik}(\cdot)}{\partial p_l^i}\right) + (\mu^i - \mu^{ik})\left(\frac{\partial S_l^{ik}(\cdot)}{\partial \mu^i}\right)\right\}$$

where $S_l^{ik}(\cdot)$ denotes the value of $S_l^i(\cdot)$ at a fixed point $(p_h^{ik}, p_l^{ik}, \mu^{ik})$. $\frac{\partial S_l^{ik}(\cdot)}{\partial p_h^i}$, $\frac{\partial S_l^{ik}(\cdot)}{\partial p_l^i}$ and $\frac{\partial S_l^{ik}(\cdot)}{\partial \mu^i}$ are the partial gradients of S_l^i at $(p_h^{ik}, p_l^{ik}, \mu^{ik})$, which can be obtained using the finite difference method, described in Appendix C. Constraint (8) in a shared capacity setting can thus be replaced by the following set of linear constraints:

$$S_l^{ik}(\cdot) + (p_h^i - p_h^{ik})\left(\frac{\partial S_l^{ik}(\cdot)}{\partial p_h^i}\right) + (p_l^i - p_l^{ik})\left(\frac{\partial S_l^{ik}(\cdot)}{\partial p_l^i}\right) + (\mu^i - \mu^{ik})\left(\frac{\partial S_l^{ik}(\cdot)}{\partial \mu^i}\right) \geq \alpha \quad \forall k \in K \quad (13)$$

Substituting the above set of constraints in place of (8), and the expressions (1) and

(2) for λ_h^i and λ_l^i results in the following quadratic programming problem (QPP) with a finite but a large number of constraints, which makes it suitable for the cutting plane method (Kelley, 1960).

$\underline{[\mathbf{PDP^i_{(K)}}]:}$

$$\max_{p_h^i,p_l^i,\mu^i} \pi^i = -(\beta_p^h+\theta_p+\gamma_p)(p_h^i)^2-(\beta_p^l+\theta_p+\gamma_p)(p_l^i)^2+2\theta_p p_h^i p_l^i$$
$$+\left\{-\beta_L^h L_h^i+\theta_L(L_l^i-L_h^i)+\gamma_L(L_h^j-L_l^j)+\gamma_p p_h^j+m^i(\beta_p^h+\gamma_p)+a^i\right\}p_h^i$$
$$+\left\{-\beta_L^l L_l^i+\theta_L(L_h^i-L_l^i)+\gamma_L(L_l^j-L_l^i)+\gamma_p p_l^j+m^i(\beta_p^l+\gamma_p)+a\right\}p_l-A^i\mu$$
$$+(\beta_L^h+\gamma_L)m^i L_h^i+(\beta_L^l+\gamma_L)L_l^i m^i-\gamma_L(L_h^j+L_l^j)m^i-\gamma_p(p_h^j+p_l^j)^i-2m^i a^i \tag{14}$$

subject to:

$$p_h^i,p_l^i,\mu^i\geq 0 \tag{15}$$
$$-(\beta_p^h+\theta_p+\gamma_p)p_h^i+\theta_p p_l^i\geq(\beta_L^h+\theta_L+\gamma_L)L_h^i-\theta_L L_l^i-\gamma_p p_h^j-\gamma_L L_h^j-a^i \tag{16}$$
$$\theta_p p_h^i-(\beta_p^l+\theta_p+\gamma_p)p_l^i\geq-\theta_L L_h^i+(\beta_L^l+\theta_L+\gamma_L)L_l^i-\gamma_p p_l^j-\gamma_L L_l^j-a^i \tag{17}$$
$$-(\beta_p^h+\gamma_p)p_h^i-(\beta_p^l+\gamma_p)p_l^i-\mu^i$$
$$<(\beta_L^h+\gamma_L)L_h^i+(\beta_L^l+\gamma_L)L_l^i-\gamma_p(p_h^j+p_l^j)-\gamma_L(L_h^j+L_l^j)-2a^i \tag{18}$$
$$\left(\frac{\partial S_l^{ik}(\cdot)}{\partial p_h^i}\right)p_h^i+\left(\frac{\partial S_l^{ik}(\cdot)}{\partial p_l^i}\right)p_l^i+\left(\frac{\partial S_l^{ik}(\cdot)}{\partial \mu^i}\right)\mu^i\geq\alpha-S_l^{ik}(\cdot)+$$
$$\left(\frac{\partial S_l^{ik}(\cdot)}{\partial p_h^i}\right)p_h^{ik}+\left(\frac{\partial S_l^{ik}(\cdot)}{\partial p_l^i}\right)p_l^{ik}+\left(\frac{\partial S_l^{ik}(\cdot)}{\partial \mu^i}\right)\mu^{ik}\qquad\forall k\in K \tag{19}$$
$$-(\beta_p^h+\theta_p+\gamma_p)p_h^i+\theta_p p_l^i-\mu^i$$
$$\leq\frac{\ln(1-\alpha)}{L_h^i}-a^i+(\beta_L^h+\theta_L+\gamma_L)L_h^i-\theta_L L_l^i-\gamma_p p_h^j-\gamma_L L_h^j \tag{20}$$

It is easy to verify that the Hessian of (14) is negative semidefinite. Therefore, $[PDP_{(K)}]$ has a quadratic concave objective function. Moreover, all its constraints are linear. Hence, Karush-Kuhn-Tucker (KKT) conditions are both necessary and sufficient for its global optimal solution (Luenberger, 1984). $[PDP^i_{(K)}]$ can be solved using any of the standard algorithms like Wolfe's Algorithm (Wolfe, 1959).

We use the matrix geometric method to numerically evaluate the sojourn time distribution, $S_l^{ik}(\cdot)$, at a given point (p_h^{ik}, p_l^{ik}, μ^{ik}) (see Appendix B). We refer the reader to Neuts (1981) for details of the matrix geometric method. Once $S_l^i(\cdot)$ is evaluated at a point (p_h^{ik}, p_l^{ik}, μ^{ik}), its gradients are obtained using the *finite difference method*, described in Appendix C. The gradients are used to generate cuts of the form (19), which are added iteratively in the cutting plane algorithm (see Appendix D).

4.2.2. The Pricing and Leadtime Decision Problem $[PLDP^i_{SC}]$

The pricing and leadtime decision problem $[PLDP^i_{SC}]$ adds an additional dimension to $[PDP^i_{SC}]$ by treating L_h^i as a decision variable, which the firm tries to jointly optimize along with p_h^i, p_l^i and μ^i for a given set strategy of firm $j\neq i$. This makes constraint (7^{SC}) non-linear, and the model substantially more challenging to solve. We use the solution to $[PDP^i_{(K)}]$ and the golden section search method (Luenberger, 1984) to solve $[PLDP^i_{SC}]$, which can be rewritten as:

$$\max_{L_h^i \in [0,\, L_l^i)} \quad f(L_h^i)$$

where $f(L_h^i)$ is a $[PDP_{SC}^i]$ for a fixed L_h^i. We have shown in a dedicated capacity setting that $f(L_h^i)$ has a unique maximum when a^i is high (see Appendix A). Our extensive numerical experiments with $f(L_h^i)$ suggests that a sufficiently large a^i guarantees that $f(L_h^i)$ has a unique maximum in a shared capacity setting as well, and hence $[PLDP_{SC}^i]$ can be solved efficiently using the golden section search method. At each step, the algorithm solves a $[PDP_{(K)}^i]$ to evaluate $f(L_h^i)$ for a given value of L_h^i.

5. Duopoly Problem

We now study the price and leadtime decisions for a duopoly problem. One basic question is to investigate whether an equilibrium exists, and if so, how will the equilibrium change under different operational settings and market characteristics. To investigate the effect of capacity strategy on equilibrium price and leadtime decisions, we study the following three scenarios: (1) DD: both firms using dedicated capacities; (ii) SS: both frims using shared capacities; (iii) DS: firms 1 using dedicated capacities while firm 2 using shared capacities. However, we study the DD scenario in a greater detail as it allows us to obtain analytical results.

Under competition, both firms simultaneously announce their price and leadtime decisions. We assume that firm $i \in \{1, 2\}$ has full knowledge of the operational setting of firm $j = 3 - i$, including its capacity strategy and also its parameters A, m and a. Firm i can thus correctly anticipate the best response of firm j to its own moves, and can hence strategically plan its own strategy. Nash Equilibrium is reached when none of the firms can do better by unilaterally deviating from its decisions. A Nash equilibrium is thus a vector of strategies (s^{i*}, s^{j*}) such that for each firm i, $\pi^i(s^{i*}, s^{j*}) = \max_{s^i} \pi^i(s^i, s^{j*})$, $i \in \{1, 2\}$ and $j = 3 - i$. In other words, the strategy used by either firm is the best response to the strategy chosen by the other.

We first study the equilibrium results under pure price competition in a DD scenario. Pure price competition is relevant to situations where the industry may face a significantly higher stickiness for the leadtime decisions compared to its ability to vary prices. A relatively higher stickiness for leadtime decisions may arise, for example, when the services are partly outsourced to a third party. In such a situation, the firms may not be able to revise their leadtime decisions as frequently as they can revise its prices, and may compete primarily based on prices, treating their delivery times as given. Other factors contributing to stickiness in leadtime decisions can be found in (Allon and Federgruen, 2007).

5.1. Pure Price Competition in a DD Setting

Proposition 1 gives the best response prices of a firm. Thus, when the leadtime decisions are fixed, such that firms compete purely using prices, equilibrium prices can be obtained in closed-form by the simultaneous solution of the 4 linear equations given by (10) and (11) (2 equations corresponding to each $i \in \{1, 2\}$).

Proposition 3. *Pure price competition in a DD setting has a Nash equilibrium. Further, if the firms are identical, then the equilibrium prices are symmetric, given by:*

$$
\begin{aligned}
p_h^*(L_h) =& \frac{(2\beta_p^l + 4\theta_p + \gamma_p)a - \{\beta_L^h(2\beta_p^l + 2\theta_p + \gamma_p) + (2\beta_p^l + \gamma_p)\theta_L\}L_h}{D_1} \\
& \frac{+\{(2\beta_p^l + \gamma_p)\theta_L - 2\beta_L^l\theta_p\}L_l}{D_1} \\
& \frac{+(2\beta_p^h\beta_p^l + 2\beta_p^h\theta_p + \beta_p^h\gamma_p + 2\beta_p^l\theta_p + 2\beta_p^l\gamma_p + 4\theta_p\gamma_p + \gamma_p^2)(A+m)}{D_1}
\end{aligned} \tag{21}
$$

$$
\begin{aligned}
p_l^*(L_h) =& \frac{(2\beta_p^h + 4\theta_p + \gamma_p)a + \{(2\beta_p^h + \gamma_p)\theta_L - 2\beta_L^h\theta_p\}L_h}{D_1} \\
& \frac{-\{\beta_L^l(2\beta_p^h + 2\theta_p + \gamma_p) + (2\beta_p^h + \gamma_p)\theta_L\}L_l}{D_1} \\
& \frac{+(2\beta_p^l\beta_p^h + 2\beta_p^l\theta_p + \beta_p^l\gamma_p + 2\beta_p^h\theta_p + 2\beta_p^h\gamma_p + 4\theta_p\gamma_p + \gamma_p^2)(A+m)}{D_1}
\end{aligned} \tag{22}
$$

where $D_1 = 4\beta_p^h\beta_p^l + 4\beta_p^h\theta_p + 2\beta_p^h\gamma_p + 4\beta_p^l\theta_p + 2\beta_p^l\gamma_p + 4\theta_p\gamma_p + \gamma_p^2$.

Proof. See Appendix E. □

The corresponding price differentiation for a given L_h is then:

$$
\begin{aligned}
p_h^*(L_h) - p_l^*(L_h) =& \frac{2(\beta_p^l - \beta_p^h)a + 2(\beta_p^l + \beta_p^h + \gamma_p)\theta_L(L_l - L_h) + \beta_L^l(2\beta_p^h + \gamma_p)L_l}{D_1} \\
& \frac{-(2\beta_p^l + \gamma_p)\beta_L^h L_h + (\beta_p^l - \beta_p^h)\gamma_p(A+m)}{D_1}
\end{aligned} \tag{23}
$$

5.1.1. Effect of Pure Price Competition in a DD Setting

We now study the effect of price competition on a firm's price and leadtime decisions in a dedicated capacity setting. We know competition generally drives prices down. But how does competition affect price differentiation/discrimination? To answer this, we compare the optimal prices of a monopolist with its equilibrium prices when it faces price competition from an identical firm. A monopolist setting can be represented using a competitive model for two identical firms, each with a market base $2a$, but with $\gamma_p = \gamma_L = 0$. This represents two identical firms operating in (geographically or otherwise) different markets such that they do not poach each other's market share. In contrast, a competitive setting represents a situation in which two firms operate in the same market, each with a market base $2a$, such that each firm's demand is affected by the relative prices of the two firms. Mathematically, this corresponds to $\gamma_p > 0$, $\gamma_L > 0$ in our competitive model.

Proposition 4. *Pure price competition in a dedicated capacity setting always results in: (a) a lower express price p_h^*, (a) a lower regular price p_l^*, and (c) a lower price differentiation ($p_h^* - p_l^*$). Further, the effects are more pronounced in presence of product substitution.*

Proof. See Appendix F □

The effect of competition on individual prices is not surprising. In fact, it is well established in theory that competition always decreases prices (Varian, 1989). A practical example is the ongoing price war between Amazon.com, Apple, Barnes & Noble, and Sony for their e-readers (Miller, 2010). However, as pointed out earlier in Section 1, the extant literature in Industrial Organization has very contradicting results on the effect of competition on price differentiation. Our model, with an important linkage between marketing decision of price discrimination and operation's capacity related decisions, provides results that concur with the traditional theory on price discrimination, which predicts that market competition decreases a firms ability to use price discrimination. Further, our results suggest that the effects of competition on individual prices as well as price discrimination are more pronounced in the presence of product substitution. This suggests that the degree of price discrimination (second degree in the presence of product substitution, and third degree in absence of product substitution) further plays a role in deciding the intensity of the effect of price competition.

5.2. Price and Leadtime Competition in a DD Setting

We now consider a more general situation where firms have flexibility in quoting the leadtimes to their express customers. We still assume there is a standard leadtime for regular customers established by the industry (Section 3 discusses the situations in which a standard leadtime for regular customers is justified). In such a situation, firms compete by strategically selecting the express leadtime in addition to the two prices. The equilibrium express leadtimes are given by the simultaneous solution of the system of 2 non-linear equations, given by (9) = 0 for $i = 1, 2$. In absence of a closed-form solution for this system of non-linear equations, we design an iterative procedure, described in Figure 2, that always converges to the equilibrium solution. We solve for an equilibrium solution assuming the game is played dynamically, starting at an initial solution, until none of the firms has an incentive to deviate from its decision unilaterally.

1. *Initialization*: For each firm i, set $p_h^i = p_l^i = m^i$, $L_h^i = 0$ or $L_h^i = L_l$.
2. *Iterative step*: Start with $i = 1$. Use the best response obtained for Firm i problem. Repeat this for $i = 2$.
3. *Convergence criteria*: Repeat step 2 until each firm's decision values differ from their previous values by less than some predetermined tolerance level ϵ.

Figure 2: Iterative Algorithm for Nash Equilibrium

Proposition 5. *The iterative algorithm given in Figure 2 converges to a Nash Equilibrium for a DD setting.*

Proof. See Appendix G. □

5.2.1. Effect of Price and Leadtime Competition

When firms use leadtime, in addition to price, as a strategic tool to attract demand and compete in the market, this leads to another question of interest: how does competition affect both price and leadtime differentiation? To answer this, we compare the equilibrium prices and leadtime decisions in a competitive setting with that under a monopolistic setting. The effect of competition on price and leadtime differentiation, in general, depends on the relative intensities of price competition (γ_p) and leadtime competition (γ_L), as well as other demand parameters. The following proposition summarizes the effect of competition for special cases.

Proposition 6. *Price and leadtime competition in a dedicated capacity setting: (a) decreases both leadtime differentiation and price differentiation when $\gamma_L = 0$; (b) increases both leadtime differentiation and price differentiation when $\gamma_p = 0$.*

Proof. See Appendix H □

The above proposition suggests that price and leadtime competition may increase or decrease price and leadtime differentiation, depending on customers' behavior. This is intuitive. When $\gamma_L = 0$, customers' choice of a firm is not influenced by the relative leadtimes but by the relative prices offered by the two firms. In such a situation, firms tend to cut prices to attract customers. At the same time, they increase their express leadtime (and hence decrease their leadtime differentiation) in order to cut their capacity cost and maintain their profit. It further follows from (23) that a smaller leadtime differentiation (and hence larger express leadtime) also results in a smaller price differentiation. On the other hand, when $\gamma_p = 0$, customers' choice of a firm is not influenced by the relative prices but by the relative leadtimes offered by the two firms. In such a situation, firms try to cut their leadtimes to attract customers. This results in a smaller express leadtime, and hence a larger leadtime differentiation. Again, it follows from (23) that a larger leadtime differentiation also allows the firms to maintain a larger price differentiation.

5.3. Effect of Capacity Strategy

We have thus far studied competition between firms that use dedicated capacities for different market segments. However, as discussed in Section 1, different firms, even in the same industry, use different capacity strategies. A natural question then is: how does firms' capacity strategy affect equilibrium price and leadtime decisions? Jayaswal et al. (2011) have addressed the same question but in a monopolistic setting. In this paper, we pose the same question in a competitive setting. To answer this, we compare the equilibrium decisions of otherwise symmetric firms under the three scenarios, described in Table **??**. The best response of a firm in a shared capacity setting, however, lacks an analytical characterization. Therefore, it is not possible to derive analytical results for equilibrium decisions when at least one of the competing firms uses shared capacities for its different market segments. We, therefore, test our models numerically under different combinations of parameter values. Our numerical results suggest that the price and leadtime competition always has a unique equilibrium, obtained using the algorithm described in Figure 2. Generalizations based on observable patterns that emerge from these numerical experiments are reported as observations.

Table 1: Market parameters used in numerical examples

Market Type ↓	β_p^h	β_p^l	θ_p	β_L^h	β_L^l	θ_L	γ_p	γ_L
TDS	0.55	0.75	0.15	0.9	0.7	0.5	0.4	0.4
PDS	0.5	0.7	0.4	0.9	0.7	0.1	0.4	0.4

Table 2: Firm-specific parameters used in numerical examples

Firm 1				**Firm 2**			
a^1	m^1	α^1	L_l^1	a^2	m^2	α^2	L_l^2
10	3	0.99	1	10	3	0.99	1

We present a small sample from our extensive numerical experiments to illustrate the effect of capacity strategy on equilibrium decisions. We use the parameter setting described in Tables 1 and 2 for two levels of capacity cost: (i) $A = 0.01$ (for small capacity cost) and (ii) $A = 0.25$ (for large capacity cost). A comparison of the equilibrium prices and leadtimes in an SS versus a DD setting is shown in Table 3, and for a DS setting is shown in Table 4. Results from the numerical experiments are summarized in the following Observation. It suggests that the effect of firms' capacity strategy in a monopolistic setting, as studied by Jayaswal et al. (2011), also extends to the competitive setting.

Table 3: Numerical Results for DD and SS settings

	A = 0.01				**A = 0.25**			
	PDS		*TDS*		*PDS*		*TDS*	
	DD	*SS*	*DD*	*SS*	*DD*	*SS*	*DD*	*SS*
L_h^*	0.079835	0.07984	0.075916	0.075927	0.40887	0.375944	0.393785	0.382392
$L_l - L_h^*$	0.920165	0.92016	0.924084	0.924073	0.59113	0.624056	0.606215	0.617608
p_h^*	8.48642	8.4861	8.487575	8.487409	8.475324	8.496719	8.397922	8.407862
p_l^*	7.4262	7.42033	6.748059	6.742129	7.536979	7.410724	6.933531	6.790338
$p_h^* - p_l^*$	1.06022	1.06577	1.739516	1.74528	0.938345	1.085995	1.464391	1.617524

Table 4: Numerical Results for DS setting

	A = 0.01				**A = 0.25**			
	PDS		*TDS*		*PDS*		*TDS*	
	D	*S*	*D*	*S*	*D*	*S*	*D*	*S*
L_h^*	0.079836	0.079838	0.075915	0.075928	0.409076	0.375427	0.393641	0.381862
$L_l - L_h^*$	0.920164	0.920162	0.924085	0.924072	0.590924	0.624573	0.606359	0.618138
p_h^*	8.486179	8.486345	8.48746	8.487527	8.469203	8.502327	8.395587	8.410847
p_l^*	7.425453	7.42108	6.747257	6.742933	7.520666	7.426818	6.914207	6.809853
$p_h - p_l^*$	1.060726	1.065265	1.740203	1.744595	0.948537	1.075509	1.481381	1.600994

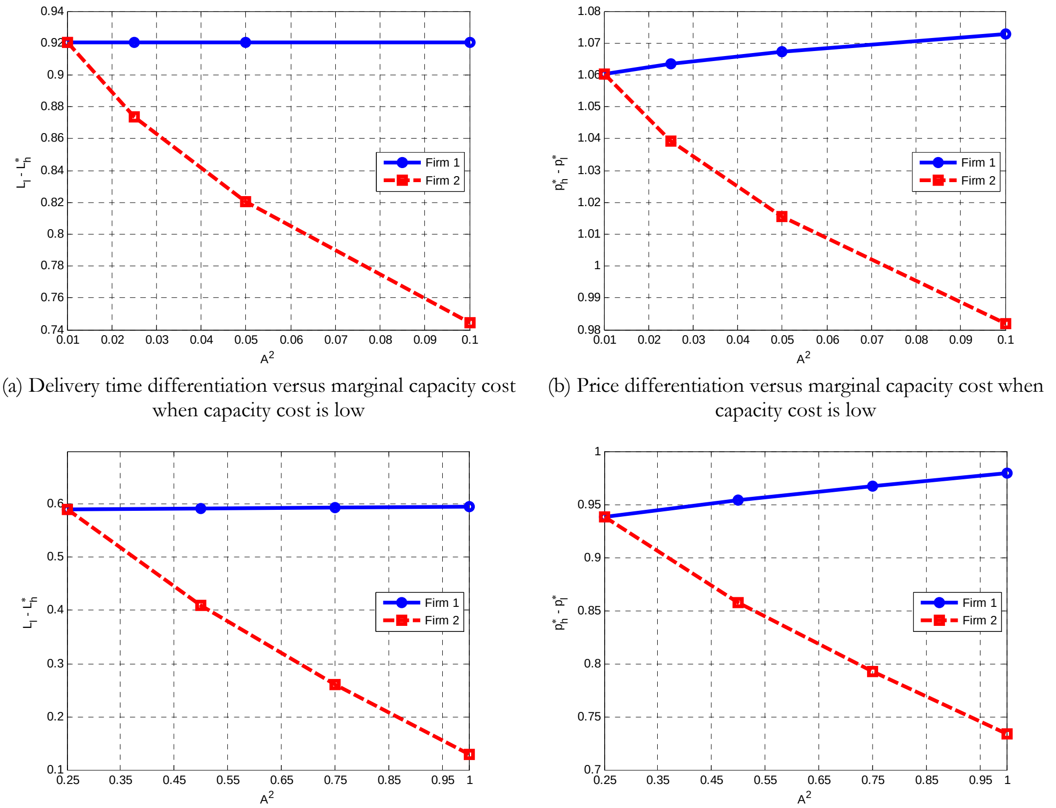
Observation 1: - *Price and leadtime competition under SS, compared to DD, results in: (a) a larger price differentiation at equilibrium, and (b) a larger leadtime differentiation at equilibrium if capacity cost is high, but a smaller leadtime differentiation when capacity cost is small.*
- *Price and leadtime competition under DS results in: (a) a larger price differentiation at equilibrium for the firm using shared capacities, and (b) a larger leadtime differentiation at equilibrium for the firm using shared capacities if capacity cost is high, but a smaller leadtime differentiation for the firm using shared capacities when capacity cost is small.*

5.4. Effect of Asymmetry Between Firms

We have thus far studied competition between firms that are symmetric with respect to their market base a, capacity cost A and operating cost m. When the competing firms are asymmetric, they will try to exploit their competitive advantage of a lower capacity cost A, or a higher market base a. We study the effects of such asymmetries on their equilibrium decisions.

5.4.1. Asymmetry in Capacity Cost

Observation 2: *If one of the firms, which are otherwise identical, has a higher capacity cost, then compared to the other firm at equilibrium:*
- in a DD setting, it has (a) a smaller leadtime differentiation, and (b) a smaller price differentiation (Refer to Figure 3).
- in an SS seting, it has (a) a smaller leadtime differentiation, and (b) a smaller price differentiation if the absolute capacity costs are very small, but a larger price differentiation if the absolute capacity costa are high (Refer to 4).



(a) Delivery time differentiation versus marginal capacity cost when capacity cost is low

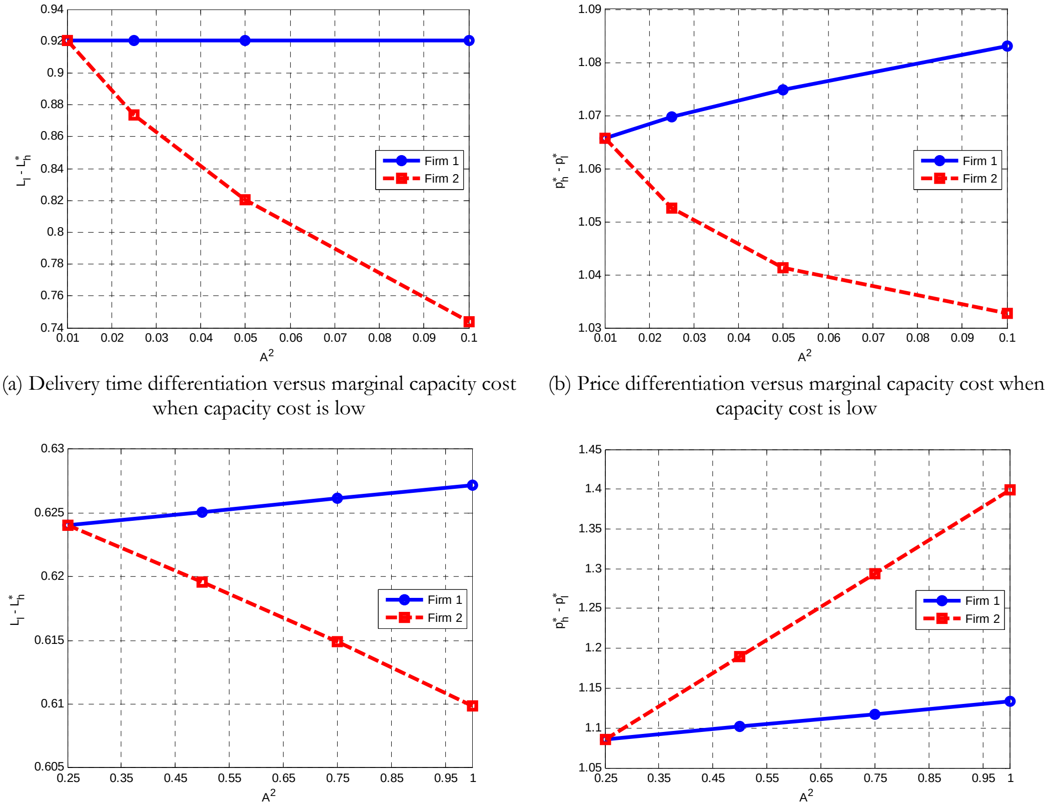
(b) Price differentiation versus marginal capacity cost when capacity cost is low

(c) Delivery time differentiation versus marginal capacity cost when capacity cost is high

(d) Price differentiation versus marginal capacity cost when capacity cost is high

Figure 3: Effects of capacity cost asymmetry on product differentiation decisions in a DD setting

Figure 3 shows the equilibrium price and leadtime differentiations of the two firms in a DD setting that differ in their capacity costs but are otherwise identical. Figure 4 shows similar plots for an SS setting. We show these plots for market parameter values shown in Table 1 corresponding to a PDS type market, although the qualitative results are independent of the specific market parameters. Firm-specific parameters are as shown in Table 2. In one set of experiments, we fix the capacity cost of firm 1, A^1, at 0.01 and vary that for firm 2, A^2, from 0.01 to 0.10. In another set of experiments, we fix A^1 at 0.25 and vary A^2 from 0.25 to 1.0. This helps us capture the effect of a larger capacity cost incurred by firm 2 on the decisions of the two firms at equilibrium. As evident from the



(a) Delivery time differentiation versus marginal capacity cost when capacity cost is low

(b) Price differentiation versus marginal capacity cost when capacity cost is low

(c) Delivery time differentiation versus marginal capacity cost when capacity cost is high

(d) Price differentiation versus marginal capacity cost when capacity cost is high

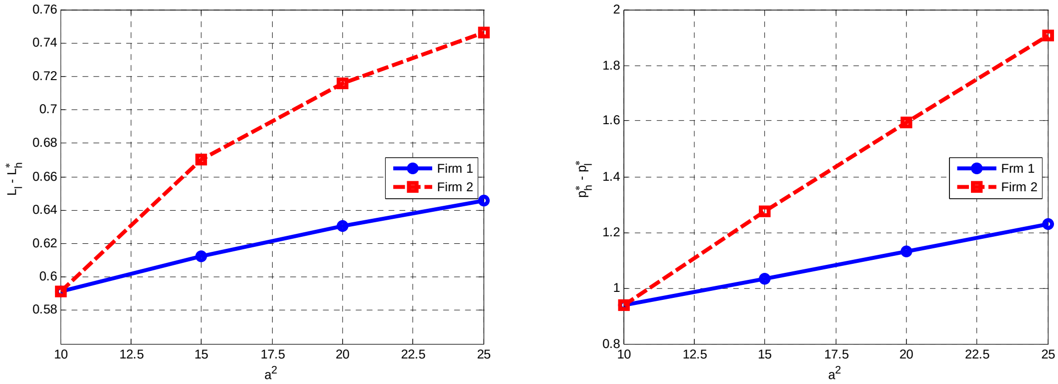
Figure 4: Effects of capacity cost asymmetry on product differentiation decisions in an SS setting

plots, when the firms are symmetric ($A^2 = A^1$), the leadtime and price differentiations of both firms coincide. Any increase in firm $2's$ capacity cost (A^2) always decreases its leadtime differentiation at equilibrium, irrespective of the capacity settings used by the two firms. An increase in A^2 also decreases firm $2's$ price differentiation at equilibrium in a DD setting. In an SS setting, an increase in A^2 decreases firm $2's$ price differentiation only when A^1 and A^2 are still small (≤ 0.1); when A^1 and A^2 are high (≥ 0.25), an increase in A^2 increases firm $2's$ price differentiation. However, the effect of an increase in firm $2's$ capacity cost A^2 may have a similar or a contrasting effect on firm 1, depending on the market parameters and the level of the capacity cost. Whatever be the effects on individual firms, when $A^2 > A^1$, firm 2 always has a smaller leadtime differentiation and a smaller price differentiation in a DD setting, but a higher price differentiation for larger absolute capacity costs in an SS setting.

5.4.2. Asymmetry in Market Base

Observation 3: *If one of the firms, which are otherwise identical, has a larger market base, then compared to the other firm at equilibrium:*
- it always has (a) a larger leadtime differentiation, and (b) a larger price differentiation, irrespective of the capacity strategy of either firm (Refer to Figures 5 and 6).

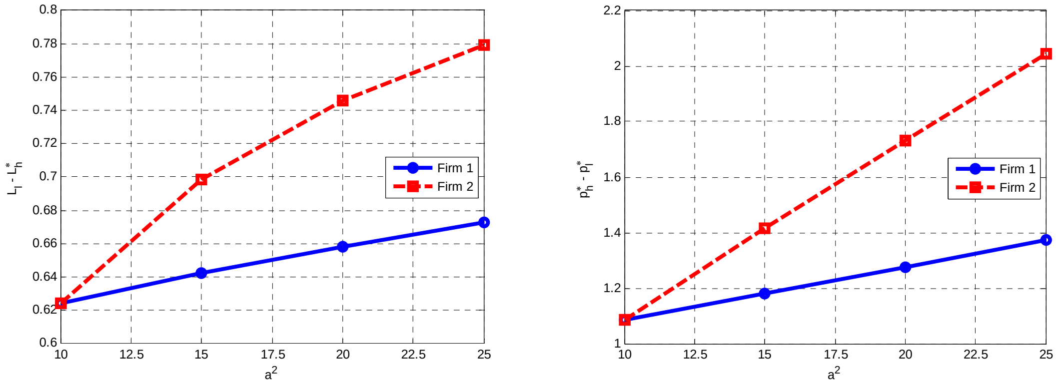
We illustrate this result using a sample from our numerical experiments. We consider two firms that have different market bases ($a^1 \neq a^2$), but are otherwise identical. Difference in the market bases of the two firms means that one firm always has a higher mean demand even if they both offer the same leadtimes at the same prices. This may be the



(a) Delivery time differentiation versus market base

(b) Price differentiation versus market base

Figure 5: Effects of asymmetry in market base on product differentiation decisions in a DD setting



(a) Delivery time differentiation versus market base

(b) Price differentiation versus market base

Figure 6: Effects of asymmetry in market base on product differentiation decisions in an SS setting

result of a difference in their brand appeal to the customers or due to a more convenient locations or a better customer experience at one of the firms. We assume the market is PDS type (parameter values shown in Table 1), although the generalizations drawn are independent of the specific market parameters. Firm specific parameters are as shown in Table 2. The market base a^1 for firm 1 is now fixed at 10, while that for firm 2 (a^2) is varied. Figures 5 and 5 show the equilibrium price and leadtime differentiations of the two firms in a DD and an SS setting, respectively. This helps us capture the effect of a larger market base of firm 2 on the decisions of the two firms at equilibrium. As evident from the plots, when the firms are symmetric ($a^2 = a^1$), the leadtime and price differentiations of both firms coincide. Any increase in firm $2's$ market base (a^2) increases its lead time differentiation as well as the price differentiation at equilibrium. Although firm $1's$ price and leadtime differentiation decisions also increase with a^2 in this case, this is specific only to this set of market parameters. In general, the behavior of firm $1's$ decisions depends on the market parameters. Whatever be the effects on individual firms, when $a^2 > a^1$, firm 2 always has a larger leadtime differentiation and a larger price differentiation, irrespective of the capacity strategy and market parameters.

6. Conclusion & Future Research

In this paper, we studied the product differentiation strategies of competing firms, which may use different capacity strategies. Our primary objective was to understand the effect of competition on a firm's product differentiation strategy since there is still no general agreement on it in the existing literature. We further investigated the effect of different (dedicated versus shared) capacity strategies used by competing firms on their product differentiation strategies at equilibrium. For this, we developed a general mathematical model, special cases of which captures a firm's best response to its competitor's decisions, depending on whether it uses dedicated or shared capacities to serve different market segments.

Our study provides insight into the effect of competition on price discrimination. We showed that when firms use dedicated capacities, pure price competition always reduces individual prices as well as price discrimination. However, when firms use leadtimes, in addition to prices, as strategic variables to compete in the market, the effect of competition on product differentiation further depends on customers' behavior. Our study also brings out the effect of firms' capacity strategy on their price and leadtime differentiation decisions. Specifically, when processing capacities are expensive, the firm with shared capacities should offer faster and more expensive product to time sensitive customers and slower and cheaper product to price sensitive customers compared to the firm using dedicated capacities. This implies that the firm with shared capacities should offer more differentiated products. Further, the above effect of the capacity strategy does not depend on any end customer characteristics or whether the products are substitutable or not. When asymmetry in capacity costs exists between firms, the way a firm should exploit its lower capacity cost further depends on its own capacity strategy and also of its competitor. Specifically, the firm with cheaper capacities should make its products more differentiated if both firms use dedicated capacities. If both firms use shared capacities, then the firm with cheaper capacities should again make its leadtimes more differentiated, but whether it should offer more homogeneous or more differentiated prices depends further on the capacity strategy used by the two frms as well as their level of capacity cost. Whereas the firm with a larger market base should always offer more differentiated products, irrespective of the capacity strategy of either firm.

There are a number of directions in which the current research can be extended. One possible extension would be to develop a good approximation for the sojourn time distribution $S_l(\cdot)$ of the low priority customers in a shared capacity setting, which can be used in the optimization model to obtain a closed-form analytical characterization of a firm's best response. This will also allow for a proof of convergence and uniqueness of the Nash Equilibrium when one of the firms uses shared capacities, which will further allow one to study the effect of competition on price discrimination even when firms use shared capacities for different market segments. Further, the mathematical model for the best response in a shared capacity setting can be extended to include delay dependent dynamic priority discipline. Another possible extension might be to include the guaranteed leadtime for the regular customers also as a decision variable.

Appendix A. Proof of Proposition 1

It is well known that at optimality, the two leadtime reliability constraints (7^{DC}) and (8^{DC}) must be binding (Palaka et al., 1998; So and Song, 1998; Boyaci and Ray, 2003).

This implies that the two service rates will be given by:

$$\mu_c^i = -\frac{\ln(1-\alpha)}{L_c^i} + \lambda_c^i \quad c \in \{l, h\}$$

As a result, $[PLDP_{DC}]$ reduces to maximizing (3) with μ_c^i as given above. The system stability conditions (6^{DC}) are automatically satisfied by the expressions for μ_c^i. Upon substituting the expressions for μ_c^i into (3), and taking it partial derivatives with respect to p_h^i and p_l^i gives the following Hessian for a fixed L_h^i:

$$\begin{pmatrix} -2(\beta_p^h + \theta_p + \gamma_p) & 2\theta_p \\ 2\theta_p & -2(\beta_p^l + \theta_p + \gamma_p) \end{pmatrix}$$

Clearly, the Hessian is negative definite. This shows that the objective function $\pi^i(L_h^i)$ is strictly concave for a fixed L_h^i, and, therefore, has a unique pair of optimal prices $p_h^{i*}(L_h^i)$ and $p_l^{i*}(L_h^i)$, which can be obtained by solving the following system of equations:

$$\frac{\partial \pi^i(L_h^i)}{\partial p_h^i} = 0; \ \frac{\partial \pi^i(L_h^i)}{\partial p_l^i} = 0$$

Substituting the optimal prices given by (10) and (11) into the objective function, and differentiating it with respect to L_h^i gives:

$$\frac{\partial \pi^i(L_h^i)}{\partial L_h^i} = -\left(\beta_L^h + \theta_L + \gamma_L\right)\left(p_h^{i*}(L_h^i) - m^i - A^i\right) + \theta_L\left(p_l^{i*}(L_h^i) - m^i - A^i\right) - \frac{A\ln(1-\alpha)}{(L_h^i)^2} \tag{A1}$$

$$\frac{\partial^2 \pi^i(L_h^i)}{\partial (L_h^i)^2} = -\left(\beta_L^h + \theta_L + \gamma_L\right)\left(\frac{\partial p_h^i(L_h^i)}{\partial L_h^i}\right) + \theta_L\left(\frac{\partial p_l^i(L_h^i)}{\partial L_h^i}\right) + \frac{2A^i ln(1-\alpha)}{(L_h^i)^3} \tag{A2}$$

$$\frac{\partial^3 \pi^i(L_h^i)}{\partial (L_h^i)^3} = -\frac{6A^i\ln(1-\alpha)}{(L_h^i)^4} \tag{A3}$$

The the first three derivatives of $\pi^i(L_h^i)$ suggests that it has the following properties: (i) As $L_h^i \to 0^+$, $\pi^i(L_h^i) \to -\infty$. (ii) $\pi^i(L_h^i)$ is increasing concave in L_h^i in the vicinity of $L_h^i = 0^+$. (iii) As L_h^i increases from 0, $\pi^i(L_h^i)$ changes from concave to convex for some $L_h^i \in (0, +\infty)$, and never becomes concave again. It is clear from the above properties of $\pi^i(L_h^i)$ that it has a unique maximum and at most one minimum in $[0, +\infty)$. The stationary points are given by the roots of (A1) in $[0, +\infty)$, and the maximum is always the smaller of the two. Further, $\frac{\partial \pi^i(L_h^i)}{\partial L_h^i}\big|_{L_h^i = L_l^i} < 0$ is sufficient to guarantee that (A1) has only one root in the interval $[0, L_l^i)$, and that it is the point of maximum. The condition simplifies to:

$$\frac{K_1 a^i + K_2 L_l^i + K_3 A^i + K_4 m^i}{2(\beta_p^h\beta_p^l + \beta_p^h\theta_p + \beta_p^l\theta_p + \beta_p^h\gamma_p + \beta_p^l\gamma_p + 2\theta_p\gamma_p + \gamma_p^2)} - \frac{A^i\ln(1-\alpha)}{(L_l^i)^2} < 0 \tag{A4}$$

where K_1, K_2, K_3, K_4 are functions only of the market parameters (β_p^c, β_L^c, θ_p, θ_L, γ_p, γ_L), and hence are constants. Further,

$$K_1 = -\left\{(\beta_p^l - \beta_p^h)\theta_L + (\beta_L^h + \gamma_L)(\beta_p^l + 2\theta_p + \gamma_p)\right\}$$

Since $\beta_p^h < \beta_p^l$, a necessary condition for (A4) to hold is a^i to be high. A sufficiently high value of a^i also guarantees $p_c^i > 0$, $p_h^i > p_l^i$ and $\lambda_c^i > 0$.

Appendix B. Matrix Geometric Method

Joint Stationary Queue Length Distribution: If we define $N_h(t)$ and $N_l(t)$ as state variables representing the number of high and low priority customers in the system at time t, then $\{\mathbf{N}(t)\} := \{N_l(t), N_h(t), t \geq 0\}$ is a continuous-time two-dimensional Markov chain with state space $\{\mathbf{n} = (n_l, n_h)\}$. The key idea we employ here is that $\{\mathbf{N}(t)\}$ is a *quasi-birth-and-death* (QBD) process, which allows us to develop a matrix geometric solution for the joint distribution of the number of customers of each class in the system. A simple implementation of the matrix geometric method, however, requires the number of states in the QBD process to be finite. For this, we treat the queue length of high priority customers (including the one in service) to be of finite size M, but of size large enough for the desired accuracy of our results. Since high priority customers are always served in priority over low priority customers, it is reasonable to assume that its queue size will always be bounded by some large number.

In the Markov process $\{\mathbf{N}(t)\}$, a transition can occur only if a customer of either class arrives or a customer of either class is served. The possible transitions are:

Table B.5: Transition rates for the priority queue

From	To	Rate	Condition
(n_l, n_h)	(n_l, n_h+1)	λ_h^i	for $n_l \geq 0$, $n_h \geq 0$
(n_l, n_h)	(n_l+1, n_h)	λ_l^i	for $n_l \geq 0$, $n_h \geq 0$
(n_l, n_h)	(n_l, n_h-1)	μ^i	for $n_l \geq 0$, $n_h > 0$
(n_l, n_h)	(n_l-1, n_h)	μ^i	for $n_l > 0$, $n_h = 0$

The infinitesimal generator Q associated with our system description is thus block-tridiagonal:

$$Q = \begin{pmatrix} B_0 & A_0 & & \\ A_2 & A_1 & A_0 & \\ & A_2 & A_1 & A_0 \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

where B_0, A_0, A_1, A_2 are square matrices of order $M+1$. These matrices can be easily constructed using the transition rates described above.

$$A_0 = \begin{pmatrix} \lambda_l^i & & & & \\ & \lambda_l^i & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \lambda_l^i \end{pmatrix}; \quad A_2 = \begin{pmatrix} \mu^i & & & & \\ & 0 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}; \quad B_0 = \begin{pmatrix} * & \lambda_h^i & & & \\ \mu^i & * & \lambda_h^i & & \\ & \mu^i & * & \lambda_h^i & \\ & & \ddots & \ddots & \ddots \\ & & & \mu^i & * \end{pmatrix}$$

where $*$ is such that $A_0\mathbf{e} + B_0\mathbf{e} = \mathbf{0}$. $A_1 = B_0 - A_2$.

We denote $\mathbf{x}$ as the stationary probability vector of $\{\mathbf{N}(t)\}$:

$$\mathbf{x} = [x_{00}, x_{01}, \ldots, x_{0M}, x_{10}, x_{11}, \ldots, x_{1M}, \ldots, \ldots, x_{n0}, x_{n1}, \ldots, x_{nM}, \ldots, \ldots]$$

The vector $\mathbf{x}$ can be partitioned by levels into sub vectors $\mathbf{x}_n$, $n \geq 0$, where $\mathbf{x}_n = [x_{n0},\ x_{n1},\ \ldots,\ x_{nM}]$ is the stationary probability of states in level n ($n_l = n$). Thus, $\mathbf{x} = [\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \ldots]$. $\mathbf{x}$ can be obtained using a set of balance equations, given in matrix form, by the following standard relations (Latouche and Ramaswami, 1999; Neuts, 1981):

$$\mathbf{x}Q = \mathbf{0}; \quad \mathbf{x}_{n+1} = \mathbf{x}_n R$$

where R is the minimal non-negative solution to the matrix quadratic equation:

$$A_0 + RA_1 + R^2 A_2 = \mathbf{0}$$

The matrix R can be computed using well known methods (Latouche and Ramaswami, 1999). A simple iterative procedure often used is:

$$R(0) = 0 \ ; \quad R(r+1) = -\left[A_0 + R^2(r)A_2\right] A_1^{-1}$$

The probabilities $\mathbf{x}_0$ are determined from:

$$\mathbf{x}_0(B_0 + RA_2) = \mathbf{0}$$

subject to the normalization equation:

$$\sum_{n=0}^{\infty} \mathbf{x}_n \mathbf{e} = \mathbf{x}_0 (I - R)^{-1} \mathbf{e} = 1$$

where $\mathbf{e}$ is a column vector of ones of size $M+1$.
Estimation of $S_l^i(\cdot)$: The leadtime W_l^i of a low priority customer is the time between its arrival to the system till it completes service. It may be preempted by one or more high priority customers for service. So it is difficult to characterize the distribution $S_l^i(\cdot)$. Ramaswami and Lucantoni (1985) present an efficient algorithm based on *uniformization* to derive the complimentary distribution of waiting times in phase-type and QBD processes. We adopt their algorithm to derive $S_l^i(\cdot)$, the distribution of the waiting time plus the time in service of low priority customers.

Consider a tagged low priority customer entering the system. The time spent by the tagged customer depends on the number of customers of either class already present in the system ahead of it, and also on the number of subsequent high priority arrivals before it completes its service. All subsequent low priority arrivals, however, have no influence on its time spent in the system. The tagged customer's time in the system is, therefore, simply the time until absorption in a modified Markov process $\{\tilde{\mathbf{N}}(t)\}$, obtained by setting $\lambda_l^i = 0$. Consequently, matrix $\tilde{A}_0$, representing transitions to a higher level, becomes a zero matrix. We define an *absorbing* state, call it state $0'$, as the state in which the tagged customer has finished its service. The infinitesimal generator for this process can be represented as:

$$\tilde{Q} = \left(\begin{array}{c|cccccc} 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline b_0 & \tilde{B}_0 & 0 & & & \\ 0 & A_2 & \tilde{A}_1 & 0 & & \\ 0 & & A_2 & \tilde{A}_1 & 0 & \\ \vdots & & & \ddots & \ddots & \ddots \end{array}\right)$$

where, $\tilde{B}_0 = B_0 + A_0$; $\tilde{A}_1 = A_1 + A_0$; and $b_0 = [\mu^i \quad 0 \quad \cdots \quad 0]^T_{M+1}$. The first row and column in $\tilde{Q}$ corresponds to the absorbing state $\acute{0}$. The time spent in system by the tagged customer, which is the time until absorption in the modified Markov process with rate matrix $\tilde{Q}$, depends on the prices (p_h^i and p_l^i), through the arrival rates (λ_h^i and λ_l^i), and the service rate μ^i. For given prices (p_h^{ik}, p_l^{ik}) and service rate μ^{ik}, the distribution of the time spent by a low priority customer in the system is $S_l^{ik}(y) = 1 - \overline{S_l^{ik}}(y)$, where $\overline{S_l^{ik}}(y)$ is the stationary probability that a low priority customer spends more than y units of time in the system. Further, let $\overline{S_{ln}^{k}}(y)$ denote the conditional probability that a tagged customer, who finds n low priority customers ahead of it, spends a time exceeding y in the system. The probability that a tagged customer finds n low priority customers is given, using the PASTA property, by $\mathbf{x}_n = \mathbf{x}_0 R^n$. $\overline{S_l^{ik}}(y)$ can be expressed as:

$$\overline{S_l^{ik}}(y) = \sum_{n=0}^{\infty} \mathbf{x}_n \overline{S_{ln}^{ik}}(y) \mathbf{e} \tag{B1}$$

$\overline{S_{ln}^{ik}}(y)$ can be computed more conveniently by uniformizing the Markov process $\{\tilde{\mathbf{N}}(t)\}$ with a Poisson process with rate γ, where

$$\gamma = \max_{0 \le m \le M} (-\tilde{A}_1)_{mm} = \max_{0 \le m \le M} - (A_0 + A_1)_{mm}$$

so that the rate matrix $\tilde{Q}$ is transformed into the discrete-time probability matrix:

$$\hat{Q} = \frac{1}{\gamma}\tilde{Q} + I = \left(\begin{array}{c|ccccc} 1 & 0 & 0 & 0 & 0 & \cdots \\ \hline \hat{b_0} & \hat{B}_0 & 0 & & & \\ 0 & \hat{A}_2 & \hat{A}_1 & 0 & & \\ 0 & & \hat{A}_2 & \hat{A}_1 & 0 & \\ \vdots & & & \ddots & \ddots & \ddots \end{array}\right)$$

where $\hat{A}_2 = \frac{A_2}{\gamma}$, $\hat{A}_1 = \frac{\tilde{A}_1}{\gamma} + I$, $\hat{b_0} = \frac{b_0}{\gamma}$. In this uniformized process, points of a Poisson process are generated with a rate γ, and transitions occur at these epochs only. The probability that r Poisson events are generated in time y equals $e^{-\gamma y}\frac{(\gamma y)^r}{r!}$. Suppose the tagged customer finds n low priority customers ahead of it. Then, for its time in system to exceed y, at most n of the r Poisson points may correspond to transitions to lower levels (i.e., service completions of low priority customers). Therefore,

$$\overline{S_{ln}^{ik}}(y) = \sum_{r=0}^{\infty} e^{-\gamma y}\frac{(\gamma y)^r}{r!} \sum_{v=0}^{n} G_v^{(r)} \mathbf{e}, \qquad n \ge 0 \tag{B2}$$

where, $G_v^{(r)}$ is a matrix such that its entries are the conditional probabilities, given that the system has made r transitions in the discrete-time Markov process with rate matrix $\hat{Q}$, that v of those transitions correspond to lower levels (i.e., service completions of low priority customers). Substituting the expression for $\overline{S_{ln}^{ik}}(y)$ from (B2) into (B1), we obtain:

$$\overline{S_l^{ik}}(y) = \sum_{r=0}^{\infty} d_r e^{-\gamma y}\frac{(\gamma y)^r}{r!} \tag{B3}$$

where, d_r is given by:

$$d_r = \sum_{n=0}^{\infty} \mathbf{x}_0 R^n \sum_{v=0}^{n} G_v^{(r)} \mathbf{e}, \qquad r \geq 0 \tag{B4}$$

Now,

$$\begin{aligned}
&\sum_{n=0}^{\infty} R^n \sum_{v=0}^{n} G_v^{(r)} \mathbf{e} \\
&= \sum_{n=0}^{r+1} R^n \sum_{v=0}^{n} G_v^{(r)} \mathbf{e} + \sum_{n=r+2}^{\infty} R^n \sum_{v=0}^{r} G_v^{(r)} \mathbf{e} && \big(\text{since } G_v^{(r)} = 0 \text{ for } v > r\big) \\
&= \sum_{v=0}^{r+1} \sum_{n=v}^{r+1} R^n G_v^{(r)} \mathbf{e} + (I-R)^{-1} R^{r+2} \mathbf{e} && \left(\text{since } \sum_{v=0}^{r} G_v^{(r)} \mathbf{e} = \mathbf{e}\right) \\
&= \sum_{v=0}^{r+1} (I-R)^{-1} (R^v - R^{r+2}) G_v^{(r)} \mathbf{e} + (I-R)^{-1} R^{r+2} \mathbf{e} \\
&= \sum_{v=0}^{r} (I-R)^{-1} R^v G_v^{(r)} \mathbf{e} + (I-R)^{-1} R^{r+1} G_{r+1}^{(r)} \mathbf{e} && \left(\text{since} \sum_{v=0}^{r+1} G_v^{(r)} \mathbf{e} = \mathbf{e}\right) \\
&= \sum_{v=0}^{r} (I-R)^{-1} R^v G_v^{(r)} \mathbf{e} && \big(\text{since } G_v^{(r)} = 0 \text{ for } v > r\big) \\
&= (I-R)^{-1} H_r \mathbf{e} && r \geq 0
\end{aligned}$$

where, $H_r = \sum_{v=0}^{r} R^v G_v^{(r)}$. Therefore,

$$S_l^{ik}(L_l^i) = 1 - \overline{S_l^{ik}}(L_l^i) = \sum_{r=0}^{\infty} e^{-\gamma L_l} \frac{(\gamma L_l)^r}{r!} \mathbf{x}_0 (I-R)^{-1} H_r \mathbf{e} \tag{B5}$$

H_r can be computed recursively as:

$$H_{r+1} = H_r \hat{A}_1 + R H_r \hat{A}_2; \quad H_0 = I$$

Therefore, for given prices (p_h^{ik}, p_l^{ik}) and service rate (μ^{ik}), $S_l^{ik}(\cdot)$ in (16) can be computed using (B5).

Appendix C. Estimation of the Gradient of $S_l^i(\cdot)$

There are several methods available in the literature to compute the gradients of $S_l^i(\cdot)$. We use a *finite difference* method as it is probably the simplest and most intuitive, and can be easily explained (Atlason et al., 2004). Using the finite difference method, the gradients can be computed as:

$$\frac{\partial S_l^{ik}(\cdot)}{\partial p_h^i} = \frac{S_l^i(\cdot)\big|_{(p_h^{ik}+dp_h^i, p_l^i, \mu^i)} - S_l^{ik}(\cdot)\big|_{(p_h^{ik}-dp_h^i, p_l^i, \mu^i)}}{2dp_h^i}$$

$$\frac{\partial S_l^{ik}(\cdot)}{\partial p_l^i} = \frac{S_l^{ik}(\cdot)\big|_{(p_h^i, p_l^{ik}+dp_l^i, \mu^i)} - S_l^{ik}(\cdot)\big|_{(p_h^i, p_l^{ik}-dp_l^i, \mu^i)}}{2dp_l^i}$$

$$\frac{\partial S_l^{ik}(\cdot)}{\partial \mu^i} = \frac{S_l^{ik}(\cdot)\big|_{(p_h^i, p_l^i, \mu^{ik}+d\mu^i)} - S_l^{ik}(\cdot)\big|_{(p_h^i, p_l^i, \mu^{ik}-d\mu^i)}}{2d\mu^i}$$

where dp_h^i, dp_l^i and $d\mu^i$ (referred to as step sizes) are infinitesimal changes in the respective variables.

Appendix D. The Cutting Plane Algorithm

We now describe the cutting plane algorithm to solve $[PDP_{(K)}]$. The algorithm fits the framework of Kelley's cutting plane method (Kelley, 1960). It differs from the traditional description of the algorithm in that we use the matrix geometric method to generate the cuts and evaluate the function values instead of having an algebraic form for the function and using analytically determined gradients to generate the cuts. Figure D.7 shows a flowchart of the cutting plane algorithm. The algorithm works as follows: We

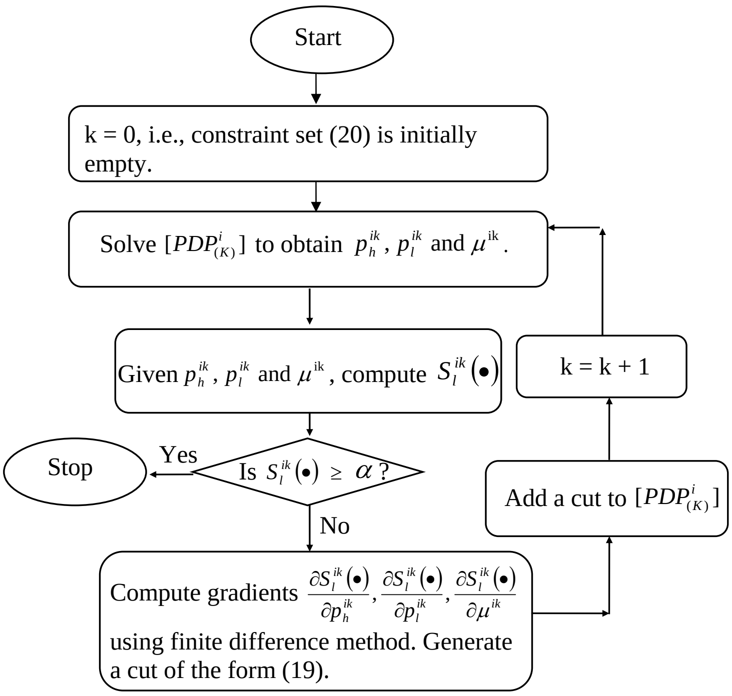


Figure D.7: Cutting Plane Algorithm

start with an empty constraint set (20), which results in a simple QPP, and obtain an initial solution $k_0 := (p_h^{i0},\ p_l^{i0},\ \mu^{i0})$. We use the matrix geometric method to compute the distribution $S_l^{ik_0}(\cdot)$ of W_l^i. If $S_l^{ik_0}(\cdot)$ meets the leadtime reliability constraint α, we stop with an optimal solution to $[PDP^i_{(K)}]$, else we add to (20) a linear constraint/cut generated using the finite difference method. The new cut eliminates the current solution but does not eliminate any feasible solution to $[PDP^i_{(K)}]$. This procedure repeats until the leadtime reliability constraint is satisfied within a sufficiently small tolerance limit ϵ such that $|S^i_l(\cdot) - \alpha| \leq \epsilon$. The method has been proved to converge (Atlason et al., 2004).

The success of the cutting plane algorithm relies on the concavity of $S^i_l(\cdot)$. We have already demonstrated, using computational results obtained by the matrix geometric method, that $S^i_l(\cdot)$ is concave in $(p^i_h,\ p^i_l)$ and separately concave in μ^i. However, it is difficult to establish the joint concavity of $S^i_l(\cdot)$ in $(p^i_h,\ p^i_l,\ \mu^i)$. If the concavity assumption is violated, then the algorithm may cut off parts of the feasible region and terminate with a solution that is suboptimal. We include a test to ensure the concavity assumption is not violated. This is done by ensuring that a new point, visited by the cutting plane algorithm after each iteration, lies below all the previously defined cuts, and that all previous points lie below the newly added cut. The test, however, cannot ensure that $S^i_l(\cdot)$ is concave unless it examines all the points in the feasible region. Still, it does help ensure that the concavity assumption is not violated at least in the region visited by the algorithm. Details of the test can be found in Atlason et al. (2004).

Appendix E. Proof of Proposition 3

The equilibrium prices are given by the simultaneous solution of the 4 linear equations given by (10) and (11) for $i \in \{1, 2\}$. The system of equations in matrix notation is given by $\mathbf{Ax} = \mathbf{b}$.

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & \frac{-(\beta^l_p+\theta_p+\gamma_p)\gamma_p}{2D} & \frac{-\theta_p\gamma_p}{2D} \\ 0 & 1 & \frac{-\theta_p\gamma_p}{2D} & \frac{-(\beta^h_p+\theta_p+\gamma_p)\gamma_p}{2D} \\ \frac{-(\beta^l_p+\theta_p+\gamma_p)\gamma_p}{2D} & \frac{-\theta_p\gamma_p}{2D} & 1 & 0 \\ \frac{-\theta_p\gamma_p}{2D} & \frac{-(\beta^h_p+\theta_p+\gamma_p)\gamma_p}{2D} & 0 & 1 \end{pmatrix} \tag{E1}$$

where $D = \beta^h_p\beta^l_p + \beta^h_p\theta_p + \beta^l_p\theta_p + \beta^h_p\gamma_p + \beta^l_p\gamma_p + 2\theta_p\gamma_p + \gamma_p^2$

$$\mathbf{x} = \begin{pmatrix} p^{1*}_h & p^{1*}_l & p^{2*}_h & p^{2*}_l \end{pmatrix}^T$$

and $\mathbf{b}$ is a 4x1 matrix of constants. $\mathbf{A}$ is symmetric and strictly diagonally dominant since we have $A_{ij} = A_{ji}\ \forall i, j$ and $\sum_{j \neq i} |A_{ij}| < A_{ii}\ \forall i$. Hence, $\mathbf{A}$ is positive definite (Horn and Johnson, 1985). This implies that $\mathbf{A}$ is full-rank, and hence the system of linear equations $\mathbf{Ax} = \mathbf{b}$ has a unique solution. This proves the uniqueness of the equilibrium.

Further, when the firms are identical, they have the same operating parameter values ($a^1 = a^2$; $m^1 = m^2$; $A^1 = A^2$; $\alpha^1 = \alpha^2$; $L^1_l = L^2_l$; $L^1_h = L^2_h$). Denote the equilibrium solution by the 2-tuple $(s^{1*}(L_h), s^{2*}(L_h))$, where $s^{i*}(L_h) := (p^{i*}_h(L_h), p^{i*}_l(L_h))$. Assume the contrary that the equilibrium solution is not symmetric, i.e., $s^{1*}(L_h) \neq s^{2*}(L_h)$. Since the two firms are identical, this implies that (s^{2*}, s^{1*}) must also be a Nash Equilibrium, which contradicts the uniqueness of the Nash Equilibrium. Hence, $s^{1*}(L_h) = s^{2*}(L_h)$. Substituting $p^{1*}_h(L_h) = p^{2*}_h(L_h) = p^*_h(L_h)$ and $p^{1*}_l(L_h) = p^{2*}_l(L_h) = p^*_l(L_h)$ in the expressions for the best response prices, given by (10) and (11), and solving the resulting system of 2 equations in 2 unknown gives (21) and (22).

Appendix F. Proof of Proposition 4

The duopoly prices under pure price competition are given by (21) and (22). The monopolist prices can be obtained from (21) and (22) by substituting $\gamma_p = \gamma_L = 0$. Comparing the monopolist prices with the duopoly prices, we get:

$$
\begin{aligned}
&p_h^{DD*}(L_h)\Big|_{duopoly} - p_h^{DC*}(L_h)\Big|_{monopoly} \\
&= \frac{-\gamma_p \left\{K_1^h a + K_2^h L_h + K_3^h L_l + K_4^h (A+m)\right\}}{4\beta_p^h\beta_p^l + 4\beta_p^h\theta_p + 4\beta_p^l\theta_p + 2\beta_p^h\gamma_p + 2\beta_p^l\gamma_p + 4\theta_p\gamma_p + \gamma_p^2}
\end{aligned} \tag{F1}
$$

$$
\begin{aligned}
&p_l^{DD*}(L_h)\Big|_{duopoly} - p_l^{DC*}(L_h)\Big|_{monopoly} \\
&= \frac{-\gamma_p \left\{K_1^l a + K_2^l L_h + K_3^l L_l + K_4^l (A+m)\right\}}{4\beta_p^h\beta_p^l + 4\beta_p^h\theta_p + 4\beta_p^l\theta_p + 2\beta_p^h\gamma_p + 2\beta_p^l\gamma_p + 4\theta_p\gamma_p + \gamma_p^2}
\end{aligned} \tag{F2}
$$

$$
\begin{aligned}
&\left(p_h^{DD*}(L_h) - p_l^{DD*}(L_h)\right)\Big|_{duopoly} - \left(p_h^{DC*}(L_h) - p_l^{DC*}(L_h)\right)\Big|_{monopoly} \\
&= \frac{-\gamma_p \left\{K_1^d a + K_2^d L_h + K_3^d L_l + K_4^d (A+m)\right\}}{4\beta_p^h\beta_p^l + 4\beta_p^h\theta_p + 4\beta_p^l\theta_p + 2\beta_p^h\gamma_p + 2\beta_p^l\gamma_p + 4\theta_p\gamma_p + \gamma_p^2}
\end{aligned} \tag{F3}
$$

where, $K_i^d = K_i^h - K_i^l$, and K_i^h and K_i^l for $i \in \{1,4\}$ are some functions of system parameters, and hence Clearly, when $\gamma_p = 0$, $p_h^{DD*}(L_h)\Big|_{duopoly} = p_h^{DC*}(L_h)\Big|_{monopoly}$ and $p_l^{DD*}(L_h)\Big|_{duopoly} = p_l^{DC*}(L_h)\Big|_{monopoly}$. For, $\gamma_p > 0$, (F1), (F2) and (F3) are dictated mainly by K_1^h and K_1^l and K_1^d, respectively since a is assumed to be large. Further,

$$K_1^h = 2(\beta_p^l)^2 + 2\beta_p^h\theta_p + 6\beta_p^l\theta_p + 8\theta_p^2 + \beta_p^l\gamma_p + 2\theta_p\gamma_p > 0$$

$$K_1^l = 2(\beta_p^h)^2 + 6\beta_p^h\theta_p + 2\beta_p^l\theta_p + 8\theta_p^2 + \beta_p^h\gamma_p + 2\theta_p\gamma_p > 0$$

$$K_1^d = (\beta_p^l - \beta_p^h)\gamma_p + 2\{(\beta_p^l)^2 - (\beta_p^h)^2\} + 4(\beta_p^l - \beta_p^h)\theta_p > 0$$

Therefore, $K_1^h > 0$, $K_1^l > 0$ and $K_1^d > 0 \Rightarrow$ (F1) < 0, (F2) < 0 and (F3) < 0, respectively if $\gamma_p > 0$. This shows that pure price competition decreases both the express and regular prices as well as the price differentiation. Further, it clearly follows from the expressions for K_1^h, K_1^l and K_1^d that the effects are more pronounced when $\theta_p > 0$, i.e., in presence of product substitution.

Appendix G. Proof of Proposition 5

Given the strategy of firm $j \in \{1,2\}$, the best response express leadtime of firm $i = 3 - j$ satisfies:

$$\frac{\partial \pi^i}{\partial L_h^i} = 0$$

Taking the total derivative of the above relation with respect to the express leadtime L_h^j of firm j, we get:

$$\frac{d}{dL_h^j}\left(\frac{\partial \pi^i}{\partial L_h^i}\right) = \frac{\partial}{\partial L_h^j}\left(\frac{\partial \pi^i}{\partial L_h^i}\right) + \frac{\partial}{\partial p_h^j}\left(\frac{\partial \pi^i}{\partial L_h^i}\right)\frac{\partial p_h^j}{\partial L_h^j} + \frac{\partial}{\partial p_l^j}\left(\frac{\partial \pi^i}{\partial L_h^i}\right)\frac{\partial p_l^j}{\partial L_h^j} + \frac{\partial}{\partial L_h^i}\left(\frac{\partial \pi^i}{\partial L_h^i}\right)\frac{dL_h^i}{dL_h^j} = 0$$

$$\Rightarrow \frac{dL_h^i}{dL_h^j} = \frac{-\left[\frac{\partial^2 \pi^i}{\partial L_h^j \partial L_h^i} + \frac{\partial^2 \pi^i}{\partial p_h^j \partial L_h^i}\frac{\partial p_h^j}{\partial L_h^j} + \frac{\partial^2 \pi^i}{\partial p_l^j \partial L_h^i}\frac{\partial p_l^j}{\partial L_h^j}\right]}{\frac{\partial^2 \pi^i}{\partial (L_h^i)^2}}$$

For a DD setting, the above relation simplifies to:

$$\frac{dL_h^i}{dL_h^j} = \frac{-\left[\gamma_p\left\{\left(\frac{\partial p_h^j}{\partial L_h^j}\right)^2 + \left(\frac{\partial p_l^j}{\partial L_h^j}\right)^2\right\} + \gamma_L\left(\frac{\partial p_h^j}{\partial L_h^j}\right)\right]}{\frac{\partial^2 \pi^i}{\partial (L_h^i)^2}} \tag{G1}$$

We know that for $L_h \leq L_h^*$:

$$\frac{\partial^2 \pi^i}{\partial (L_h^i)^2} < 0$$

The numerator in RHS of (G1) consists of terms that are functions only of the market parameters, and hence is a constant for a given parameter setting. Further,

$$\gamma_p\left\{\left(\frac{\partial p_h^j}{\partial L_h^j}\right)^2 + \left(\frac{\partial p_l^j}{\partial L_h^j}\right)^2\right\} > 0 \; and \; \gamma_L\left(\frac{\partial p_h^j}{\partial L_h^j}\right) < 0$$

Therefore, we have:

$$\frac{dL_h^i}{dL_h^j} \geq 0 \; if \; \gamma_p\left\{\left(\frac{\partial p_h^j}{\partial L_h^j}\right)^2 + \left(\frac{\partial p_l^j}{\partial L_h^j}\right)^2\right\} \geq \gamma_L\left(\frac{\partial p_h^j}{\partial L_h^j}\right) \tag{G2}$$

$$\frac{dL_h^i}{dL_h^j} < 0 \; if \; \gamma_p\left\{\left(\frac{\partial p_h^j}{\partial L_h^j}\right)^2 + \left(\frac{\partial p_l^j}{\partial L_h^j}\right)^2\right\} < \gamma_L\left(\frac{\partial p_h^j}{\partial L_h^j}\right) \tag{G3}$$

This suggests that if the market parameters are such that (G2) holds, firm i always increases (decreases) its express leadtime L_h^i in response to a corresponding increase (decrease) in firm $j's$ express leadtime L_h^j. We let $p_h^i(n)$, $p_l^i(n)$ and $L_h^i(n)$ be the best response decisions of firm i at the n^{th} iteration of the procedure. If $L_h^i(0) = 0$, then $L_h^i(n) \geq L_h^i(0)$ for all n. We will show that if (G2) holds, $L_h^i(n)$ is increasing in n for $i \in \{1, 2\}$. As L_h^i is bounded above ($L_h^i < L_l$), for $i \in \{1, 2\}$, this will establish that the iterative procedure converges. We prove the convergence by induction as follows:

1. (Step $n = 1$): We know that $L_h^i(1) \geq L_h^i(0)$ for $i \in \{1, 2\}$.
2. (Step $n - 1$): Assume that $L_h^i(n-1) \geq L_h^i(n-2)$ for $i \in \{1, 2\}$.
3. (Step n): Given the inductive assumption from Step $n - 1$, (G2) implies that

$L_h^i(n) \geq L_h^i(n-1)$ for $i \in \{1,2\}$.

This completes our induction. In case (G3) holds, convergence of the algorithm can proved similarly by letting $L_h^1(0) = L_l$ and $L_h^2(0) = 0$ and by showing that $L_h^1(n)$ is decreasing in n while $L_h^2(n)$ is increasing in n.

Appendix H. Proof of Proposition 6

The effect of competition on the express leadtime when firms use dedicated capacities is given by:

$$\left.\frac{\partial \pi(L_h)}{\partial L_h}\right|_{duopoly} - \left.\frac{\partial \pi(L_h)}{\partial L_h}\right|_{monopoly}$$
$$= \frac{-\{K_1 a + K_2 L_h + K_3 L_l + K_4(A+m)\}}{2(4\beta_p^h\beta_p^l + 4\beta_p^h\theta_p + 4\beta_p^l\theta_p + 2\beta_p^h\gamma_p + 2\beta_p^l\gamma_p + 4\theta_p\gamma_p + \gamma_p^2)(\beta_p^h\beta_p^l + \beta_p^h\theta_p + \beta_p^l\theta_p)} \quad \text{(H1)}$$

where, K_1, K_2, K_3 and K_4 are some functions only of the system parameters, and hence are constants. For large a, (H1) is dictated mainly by K_1, which is given by:

$$\begin{aligned} K_1 = &\left\{4\beta_p^h(\beta_p^l)^2 + 4(\beta_p^l)^2\theta_p + 8\beta_p^h\theta_p^2 + 8\beta_p^l\theta_p^2 + 12\beta_p^h\beta_p^l\theta_p\right\}\gamma_L \\ &- \left\{\beta_L^h\beta_p^l + 2\beta_L^h\theta_p + [(\beta_p^l)^2 - (\beta_p^h)^2]\theta_L\right\}\gamma_p^2 \\ &- \left\{2\beta_p^h\beta_L^h\theta_p + 2(\beta_p^l)^2\beta_L^h + 4\beta_p^l\theta_L + 8\beta_L^h\theta_p^2\right\}\gamma_p \\ &- \left\{[6\beta_p^l\beta_L^h - 4\beta_p^h\theta_L]\theta_p + 2[(\beta_p^l)^2 - (\beta_p^h)^2]\theta_L\right\}\gamma_p \\ &+2\left\{\beta_p^h\beta_p^l + \beta_p^h\theta_p + \beta_p^l\theta_p\right\}\gamma_L\gamma_p \end{aligned} \quad \text{(H2)}$$

Clearly, the effect of competition on L_h, and hence on leadtime differentiation, depends on the relative intensities of price competition (γ_p) and leadtime competition (γ_L), as well as other demand parameters. $\gamma_p = 0$ and $\gamma_L > 0$ results in (H2) > 0, and hence (H1) < 0. Further, $\pi(L_h)$ is increasing concave in L_h for $L_h \leq L_h^{DC*}$ (see Appendix A). This, together with (H1) < 0, implease that:

$$L_h^*|_{duopoly} := \{L_h|_{duopoly} : \left.\frac{\partial \pi(L_h)}{\partial L_h}\right|_{duopoly} = 0\} < L_h^*|_{monopoly} := \{L_h|_{monopoly} : \left.\frac{\partial \pi(L_h)}{\partial L_h}\right|_{monopoly} = 0\}$$

This implies that L_h is smaller under competition when $\gamma_p = 0$. Further, (F1), (F2) and (F3) suggest that for a given L_h, the equilibrium prices as well as the price differentiation coincide with the monopolist prices and price differentiation for $\gamma_p = 0$. However, a smaller L_h under monopoly compared to duopoly results in a larger price differentiation for $\gamma_p = 0$. $\gamma_p > 0$ and $\gamma_L = 0$, on the other hand, results in (H2) < 0, and hence (H1) > 0. Thus, L_h is larger under competition. A larger L_h results in a smaller price differentiation.

Allon, G., Federgruen, A., 2007. Competition in service industries. Operations Research, 53, 37–55.

Allon, G., Federgruen, A., 2008. Competition in service industries wit segmented markets. Working paper. Graduate School of Business, Columbia University, NY, USA.

Armony, H., Haviv, M., 2003. Price and delay competition between two service providers. European Journal of Operational Research, 147, 32–50.

Atlason J., Epelman M.A., Henderson S.G., 2004. Call center staffing with simulation and cutting plane methods. Annals of Operations Research, 127, 333–358.

Blackburn, J.D., 1991. Time-Based Competition: The Next Battleground in American Manufacturing. Business One Irwin, Homewood, IL, USA.

Borenstein, S., 1989. Hubs and high fares: Dominance and market power in the U.S. airline industry. The RAND Journal of Economics, 20, 344–365.

Borenstein, S., Rose, N., 1994. Competition and price dispersion in the U. S. airline industry. The Journal of Political Economy, 102, 653–683.

Boyaci, T., Ray, S., 2003. Product differentiation and capacity selection cost interaction in time and price sensitive markets. Manufacturing and Service Operations Management, 5, 18–36.

Boyaci, T., Ray, S., 2006. The impact of capacity costs on product differentiation in delivery time, delivery reliability and price. Production and Operations Management, 15, 179-197.

Cachon, G.P., Harker, P.T., 2002. Competition and outsourcing with scale economies. Management Science, 48, 1314–1333.

Chang, W., 1965. Preemptive priority queues. Operations Research, 13, 820–827.

Chen, H., Wan, Y., 2003. Price competition of make-to-order firms. IIE Transactions, 35, 817–832.

Gale, I., 1993. Price dispersion in a market with advanced purchases. Review of Industrial Organization, 8, 451–464.

Gerardi, K.S., Shapiro, A.H., 2007. Does competition reduce price discrimination? New evidence from the airline industry. Working Paper No. 07-7. Federal Reserve Bank of Boston.

Horn, R.A., Johnson, C.R., 1985. Matrix Analysis. Cambridge University Press, Cambridge, UK.

Jayaswal, S., Jewkes, E.M., Ray, S., 2011. Product differentiation and operations strategy in a capacitated environment. European Journal of Operational Research, 210, 716–728.

Kelley, J.E. Jr., 1960. The cutting plane method for solving convex programs. Journal of the Society for Industrial and Applied Mathematics, 8, 703–711.

Latouche, G., Ramaswami, V., 1999. Introduction to matrix analytic methods in stochastic modeling. The American Statistical Association and the Society for Industrial and Applied Mathematics, Philadelphia, USA.

Lederer, P.J., Li, L., 1997. Pricing, production, scheduling and delivery-time computation. Operations Research 45, 407–420.

Levhari, D., Luski, I., 1978. Duopoly pricing and waiting lines. European Economic Review 11, 17–35.

Loch, C., 1991. Pricing in markets sensitive to delay. Ph.D. thesis, Stanford University, CA, USA.

Luenberger, D.G., 1984. Linear and Nonlinear Programming. Addison-Wiley Publishing Company, CA, USA.

Maltz, A., Maltz, E., 1998. Customer service in the distribution channel: Empirical findings. Journal of Business Logistics, 19, 103–120.

Miller, C.C., 2010. In price war, new kindle sells for $139. The New York Times, July 28.

Neuts, F.M., 1981. Matrix Geometric Solutions in Stochastic Models: An Algorithmic Approach. Dover Publications, Mineola, NY, USA.

Pekgun, P., 2007. An analysis of pricing and leadtime policies within the marketing/operations interface. PhD thesis. H. Milton School of Industrial Engineering Georgia Institute of Technology, Atlanta, Georgia, USA.

Pekgun, P., Griffin, P.M., Keskinocak, P., 2006. Centralized vs. decentralized competition for price and lead-time sensitive demand. Working paper. H. Milton School of Systems and Industrial Engineering, Georgia Institute of Technology, Atlanta, GA, USA.

Ramaswami, V., Lucantoni, D.M., 1985. Stationary waiting time distribution in queues with phase type service and in quasi-birth-and-death processes. Stochastic Models 1, 125–136.

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, 769–781.

Rochet, J., Stole, L., 2002. Nonlinear pricing with random participation. The Review of Economic Studies, 69, 277–311.

Ross, S.M., 2010. Introduction to Probability Models, tenth ed. Academic Press, San Diego, USA.

Schmidt, M.S., Aschkenase, S., 2004. The building blocks of service. Supply Chain Management Review, 8, 34–40.

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

Stole, L., 1995. Nonlinear pricing and oligopoly. Journal of Economics & Management Strategy, 44, 529–562.

Talluri, K.T., Van Ryzin, G.J., 2004. The Theory and Practice of Revenue Management. Kluwer Academic Publishers, Boston, USA.

Tirole, J., 1998. The Theory of Industrial Organization. MIT Press, Cambridge, MA, USA.

Tsay, A., Agrawal, A.N., 2000. Channel dynamics under price and service competition. Manufacturing and Service Operations Management, 2, 272–391.

Varian, H.R., 1985. Price discrimination and social welfare. American Economic Review, 75, 870–875.

Varian, H.R., 1989. Price discrimination. eds, Schmalensee, R., Willig, R. Handbook of Industrial Organization, Elsevier Science Publishers, North-Holland, Amsterdam.

Wolfe, P., 1959. The simplex method for quadratic programming. Econometrica, 27, 382–398.

Zhao, X., Stecke, K.E., Prasad, A., 2008. Lead time and price quotation mode selection: Uniform or differentiated? Working paper. The School of Management, The University of Texas at Dallas, TX, USA.