ISE Industrial and Systems Engineering

Capacity Rationing for Contract Manufacturers Serving Multiple Markets

Banu Gemici Ozkan
Lehigh University

S. David Wu
Lehigh University

Report: 09T-015

Capacity Rationing for Contract Manufacturers Serving Multiple Markets

Banu Gemici Ozkan, S. David Wu

bag4@lehigh.edu, sdw1@lehigh.edu

Industrial and Systems Engineering Department,

Lehigh University, Bethlehem, PA

January, 2009

1 Introduction

The purpose of this research is to provide model, analyze and develop decision tools for contract manufacturers in high-tech industries, and facilitate their decisions related to capacity and supply planning. Despite its enormous development potentials, the issues facing the contract manufacturers are not well studied. We aim to close this gap by providing an analysis of capacity allocation problems in two different settings and proposing methodologies at operational and tactical levels.

Contract manufacturing in electronic industry has grown from a few billion dollar industry in the early 1990's to over $300 billion in 2008. The share of manufacturing done on contract basis in US was well over 50% (Gartner, 2004 [13]), and contract manufacturers in electronics industry represent a dominating force and economic driver for regions such as Mainland China, Taiwan, Korea and Malaysia.

By consolidating demands from different brand carrying customers, the contract manufacturers are able to realize a much higher utilization on their equipment, thereby reducing unit costs. Thus, the contract manufacturers can offer their customers a greater variety of products at a significantly lower cost. During the booms time of late 1990's the average utilization rate in electronics contract manufacturing business has been between 75%-100% while after the economic downturn the utilization rates fell to less than 50% (Benson-Armer et al. (2004) [4]). Lower utilization rates drive high level of competition among contract manufacturers, resulting in razor thin profit margins for the services they deliver to their brand carrying customers. Since the production capacity expansion requires long lead-time and capital-intensive investment, it is imperative to use the capacity most efficiently for the profitability of the contract manufacturing business.

In high-tech industry, it is very typical for branded companies to announce product release date beforehand, i.e. Apple's announcement of first generation iPhones months before the actual release date in 2007, or Toyota's announcement of first mass-market hybrid car model Prius to U.S. market, in 2000. The introduction time, can be set by managers with strategic concerns of better competing with incumbents, obtaining a higher market share or being the first entrant. However, the limitations on production logistics may hamper expected profits. This research is posed to aid in planning of production logistics from the contract manufacturer's perspective when the release date is known beforehand.

In general, contract manufacturers serve multiple branded customers and have to meet the supply needs of a portfolio of high-tech products. A technological product, differing from the traditional stationary demand models, exhibits a bell shaped pattern due to dynamics of the high-tech industry and purchasing behavior of consumers. Consequently, the capacity planning for the high-tech products requires innovative approaches rather than the traditional analysis of capacity planning for stationary demands.

In this chapter, we propose two models to address the contract manufacturers' capacity planning requirement. First model is applicable to make-to-stock systems where the products are similar to each other or share similar characteristics that the same production capacity could be used to produce these products. The second model is applicable to make to order systems where the products are highly differentiated or there are regulations that the same production capacity cannot be used for these products interchangeably. We provide methodologies and solutions to allocate the capacity among these products in the two settings. The methodologies and solutions techniques are easy to implement and they could be utilized in wide variety of demand capacity situations.

In the next section, we give a brief explanation of models used for technological forecasting.

1.1 Demand Modeling in High-Tech Industries

The life span of a high-tech product is mostly determined with the pace of technological innovation. As the extent of technological innovation accelerates, more products are driven to market in shorter periods of time. Due to this fast paced environment, more products become obsolete even before the technology used to develop these products mature.

In general, high-tech products, after being introduced to the market generates an increasing demand profile until it penetrates the market, then sketches decreasing demand shape until the time the product becomes obsolete or its substitution is released to the market. Traditionally, the market penetration level of the high-tech products has been modeled with sigmoid functions. Sigmoid functions produce S shaped curves and include certain families of functions, i.e. Logistics, Weibull, Gaussian ... etc. The derivative of these sigmoid curves display bell shaped pattern and are used to forecast the demand rate within the life span of the product.

Meade and Islam (1998) [22] document 29 different functions that have been empirically used in practice for forecasting the demand of technological products. Among these, the most famous and widely used model is the Bass model, which was developed by F. M Bass in 1969. Bass models the product demand as a product of two driving sources: (1) impact of mass media influences and (2) word out mouth effect of previous purchasers. According to Bass (1969) [2], the demand rate at an instantaneous time is given by:

$$d(t) = (p + \frac{q}{m}D(t))(m - D(t))$$

where p is the coefficient of innovation (mass media effect), q is the coefficient of imitation (word of mouth effect), and m is the size of potential market. $D(t)$ is the cumulative demand observed up to time t where $\frac{dD(t)}{dt} = d(t)$ and $D(t) \leq m$. The closed form of the demand rates are:

$$D(t) = mp\frac{1 - e^{-(p+q)t}}{p + qe^{-(p+q)t}} \qquad d(t) = mp\frac{(p+q)^2 e^{-(p+q)t}}{(p + qe^{-(p+q)t})^2}$$

While Bass and most other life cycle growth models was originally built in continuous time environment, the discrete analogue of the Bass and other models exist and conveniently used in periodic time settings. In the discrete version, the periodic demand at time t is given by: $d_t = pm + (q-p)D_{t-1} - \frac{q}{m}D_{t-1}^2$ where cumulative demand up to time t is given by: $D_t = \sum_{t=0}^{t} d_t$ and $d_0 = pm$.

In this chapter, we use both version of the life cycle models. In Section 3 we use life cycle models in periodic time while in Section 4 we use the life cycle models in continuous time.

In the following sections, we will model a contract manufacturer (CM) who serves a customer base using her availability capacity. The CM knows the expected demand distribution of this product over the planning horizon. We refer to this product as the *primary product* and its customer base as the *primary market*. At a certain time within the planning horizon the CM accepts a another project proposal from a branded customer, or enters a new market. The CM also has the information on the expected demand distribution of the second product. We refer to this product as the *secondary product*, its customer base as the *secondary market* and the time when the demand for secondary product will be available to CM as the *entry time*. Knowing the expected demand distribution of these products and entry time we would like to analyze the CM's optimal production and allocation policy. We will answer this research question in two different settings; Make-to-Stock systems and Make-to-Order systems and then will provide methodologies and solutions to determine optimal allocation policies.

The reminder of this chapter is organized as follows. We summarize the relevant literature of capacity planning for differentiated and similar products, then in Section 3 we develop a capacity planning framework for make-to-stock systems, in Section 4 we formulate and solve a decision model for make-to-order systems. Both sections include discussions of the results and numerical examples to support the findings. The chapter finishes with conclusions and summary of results.

2 Literature Review

There is a rich literature on modeling and describing demand for high-tech products. Meade and Islam (1998) [22] document 29 different growth curves found in the literature. One of the earliest and most widely used model among the life cycle models is the famous Bass model. Bass (1969) [2] describes the demand for a new product by the theory of adoption and diffusion, and the model is very powerful in estimating the magnitude and timing of the peak sales when the parameters are appropriately estimated.

After the introduction of Bass model, a large body of literature revisiting the structural and conceptual assumptions together with the research on estimation issues has been formed. Mahajan et al.(1990) [21] provide an excellent survey, and categorize these developments until 1990 into five subareas; basic diffusion models, parameter estimation considerations, flexible diffusion models, refinement and extensions, and use of diffusion models. After its introduction, Bass model has been traditionally used for sales forecasting, however one of the other useful applications of the Bass model is to select optimal marketing mix strategies to maximize profitability over planning horizon considering the life-cycle dynamics. Mahajan et al. (1990) [21] summarize the use of the Bass model. More recently, Kumar and Swaminathan (2003) [20] consider the production and sales decisions of a single item in a capacitated system under life-cycle dynamics. They assume that the demand of the item follows a Bass type model with supply constraints. In their model the demand at an instantaneous time depends on the cumulative sales up to that time instead of cumulative demand. They conclude that, myopic sales plan is not necessarily optimal, and inventory build up heuristic is a robust approximation of the optimal sales plan. They compare the performance measures of myopic and build up policies over a wide range of parameters. In a special setting, they prove the optimality of build up policy when there is no initial inventory in the system. In a similar setting, Ho et al. (2002) [17] analyze capacity, time to market and demand fulfillment decisions jointly. Their demand model is same as Kumar and Swaminathan (2003) [20]. They show that delaying the product launch can act as a substitute for capacity. One interesting result of their study is that for fixed values of capacity and product launch time, when faced between the choice of selling an available unit immediately versus delaying the sale in order to reduce future shortages, the firm should favor the first choice. This is due to time benefit of the immediate cash flow outweighs the negative effect of customer loss due to demand acceleration. Our problem setting is similar to Kumar and Swaminathan (2003) [20] and, Ho et al. (2002) [17]. However we significantly depart from their work by considering the allocation of fixed capacity among two product groups over the planning horizon. These are not considered in the previous two works.

The CM's allocation problem in make-to-stock systems is classified as *capacitated production-inventory problem*, if the primary and the secondary products are different. In capacitated production inventory problems the objective is to minimize the total discounted or average costs over a finite or infinite horizon under limited capacity. The earliest formulation and the analysis of this problem was by Evans (1967) [8]. He formulates the problem as a dynamic program under stationary demand assumption. In cases where the capacity is not binding, he was able to characterize the optimal policy in which, each item attains a maximal stock level. However, when the capacity is binding he was unable to characterize the optimal policy. Federgruen and Zipkin (1986a, 1986b) [9] [10] consider capacitated single item, periodic review inventory model under stochastic demand. They prove the optimality of modified base-stock policy for both discrete and continuous demand distribution assuming stationary data and convex one period cost function. The modified base-stock policy is described as; follow a base-stock policy whenever possible, and produce to capacity when the prescribed production would exceed the capacity. Glasserman (1996) [12] addresses the capacitated, multi-item production-inventory system with continuous review. Under a subclass of allocation policies in which, some fraction of the total capacity is permanently dedicated to each of the items throughout the planning horizon, production of an item follows a base-stock policy. He also presents procedures for choosing asymptotically optimal base-stock levels and capacity allocations. DeCroix and Arreola-Risa (1998) [6] generalize Glasserman (1996) [12] for periodic review systems and show the optimality of modified base-stock policy for capacitated, multi-item production-inventory system. When the products are homogenous (i.e identical demand distributions and cost parameters), they show that symmetric resource allocation policy is optimal for both finite-horizon and infinite-horizon problem. Kapuściński and Tayur (1998) [18] study the similar system as Federgruen and Zipkin (1986a, 1986b) [9] [10] studied, with an exception. They consider, stochastic and cyclic demand. They provide the optimal policy for finite-horizon, discounted infinite-horizon, and infinite-horizon average costs. The optimal policy is the modified base-stock policy as described by Federgruen and Zipkin (1986a, 1986b) [9] [10], but with different base-stock levels for each period. For the similar system as Kapuściński and Tayur (1998) [18], with a different method, Aviv and Federgruen (1997) [1] independently proved the optimality of modified base-stock level for the infinite-horizon case.

If the primary and the secondary product demands are similar, then our problem in make to stock systems can be classified as *rationing problem.* In this problem, the decision maker first decides the production or order quantity of product, later when the demands are realized, the decision maker allocates the products to the different classes of customers according to the *rationing rule*, which is to be determined. One of the earliest formulations of this problem for an uncapacitated discrete time system is made by Veinott (1965) [27]. His focus is to find how much and when to replenish the orders. In a nonstationary environment he proves the optimality of a base-stock policy, but he does not consider any rationing levels. Later, Topkis (1968) [26] extends the results of Veinott (1965) [27] by considering how inventory should be allocated in a single period of a periodic review model. He does the analysis by breaking down the single period into finite number of subperiods, and as the demands are realized, he decides between satisfying the demand now, or reserving the inventory to fill higher-class demands in the subsequent periods. He proves that, for each review subperiod, there exist optimal nonnegative rationing levels for each demand class such that one satisfies demand of a given class only if there is no unsatisfied demand of higher class, and inventory level is above the rationing level for that class. However, he does not allow replenishment of inventory within the subperiods. Most of the rationing problem literature deals with the continuous time environment such as Nahmias and Demmy (1981) [23], Desphande et al. (2003) [7], or the queueing control environment such as Ha (1997a, 1997b, 2000) [14], [15], [16]. Rationing problem is also similar to *assortment problem* and *substitution problem.* In the assortment problem, a firm has the ability to produce n different items, where each item has its own demand, but the firm must satisfy all the demand while producing m ($m < n$) items. In this problem demand for any inferior item can be substituted by a superior or more costly item, and the objective is to minimize all the costs. Pentico (1974) [25] analyzes the problem with stochastic demands and in multi-period environment. By making some assumptions on the pattern of demand, he was able to simplify the problem. In generalized version of the single period newsvendor problem, Parlar and Goyal (1984) [24] investigated two-way substitution problem. Gerchak et al. (1996) [11] study single period production systems with random yield and downward substitutable demand. They prove that expected profit functions are concave and derive the optimality conditions. Bassok et al. (1999) [3] analyze single period, multi-product substitution problem with downward substitution. They prove that greedy allocation policy is optimal.

In our research, the CM serves the primary and the secondary market with the same or similar products, which use the same technology and facility. Unlike the classical *capacitated production-inventory problem* or the classical *rationing problem* the CM needs to consider the allocation scheme under life cycle demand models.

3 Capacity Rationing for Make-to-Stock Systems

Make-to-stock manufacturing in high-tech industry is frequently employed especially if the depreciation on the stock values could be tolerated, or when the planning horizons are rather short. Building inventory acts as an extra capacity. Due to inability to develop capacity in short period of time through capital investment, having available stocks when the product is demanded could have been proved to be very profitable; even more than the profits obtained from physical value of the product. For example brands that sell famous home-video-games such as Sony and Nintendo pile large amounts of PS3 and Wii stocks before the products are commercially available. Although it is not frequently observed that a Wii fan buying PS3 console after desperately long waiting times for the product availability, a shopper for mp3 players may choose to buy the competitor's brand to prevent long waiting times. The methodology proposed in this section formally addresses these issues and proposes policies to prevent the lost demand due to unavailability. In the next section, we state model assumptions and formulate our problem.

3.1 Problem Setting and Assumptions

A high-tech CM with finite production capacity currently uses his capacity to produce a product referred as the primary product. The CM anticipates a certain time, referred as entry time, t^e, at which she will use her capacity to produce a second type of product, referred as secondary product. The primary and the secondary products are very similar to each other and same production facilities or precesses could be used to produce these products. The CM's maximum production rate, C is fixed and each type of product require one unit of capacity for production.

Demands of each product is described with life cycle growth models and observed in periodic intervals. Actual demands of the products are stochastic and modeled as the expected demand plus an i.i.d. error terms for each period. Formally to fix the notation, let the expected primary demand be at time t is given by μ_t^p. Then actual demand observed at time t is given by $d_t^p = \mu_t^p + \xi_1$ where ξ_1 is i.i.d random variable with the pdf and the cdf $f(\xi_1)$ and $F(\xi_1)$, respectively. The expected demand of the secondary product at time t, μ_t^s is described in relation to μ_t^p as $\mu_t^s = \gamma\mu_{t-k}^p$ where γ is the scale parameter and $k > 0$ is the integer lag parameter. The actual demand of the secondary product observed at time t is modeled as $d_t^s = \mu_t^s + \xi_2$ where ξ_2 is i.i.d random variable where pdf and cdf are $f(\xi_2)$ and $F(\xi_2)$, respectively.

The planning horizon includes T units of equally spread decision points of one unit length. At each decision point of time, t the CM needs to determine (1) how much to produce, (2) how much to allocate the on hand inventory for the primary market if $t < t^e$ and (3) how much to allocate the on hand inventory for the primary and the secondary market if $t \geq t^e$. The production cost, c is fixed per unit per period. The selling price for the primary product, $\pi_t^p(d_t^p)$ is a function of quantity demanded from the primary market, with $\frac{\partial \pi_t^p(d_t^p)}{\partial d_t^p} < 0$. As the primary market demand increases selling price of per unit product decreases. This structure is commonly observed in spot market where there is a high level of competition among the suppliers. The selling price of the secondary product, π^s is per unit and constant. The unsatisfied demand from the primary and the secondary market is lost incurring a penalty per unit cost of p^p and p^s per period, respectively. There is a fixed per unit holding cost of h per period, and periodic discount factor, $\alpha \in (0, 1]$ is used to help determine net present valuation.

We further impose the following assumptions;

1. $\pi_t^p(d_t^p) + p^p > c$ for all possible realizations of d_t^p, and similarly $\pi^s + p^s > c$. This assumption ensures that, it is not optimal to never satisfy current period's demand and accumulate penalty costs.

2. $\alpha^i(\pi^p_{t+i}(d^p_{t+i}) + p^p) - (\pi^p_t(d^p_t) + p^p) < \sum_{k=0}^{i-1} \alpha^k h \qquad \forall i \in \{1, 2, ..., T-t\}$. This assumption ensures that, it is not optimal to cut current period's primary market demand to satisfy the future period's primary market demand at a higher price. The assumption is practically relevant when the CM only serves the primary market. This assumption automatically holds for the secondary market, since π^s is constant.

Note that we are not imposing any assumptions for the cross benefit case (i.e. cutting current period's primary demand, to satisfy future secondary market demand at a higher price, or vice versa). Since, these allocation decisions are critical to the CM's profitability.

We assume that production decisions take place at the beginning of each period and unsatisfied demand is lost. The decisions that the CM faces throughout the planning horizon are described as follows:

1. y_t: the inventory level after production at the beginning of period t. I_t being the on hand inventory level before the production decision takes place, the production can be at most the production capacity; $y_t - I_t \leq C$, and the inventory level after the production cannot be less than the beginning onhand inventory. $I_t \leq y_t$.

2. x_t: Inventory level below which the CM cannot use the on hand inventory to satisfy the primary market demand. ($x_t \leq y_t$).

3. z_t: Inventory level below which the CM cannot use the on hand inventory to satisfy the secondary market demand. In other words z_t is the minimum reserved inventory level for the next period. ($0 \leq z_t \leq x_t$).

Note that, the order of shipments, does not affect the problem structure, since for any given demand, the CM cannot satisfy more than the difference of two inventory levels assigned for that type of demand. In any given period, CM will supply $\min\{y_t - x_t, d^p_t\}$ to the primary market. After satisfying the primary market demand, the remaining inventory level is $\max\{x_t, y_t - d^p_t\}$. If $t < t^e$ then $x_t = z_t$, that is if the entry has not been realized yet, there is no need to reserve capacity for the secondary market. If $t \geq t^e$ then, then CM will deliver $\min\{\max\{x_t, y_t - d^p_t\} - z_t, d^s_t\}$ to the secondary market, and then the remaining inventory level at the end of period would be $\max\{\max\{y_t - d^p_t, x_t\} - d^s_t, z_t\}$, which is the inventory level at the beginning of the next period.

Table 1: Summary of notation for Make-to-Stock system capacity modeling

Notation	Description
$\pi_t^p(d_t^p)$	Unit selling price to the primary market
π^s	Unit selling price to the secondary market
c	Unit production cost
p^p	Unit penalty cost of not satisfying demand from primary market
p^s	Unit penalty cost of not satisfying demand from secondary market
h	Unit holding cost per time period
I_t	Inventory level at the beginning of period t
C	Maximum periodic production capacity
y_t	Inventory level after production,
x_t	Minimum inventory level after satisfying the primary market's demand
z_t	Minimum inventory level after satisfying secondary market's demand
d_t^p, d_t^s	Realizations primary and secondary market demands respectively
k	Integer lag parameter for the secondary market demand
γ	Scale parameter for the secondary market demand
$f(\cdot), F(\cdot)$	pdf and cdf of i.i.d random variables ξ_1 and ξ_2
α	Periodic discount factor

CM's problem is to minimize total expected cost of the planning horizon. At the end of the planning horizon, the remaining on hand inventory can be salvaged at the production cost, c. Table 1 summarizes the notation and description of the problem parameters and decision variables. The following functions will be useful in modeling the CM's problem:

1. $H_t^0(I_t)$ = minimum expected net discounted cost in periods $1, 2, ..., t$, given that period t begins with inventory level I_t and $t < t^e$

2. $H_t^1(I_t)$ = minimum expected net discounted cost in periods $t, t+1, ..., T$, given that period t begins with inventory level I_t, and $t \geq t^e$

$H_t^1(I_t)$ is the cost-to-go function of CM's after entry problem and $H_t^0(I_t)$ is the minimum expected cumulative cost of before entry. $H_t^0(I_t)$ is found according to following formulation:

$$\begin{aligned} H_t^0(I_t) &= \sum_{i=1}^{t} \alpha \min_{y_i, x_i} E[g_i^0(I_i, y_i, x_i)] \\ \text{subject to} \quad & 0 \leq x_i \leq y_i \leq I_i + C, \quad I_i \leq y_i \\ & I_{i+1} = \max\{y_i - d_i^p, x_i\}, \qquad t < t^e \end{aligned} \tag{1}$$

where, $g_t^0(I_t, y_t, x_t)$ is the periodic cost function and found by:

$$c(y_t - I_t) - \pi_t^p(d_t^p)\min\{y_t - x_t, d_t^p\} + p^p(y_t - x_t - d_t^p)^- + h\max\{y_t - d_t^p, x_t\}$$

The first term above is the production cost, the second term is the revenues obtained from primary market, the third term is the penalty cost of unsatisfied primary market demand, the fourth term is the holding cost of the inventory carried to the next period. $H_t^1(I_t)$ is found according to the following dynamic programming formulation:

$$\begin{aligned} H_t^1(I_t) &= \min_{y_t, x_t, z_t} E[g_t^1(I_t, y_t, x_t, z_t) + \alpha H_{t+1}^1(I_{t+1})] \\ \text{subject to} \quad & 0 \le z_t \le x_t \le y_t \le I_t + C, \quad I_t \le y_t \\ & I_{t+1} = \max\{\max\{y_t - d_t^p, x_t\} - d_t^s, z_t\} \qquad t \ge t^e \end{aligned} \tag{2}$$

where, $g_t^1(I_t, y_t, x_t, z_t)$ is the periodic cost function and found by:

$$\begin{aligned} & c(y_t - I_t) - \pi_t^p(d_t^p)\min\{y_t - x_t, d_t^p\} + p^p(y_t - x_t - d_t^p)^- \\ & - \pi^s\min\{\max\{y_t - d_t^p, x_t\} - z_t, d_t^s\} + p^s(\max\{y_t - d_t^p, x_t\} - z_t - d_t^s)^- \\ & + h\max\{\max\{y_t - d_t^p, x_t\} - d_t^s, z_t\} \end{aligned}$$

The first term above is the production cost, the second term is the revenues obtained from primary market, the third term is the penalty cost of unsatisfied primary market demand, the fourth term is the revenues obtained from the secondary market, the fifth terms is the penalty cost of unsatisfied secondary market demand, and the last term is the holding cost of the inventory carried to the next period. The boundary conditions for Formulations 1 and 2 are $H_0^0(I_0) = 0$ and $H_{T+1}^1(I_{T+1}) = -cI_{T+1}$, where I_{T+1} are found from the recursions in 2.

Before the entry time CM needs to optimize the production and supply decisions considering the after entry demands from the primary and the secondary markets. For a given entry time the CM's optimal cost during planning horizon is:

$$\begin{aligned} V(I_0) =& H_{t^e-1}^0(I_{t^e-1}) + H_{t^e}^1(I_{t^e}) \\ \text{subject to } & I_{t+1} = \max\{\max\{y_t - d_t^p, x_t\} - d_t^s, z_t\} \qquad \forall t \ge t^e \\ & I_{t+1} = \max\{y_t - d_t^p, x_t\} \qquad \forall t < t^e \\ & I_t \le y_t \\ & 0 \le z_t \le x_t \le y_t \le I_t + C \end{aligned} \tag{3}$$

3.2 Model Analysis

The optimization problem 3 is quite complex to solve. Without the capacity constraint, it can be easily shown that $H_t^0(I_t)$ is convex in $I_t \forall t \in \{1..t^e - 1\}$ hence, the optimal policy is the base-stock policy for the case when the CM is serving the primary market only. With the capacity constraints, the optimal policy is a modified base-stock policy. At each period one brings the on hand inventory to the base stock level as close as possible. If on hand inventory is less than base stock level the CM produces up to base stock level and if on hand inventory is more than the base stock level, the CM produces nothing.

The analysis of $H_t^1(I_t)$, even without the capacity constraints, is complicated. In this situation, the CM needs to decide the inventory level after production, y_t, and the minimum on hand inventory level for the secondary market, x_t, in each period. For the unlimited capacity case, there is no need to produce more than the demand needed in current period (i.e. $z_t = 0$). Then, the cost function in each period is $H_t^1(I_t)$ in formulation 2 except the upper bound on $y_t \leq I_t + C$.

If we explicitly analyze $H_t^1(I_t)$ for the unlimited case starting from the last period T. The underlying cost-to-go function is:

$$
\begin{aligned}
H_T^1(I_T) = & -cI_T + \min_{y_T, x_T} cy_T + \int_0^{y_T - x_T} \Bigg[-\pi_T^p(d_T^p) d_T^p - \pi^s \Bigg[\int_0^{y_T - d_T^p} d_T^s \mathrm{d}F(d_T^s) \\
& + \int_{y_T - d_T^p}^{\infty} (y_T - d_T^p) \mathrm{d}F(d_T^s) \Bigg] + p^s \int_{y_T - d_T^p}^{\infty} (d_T^p + d_T^s - y_T) \mathrm{d}F(d_T^s) \\
& + (h - \alpha c) \int_0^{y_T - d_T^p} (y_T - d_T^p - d_T^s) \mathrm{d}F(d_T^s) \Bigg] \mathrm{d}F(d_T^p) \\
& + \int_{y_T - x_T}^{\infty} \Bigg[-\pi_T^p(d_T^p)(y_T - x_T) + p^p(d_T^p - (y_T - x_T)) \\
& - \pi^s \Bigg[\int_0^{x_T} d_T^s \mathrm{d}F(d_T^s) + \int_{x_T}^{\infty} x_T \mathrm{d}F(d_T^s) \Bigg] + p^s \int_{x_T}^{\infty} (d_T^s - x_T) \mathrm{d}F(d_T^s) \\
& + (h - \alpha c) \int_0^{x_T} (x_T - d_T^s) \mathrm{d}F(d_T^s) \Bigg] \mathrm{d}F(d_T^p) \\
& \text{subject to} \\
& 0 \leq x_T \leq y_T, \quad I_T \leq y_T
\end{aligned}
\tag{4}
$$

First three lines of the above equation is the cost of last period, when observed demand from the primary market is less than the amount of inventory reserved for it (i.e. $d_T^p \leq y_T - x_T$). Note that, there is no penalty cost for not satisfying the primary market demand in this part, since the realized demand does not overshoot the on hand inventory for the primary market. Last three lines of Equation 4 represents the cost of last period, when observed demand from the primary market is more than on hand inventory reserved for the primary market. (i.e. $d_T^p > y_T - x_T$). After combining the terms and making simplifications (see Appendix A), Equation 4 becomes:

$$
\begin{aligned}
H_T^1(I_T) = & -cI_T + p^p\mu_T^p + p^s\mu_T^s + \\
& \min_{y_T, x_T} cy_T - (p^p + \pi^p)(y_T - x_T) - (p^s + \pi^s)x_T \\
& + (p^p + \pi^p - p^s - \pi^s)\int_0^{y_T - x_T} F(d_T^p)\mathrm{d}d_T^p \\
& + (p^s + \pi^s + h - \alpha c)\left[\int_0^{y_T - x_T}\int_0^{y_T - d_T^p} F(d_T^s)\mathrm{d}d_T^s\mathrm{d}F(d_T^p)\right. \\
& \left. + \int_{y_T - x_T}^{\infty}\int_0^{x_T} F(d_T^s)\mathrm{d}d_T^s\mathrm{d}F(d_T^p)\right]
\end{aligned}
$$

where μ_T^p and μ_T^s are the mean levels of the primary market and the local market demand at period T. Unfortunately, the function in the minimization part is neither convex nor strictly quasiconvex.(See Appendix A) Even in the uncapacitated case, there is no analytical optimization for the periodic cost function, $g_t^1(I_t, y_t, x_t, z_t)$

Inability to find an analytical solution to a dynamic programming (DP) problem is quite a common issue. In most cases a numerical solution is necessary. However, the computational requirements for this are often overwhelming, and for many problems a complete solution of the problem by DP is impossible. The reason lies in what Bellman has called the "curse of dimensionality", which refers to an exponential increase of the required computation as the problem's size increases. For $H_t^1(I_t)$ (for the capacitated case), state, control and the disturbance spaces are $(\Re^+)^1, (\Re^+)^3, (\Re^+)^2$, respectively. In a straightforward numerical approach, these spaces are discretized. Taking d discretization points per state axis results in a state space grid with d points. For each of these points, the minimization must be carried out numerically, which involves comparison of d^3 numbers, and to calculate these numbers, one must calculate an expected value of over the disturbance, which is the weighted sum of d^2 numbers. Also, calculation must be done for each of the T stages. Thus, number of computational operations can be as much as Td^6. If $T = 10$ and $d = 100$, then we the number of computations is 10^{13}. Even a computer, which can perform 1000 operations/sec., then finding a numerical solution would take 317 years.

Often in practice one settles for a suboptimal control scheme that finds a reasonable balance between convenient implementation and adequate performance. For this problem, we propose to apply *Certainty Equivalent Controller* (CEC) to simplify the computational requirements.

CEC is a suboptimal control scheme that is inspired by linear-quadratic control theory. At each stage CEC finds an optimal decision if some or all the uncertain quantities were fixed at some "typical" values. Bertsekas (2000) [5]. In our problem uncertain quantities are the periodic demands of the primary market and the secondary market. *Typical* values for these random quantities would be their expected values, μ_t^p and μ_t^s $\forall t \in \{1..T\}$, for the primary market demand and the secondary market demand, respectively. The CEC approach often performs well in practice and yields near optimal policies. In fact, for the linear quadratic problems where there is no constraint on the selection of decision variables, the CEC produces the optimal policy. However, this is not the case for our problem environment, since we have capacity constraints on our decision variables even though we formulate a quadratic problem.

In the subsequent parts, we will find the optimal CEC solution to $H_t^0(I_t)$ and $H_t^1(I_t)$. In order to prevent confusion, we will denote the deterministic functions as $\hat{H}_t^0(I_t)$ and $\hat{H}_t^1(I_t)$, and optimal decisions as $\hat{y}_t, \hat{x}_t, \hat{z}_t$ $\forall t \in \{1..T\}$ that are generated by CEC.

With CEC, our problem reduces to find an optimal policy to two deterministic problems. The first problem is to find $\hat{H}_{t^e}^1(I_{t^e})$ that is an optimal order up to level $(\hat{y}_t)$, optimal supply to the primary market $(\hat{y}_t - \hat{x}_t)$, and optimal supply to the secondary market $(\hat{x}_t - \hat{z}_t)$ $\forall t \geq t^e$. The second problem is to find an optimal order up to level $(\hat{y}_t)$ and optimal supply to the primary market $(\hat{y}_t - \hat{x}_t)$ $\forall t < t^e$ in anticipation of $\hat{H}_{t^e}^1(I_{t^e})$.

In the next sections we re-formulate the above problems and develop exact algorithms to find to optimal solutions to CEC problems. We will use the same notation except we add ^ to the notation in order to differentiate from the originals.

3.2.1 Optimal CEC Policies: After Entry

In this section, we will develop an exact algorithm to find the optimal solution to $\hat{H}_{t^e}^1(I_{t^e})$, given the inventory level at the beginning of the entry period is I_{t^e}. The demand expected demand vectors are $\{\mu_{t^e}^p, \mu_{t^e+1}^p ... \mu_T^p\}$ and $\{\mu_{t^e}^s, \mu_{t^e+1}^s ... \mu_T^s\}$ for the primary and the secondary market. Note that when the BC demand is known at a given period we also know the selling price for the unit primary market demand. For a given vector of primary market demand we can associate a price vector $\{\pi^p_{t^e}, \pi^p_{t^e+1}, ..., \pi^p_T\}$.

Let the periodic cost of CM after entry is represented as:

$$\begin{aligned}
\hat{g}^1_t(I_t, z_t, x_t, y_t) =& (\frac{h}{\alpha} - c)I_t + cy_t - \pi^p_T \text{min}\{y_t - x_t, \mu^p_t\} + p^p \text{max}\{\mu^p_t - (y_t - x_t), 0\} \\
& - \pi^s \text{min}\{\text{max}\{y_t - \mu^p_t, x_t\} - z_t, \mu^s_t\} \\
& + p^s \text{max}\{\mu^s_t + z_t - \text{max}\{y_t - \mu^p_t, x_t\}, 0\} \\
=& (\frac{h}{\alpha} - c)I_t - \pi^p_t \mu^p_t - \pi^s \mu^s_t + cy_t + (\pi^p_t + p^p)\text{max}\{\mu^p_t - (y_t - x_t), 0\} \\
& + (\pi^s + p^s)\text{max}\{\mu^s_t + z_t - \text{max}\{y_t - \mu^p_t, x_t\}, 0\}
\end{aligned}$$

The boundary condition is $\hat{H}^1_{T+1}(I_{T+1}) = (h/\alpha - c)I_{T+1}$ where I_{T+1} is obtained from the recursive formula $I_{t+1} = \max\{z_t, \max\{y_t - \mu^p_t, x_t\} - \mu^s_t\}$. CEC cost-to-go function:

$$\begin{aligned}
\hat{H}^1_t(I_t) =& \min_{y_t, x_t, z_t} \hat{g}^1_t(I_t, z_t, x_t, y_t) + \alpha \hat{H}^1_{t+1}(I_{t+1}) \\
\text{subject to } \; I_{t+1} =& \text{max}\{z_t, \text{max}\{y_t - \mu^p_t, x_t\} - \mu^s_t\}, \\
& 0 \le z_t \le x_t \le y_t \le C + I_t, \; \forall t \in \{1, 2, ..T\}
\end{aligned} \tag{5}$$

Although the Problem 5 is a shortest path problem, continuity of state and decision variables make it intractable to enumerate the optimal decisions for every possible state variable. However, by analyzing the characteristics of the optimal solutions we are able to provide an algorithm to solve Problem 5 exactly.

Proposition 3.1. *At period T, optimal value $\hat{z}_T = 0$.*

Proof. We will prove our claim by contradiction. After excluding the constant terms, the minimization problem to find the last stage cost-to-go function can be written as:

$$\begin{aligned}
\min_{y_T, x_T, z_T} \; & cy_T + (\pi^p_T + p^p)\text{max}\{\mu^p_t - (y_T - x_T), 0\} \\
& + (\pi^s + p^s)\text{max}\{\mu^s_T + z_T - \text{max}\{y_T - \mu^p_T, x_T\}, 0\} \\
& + (h - \alpha c)\text{max}\{z_T, \text{max}\{y_T - \mu^p_T, x_T\} - \mu^s_T\} \\
\min_{y_T, x_T, z_T} \; & cy_t + \frac{\pi^p_T + p^p}{2}\Big(|\mu^p_T + x_T - y_T| + (\mu^p_T + x_T - y_T)\Big) + 0.5(\pi^s + p^s + h - \alpha c)z_T \\
& + \frac{\pi^s + p^s + h - \alpha c}{2}\left|\mu^s_T + z_T - \frac{1}{2}\Big(|\mu^p_T + x_T - y_T| + (x_T + y_T - \mu^p_T)\Big)\right| \\
& + \frac{\pi^s + p^s - (h - \alpha c)}{2}\Big(\mu^s_T - \frac{1}{2}(|\mu^p_T + x_T - y_T| + (x_T + y_T - \mu^p_T))\Big)
\end{aligned} \tag{6}$$

Third, fourth and fifth lines are obtained from using $\max\{a,b\} = 1/2(|b-a| + (a+b))$ and $\min\{a,b\} = 1/2(-|a-b| + (b-a))$s. Let $\hat{y}_T$, $\hat{x}_T$ $\tilde{z}_T$ be the optimal solution to the minimization problem 6 satisfying $0 < \tilde{z}_T \leq \hat{x}_T \leq \hat{y}_T \leq C + I_T$, and $\tilde{H}^1_T(I_T)$ is the cost obtained from this solution. We will analyze $\tilde{H}^1_T(I_T)$ in two cases:

Case 1 $\mu^p_T + \mu^s_T > C + I_T$. Then, depending on the value of $\hat{x}_T$ and $\hat{y}_T$ the outer absolute value in the fourth line simplifies to either $|\mu^s_T + \tilde{z}_T - \hat{x}_T|$ (if, $\hat{y}_T \leq \mu^p_T + \hat{x}_T$) or, $|\mu^s_T + \tilde{z}_T + \mu^p_T - \hat{y}_T|$ (if, $\hat{y}_T > \mu^p_T + \hat{x}_T$). In the first case,

$$\mu^s_T + \tilde{z}_T - \hat{x}_T \geq \mu^s_T + \tilde{z}_T + \mu^p_T - \hat{y}_T > C + I_T + \tilde{z}_T - \hat{y}_T > 0$$

The first, inequality is obtained from using $\hat{y}_T \leq \mu^p_T + \hat{x}_T$, second equality is obtained using the fact that $\mu^p_T + \mu^s_T > C + I_T$ and the last inequality is obtained using $\tilde{z}_T > 0$ and $\hat{y}_T \leq C + I_T$. Thus in either case the function in the outer absolute value (on the fourth line) is positive. Since $\pi^s + p^s + h - \alpha c > 0$ (from our first model assumption) $\tilde{z}_T > 0$ cannot be optimal.

Case 2 $\mu^p_T + \mu^s_T \leq C + I_T$. From the Case 1 discussion, it is sufficient to analyze the sign of $\mu^s_T + \tilde{z}_T + \mu^p_T - \hat{y}_T$. Depending on the value of $\hat{y}_T$, $\mu^s_T + \tilde{z}_T + \mu^p_T - \hat{y}_T$ can be positive, 0, or negative. When it is positive Case 1 applies. When it is 0, $\tilde{H}^1_T(I_t)$ becomes $c\hat{y}_T + 0.5(\pi^s + p^s + h - \alpha c)\tilde{z}_T + 0.5(\pi^s + p^s - h + \alpha c)(\mu^s_T + \mu^p_T - \hat{y}_T)$. Since $(\pi^s + p^s + h - \alpha c) > 0$ $\tilde{z}_T > 0$ can not be optimal. When $\mu^s_T + \tilde{z}_T + \mu^p_T - \hat{y}_T$ is negative, then $\tilde{H}^1_T(I_T)$ becomes, $c\hat{y}_T + (h - \alpha c)(\hat{y}_T - \mu^s_T - \mu^p_T)$. Since $h + c - \alpha c > 0$ we can use $\hat{y}_T > \mu^p_T + \mu^s_T + \tilde{z}_T$ to have:

$$\tilde{H}^1_T(I_T) > c(\mu^p_T + \mu^s_T) + (h + c - \alpha c)\tilde{z}_T$$

Thus, $\tilde{z}_T > 0$ cannot be optimal.

□

Corollary 3.1. $\hat{y}_T = max\{I_T, min\{\mu^p_T + \mu^s_T, C + I_T\}\}$,

$$\hat{x}_T = \begin{cases} \begin{cases} max\{\hat{y}_T - \mu^p_T, 0\} & \pi^p_T + p^p \geq \pi^s + p^s \\ min\{\mu^s_T, \hat{y}_T\} & \pi^p_T + p^p < \pi^s + p^s \end{cases} & \mu^p_T + \mu^s_T > I_T \\ \hat{y}_T - \mu^p_T & \mu^p_T + \mu^s_T \leq I_T \end{cases}$$

Proof. Using the result of Proposition 3.1, we insert $\hat{z}_T = 0$ into Problem 6. If $\mu_T^p + \mu_T^s < I_T$, $\hat{y}_T = I_T$ and $\hat{x}_T = \hat{y}_T - \mu_T^p$. For the other conditions, we will analyze the cost-to-go function , in four cases:

Case 1 $\mu_T^p > C + I_T$ and $\mu_T^s > C + I_T$

$$\begin{aligned}
&\min_{y_T, x_T} cy_T + (\pi_T^p + p^p)\max\{\mu_T^p + x_T - y_T, 0\} \\
&\qquad + (\pi^s + p^s)\max\{\mu_T^s - \max\{y_T - \mu_T^p, x_T\}, 0\} \\
&\qquad + (h - \alpha c)\max\{0, \max\{y_T - \mu_T^p, x_T\} - \mu_T^s\} \\
&\min_{y_T, x_T} cy_T + (\pi_T^p + p^p)(\mu_T^p + x_T - y_T) + (\pi^s + p^s)(\mu_T^s - x_T) \\
&\quad \Rightarrow (\pi_T^p + p^p)\mu_T^p + (\pi_T^s + p^s)\mu_T^s + \min_{y_T, x_T} cy_T - (\pi_T^p + p^p)y_T + (\pi_T^p + p^p - \pi_T^s - p^s)x_T
\end{aligned}$$

Third line is obtained by the fact that $\mu_T^p + x_T - y_T > 0$ (since $\mu_T^p > C + I_t \geq y_T$ and $x_T \geq 0$), $\max\{y_T - \mu_T^p, x_T\} = x_T$ (since $y_T - \mu_T^p < 0$) leading to $\max\{\mu_T^s - x_T, 0\} = \mu_T^s - x_T$ (since $\mu_T^s > C + I_T \geq y_T \geq x_T$). The $h - \alpha c$ term becomes 0, since $x_T - \mu_T^s < 0$.
For the minimization problem in the last line, since $\pi_T^p + p^p > c$, the minimization leads to $\hat{y}_T = C + I_T$, and if $\pi_T^p + p^p > \pi^s - p^s$, $\hat{x}_T = 0$, else $\hat{x}_T = \hat{y}_T = C + I_T$.

Case 2 $\mu_T^p \leq C + I_T$ and $\mu_T^s > C + I_T$

$$\begin{aligned}
&\min_{y_T, x_T} cy_T + (\pi_T^p + p^p)\max\{\mu_T^p + x_T - y_T, 0\} \\
&\qquad + (\pi^s + p^s)\max\{\mu_T^s - \max\{y_T - \mu_T^p, x_T\}, 0\} \\
&\qquad + (h - \alpha c)\max\{0, \max\{y_T - \mu_T^p, x_T\} - \mu_T^s\} \\
\Rightarrow &\min_{y_T, x_T} cy_T + (\pi_T^p + p^p)\max\{\mu_T^p + x_T - y_T, 0\} \\
&\qquad + (\pi^s + p^s)(\mu_T^s - \min\{\mu_T^s, \max\{y_T - \mu_T^p, x_T\}\}) \\
&\qquad + (h - \alpha c)(\max\{y_T - \mu_T^p, x_T\} - \min\{\mu_T^s, \max\{y_T - \mu_T^p, x_T\}\}) \\
&\Rightarrow (\pi_T^s + p^s)\mu_T^s + \min_{y_T, x_T} cy_T - (\pi^s + p^s)\max\{y_T - \mu_T^p, x_T\} \\
&\qquad + (\pi_T^p + p^p)\max\{\mu_T^p + x_T - y_T, 0\} \\
&\Rightarrow \min_{y_T, x_T} cy_T - (\pi^s + p^s)\max\{y_T - \mu_T^p, x_T\} - (\pi_T^p + p^p)\min\{y_T - \mu_T^p, x_T\} \\
&\qquad + (\pi_T^p + p^p)x_T
\end{aligned}$$

Third line is obtained by using the relation $\max\{0, b - a\} = b - \min\{a, b\}$ with $b = \mu_T^s$ and $a = \max\{y_T - \mu_T^p, x_T\}$ for the $(\pi^s + p^s)$ term, and with $a = \mu_T^s$ and $b = \max\{y_T - \mu_T^p, x_T\}$ for the $(h-\alpha c)$ term. $(h-\alpha c)$ term is canceled since $\min\{\mu_T^s, \max\{y_T-\mu_T^p, x_T\}\}$=$\max\{y_T-\mu_T^p, x_T\}$. Last relation is obtained (after omitting the constant term) by using the relation $\max\{0, b-a\} = b-\min\{a,b\}$ with $b = x_T$ and $a = y_T - \mu_T^p$. If $\pi^s + p^s > \pi_T^p + p^p$ then last relation is minimized at, $\hat{x}_T = C + I_T$ and $\hat{y}_T = C + I_T$, else the solution is $\hat{y}_T = C + I_T$ and $\hat{x}_T \in [0, \hat{y}_T - \mu_T^p]$.

Case 3 $\mu_T^p > C + I_T$ and $\mu_T^s \leq C + I_T$

$$\begin{aligned}
&\min_{y_T, x_T} cy_T + (\pi_T^p + p^p)\max\{\mu_T^p + x_T - y_T, 0\} \\
&\quad + (\pi^s + p^s)\max\{\mu_T^s - \max\{y_T - \mu_T^p, x_T\}, 0\} \\
&\quad + (h-\alpha c)\max\{0, \max\{y_T - \mu_T^p, x_T\} - \mu_T^s\} \\
&\Rightarrow \min_{y_T, x_T} cy_T + (\pi_T^p + p^p)(\mu_T^p + x_T - y_T) + (\pi^s + p^s)\max\{\mu_T^s - x_T, 0\} \\
&\quad + (h-\alpha c)\max\{0, x_T - \mu_T^s\} \\
&\Rightarrow (\pi_T^p + p^p)\mu_T^p + (\pi_T^s + p^s)\mu_T^s + \min_{y_T, x_T} cy_T - (\pi_T^p + p^p)y_T + (\pi_T^p + p^p + h - \alpha c)x_T \\
&\quad - (\pi^s + p^s + h - \alpha c)\min\{x_T, \mu_T^s\}
\end{aligned}$$

Above function is minimized at $\hat{y}_T = C + I_T$ and $\hat{x}_T = 0$, (if $\pi_T^p + p^p > \pi^s + p^s$) or $\hat{x}_T = \mu_T^s$ (if $\pi_T^p + p^p < \pi^s + p^s$).

Case 4 $\mu_T^p + \mu_T^s \leq C + I_T$

This part is straightforward. $\hat{y}_T = \mu_T^p + \mu_T^s$ and $\hat{x}_T \in [0, \hat{y}_T - \mu_T^p]$. There is no need to produce more, since cost of carrying unit inventory $h + c - \alpha c > 0$.

□

Proposition 3.2. *Let j be the smallest period in $\{t^e, t^e+1, .., T\}$ such that $\mu_i^p + \mu_i^s \leq C + I_i$ $\forall i \in \{j, j+1, ..., T\}$, then $\hat{z}_i = 0$ $\forall i \in \{j-1, j, ..., T\}$.*

Proof. The proposition states that on hand inventory is enough to satisfy both the primary and secondary market demands , then there is no need to carry inventory into the future. The proof is straightforward, and done by starting with the last period and using the results of Proposition 3.1 recursively. Then we have:

$$\begin{aligned}
\hat{y}_i &= \mu_i^p + \mu_i^s \ \forall i \in \{j, j+1, ..., T\} \\
\hat{x}_i &= \hat{y}_i - \mu_i^p = \mu_i^s \ \forall i \in \{j, j+1, ..., T\} \\
\hat{z}_i &= \hat{x}_i - \mu_i^s = 0 \ \forall i \in \{j-1, j, ..., T\}
\end{aligned}$$

□

When the capacity is binding at a given period, it might be still optimal to carry inventory and increase the supply of the future period demands. These are the cross benefit cases. i.e. Cutting primary market's (secondary market's) demand today to satisfy more secondary market (primary market) demand tomorrow. Proposition 3.3 gives the conditions when switching is yields lower costs.

Proposition 3.3. *Let (i, j), $(j > i)$ be a pair of two periods satisfying $\mu_i^p + \mu_i^s > C + I_i$, $\mu_j^p + \mu_j^s > C + I_j$. Assume an allocation satisfying $y_i - x_i < \mu_i^p$, $x_i - z_i < \mu_i^s$ and $y_j - x_j < \mu_j^p$, $x_j - z_j < \mu_j^s$. Then a better allocation can be made if $(\pi_i^b + p^p) + \sum_{k=0}^{i-j-1} \alpha^k h < (\pi^s + p^s)\alpha^{j-i}$ or $(\pi^s + p^s) + \sum_{k=0}^{i-j-1} \alpha^k h < (\pi_j^p + p^p)\alpha^{j-i}$ holds.*

Proof. We will only show the case, cutting primary market supply at period i to increase secondary market supply at period j improves cost. The other way follows exactly the same logic. The unit cost of cutting primary market demand is $\pi_i^b + p^p$, and cost carrying unit inventory for $i - j$ periods is $\sum_{k=0}^{i-j-1} \alpha^k h$. Note that production cost is not considered since, no extra production is made at i. Current benefit of supplying one more unit to the secondary market at period j is $(\pi^s + p^s)\alpha^{j-i}$. If the relation in the proposition holds, then one can reduce $y_i - x_i$ (without affecting the secondary market supply at period i) and increase $x_j - z_j$ (without affecting the primary market supply at period j) and reduce the overall cost. □

Proposition 3.4. *Suppose at current period $\mu_t^p + \mu_t^s < C + I_t$ and $\mu_{t+i}^p + \mu_{t+i}^s > C + I_{t+i}$ for some $i \in \{1, 2, ..., T - t\}$. If $c + \sum_{j=0}^{i-1} \alpha^j h < (\pi_{t+i}^p + p^p)\alpha^i$ or $c + \sum_{j=0}^{i-1} \alpha^j h < (\pi^s + p^s)\alpha^i$ then sum of costs from t to T is reduced by carrying inventory for i periods.*

Proof. $c + \sum_{j=0}^{i-1} \alpha^j h$ is the cost of producing one more unit product today (without cutting supply of the current demand) and carrying it for i periods. $(\pi_{t+i}^b + p^p)\alpha^i$ is the benefit obtained from satisfying one more unit of primary product demand i periods later. If $c + \sum_{j=0}^{i-1} \alpha^j h \geq (\pi_{t+i}^p + p^p)\alpha^i$ then carrying inventory for i periods will not reduce the total cost at t. The same analogy applies to the secondary product case. □

Below, we propose our algorithm to solve CM's primary market and secondary market Problem. We first start with an initial solution, then improve the solution until no improvement can be made. To improve the initial solution, we will not only check the periods where there is available production capacity, but also the periods where the capacity is binding. For the periods in which the capacity is binding, by cutting low priced customer's demand and reserving inventory, the CM might be better off when she satisfies a higher priced customer's demand in the future. We call this action *Switching*. After, switching the algorithm proceeds by checking the periods where there is available production capacity.

The After Entry Algorithm

Step 1 For all $t \in \{t^e, t^e+1, ..., T\}$, assign $y_t = \max\{I_t, \min\{\mu_t^p + \mu_t^s, C + I_t\}\}$. If $\mu_t^p + \mu_t^s < I_t$, assign

$$x_t = \begin{cases} \max\{y_t - \mu_t^p, 0\} & \text{if, } \pi_t^p + p^p \geq \pi^s + p^s \\ \min\{\mu_t^s, y_t\} & \text{if, } \pi_t^p + p^p < \pi^s + p^s \end{cases}$$

else, assign $x_t = y - \mu_t^p$. For all t, assign $z_t = \max\{x_t - \mu_t^s, 0\}$

Step 2 Switching. Find all $t \in \{t^e, t^e+1, ..., T\}$ such that $\mu_t^p + \mu_t^s > C + I_t$ and $\mu_{t+i_t}^p > y_{t+i_t} - x_{t+i_t}$ or $\mu_{t+i_t}^s > x_{t+i_t} - z_{t+i_t}$ for some $i_t \in \{1, 2, ..T - t\}$. If no such t is found go to *Step 4*, else put them in a list and go to *Step 3*.

Step 3 If the list is empty go to *Step 4*, else pick the largest t, delete from the list. While, $y_t - x_t > 0$ or $x_t - z_t > 0$ find i_t^* according to:

$$\begin{aligned} i_t^{s*} =& \text{argmax}\{(\pi^s + p^s)\alpha^{i_t^s} - (\pi_t^p + p^p + \sum_{j=0}^{i_t^s - 1} \alpha^j h)\} \quad \forall i_t^s \in \{1, 2, ...T - t\} \\ & \text{subject to } y_t - x_t > 0 \text{ and } \mu_{t+i_t^s}^s - (x_{t+i_t^s} - z_{t+i_t^s}) > 0 \\ i_t^{p*} =& \text{argmax}\{(\pi_{t+i_t^p}^p + p^p)\alpha^{i_t^p} - (\pi^s + p^s + \sum_{j=0}^{i_t^p - 1} \alpha^j h)\} \quad \forall i_t^p \in \{1, 2, ...T - t\} \\ & \text{subject to } x_t - z_t > 0 \text{ and } \mu_{t+i_t^p}^p - (y_{t+i_t^p} - x_{t+i_t^p}) > 0 \\ i_t^* =& \text{argmax}\{\max\{(\pi^s + p^s)\alpha^{i_t^{s*}} - (\pi_t^p + p^p + \sum_{j=0}^{i_t^{s*} - 1} \alpha^j h), (\pi_{t+i_t^{p*}}^p + p^p)\alpha^{i_t^{p*}} \\ & - (\pi^s + p^s + \sum_{j=0}^{i_t^{p*} - 1} \alpha^j h)\}\} \end{aligned}$$

If the function in the last argmax is negative or no i_t^* is found, then go to the beginning of *Step 3*, else do the following; If the function in the last argmax is maximized by the first component, then increase y_j, $(\forall j \in \{t+1, ..., t+i_t^*\})$, x_j $(\forall j \in \{t, ..., t+i_t^*\})$ and z_j $(\forall j \in t, ..., t+i_t^* - 1)$

by $\min\{y_t - x_t, \mu^s_{t+i^*_t} - (x_{t+i^*_t} - z_{t+i^*_t})\}$. If function in the last argmax is maximized by the second component, then increase y_j, $(\forall j \in \{t+1, ..., t+i^*_t\})$, x_j $(\forall j \in \{t+1, ..., t+i^*_t - 1\})$ and z_j $(\forall j \in t, ..., t+i^*_t - 1)$ by $\min\{x_t - z_t, \mu^p_{t+i^*_t} - (y_{t+i^*_t} - x_{t+i^*_t})\}$

Step 4 Find all $t \in \{t^e, t^e+1, ..., T\}$ such that $\mu^p_t + \mu^s_t < C + I_t$ and $\mu^p_{t+k_t} + \mu^s_{t+k_t} > C + I_{k_t}$ for some $k_t \in \{t^e - t, t^e - t + 1, ..., T - t\}$. If there is no such t, end the algorithm, else put them in a list and go to *Step 5*.

Step 5 If the list is empty, end the algorithm else, pick the largest t, delete from the list and find i^*_t as follows:

$$i^{p*}_t = \text{argmax}\{\alpha^{i^p_t}(\pi^p_{t+i^p_t} + p^p) - c - \sum_{j=0}^{i^p_t - 1} \alpha^j h\} \quad \forall i^p_t \in \{t^e - t, t^e - t + 1, ..., T - t\}$$

$$\text{subject to } \mu^p_{t+i^p_t} - (y_{t+i^p_t} - x_{t+i^p_t}) > 0$$

$$i^{s*}_t = \text{argmax}\{\alpha^{i^s_t}(\pi^s_{t+i^s_t} + p^s) - c - \sum_{j=0}^{i^s_t - 1} \alpha^j h\} \quad \forall i^s_t \in \{t^e - t, t^e - t + 1, ..., T - t\}$$

$$\text{subject to } \mu^s_{t+i^s_t} - (x_{t+i^s_t} - z_{t+i^s_t}) > 0$$

$$i^*_t = \text{argmax}\{\max\{\alpha^{i^{p*}_t}(\pi^p_{t+i^{p*}_t} + p^p) - c - \sum_{j=0}^{i^{p*}_t - 1} \alpha^j h,$$

$$\alpha^{i^{s*}_t}(\pi^s_{t+i^{s*}_t} + p^s) - c - \sum_{j=0}^{i^{s*}_t - 1} \alpha^j h\}\}$$

If the function in the last argmax is negative or no i^*_t is found, then go to *Step 4*, else do the following; If the function in the last argmax is maximized by the first component, then increase y_j, $(\forall j \in \{t, ..., t+i^*_t\})$, x_j $(\forall j \in \{t, ..., t+i^*_t - 1\})$ and z_j $(\forall j \in t, ..., t+i_t - 1)$ by $\min\{\mu^p_{t+i^*_t} - (y_{t+i^*_t} - x_{t+i^*_t}), C + I^*_t - y_t\}$. If the function in the last argmax is maximized by the second component, then increase y_j, $(\forall j \in \{t, ..., t+i^*_t\})$, x_j $(\forall j \in \{t, ..., t+i^*_t - 1\})$ and z_j $(\forall j \in t, ..., t+i^*_t - 1)$ by $\min\{\mu^s_{t+i^*_t} - (x_{t+i^*_t} - z_{t+i^*_t}), C + I_t - y_t\}$.

Theorem 3.2. *The after entry algorithm terminates after finite number of steps, and at termination values of y_t, x_t and z_t are optimal. (i.e. $\hat{y}_t, \hat{x}_t, \hat{z}_t$)*

Proof. In *Step 3*, the size of the list can be at most $T - 1$. For every t in the list, the algorithm searches for a better allocation than the one assigned in *Step 1*. For a given t, the result of a search can be either negative (when no i_t^* is found or cost cannot be improved for i_t^*), or positive (when i_t^* improves overall cost). In the first case, t is removed and is not considered in this list, again. In the second case, a better allocation is made (by reducing current allocation of one customer and increasing future allocation of the other customer). After the allocation, either current allocation of the customer becomes 0, or future allocation of the customer is at its demanded quantity. In the latter case, algorithm searches for the next best i_t^*. Again, if the search for i_t^* is positive, a better allocation is made. At some point, algorithm will stop the search for next best i_t^*, since current allocation for the customer will hit 0, and in this case other t's in the list will be considered. After switching step is done, the algorithm will search for cost improvements by producing more, and this step proceeds exactly same as the improvement step in the primary market Algorithm. Hence, the algorithm will end after finite number of steps.

To prove the optimality of the allocation, first we will prove that, if no t is found in *Step 4*, then the allocation is optimal. If the algorithm cannot find any t at *Step 4*, then either $\mu_t^p+\mu_t^s < C+I_t \; \forall t$, or $\mu_t^p+\mu_t^s > C+I_t \; \forall t$. In the first case, then *Step 1* assigns $y_t = \max\{I_t, \mu_t^p+\mu_t^s\}$, $x_t = y_t - \mu_t^p$ and $z_t = \max\{x_t - \mu_t^s, 0\}$. If initial inventory $I_1 \leq \mu_1^p + \mu_1^s$, then $y_t = \mu_t^p + \mu_t^s$, $x_t = y_t - \mu_T^p = \mu_t^s$ and $z_t = 0$. Then using Proposition 3.2, (using j=1), y_t, x_t and z_t are optimal. If initial inventory $I_1 > \mu_1^p + \mu_1^s$, then algorithm imposes no production until all beginning inventory is used to satisfy demand. One can never reduce costs acting otherwise, since all the demands can be satisfied from that period, and carrying would incur extra positive holding cost. Thus, the allocation at *Step 1* is optimal, when $\mu_t^p + \mu_t^s < C + I_t \; \forall t$. The algorithm does not find any t at *Step 2*, since all the demands have been satisfied. In the second case, ($\mu_t^p + \mu_t^s > C + I_t \; \forall t$) *Step 1* assigns all $y_t = C + I_t$, $z_t = 0$, and selects x_t depending on the profitability of the customers. However, this allocation may or may not be optimal. *Step 3*, searches for a better allocation. In this case, all the periods (except the terminal period) will be in the list. Because, the total supply is less than the total demand at each period. By going backwards, *Step 3* checks whether cost can be decreased by cutting the supply of one customer at that period and increasing the supply of the other customer for all the future periods. If no such period is found then the current allocation is optimal for that period. If a period is found, then a better allocation is made. This procedure repeats itself until no improvement can be found for that period. Since, every period is checked, *Step 3* produces optimal cost-to-go functions for each period. When *Step 3* ends no t is found at *Step 4*, since *Step 3* does not reduce production level of a given period. (Still each period use its capacity in full).

For all the other demand cases, every step of the algorithm will be executed. After the allocation at the initial step, the algorithm will search for switching. *Step 3* checks only the periods that use its capacity in full. Allocation after switching is at least as good as the allocation made in *Step 1*. Since, the periods that have available capacity haven't been checked, the algorithm performs this at *Step 4*. Thus, all of the periods are checked by the algorithm, starting from the largest periods. □

In Theorem 3.2 we showed that the after entry algorithm yields optimal values of $\hat{y}_t$, $\hat{x}_t$, and $\hat{z}_t$ $\forall t \in \{t^e, t^e+1, ..., T\}$, hence we have $\hat{H}^1_{t_e}(I^e_t)$. In order to solve Problem 3, we need to determine $\hat{y}_t$ and $\hat{x}_t$ $\forall t \in \{1, 2, ..., t^e-1\}$ to find $\hat{H}^0_{t^e-1}(I_0)$, I_{t^e} considering the after entry cost to go function $\hat{H}^1_{t_e}(I^e_t)$. Next, we show that by modifying *Step 1*, *Step 4* and *Step 5* of the After Entry Algorithm we have the optimal CEC policy for the CM's Problem 3 and obtain $\hat{V}(I_0)$.

3.2.2 Optimal CEC Policies

The CM needs to find the optimal starting inventory level, I_{t^e} that is available to her once she enters the secondary market. In order for that the CM needs to decide whether to carry extra inventory or not to satisfy the primary and secondary market demands after the entry time. Below we propose three minor modifications to the after entry algorithm. Theorem 3.3 shows the optimality of the modified after entry algorithm for the CM's problem 3.

Modification to the After Entry Algorithm

1. In *Step 1*, additionally assign $y_t = \max\{I_t, \min\{\mu^p_t, C + I_t\}\}$ and $x_t = \max\{y_t - \mu^p_t, 0\}$ for all $t \in \{0, 1, ...t^e - 1\}$.

2. In *Step 2*, include into the list all $t \in \{1, 2, ..., t^e - 1\}$ such that $\mu^p_t > C + I_t$ and $\mu^s_{t+j_t} > x_{t+j_t} - z_{t+j_t}$ for some $j_t \in \{1, 2, ..., T - t\}$

3. In *Step 4*, include into the list all $t \in \{1, 2, ..., t^e-1\}$ such that $\mu^p_t < C+I_t$ and $\mu^p_{t+i_t} > C+I_{t+i_t}$ for some $i_t \in \{1, 2, ..., t^e-1-t\}$ or $\mu^p_{t+j_t}+\mu^s_{t+j_t} > C+I_{t+j_t}$ for some $j_t \in \{t^e-t, t^e-t+1, ..., T-t\}$

4. In *Step 5*, Change the domain to $\forall i^p_t \in \{1, 2, ..., T\}$ for the search of i^{b*}_t

Theorem 3.3. *The modified after entry algorithm ends after finite number of steps and at termination values of y_t, x_t and z_t are optimal for the CM's planning horizon.*

Proof. Termination in finite number of steps is a direct result of Theorem 3.2. Also by Theorem 3.2, y_t, x_t and z_t are optimal for all $t \in \{t^e, t^e+1, ..., T\}$. It only remains to show that y_t and x_t are optimal for all $t \in 1, 2, ...t^e - 1\}$. The proof is similar to the proof of Theorem 3.2. We will summarize here briefly. In *Step 1* the we assign the myopic solution as the initial solution for all $t \in 1, 2, ...t^e - 1\}$. There are only two possible ways to improve the myopic solution. First, with the cross benefits case. The CM may benefit from lowering the current periods primary market demand and build inventory for the unsatisfied secondary market demand, second the CM may benefit from build up inventory for the future if there is ample capacity is the current period. The first type of benefits are checked by the modifications done in *Step 2* and the second types of benefits are checked by the modification on *Step 4* and *Step 5*. Since the algorithm proceeds backwards, each periodic update done first in *Step 3* and later in *Step 5* yields the optimal solution for the periods. □

We showed that the modified entry algorithm solves the CM's Problem 3. Next we present a numerical example to illustrate the findings in this section.

3.3 Numerical Example

Consider that both the primary market and the secondary market demands are modeled by Bass diffusion models with parameters; $p = 0.025$, $q = 0.37$, $k = 2$, $m = 1000$ and $\gamma = 0.85$. The length of the planning horizon is $T = 14$ periods. The maximum capacity level is, $C = 75$. The fixed entry entry cost is $K = \$50$. The primary product selling price at time t, π_t^p depends on the quantity demanded at t. The specific relation ship os given by $\pi_t^p = \frac{\theta(1+\ln(\mu_t^p))}{\mu_t^p}$ and $\theta = 15$. The secondary product selling price is $\pi^s = \$2$ per unit, and production cost is $c = \$1$ per unit per time period. The penalty cost of overage demand is same for the primary and the secondary products and $p^p = p^s = \$3$ per unit per time period. Inventory holding cost $h = \$0.25$ per unit per time period. Initial inventory level, $I_1 = 0$. Periodic discount factor is $\alpha = 0.9$. What is the optimal levels of y_t, x_t and z_t and optimal cost if entry time is given $t^e = 2$, $t^e = 6$?

Solution:

Table 2 summarize the demand and price relationship for the primary and secondary market given problem parameters above. Note that the selling price of the secondary market is constant throughout the period however the selling price of the primary market is changing in regard to the quantity demanded from the primary market. According to the problem parameters given above,

Table 2: Price and demand relationship for the primary and the secondary market, numerical example

t	μ_t^b	μ_t^m	$\sum_{i=1}^t \mu_i^p$	$\sum_{i=1}^t \mu_i^s$	π_t^p
1	34.9	0.0	34.9	0.0	1.95
2	47.6	0.0	82.5	0.0	1.53
3	62.5	29.7	145.0	29.7	1.23
4	78.4	40.4	223.4	70.1	1.03
5	92.9	53.1	316.3	123.3	0.89
6	102.7	66.7	419.0	189.9	0.82
7	105.3	78.9	524.3	268.9	0.81
8	99.9	87.3	624.2	356.2	0.84
9	88.1	89.5	712.3	445.7	0.93
10	72.8	84.9	785.1	530.6	1.09
11	57.0	74.9	842.1	605.5	1.33
12	42.8	61.9	884.9	667.3	1.67
13	31.1	48.4	916.0	715.8	2.14
14	22.1	36.4	938.1	752.1	2.78

the secondary product is more profitable than the primary product in the middle of the planning horizon and less profitable at the beginning and end of the planning horizon.

For $t^e = 2$, the modified after entry algorithm produces the results summarized in Table 3. According to this $y_t - x_t$ is the supply level for the primary product, $x_t - z_t$ is the supply level for the secondary product and z_t is the level of inventory carried to the next period. The final column in Table 3 displays the optimal costs from period t through the end of horizon.

Figure 1 displays the optimal production and supply quantities when the entry time is $t^e = 2$. Note that, during the initial periods (periods 1 through 3) the on hand inventory is used to satisfy the primary market demand. However, later the secondary product becomes more profitable and the CM's cuts primary products supply to satisfy more secondary product demand. (Period 6). Later in the planning horizon the primary product becomes more profitable and this time CM lowers the secondary product's supply to be able to carry inventory to satisfy primary product demand in the future.(Period 13)

For $t^e = 6$, the modified after entry algorithm produces the results summarized in Table 4. Figure 2 displays the optimal production and supply quantities when the entry time is $t^e = 6$. Note that, the cost to go function in period 9 through 14 is the same as cost to go functions for the same periods when $t^e = 2$. Since it is more profitable to satisfy the secondary market demand in the mid

Table 3: Solution to numerical example of make-to-stock system when $t^e = 2$

t	y_t	x_t	z_t	$y_t - x_t$	$x_t - z_t$	$V_t(I_t)$
1	75	40.1	40.1	34.9	0.0	578.1
2	115.1	67.5	67.5	47.6	0.0	623.7
3	142.5	80.0	50.3	62.5	29.7	672.0
4	125.3	46.9	6.4	78.4	40.4	800.9
5	81.4	53.1	0.0	28.3	53.1	984.0
6	75	70.6	3.94	4.4	66.7	940.9
7	78.9	78.9	0.0	0.0	78.9	785.5
8	75.0	75.0	0.0	0.0	75.0	613.9
9	75.0	75.0	0.0	0.0	75.0	391.5
10	75.0	75.0	0.0	0.0	75.0	176.3
11	75.0	74.9	0.0	0.1	74.9	3.6
12	75.0	61.9	0.0	13.1	61.9	-102.3
13	75.0	43.9	0.0	31.1	43.9	-133.9
14	58.4	36.3	0.0	22.1	36.3	-75.7

Figure 1: Solution to numerical example of make-to-stock system when $t^e = 2$

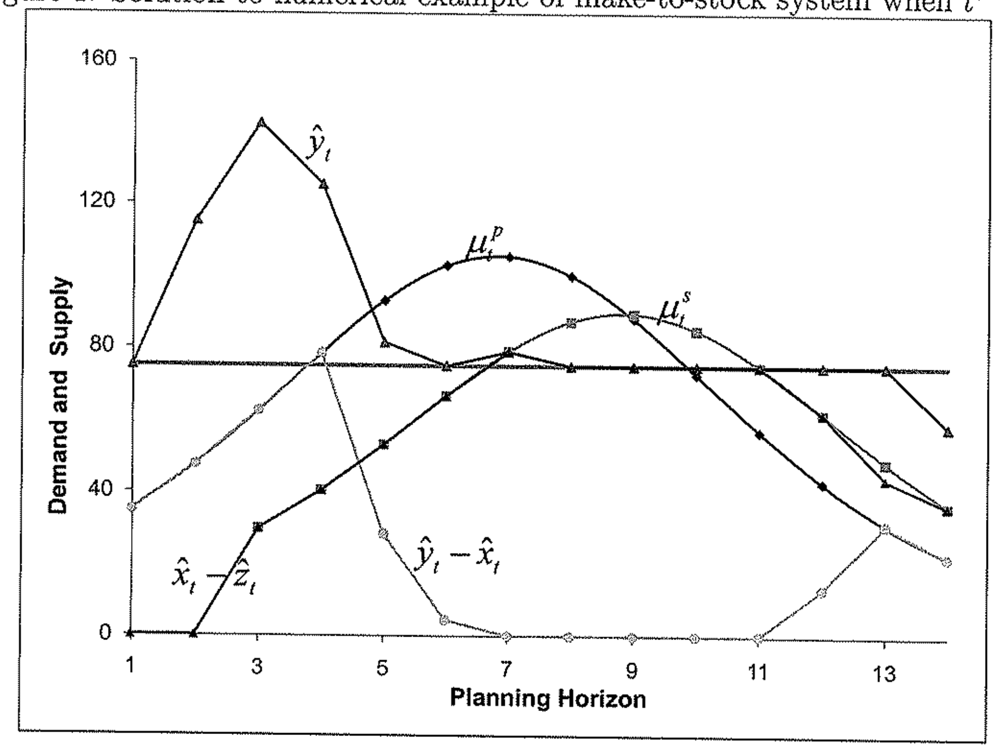


Table 4: Solution to numerical example of make-to-stock system when $t^e = 6$

t	y_t	x_t	z_t	$y_t - x_t$	$x_t - z_t$	$V_t(I_t)$
1	75	40.1	40.1	34.9	0.0	652.5
2	115.1	67.5	67.5	47.6	0.0	706.4
3	142.5	80.0	80.0	62.5	0.0	763.9
4	125.3	93.6	93.6	31.7	0.0	828.8
5	128.1	53.1	53.1	75.0	0.0	725.0
6	128.1	123.7	57.1	4.4	66.7	767.3
7	132.1	91.2	12.3	40.8	78.9	577.8
8	87.3	87.3	0.0	0.0	87.3	552.5
9	75.0	75.0	0.0	0.0	75.0	391.5
10	75.0	75.0	0.0	0.0	75.0	176.3
11	75.0	74.9	0.0	0.1	74.9	3.6
12	75.0	61.9	0.0	13.1	61.9	-102.3
13	75.0	43.9	0.0	31.1	43.9	-133.9
14	58.4	36.3	0.0	22.1	36.3	-75.7

periods, the CM carries inventory in the first periods to satisfy the secondary market demand in the middle periods. The primary market demand is only satisfied myopically when it is not optimal to carry the inventory for the secondary market. As seen from the Table 4 costs are higher when entry is at $t^e = 6$.

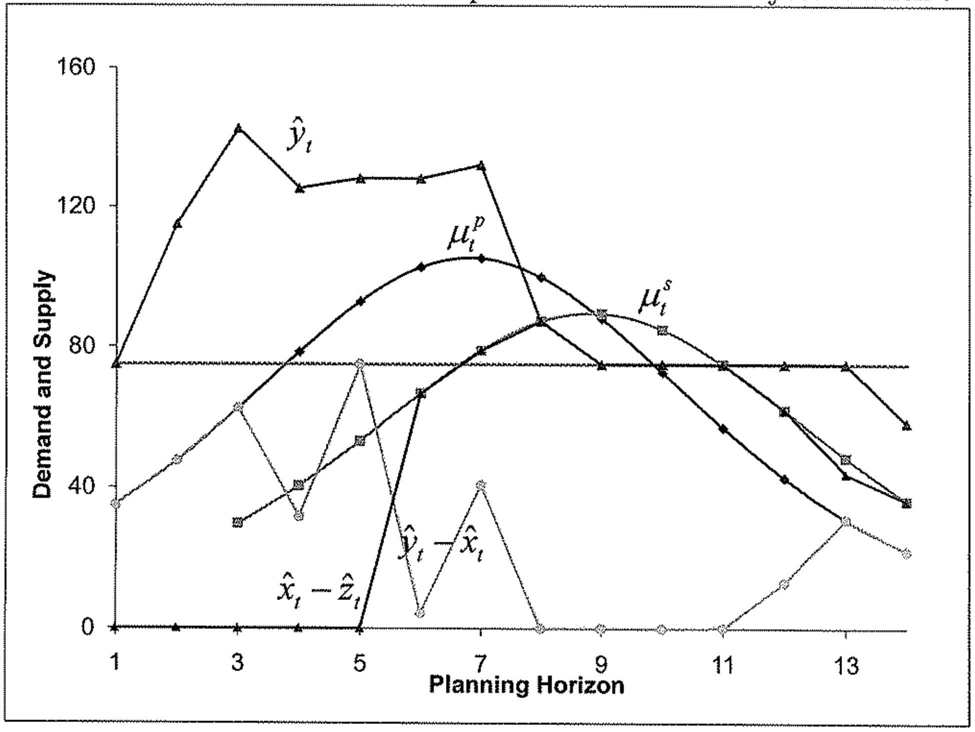
3.4 Discussion

In this section, we formulated and solved a typical contract manufacturer's production and capacity allocation problem in make-to-stock systems. The solution methodology is developed for two products or two families of products, where the production capacity could be used for these products interchangeably. We extend the results of Kumar and Swaminathan (2003) [20] into two products case, and in addition to their results we have the following contributions;

- It is not optimal to build up inventory (capacity) for the same product type in the expense of current demand, however it may be optimal to build up inventory for the other product group in the expense of current demand. That is the CM would benefit from shifting allocation to satisfy more profitable product in the future.

- Myopic policy is only optimal when the value of capacity expansion by building up inventory is not compensate the cost of capacity expansion. i.e. there is no cross benefits opportunity

Figure 2: Solution to numerical example of make-to-stock system when $t^e = 6$



and production capacity is binding throughout the planning horizon.

- We developed a decision tool that can be used in a wide range of descriptive demand relationships between two product families. The techniques used to develop the decision tool can be easily extend to more than two product situations

- We demonstrated the applicability of the proposed solution with numerical examples

One drawback of the developed solution technique is that it does not account the stochasticity of the demand models for each market. However, the limitations could easily be diminished by sampling from the the primary and secondary market's demand distribution and take the weighted average of the optimal decisions for each sample. The approach is widely used in stochastic programming literature and recent theoretical and and empirical evidence show that accurate answer to the true problem can be obtained by surprisingly few number of sample sizes. [19]

Next, we formulate and solve the contract manufacturers' capacity allocation problem for the make-to-order systems.

4 Capacity Rationing for Make-to-Order Systems

Make-to-Order manufacturing systems in high-tech industries are observed in variety of settings. If depreciation rates are high or life cycles are short Make-to-Order manufacturing system could be utilized. There might also be physical constraints, or regulations that prevent stocking piles of inventory. In these situations, products are manufactured (or assembled) as the demand is observed.

In this section, we will address the allocation problem of a contract manufacturer among two products or product families, when the products are highly differentiated. Similar to the research questions posed in the previous section we would like to find an allocation level when the CM knows the entry date for a secondary market. However, the allocation scheme in this setting will be fixed due to inability to substitute the same capacity for each product groups. Hence the allocation problem in this section will be addressed from a tactical planning level. We will provide optimal allocation schemes for any entry time and demonstrate our findings with numerical examples. Next, we state the assumptions of the problem and formulate our model.

4.1 Problem Setting and Assumptions

A high-tech with finite capacity currently uses her entire capacity to satisfy demands of a customer group, referred as the primary market. The CM anticipates a certain time point, referred as entry time, at which a certain group of customers' demand will be available for a different product she has to produce. We call the second type product as the secondary product and its market as the secondary market. The primary and the secondary products are highly differentiated and same production facilities or precesses cannot be used to produce these products. The CM's maximum production rate, c is fixed and each type of product require one unit of capacity for production.

Demands of each product is described with life cycle growth models and observed continuously. The CM uses expected demand distribution for the planning purposes of her capacity allocation levels. The expected demand rates for primary and secondary product at time t is represented as, $d^p(t)$ and $d^s(t)$, respectively. The relationship between the two demand models are described as $d^s(t) = d^p(t - \delta)$, where δ is a lag parameter. Both the primary and secondary markets have the same potential customer bases. For the modeling simplicity, we assume that $d^p(t) + d^s(t)$ has a unique maximum, which inserts an upper bound on δ. (See Appendix B). The length of the planning horizon is T. The planning horizon is long enough to observe the increasing and diminishing patterns of both type of demand curves. Let t^* be the time point at which the maximum demand rate of the primary product is expected. Then the maximum demand rate of the secondary product is expected to be observed at $t^* + \delta$.

Unit selling price for the primary and secondary products are fixed and represented by π^p and π^s respectively. Also, there is a fixed penalty cost rate, p^p and p^s for unsatisfied primary and secondary product demands. Before the known entry time, t^e the CM's uses her production capacity to serve primary market demand. Knowing the fixed entry time, t^e she wants to determine an allocation level to minimize costs over the planning horizon. The allocation level to be determined, q for the primary product will be fixed throughout the planning horizon. The remaining portion of the capacity $c-q$ would be used to satisfy the secondary market demand until the end of T. We assume that $c < d^p(t) + d^s(t)$ is satisfied for some time in the planning horizon, otherwise there would no need for an allocation decision. Table 5 summarizes the notation that will be used for this problem.

For a given entry time, CM's cost minimizing problem can be stated as:

$$\begin{aligned} C(t^e) = \min_q & - \pi^p \int_0^{t^e} \min\{c, d^p(u)\}du + p^p \int_0^{t^e} (d^p(u) - c)^+ du \\ & - \int_{t^e}^{T} \Big(\pi^p \min\{q, d^p(u)\} + \pi^s \min\{c - q, d^s(u)\} \Big) du \\ & + \int_{t^e}^{T} \Big(p^p (d^p(u) - q)^+ du + p^s (d^s(u) - (c-q))^+ \Big) du \\ \text{subject to} \quad & q \le c \end{aligned} \tag{7}$$

The first two terms measures the profits from the primary product before the entry time and is irrelevant in optimization problem. The terms in second line are the revenues obtained from both products after the entry and the terms in the third lines are the penalty costs of unsatisfied demands.

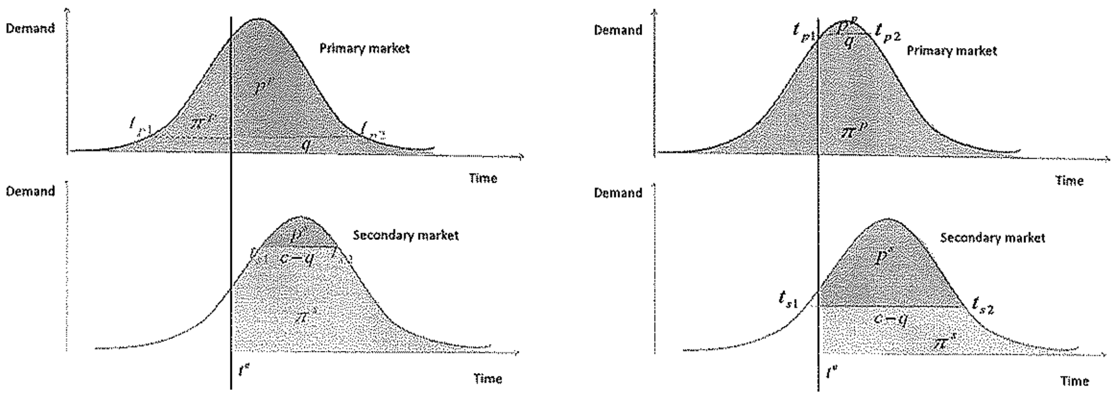
Before the analysis of Problem 7 we will introduce a formal definition for the time periods where the allocated capacity is binding for the two demands. Let the small and the large roots of equation $d^s(t) + d^p(t) = c$ be t_{c1} and t_{c2} respectively. The production capacity c is more than the total demand, $d^s(t) + d^p(t)$ during $[0, t_{c1})$, binding during $[t_{c1}, t_{c2}]$ and again the production capacity os more than the total demand during $(t_{c2}, T]$. If entry time is after t_{c2}, the allocation decision, q is trivial.

Table 5: Summary of notation for Make-to-Order system capacity modeling

Notation	Description
$d^p(t)$	Instantaneous primary market demand at time t
$d^s(t)$	Instantaneous secondary market demand at time t
T	Length of the planning horizon
δ	Time lag between the primary and the secondary demand
c	CM's maximum production capacity rate
p^p	Penalty cost paid for each unit of primary market demand over the capacity
p^s	Penalty cost paid for each unit of secondary market demand over the capacity
π^p	Unit selling price for the primary product
π^s	Unit selling price for the secondary product
q	CM's decision of allocation level

Similarly, let t_{p1} and t_{p2} be the small and the large roots of $d^p(t) = q$, and t_{s1} and t_{s2} be the small and the large roots of $d^s(t) = c - q$. The range $[t_{p1}, t_{p2}]$ represents the duration where the primary product demand is more than the allocated capacity q. So if the given entry time $t^e \leq t_{p1}$ then the CM would pay per unit penalty cost, p^p, for the primary product during $[t_{p1}, t_{p2}]$. If the entry time satisfied $t_{p1} < t^e < t_{p2}$, upon entry the CM would pay per unit penalty cost for the primary product during $[t^e, t_{p2}]$. The same argument applies to the secondary product. The range $[t_{s1}, t_{s2}]$ represents the duration where the secondary product demand is more than the allocated capacity $c - q$. If the given entry time satisfies $t^e \leq t_{s1}$, the CM would pay per unit penalty cost for the secondary product during $[t_{s1}, t_{s2}]$, but if the entry time lies in this range, then the CM would pay the penalty cost during $[t^e, t_{s2}]$.

For a given entry time if the CM increases q, then $t_{p2} - t_{p1}$ would decrease and $t_{s2} - t_{s1}$ would increase. If the CM sets q too high, then there is less opportunity to utilize from the secondary market demand. If the CM sets q too low, then the CM might loose unnecessary primary product demand. The CM's problem is to find the right balance between the two market. Figure 3 represents the trade of between setting the allocation level too high and too small.



(a) High level of q (b) Low level of q

Figure 3: The tradeoff between setting the allocation level too high and too low

4.2 Model Analysis

If we use the property $(b-a,0)^+ = b - \min\{a,b\}$ in Equation 7 and eliminate the constant terms, the reduced form of the objective function becomes:

$$\min_q(\pi^p + p^p)\int_{t^e}^{T}(d^p(u)-q)^+du + (\pi^s+p^s)\int_{t^e}^{T}(d^s(u)-(c-q))^+du \tag{8}$$

The value of the integrals above depend on the primary and secondary market demand properties at the entry time. Below Propositions 4.1-4.4 summarize the properties of the optimal allocation level for a given entry time and demand relations.

Proposition 4.1. *If both the primary and secondary market demands are decreasing in $[t^e, T]$ then the optimal value of q satisfies $(\pi^p + p^p)(t_{p2} - t^e) = (\pi^s + p^s)(t_{s2} - t^e)$. Furthermore, q satisfying this property is unique.*

Proof. If both demands are decreasing in $[t^e, T]$ then Equation 8 becomes

$$\min_q(\pi^p + p^p)(D^p(t_{p2}) - D^p(t^e) - q(t_{p2}-t^e)) + (\pi^s+p^s)(D^s(t_{s2}) - D^s(t^e) - (c-q)(t_{s2}-t^e))$$

If we take the first derivative of the objective function with respect to q:

$$\begin{aligned}\Rightarrow &(\pi^p+p^p)(d^p(t_{p2})\frac{\partial t_{p2}}{\partial q} - q\frac{\partial t_{p2}}{\partial q} - (t_{p2}-t^e))\\ &+ (\pi^s+p^s)(d^s(t_{s2})\frac{\partial t_{s2}}{\partial q} - (c-q)\frac{\partial t_{s2}}{\partial q} + (t_{s2}-t^e))\\ \Rightarrow &-(\pi^p+p^p)(t_{p2}-t^e) + (\pi^s+p^s)(t_{s2}-t^e)\end{aligned}$$

Since $d^p(t_{p2}) = q$ and $d^s(t_{s2}) = c - q$, the terms containing $\frac{\partial t_{p2}}{\partial q}$ and the terms containing $\frac{\partial t_{s2}}{\partial q}$ cancel each other. If we take the second derivative of the objective function with respect to q, we have:

$$-(\pi^p + p^p)\frac{\partial t_{p2}}{\partial q} + (\pi^s + p^s)\frac{\partial t_{s2}}{\partial q} > 0$$

since $\frac{\partial t_{p2}}{\partial q}$ is negative and $\frac{\partial t_{s2}}{\partial q}$ is positive, the objective function is convex in q. If $q = d^p(t^e)$ then the first order derivative is positive and if $q = c - d^s(t^e) < d^p(t^e)$ the first order derivative is negative. Since the objective function is convex in q there exists a unique level of q that satisfies $(\pi^p + p^p)(t_{p2} - t^e) = (\pi^s + p^s)(t_{s2} - t^e)$, which is optimal by the first order condition. □

Proposition 4.2. *If either one of the following is true then the optimal q satisfies $(\pi^p + p^p)(t_{p2} - t_{p1}) = (\pi^s + p^s)(t_{s2} - t^e)$. Furthermore the q satisfying this property is unique*

(a) $t^e \in [t_{c1}, t^*]$ *and* $2(\pi^p + p^p)(t^* - t^e) > (\pi^s + p^s)(t^e_{s2} - t^e)$

(b) $t^e \in [0, t_{c1})$ *and* $(\pi^p + p^p)(t^e_{p2} - t^e_{p1}) > 2(\pi^s + p^s)(t^* + \delta - t^e)$

where t^e_{s2} is the large root of $d^s(t) = c - d^p(t^e)$, and t^e_{p1}, t^e_{p2} are the small and the large roots of $d^p(t) = c - d^s(t^e)$.

Proof. If either one of the above cases is true then Equation 8 becomes:

$$\min_q (\pi^p + p^p)(D^p(t_{p2}) - D^p(t_{p1}) - q(t_{p2} - t_{p1})) + (\pi^s + p^s)(D^s(t_{s2}) - D^s(t^e) - (c - q)(t_{s2} - t^e))$$

If we take the first derivative of the objective function with respect to q:

$$\begin{aligned}
\Rightarrow &(\pi^p + p^p)(d^p(t_{p2})\frac{\partial t_{p2}}{\partial q} - d^p(t_{p1})\frac{\partial t_{p1}}{\partial q} - q(\frac{\partial t_{p2}}{\partial q} - \frac{\partial t_{p1}}{\partial q}) - (t_{p2} - t_{p1})) \\
&+ (\pi^s + p^s)(d^s(t_{s2})\frac{\partial t_{s2}}{\partial q} - (c - q)\frac{\partial t_{s2}}{\partial q} + (t_{s2} - t^e)) \\
\Rightarrow &- (\pi^p + p^p)(t_{p2} - t_{p1}) + (\pi^s + p^s)(t_{s2} - t^e)
\end{aligned}$$

Since $d^p(t_{p2}) = d^p(t_{p1}) = q$ and $d^s(t_{s2}) = c - q$, the terms containing $\frac{\partial t_{p2}}{\partial q}$, $\frac{\partial t_{p1}}{\partial q}$ and the terms containing $\frac{\partial t_{s2}}{\partial q}$ cancel each other. If we take the second derivative of the objective function with respect to q, we have:

$$-(\pi^p + p^p)(\frac{\partial t_{p2}}{\partial q} - \frac{\partial t_{p1}}{\partial q}) + (\pi^s + p^s)\frac{\partial t_{s2}}{\partial q} > 0$$

since $\frac{\partial t_{p2}}{\partial q}$ is negative, $\frac{\partial t_{p1}}{\partial q}$ and $\frac{\partial t_{s2}}{\partial q}$ are positive, the objective function is convex in q. If $q = d^p(t^*)$ then the first order derivative is positive and if $q = c - d^s(t^e) < d^p(t^e)$ the first order derivative is negative. Since the objective function is convex in q there exists a unique level of q that satisfies $(\pi^p + p^p)(t_{p2} - t_{p1}) = (\pi^s + p^s)(t_{s2} - t^e)$ which is optimal by the first order condition. $\square$

Proposition 4.3. *If either one of the following is true, then the optimal q satisfies $(\pi^p + p^p)(t_{p2} - t^e) = (\pi^s + p^s)(t_{s2} - t_{s1})$. Furthermore the q satisfying this equation is unique*

(a) $t^e \in [t_{c1}, t^* + \delta]$ and $(\pi^p + p^p)(t^e_{p2} - t^e) \leq 2(\pi^s + p^s)(t^* + \delta - t^e)$

(b) $t^e \in [0, t_{c1})$ and $2(\pi^p + p^p)(t^* - t^e) \leq (\pi^s + p^s)(t^e_{s2} - t^e_{s1})$

where t^e_{s1}, t^e_{s2} are the small and the large roots of $d^s(t) = c - d^p(t^e)$, and t^e_{p2} is the large root of $d^p(t) = c - d^s(t^e)$.

Proof. If either one of the above cases is true then Equation 8 becomes:

$$\min_q (\pi^p + p^p)(D^p(t_{p2}) - D^p(t^e) - q(t_{p2} - t^e)) + (\pi^s + p^s)(D^s(t_{s2}) - D^s(t_{s1}) - (c - q)(t_{s2} - t_{s1}))$$

If we take the first derivative of the objective function with respect to q:

$$\begin{aligned}
\Rightarrow &(\pi^p + p^p)(d^p(t_{p2})\frac{\partial t_{p2}}{\partial q} - q\frac{\partial t_{p2}}{\partial q} - (t_{p2} - t^e)) \\
&+ (\pi^s + p^s)(d^s(t_{s2})\frac{\partial t_{s2}}{\partial q} - d^s(t_{s1})\frac{\partial t_{s1}}{\partial q} - (c - q)(\frac{\partial t_{s2}}{\partial q} - \frac{\partial t_{s1}}{\partial q}) + (t_{s2} - t_{s1})) \\
\Rightarrow &- (\pi^p + p^p)(t_{p2} - t^e) + (\pi^s + p^s)(t_{s2} - t_{s1})
\end{aligned}$$

Since $d^p(t_{p2}) = q$ and $d^s(t_{s2}) = d^s(t_{s1}) = c - q$, the terms containing $\frac{\partial t_{p2}}{\partial q}$, and the terms containing $\frac{\partial t_{s2}}{\partial q}$ and $\frac{\partial t_{s1}}{\partial q}$ cancel each other. If we take the second derivative of the objective function with respect to q, we have:

$$-(\pi^p + p^p)\frac{\partial t_{p2}}{\partial q} + (\pi^s + p^s)(\frac{\partial t_{s2}}{\partial q} - \frac{\partial t_{s1}}{\partial q}) > 0$$

since $\frac{\partial t_{p2}}{\partial q}$ and $\frac{\partial t_{s1}}{\partial q}$ are negative, and $\frac{\partial t_{s2}}{\partial q}$ is positive, the objective function is convex in q. If $q = c - d^s(t^e)$ then the first order derivative is negative and if $q = d^p(t^e)$ the first order derivative is positive. Since the objective function is convex in q there exists a unique level of q that satisfies $(\pi^p + p^p)(t_{p2} - t^e) = (\pi^s + p^s)(t_{s2} - t_{s1})$ which is optimal by the first order condition. $\square$

Proposition 4.4. *If both of the following is true, then the optimal q satisfies $(\pi^p + p^p)(t_{p2} - t_{p1}) = (\pi^s + p^s)(t_{s2} - t_{s1})$. Furthermore the q satisfying this equation is unique*

(a) $t^e \in [0, t_{c1})$ and $(\pi^p + p^p)(t^e_{p2} - t^e_{p1}) \leq 2(\pi^s + p^s)(t^* + \delta - t^e)$

(b) $t^e \in [0, t_{c1})$ *and* $2(\pi^p + p^p)(t^* - t^e) > (\pi^s + p^s)(t^e_{s2} - t^e_{s1})$

where t^e_{p1}, t^e_{p2} *are the small and the large roots of* $d^p(t) = c - d^s(t^e)$*, and* t^e_{s1}, t^e_{s2} *are the small and the large roots of* $d^s(t) = c - d^p(t^e)$*.*

Proof. If both of the above cases is true then Equation 8 becomes

$$\min_q (\pi^p + p^p)(D^p(t_{p2}) - D^p(t_{p1}) - q(t_{p2} - t_{p1})) + (\pi^s + p^s)(D^s(t_{s2}) - D^s(t_{s1}) - (c-q)(t_{s2} - t_{s1}))$$

If we take the first derivative of the objective function with respect to q;

$$\begin{aligned}
\Rightarrow&(\pi^p + p^p)(d^p(t_{p2})\frac{\partial t_{p2}}{\partial q} - d^p(t_{p1})\frac{\partial t_{p1}}{\partial q} - q(\frac{\partial t_{p2}}{\partial q} - \frac{\partial t_{p1}}{\partial q}) - (t_{p2} - t_{p1})) \\
&+ (\pi^s + p^s)(d^s(t_{s2})\frac{\partial t_{s2}}{\partial q} - d^s(t_{s1})\frac{\partial t_{s1}}{\partial q} - (c-q)(\frac{\partial t_{s2}}{\partial q} - \frac{\partial t_{s1}}{\partial q}) + (t_{s2} - t_{s1})) \\
\Rightarrow& - (\pi^p + p^p)(t_{p2} - t_{p1}) + (\pi^s + p^s)(t_{s2} - t_{s1})
\end{aligned}$$

Since $d^p(t_{p2}) = d^p(t_{p1}) = q$ and $d^s(t_{s2}) = d^s(t_{s1}) = c - q$, the terms containing $\frac{\partial t_{p2}}{\partial q}$ and $\frac{\partial t_{p1}}{\partial q}$, and the terms containing $\frac{\partial t_{s2}}{\partial q}$ and $\frac{\partial t_{s1}}{\partial q}$ cancel each other. If we take the second derivative of the objective function with respect to q, we have;

$$-(\pi^p + p^p)(\frac{\partial t_{p2}}{\partial q} - \frac{\partial t_{p1}}{\partial q}) + (\pi^s + p^s)(\frac{\partial t_{s2}}{\partial q} - \frac{\partial t_{s1}}{\partial q}) > 0$$

since $\frac{\partial t_{p2}}{\partial q}$ and $\frac{\partial t_{s1}}{\partial q}$ are negative, and $\frac{\partial t_{s2}}{\partial q}$ and $\frac{\partial t_{p1}}{\partial q}$ are positive, the objective function is convex in q. If $q = c - d^s(t^e)$ then the first order derivative is negative and if $q = c - d^p(t^e)$ the first order derivative is positive. Since the objective function is convex in q there exists a unique level of q that satisfies $(\pi^p + p^p)(t_{p2} - t_{p1}) = (\pi^s + p^s)(t_{s2} - t_{s1})$ which is optimal by the first order condition. □

Table 7 summarizes the optimality conditions for each entry time and demand price relationship described in Proposition 4.1 through Proposition 4.4. At optimality, marginal earning from the primary market at level q is equal to marginal loss from the secondary market during planning horizon.

Remark 4.1. *As* $\pi^s + p^s \to \infty,\ q \to 0$*, and as* $\pi^p + p^p \to \infty,\ c - q \to 0$

Theorem 4.1. *The conditions defining the optimal level of* q *in Propositions 4.1-4.4 are mutually exclusive and exhaustive.*

Table 6: Proof of Theorem 4.1

Code	Stands for
A	$t^e \in [0, t_{c1})$
A^{-1}	$t^e \in [t_{c1}, t^* + \delta)$
a	$2(\pi^p + p^p)(t^* - t^e) > (\pi^s + p^s)(t^e_{s2} - t^e)$
a^{-1}	$2(\pi^p + p^p)(t^* - t^e) \leq (\pi^s + p^s)(t^e_{s2} - t^e)$
b	$(\pi^p + p^p)(t^e_{p2} - t^e_{p1}) > 2(\pi^s + p^s)(t^* + \delta - t^e)$
b^{-1}	$(\pi^p + p^p)(t^e_{p2} - t^e_{p1}) \leq 2(\pi^s + p^s)(t^* + \delta - t^e)$
$c = a^{-1}$	$(\pi^p + p^p)(t^e_{p2} - t^e) \leq 2(\pi^s + p^s)(t^* + \delta - t^e)$
d	$2(\pi^p + p^p)(t^* - t^e) \leq (\pi^s + p^s)(t^e_{s2} - t^e_{s1})$
d^{-1}	$2(\pi^p + p^p)(t^* - t^e) > (\pi^s + p^s)(t^e_{s2} - t^e_{s1})$

Proof. The optimality conditions in the second column of Table 7 applies to entry times that satisfy $t^e < t_{c2}$. Because after during $[t_{c2}, T]$ the CM's capacity is more than enough to satisfy both the primary and the secondary market demand. Coding the conditions with simple letters will be helpful to prove our claim. Table 6 summarizes the code of the conditions in Table 7.

According to this in Table 7 we have $A^{-1}a$, Ab, $A^{-1}c = A^{-1}a^{-1}$, Ad, $Ab^{-1}d^{-1}$. It is easy to see that these conditions are mutually exclusive and exhaustive. □

We demonstrated that for any demand and capacity relation there exist a unique optimal allocation level q. At optimal level of q marginal cost of per unit allocation quantity is equal to its marginal benefit. Below we demonstrate our findings with a numerical example.

4.3 Numerical Example

Consider that the both the primary market and the secondary market demand are modeled by logistics curves with parameters; $a = 200$, $b = 1$, $\delta = 2$ and $m = 1000$. The maximum capacity level is, $c = 300$. The per unit penalty cost of overshooting the primary market demand and the secondary market demand are the same and $p^p = p^s = \$8$. The selling price of the primary product is $\pi^p = 1$ and the selling price of the secondary product is $\pi^s = 2$. What would be the CM's optimal allocation level for an entry time $t^e = 0$, $t^e = 2$, $t^e = 4$, $t^e = 6$, $t^e = 8$?

Solution: If we solve the equation $d^p(t) + d^s(t) = c$, we obtain $t_{c1} = 4.74$ and $t_{c2} = 7.85$. For each entry time given we need to find which of the demand price conditions hold in Table 7 and set the q accordingly. Table 8 summarizes the optimal allocation level and the optimal cost for each entry time.

Table 7: Optimality conditions for allocation level q

Optimal q satisfies	Conditions on Demand, Price and Entry Time
$(\pi^p + p^p)(t_{p2} - t^e) = (\pi^s + p^p)(t_{s2} - t^e)$	$t^e \in [t^* + \delta, t_{c2}]$
$(\pi^p + p^p)(t_{p2} - t_{p1}) = (\pi^s + p^s)(t_{s2} - t^e)$	$t^e \in (t_{c1}, t^* + \delta)$ and $2(\pi^p + p^p)(t^* - t^e) > (\pi^s + p^s)(t^e_{s2} - t^e)$ or $t^e \in [0, t_{c1})$ and $(\pi^p + p^p)(t^e_{p2} - t^e_{p1}) > 2(\pi^s + p^s)(t^* + \delta - t^e)$
$(\pi^p + p^p)(t_{p2} - t^e) = (\pi^s + p^s)(t_{s2} - t_{s1})$	$t^e \in (t_{c1}, t^* + \delta)$ and $(\pi^p + p^p)(t^e_{p2} - t^e) \leq 2(\pi^s + p^s)(t^* + \delta - t^e)$ or $t^e \in [0, t_{c1})$ and $2(\pi^p + p^p)(t^* - t^e) \leq (\pi^s + p^s)(t^e_{s2} - t^e_{s1})$
$(\pi^p + p^p)(t_{p2} - t_{p1}) = (\pi^s + p^s)(t_{s2} - t_{s1})$	$t^e \in [0, t_{c1})$ and $(\pi^p + p^p)(t^e_{p2} - t^e_{p1}) \leq 2(\pi^s + p^s)(t^* + \delta - t^e)$ and $t^e \in [0, t_{c1})$ and $2(\pi^p + p^p)(t^* - t^e) > (\pi^s + p^s)(t^e_{s2} - t^e_{s1})$

Table 8: Optimal allocation level for the Make-to-Order model numerical example

t^e	q	t_{p1}	t_{p2}	t_{s1}	t_{s2}	Optimality condition	Optimal Cost
0	142.5	3.72	6.87	5.88	8.72	$(\pi^p + p^p)(t_{p2} - t_{p1}) = (\pi^s + p^s)(t_{s2} - t_{s1})$	$1086.23
2	142.5	3.72	6.87	5.88	8.72	$(\pi^p + p^p)(t_{p2} - t_{p1}) = (\pi^s + p^s)(t_{s2} - t_{s1})$	$1086.23
4	134.2	$t^e = 4$	6.96	5.96	8.63	$(\pi^p + p^p)(t_{p2} - t^e) = (\pi^s + p^s)(t_{s2} - t_{s1})$	$1103.80
6	80.2	$t^e = 6$	7.64	6.55	8.03	$(\pi^p + p^p)(t_{p2} - t^e) = (\pi^s + p^s)(t_{s2} - t_{s1})$	-$734.45
8	N/A	N/A	N/A	N/A	N/A	N/A	-$1146.3

- For $t^e = 0$ and $t^e = 2$, $t^e < t_{c1}$ and the conditions in the last row of Table 7 holds. If we set $q = 142.5$ and $c - q = 157.5$ we have $(\pi^p + p^p)(t_{p2} - t_{p1}) = (\pi^s + p^s)(t_{s2} - t_{s1})$ and an optimal cost of $C(0) = C(2) = \$1086.23$.

- For $t^e = 4$, $t^e < t_{c1}$ and the conditions in the third row of Table 7 hold. If we set $q = 134.2$ and $c - q = 165.8$ we have $(\pi^p + p^p)(t_{p2} - t^e) = (\pi^s + p^s)(t_{s2} - t_{s1})$ and an optimal cost of $C(4) = \$1103.80$.

- For $t^e = 6$, $t^e \in (t_{c1}, t^* + \delta)$ and the conditions in the third row of Table 7 hold. If we set $q = 80.2$ and $c - q = 219.8$ we have $(\pi^p + p^p)(t_{p2} - t^e) = (\pi^s + p^s)(t_{s2} - t_{s1})$ and an optimal cost of $C(4) = -\$734.45$.

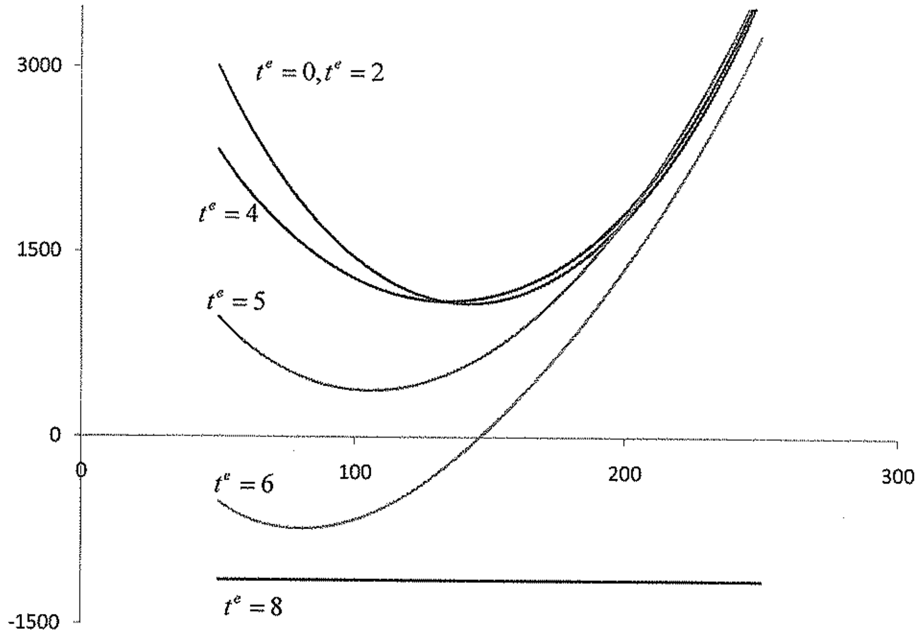
- For $t^e = 8$, $t^e > t_{c2}$ and there is no need to allocate the capacity because the capacity is more than the sum of the primary and the secondary market demands at any time. The optimal cost is $C(8) = -1146.3$

Figure 4.3 illustrate the CM's cost function in allocation level, q for the given entry times above. As proved in previous sections the CM's cost function is convex in the allocation level q. For this specific example, late entry yields lower costs to the CM due to the high penalty costs of unsatisfied demands. Also, the optimal costs are more sensitive to the later entry times rather than the early entry times. The cost function is constant for $t^e = 8$ because the $t^e > t_{c2}$ and allocation is meaningless.

4.4 Discussion

In this section, we formulated and solved a typical contract manufacturer's capacity allocation problem in make-to-order systems. The solution for the allocation level is described for two products or two families of products, where the products are highly differentiated and the same capacity cannot be used for these products interchangeably. We modeled the problem as a fixed capacity allocation problem, and showed that for any entry time there exists a fixed allocation level of the capacity. The conditions on the optimal allocation level changes with the entry time and according to demand capacity relations during after entry planning horizon. For any entry time the optimal allocation level is found at the point where marginal revenue of supplying one more unit capacity

Figure 4: CM's cost function in q for varying entry times



to primary market is equal to the marginal cost of diminishing the secondary market's allocation by one unit. The model developed in this section addresses the contract manufacturer's allocation problem at a tactical level. The results could be used to determine a size of manufacturing facility to be built or dividing a facility into two parts to use them in meeting demands of two distinct products.

5 Conclusions

As is typical in the high-tech environments, the contract manufacturers faces a dynamic and volatile market. On the other hand, utilization of the production capacity is very critical in profitability of the contract manufacturing business and it requires an extensive time frame and high level of capital investment to build up production capacity. By planning capacity level and its allocation among multiple product groups carefully, the contract manufacturers can advance among their competitors. In this chapter, we proposed methodologies and solution techniques for high-tech contract manufacturers to plan on their allocation decisions in Make-to-Stock and Make-to-Order settings. In Section 3:

- We developed an algorithm which generates the optimal production and capacity allocation levels periodically over the planning horizon. The systems anticipates the entry time of a secondary market and depending on the model parameters, directs the contract manufacturers to build up capacity or not for the future

- The algorithm considers the cross benefits cases and analyzing the discounted future price of the products determines the current supply and inventory levels

- The algorithm can be applied in a wide range of descriptive demand capacity relationships and the algorithm provides solutions from the operational planning perspective

In Section 4:

- We show the conditions to find the optimal fixed allocation level among the two products in a continuous time environment.

- The optimal allocation level is at the point where marginal revenue of supplying one more unit capacity to primary market is equal to the marginal cost of diminishing the secondary market's allocation by one unit.

- The solution can be applied in a wide range of descriptive demand capacity relationships and it addresses the contract manufacturing allocation problem from a tactical planning perspective.

We demonstrated the techniques in Section 3 and 4 with numerical examples.

References

Miscellaneous

[1] Y. Aviv and A. Federgruen, *Stochastic Inventory Models with Limited Production Capacity and Varying Parameters*, Probability in the Engineering and Information Science **11** (1997), 107–135.

[2] F. M. Bass, *A New Product Growth for Model Consumer Durables*, Management Science **15** (1969), 215–227.

[3] Y. Bassok, R. Anupindi, and R. Akella, *Single Period Multi-Product Models with Substition*, Operations Reserach **47** (1999), 632–642.

[4] R. J. Benson-Armer, D. L. Dean, and J. C. Kelleher, *Getting contract manufacturers back on track*, The McKinsey Quarterly (2004).

[5] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, Athena Scientific, Belmont, MA, 2000.

[6] G. DeCroix and T. Arreola-Rissa, *Optimal Production and Inventory Policy for Multiple Products under Resource Constraints*, Management Science **44** (1998), 950–961.

[7] V. Desphande, A. Cohen, and K. Donohue, *A Threshold Inventory Rationing Policy for Service-Differentiated Demand Classes*, Management Science **49** (2003), no. 6, 683–703.

[8] R. V. Evans, *Inventory Control of a Multi-production System with a Limited Production Resource*, Naval Research Logistics Quarterly **14** (1967), 173–184.

[9] A. Federgruen and P. Zipkin, *An Inventory Model with Limited Production Capacity and Uncertain Demands I: Average Cost Criterion*, Mathematics of Operations Research **11** (1986a), 193–207.

[10] ———, *An Inventory Model with Limited Production Capacity and Uncertain Demands II: Average Cost Criterion*, Mathematics of Operations Research **11** (1986b), 208–215.

[11] Y. Gerchak, A. Tripathy, and K. Wang, *Co-production Models with Random Yields*, IIE Transactions **28** (1996), no. 5, 391–403.

[12] P. Glasserman, *Allocating Production Capacity among Multiple Products*, Operations Research **44** (1996), no. 5, 724–734.

[13] Gartner Group, *Semiconductor Manufacturing and Design Predictions for 2004*, Research Report COM-21-7334, Gartner Group, 2003.

[14] A. Y. Ha, *Inventory Rationing in a Make-to-Stock Production System with Several Demand Classes*, Management Science **43** (1997), 1093–1103.

[15] ———, *Stock Rationing Policy for a Make-to-Stock Production System with Two Priority Classes and Backordering*, Naval Research Logistics **44** (1997), 457–472.

[16] ———, *Stock Rationing in an $M/E_k/1$ Make-to-Stock Queue*, Management Science **46** (2000), no. 1, 77–87.

[17] T. Ho, S. Savin, and C. Terwiesch, *Managing Demand and Sales Dynamics in New Product Diffusion Under Supply Constraint*, Management Science **48** (2002), no. 2, 187–206.

[18] R. Kapuściński and S. Tayur, *A Capacitated Production Inventory Model with Periodic Demand*, Operations Research **46** (1998), 899–911.

[19] A. Kleywegt, A. Shapiro, and T. Homem-de-Mello, *The sample average approximation method for stochastic discrete optimization*, SIAM Journal on Optimization **2** (2001), 479–502.

[20] S. Kumar and J. M. Swaminathan, *Diffusion of Innovations under Supply Constraints*, Operations Research **51** (2003), no. 6, 866–879.

[21] V. Mahajan, E. Muller, and F. M. Bass, *New product diffusion models in marketing: A review and directions for research*, Journal of Marketing **54** (1990), 1–26.

[22] N. Meade and T. Islam, *Technological Forecasting - Model Selection, Model Stability, and Combining Models*, Management Science **44** (1998), no. 8, 1115–1130.

[23] S. Nahmias and S. Demmy, *Operating Characteristics of an Inventory System with Rationing*, Management Science **27** (1981), no. 11, 1236–1245.

[24] M. Parlar and S. Goyal, *Optimal Ordering Decisions for Two Substitutable Products with Stochastic Demands*, Opsearch **21** (1984), no. 1, 1–15.

[25] D. W. Pentico, *The Assortment Problem with Probabilistic Demands*, Management Science **21** (1974), no. 3, 286–290.

[26] D. M. Topkis, *Optimal Ordering and Rationing Policies in a Nonstationary Dynamic Inventory Model with N Demand Classes*, Management Science **15** (1968), no. 3, 160–176.

[27] A. F. Veinott, *Optimal Policy in a Dynamic, Single Product, Nonstationary Inventory Model with Several Demand Classes*, Operations Research **13** (1965), no. 5, 761–778.