Ship-Pack Optimization in a Two-Echelon Distribution System

Naijun Wen • Stephen C. Graves • Z. Justin Ren

Singapore-MIT Alliance, National University of Singapore, naijun@gmail.com
MIT Leaders for Global Operations Program and Sloan School of Management, sgraves@mit.edu
Boston University School of Management and MIT Sloan School of Management, ren@bu.edu

December 2011

Abstract

In large distribution systems, distribution centers (DC) deliver some merchandize to their retail stores in size-specific packages, also called ship-packs. These ship-packs include cases (e.g., cartons containing 24 or 48 units), inners (packages of 6 or 8 units) or eaches (individual units). For each Stock Keeping Unit (SKU), a retailer can decide which of these ship-pack options to use when replenishing its retail stores. Working with a major US retailer, we have developed a cost model that balances DC handling costs, store handling costs and inventory-related costs at both the DC and the stores, and therefore can help to determine the optimum warehouse ship-pack for each SKU. We implement our model for a sample of 529 SKUs, and show that by changing ship-pack size for about 30 SKUs, the retailer can reduce its total cost by 0.3% - 0.4%. Interestingly, we find that most of the cost savings occurs at the DC level.

Keywords: ship-pack, distribution system, retail, replenishment, inventory optimization

I. Introduction

There has been considerable research effort spent on optimizing inventory levels in a two-echelon distribution system (Hopp and Spearman, 1996). However, one important factor is often ignored: the choice of pack size that is to be shipped from the distribution center (DC) to the retail stores for a particular item (Wagner 2002; Van Zelst et al. 2006).

This research is motivated by such a real problem of choosing the right ship-pack quantity for a major US-based retailer (which we refer to as Beta hereafter). The ship-pack quantity can typically be one of three choices: an “each” or individual unit, an “inner” (a packaged set of eaches, on the order of 6 to 8 units), or a case (e.g., a box of 24 units). The DC incurs a greater handling cost when it replenishes with eaches or inners rather than full cases for two reasons. First, warehouse associates need to spend time cutting open cases so as to replenish the picking area for either inners or eaches. Second, each replenishment order from the store entails more work picking the packages. However, replenishing with cases could pose many problems for stores as well as DCs. First, the store inventory holding cost may increase since the order amount has to be a multiple of the case quantity, which could result in more store inventory. Second, this additional inventory may occasionally exceed the available shelf space at a store. When this happens, a store must put the extra units in a backroom or high-level shelf. This practice results in extra handling and additional labor cost, and can also increase the chances of pilferage and damage. Finally, the DC sees larger demand variability when stores are replenished in cases, and as a consequence, the DC has to carry more safety stock. Thus, it is of both the DC's and the stores' interest to find the optimal ship-pack that balances the DC handling cost, the store handling cost and the inventory-related costs at both the DC and the stores. This constitutes the main goal of this study.

In this research, we develop a cost model that can be used to evaluate and optimize the costs associated with a warehouse ship-pack in the two-echelon distribution system. Our cost model has the following contributions. First of all, it is store-specific. Currently, Beta uses an Excel model that is based on an EOQ formulation to determine the optimal ship-pack; this model calculates the cost at an "average" store, namely a store with the average demand rate, and thus, ignores the wide variation in demand rates across the retail stores served by a DC. We improve upon this model by developing a comprehensive model that generates ship-pack recommendations that account for the individual-store demand characteristics for all of the stores within the distribution system. Our model is also capable of including weekly forecasts over a planning horizon, say 26 or 52 weeks. Lastly, based on inputs from Beta, we include extra-handling cost at the store level that is absent in their current calculation. This extra-handling cost accounts for the labor required first to find storage space for items that cannot fit onto store shelves and then later to retrieve them. As far as we know, such a cost has never been considered in the literature on inventory replenishment. In sum, the contribution of this research lies in the level of detail that we incorporate into the model, based on the business practices at Beta.

This paper is organized as follows. After literature review (section II), we introduce our research setting at Beta (section III). We then model the total cost in this system (Section IV), and with the data provided calculate the optimal ship-pack decisions for 529 SKUs, as well as total cost savings expected (section V). We then extend our model to consider the optimal inner-pack size choices (section VI). Finally, we conclude (section VII).

II. Literature

The economic order quantity (EOQ) problem is a century-old research topic that traces its root to a 1913 article by Ford Whitman Harris in *Factory: The Magazine of Management* (Erlenkotter, 1990). Today, the EOQ formula has become a pervasive textbook formula which every supply chain student has to learn. Traditional EOQ model assumes instant and infinite availability of products, deterministic and constant demand, constant fixed order cost and no shortages allowed (Hopp and Spearman, 1996). Three basic components are incorporated in the model: a fixed order or setup cost, a holding cost and a variable order or unit production cost. Later variations of the EOQ model have relaxed some of the assumptions. The Economic Production Lot size (EPL) model assumes a finite and fixed production rate; the Wagner-Whitin model relaxes the assumption on constant demand rate; and a variant of EOQ allows shortages and considers a back-order cost.

Although a great deal of academic literature exists on the EOQ model and its variants, very few studies have been done relating to pack size restrictions. Wagner (2002) acknowledges that the pack size could affect the order quantity in the real world. Silver et al. (1998) suggest a simple way of dealing with the pack sizes based on the form of the total cost curve in classical EOQ model. Since the total cost curve is convex, the best integral multiple of the pack sizes must be one of the two possible values surrounding the optimal continuous Q. However, a critical factor is ignored in the classical EOQ model: the handling cost of dealing with different case packs (including the individual unit which is essentially a case pack of one) both in the DC and the stores.

Van Zelst et al. (2006) recognize shelf stacking process as the largest driver of the store operational cost. Moreover, the paper also demonstrates that the case pack size is the most important driver for stacking efficiency and concludes that increasing the case pack size could increase the stacking efficiency. However, Broekmeulen et al. (2007) later develop a regression model to show that high case pack sizes tend to cause shelf space shortages. Ordering behaviors from store managers are also significantly affected by the case pack size. The larger the case pack size for an SKU is, the more the store managers tend to deviate from system generated orders (van Donselaar et al. 2006). Thus, it is difficult to decide the best case pack size even at the store level.

Besides analysis that focuses on the impact of the case pack size on the retail level, some papers have extended such studies onto the DC level. A few papers show that pack size constraints could cause a bullwhip effect in the supply chain system, which consequently increases the total system cost

(Geary et al. 2006, Lee et al., 1997a). This is in line with our modeling that larger ship-pack size induces larger demand variances at the DC level. Yan et al. (2009) address the problem of whether large case packs should be split prior to the retail level. They consider a two-echelon supply chain with a single distributor and multiple retailers under a periodic review inventory system. Assuming retail demand from an equicorrelated multivariate Poisson distribution, Yan et al. designed a factorial experiment with eight parameters including the number of retailers, the average retailer demand, heterogeneity of the retailer demand, the spatial correlation between retailer demands, the delivery pack size, the inventory safety factor, the review period at the retailer level and the critical protection period at the distributor level. Each parameter has three values that represent low, medium and high levels respectively. It is worth noting that the three pack sizes experimented are 1, 6 and 24, since these three pack sizes are also the most common among Beta's SKUs. Through simulation and ANOVA analysis, they find that of the eight parameters, the pack size has the most significant effect on amplifying demand variance up the supply chain, and it is also one of the most significant factors that result in larger stock-on-hand and back-orders at retailer level. Thus, the recommendation is to split packs at the distributor level. However, the paper ends on a cautionary note that soft costs such as breakage, pilferage and increased labor costs should be considered by management before any decision is made. It also suggests future research to include such financial implications, which is what this project does.

III. Research Setting: A Two-Echelon Distribution System with (R,s,S) Policy

Beta is a major retail company with over 1,500 stores in the United States that are supplied by a handful of regionally-located DC's. It carries approximately 12,000 SKUs. Each store is assigned to a DC; the SKUs carried by a store are replenished either from a DC or directly from the vendor (or supplier) by a flow-through policy. Under the flow-through policy, goods from the vendor are received at the DC and then directly sent to respective picking locations, from which store orders are fulfilled. Thus, the stores receive virtually everything from the DC.

As the choice of ship-pack quantity is made at each DC, we focus on a single two-echelon distribution system as depicted in Figure 1. For Beta each DC serves between 200 and 400 stores.

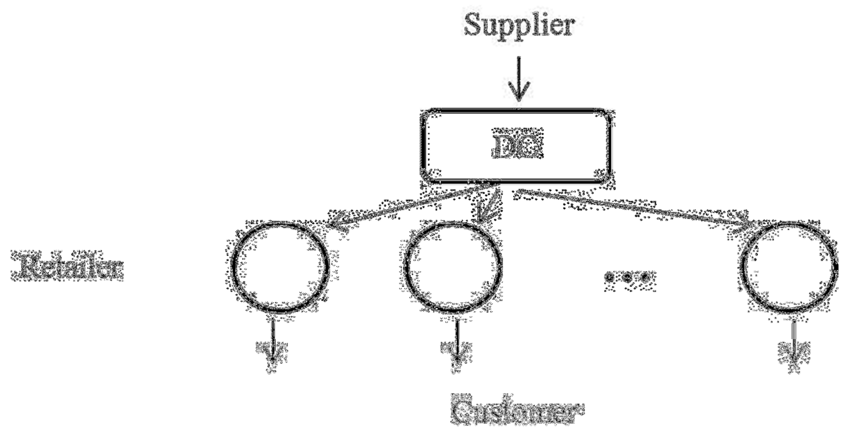


Figure 1: Two-echelon distribution system with single warehouse and multiple retailers

Each store is replenished on a regular weekly schedule. Low volume stores are replenished once a week on a fixed day; higher volume stores are replenished two to five times a week, also on fixed days. Beta follows the (R, s, S) inventory control policy. At each review period R, the inventory control system checks the Inventory Position (IP) of all Stock Keeping Units (SKUs) at the store. If IP $\leq s$ (the reorder point ROP) for an SKU, then an order will be placed for that SKU to bring its inventory level to at least S (the order-up-to-level OUTL).

IV. Cost Model

1. Notation and Assumptions

Our goal is to develop a cost model that captures the relevant cost components affected by the ship-pack size for an SKU in the two-echelon distribution system. With such a cost model, we can determine the total system cost for an SKU for each choice of ship-pack, i.e., eaches, inners and cases. Then the ship-pack with the lowest total cost is the optimal decision. More specifically, since we consider a single DC and multiple non-identical retailers, we will model the expected cost for an SKU for a store as a function of the ship-pack quantity, then sum up the costs for all stores plus an additional DC inventory cost to obtain the total system cost.

In general, there are three major categories of costs to consider based on our observation and analysis of the existing operations: the DC handling cost, the store handling cost and inventory-related costs.

- *DC Handling Cost*: As introduced earlier, the picking and replenishing activities at the DC differ substantially for full cases (FCs) and for break packs (BPs, including eaches and inners). Within the DC there are three types of handling: *receiving*, *replenishment* and *picking*. The receiving activity entails the handling associated with receipt of pallet loads from the vendor into reserve storage at the DC; the handling cost here does not depend on the ship-pack quantity and is not included in the model. The DC replenishment cost is the handling cost associated with replenishing the picking area in the DC from the reserve storage area. The replenishment cost for cases is much less than that for break packs: the cost to replenish break packs entails not just moving smaller quantities, but also the need to open cases and empty their contents into the picking racks. The DC picking cost is the handling cost for filling the store orders; picking break packs is less efficient on a per unit basis and hence more costly than picking cases.
- *Store Handling Cost*: The store handling cost includes the normal receiving cost, plus any extra handling cost. If an item does not fit onto the shelf during the regular shelf-stacking process, then it has to go to a top-tier shelf or to the backroom. Either way extra labor is needed to retrieve that item and put it back onto the shelf at a later time. .
- *Inventory-related Costs*: A larger ship-pack quantity will tend to result in larger order quantities by the stores. As with the EOQ model, a larger order quantity results in more inventory at the store but fewer orders and less fixed ordering cost. A larger ship-pack quantity can also induce the bullwhip effect in the supply chain, as shown by Yan et al. (2006). As a larger ship-pack quantity makes the replenishment orders from the stores less frequent and larger, the total demand seen by the DC (equal to the sum of the demand processes from the stores) will be more variable; thus we expect the DC to need more safety stock for a fixed service level.

More specifically, we model six cost elements for an (SKU k, store i) pair: (1) the DC replenishment cost of SKU k attributed to store i, (2) the DC picking cost of SKU k attributed to store i, (3) the normal receiving cost at store i, (4) the extra handling cost at store i if there are units that do not fit onto the shelf, (5) the average inventory cost of SKU k at store i, and (6) the fixed order cost of SKU k for store i. The total system cost for SKU k includes the summation of all the expected (SKU k, store i) costs across all stores together with (7) the expected DC inventory cost of SKU k.

Before we explain how we model each cost component, let us first introduce the notation and assumptions used in the model (Table 1):

$d_{i,k,t}$	forecast of demand of SKU k at store i in week t (units)
c_k	unit cost of SKU *k* ($ per unit)
K	fixed order cost ($ per order)
$\tilde{n}_{i,k,t}^{sp}$	expected number of ship-pack sp to be shipped from DC to store i for SKU k in week t
$numstore$	total number of retail stores
$Q_{i,k,t}^{sp}$	the order quantity for store *i*, SKU *k* in week *t* in ship-pack sp (units)
$\tilde{Q}_{i,k,t}^{sp}$	expected order quantity for store *i*, SKU *k* in week *t* in ship-pack sp (units)
$SPQ_{sp,k}$	the pack size for ship-pack *sp* of SKU k (units per pack)
$replen_{sp}$	cost of replenishing ship-pack *sp* at DC ($ per case)
$pick_{sp}$	cost of picking ship-pack *sp* at DC ($ per line or per case), depending on *sp*, where *sp* = case, inner, each. A line corresponds to a store replenishment order.
$OUTL_{i,k,t}$	order-up-to-level for store *i*, SKU *k* in week *t* (units)
$OUTLdays_{i,k}$	order-up-to-level for store *i*, SKU *k* (in days of demand)
$MaxShelf_{i,k,t}$	the shelf capacity for store *i*, SKU *k* in week *t* (units)
$IP_{i,k}$	a random variable to denote the inventory position of SKU k at store *i* when the store replenishes the SKU (units)
$ROP_{i,k}$	re-order point of SKU *k* for store *i* (units)[1]
HC	normal receiving cost at store ($ per unit)
$ExtraHC$	extra handling cost at store ($ per unit)
$ExtraUnits$	expected number of extra units that need to go to mid/top section or the backroom at store (units)
ICC_{st}	store inventory carrying cost (% per year)
ICC_{dc}	DC inventory carrying cost (% per year)
$D_{dc,k}(t, t+L)$	random variable to denote the demand at the DC over time interval (*t, t*+*L*) where *L* is the replenishment lead time for the DC for SKU k
$D_{system,k}(t, t+L)$	random variable to denote the demand for the system over time interval (*t, t*+*L*), i.e., the total demand across all stores served by the DC in our two-echelon distribution system
z	the safety factor used in the DC

Table 1: Notation

Our costing model of optimal ship-pack size is based on the following assumptions:

ASSUMPTION 1. The fixed order cost *K* is assumed to be constant regardless of SKUs or stores. However, our model can be easily modified to account for different fixed order costs if necessary.

[1] The $ROP_{i,k}$ is not defined for week t because it is not subject to change over time according to Beta's practice.

ASSUMPTION 2. We distinguish the relevant costs for the two types of a break packs, i.e., inners and eaches. Thus, there are three possible values for both the DC replenishment cost and the picking cost. When *sp* = inner or each, $pick_{sp}$ is on a per line basis. A line refers to a physical aisle where products are stored in the break-pack form. When we say per line basis, we assume the picking cost is independent of the number of ship-packs a warehouse associate takes from the picking area but depends on the number of orders; the physical act of going to the aisle and locating the desired item constitutes the major portion of the cost for one store replenishment order, while taking one item or two does not matter much. This assumption works well when the number of ship-packs picked, be it eaches or inners, is relatively small. When the number picked (eaches or inners) is large, cases will clearly be a superior choice. Thus, we believe this assumption is good enough. When *sp* = case, the picking cost is proportional to the number of cases picked.

ASSUMPTION 3. Based on input from Beta, we assume that the following relationship between $MaxShelf_{i,k,t}$ and $OUTL_{i,k,t}$ holds:

$$MaxShelf_{i,k,t} = 1.25 \times OUTL_{i,k,t} \tag{1}$$

It seems counter-intuitive that the shelf capacity changes with time. However, one way of interpreting a changing shelf capacity is that stores will allocate more space for a particular SKU when demand increases, for example, when a sales promotion is on. Our model can also be easily modified to accommodate a fixed time-independent $MaxShelf_{i,k}$ for store i and SKU k.

Beta provided both the functional form and the coefficient of 1.25 for modeling the shelf capacity. Our understanding is that (1) is a rule of thumb at Beta, which represents the average relationship between the order-up-to level and the shelf space, averaged over stores and over items. Beta deemed that this assumption was fine for our purposes, asserting that the effort required to collect store-specific and/or SKU-specific coefficients exceeded any possible benefits.

ASSUMPTION 4. We assume the inventory position (IP) of an SKU when a store makes a replenishment order is a random variable that follows a discrete uniform distribution with the lower bound being zero and the upper bound equal to the ROP. Thus, it is equally likely for IP to be any integer in [0, ROP]. This assumption allows for a relatively simple way to determine the expected number of ship-packs $\tilde{n}^{sp}_{i,k,t}$.

ASSUMPTION 5. We assume that demand *within* each week occurs at a constant rate; this rate can change from week to week. We rely on this assumption to calculate the average store inventory.

ASSUMPTION 6. We assume the lead time for the store replenishment is zero. We only need this assumption for estimating the extra units that cannot fit on the shelf; this assumption results in an over-estimate of the extra units, as we ignore any store demand during the lead time.

ASSUMPTION 7. The transportation cost is assumed to be constant regardless of the ship-pack choices; with this assumption we do not need to include transportation costs in the model. This assumption is based on Beta's fixed schedule for the store deliveries and therefore a fixed cost for the overall transportation cost.

2. Model Formulation

Our model has seven cost components. Below we formulate each in detail.

(1) Fixed Order Cost

The expected annual fixed order cost for SKU k and store i is as follows:

$$Fixed\ Order\ Cost = K\,\frac{52\times d_{i,k,t}}{\widetilde{Q}^{sp}_{i,k,t}}, \tag{2}$$

where $d_{i,k,t}$ is weekly demand, and $\widetilde{Q}^{sp}_{i,k,t}$ is the expected order quantity. We annualize this cost by multiplying by 52.

To compute $\widetilde{Q}^{sp}_{i,k,t}$, we need to first determine $\tilde{n}^{sp}_{i,k,t}$, the expected number of ship-pack sp to be shipped from the DC to each store, since $\widetilde{Q}^{sp}_{i,k,t} = \tilde{n}^{sp}_{i,k,t} \times SPQ_{sp,k}$. We estimate this quantity by the following expression:

$$\tilde{n}^{sp}_{i,k,t} = E(\left\lceil \frac{OUTL_{i,k,t} - IP_{i,k}}{SPQ_{sp,k}} \right\rceil), \tag{3}$$

where $\lceil . \rceil$ is the integer ceiling function operator. In other words, it equals the expected number of ship-packs to bring the inventory position up to at least the order up to level $OUTL_{i,k,t}$.. Here the expectation is taken over the possible values for the inventory position at the time of order. By

Assumption 4, we assume that the $IP_{i,k}$ at the time of the order follows a discrete uniform distribution in the interval [0, $ROP_{i,k}$]. $OUTL_{i,k,t}$ is calculated based on $OUTLdays_{i,k}$. More specifically, for store i, SKU k and week t,

$$OUTL_{i,k,t} = \frac{OUTLdays_{i,k}}{7} * d_{i,k,t}. \tag{4}$$

(2) DC Replenishment Cost

The total replenishment cost attributed to store i is the replenishment cost per case multiplied by the total number of cases replenished at the DC attributed to store i.

$$\text{DC Replenishment Cost} = replen_{sp} \frac{52 \times d_{i,k,t}}{SPQ_{case,k}} \tag{5}$$

(3) DC Picking Cost

When the ship-pack is each or inner, the picking cost is equal to the picking cost per line multiplied by the number of orders for the SKU for the store, whereas when the ship-pack is case, the picking cost is simply the picking cost per case times the total number of cases picked for the SKU for the store.

$$\text{DC Picking Cost} = pick_{sp} \times \begin{cases} \dfrac{52 \times d_{i,k,t}}{\tilde{Q}^{sp}_{i,k,t}}, & if\ sp = each\ or\ inner \\ \dfrac{52 \times d_{i,k,t}}{case}, & if\ sp = case \end{cases} \tag{6}$$

(4) Store Receiving Cost

For each unit received at the store, a normal store receiving cost is incurred. So the expected normal receiving cost is the normal handling cost multiplied by the annual demand.

$$\text{Store Receiving Cost } = HC \times 52 \times d_{i,k,t} \tag{7}$$

(5) Store Extra Handling Cost

The store extra handling cost is equal to the extra handling cost per item times the expected number of extra units times the expected number of orders per year.

$$\text{Store Extra Handling Cost } = ExtraHC \times \ E(ExtraUnits) \times \frac{52 \times d_{i,k,t}}{\tilde{Q}_{i,k,t}^{sp}} \tag{8}$$

To determine the expected number of extra units that do not fit onto the shelf during regular shelf-stacking process, i.e., $E(ExtraUnits)$, we need to know the shelf space allocated for the SKU at the stores. The shelf space is estimated as described in the Assumption 3.

$$\begin{aligned} E(ExtraUnits) &= E\left(\max\left(0, IP_{i,k} + Q_{i,k,t}^{sp} - MaxShelf_{i,t,k}\right)\right) \\ &= E\left(\max\left(0, IP_{i,k} + \left\lceil \frac{OUTL_{i,k,t} - IP_{i,k}}{SPQ_{sp,k}} \right\rceil \times SPQ_{sp,k} - MaxShelf_{i,t,k}\right)\right) \end{aligned} \tag{9}$$

Again $IP_{i,k}$ follows a discrete uniform distribution in the interval [0,ROP].

(6) Store Inventory Cost

To derive the expected inventory level at a store, let us first illustrate the store inventory dynamics with a constant demand rate in Figure 2. For simplicity, the subscripts are dropped in the figure. We use IP_1 and IP_2 to represent the inventory position at the two successive store replenishment orders. The shaded area is the store inventory.

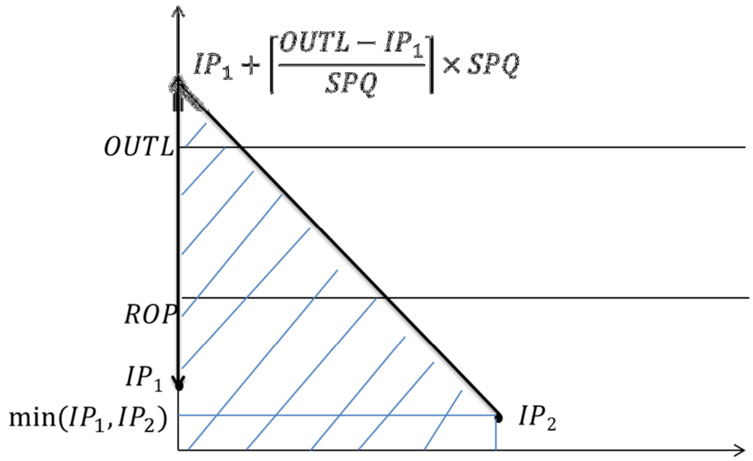


Figure 2: Store inventory illustration

The expected store inventory is

$$E(store\ inventory) = E\left(\min\left(IP_{i,k}^1, IP_{i,k}^2\right) + \frac{1}{2}\left(IP_{i,k}^1 + \left\lceil \frac{OUTL_{i,k,t} - IP_{i,k}^1}{SPQ_{sp,k}} \right\rceil \times SPQ_{sp,k} - \min\left(IP_{i,k}^1, IP_{i,k}^2\right)\right)\right)$$
$$= E\left(\frac{1}{2}\min\left(IP_{i,k}^1, IP_{i,k}^2\right) + \frac{1}{2}\left(IP_{i,k}^1 + \left\lceil \frac{OUTL_{i,k,t} - IP_{i,k}^1}{SPQ_{sp,k}} \right\rceil \times SPQ_{sp,k}\right)\right) . \tag{10}$$

Then

$$\begin{aligned} &\mathrm{E(Store\ Inventory\ Cost)} \\ &\quad = ICC_{st} \times c_k \times E\left(\frac{1}{2}\min\left(IP_{i,k}^1, IP_{i,k}^2\right) + \frac{1}{2}\left(IP_{i,k}^1 + Q_{i,k,t}^{sp}\right)\right) \\ &\quad = ICC_{st} \times c_k \times \left(\frac{1}{2}\tilde{Q}_{i,k,t}^{sp} + \frac{1}{4}ROP_{i,k} + \frac{2ROP_{i,k}^3 + 3ROP_{i,k}^2 + ROP_{i,k}}{12\left(ROP_{i,k} + 1\right)^2}\right) \end{aligned} \tag{11}$$

Appendix 1 provides details for the derivation of the above equation.

(7) DC Inventory Cost

We model the safety stock needed by the DC as being proportional to the standard deviation of the lead-time demand, namely:

$$z\sqrt{Var(D_{dc,k}(t,t+L)}. \tag{12}$$

To find $Var\big(D_{dc,k}(t,t+L)\big)$, we use the following inventory balance equation.

$$D_{dc,k}(t,t+L) = D_{system,k}(t,t+L) + \sum_{i=1}^{numStore} IP_{i,k}(t+L) - \sum_{i=1}^{numStore} IP_{i,k}(t) \tag{13}$$

That is, the demand seen by the DC over a time interval equals the demand at all of the stores over this interval plus any change in the inventory position at the stores between time t and time t+L.

To simplify the equation, let

$$IP(t) = \sum_{i=1}^{numStore} IP_{i,k}(t) \tag{14}$$

We use the above equation to approximate the variability of demand over the lead time L at the DC. Namely, we develop the approximation from:

$$Var(D_{dc,k}) = Var(D_{system,k}) + Var(IP(t+L) - IP(t)) \tag{15}$$

Here we assume the system demand over the lead time is independent of the change in IP for the stores. It is reasonable because L is likely to be much larger than the replenishment frequency at the stores. For Beta L is typically in the range of 2 to 12 weeks. Since we can approximate $Var(D_{system})$

from their purchase projections, we assume it is known and given. This leaves us the job to compute the variance of the difference in IP. We take the following steps:

$$\begin{aligned} &Var\big(IP(t+L)-IP(t)\big) \\ &\qquad = Var\big(IP(t+L)\big)+Var\big(IP(t)\big)-2Cov\big(IP(t+L),IP(t)\big) \\ &\qquad \cong 2Var(IP(t)). \end{aligned} \tag{16}$$

Here we assume that $IP(t+L)$ is independent of $IP(t)$. The larger the lead time is, the more accurate the above approximation is[2]. Moreover, we assume that at each store the inventory position IP_i is uniformly distributed between $(x, x+\ Q_{i,k,t}^{sp})$ for some value of *x*; we don't need to specify *x* as it does not affect our model. We also assume that the inventory positions for any pair of stores are independent of each other.

Thus, we have

$$Var\big(IP(t)\big) = \sum Var\left(IP_{i,k}(t)\right) = \frac{\sum \tilde{Q}_{i,k,t}^{sp\ \ 2}}{12}. \tag{17}$$

Recall that $\tilde{Q}_{i,k,t}^{sp}$ is the expected order quantity for store *i*, SKU *k* in week *t* in ship-pack sp (table 1). Hence,

$$Var\big(D_{dc,k}\big) \cong Var\big(D_{system,k}\big) + 2\frac{\sum \tilde{Q}_{i,k,t}^{sp\ \ 2}}{12} \tag{18}$$

We see for (18) a realization of the bullwhip, as the demand variability seen by the DC is an amplification of that at the stores; the extent of the bullwhip grows with the order quantities at the stores.

In summary, the DC inventory cost is given by the following equation:

---

[2] If $IP(t+L)$ were positively correlated with $IP(t)$, the approximation in equation (16) results in an upper bound on. $Var\big(IP(t+L)-IP(t)\big)$, which then yields an upper bound on the DC's inventory costs.

$$DC\_InvCost_{k,t}(sp) = ICC_{dc} \times c_k \times z \times \sqrt{Var\left(D_{system,k}\right) + \frac{\sum \tilde{Q}_{i,k,t}^{sp\;2}}{6}} \qquad (19)$$

(8) Total System Cost

The expected annualized cost for week t for an SKU for store k is the summation of all the cost components described in equations (1) -(13):

$$\begin{aligned} Cost_{i,k,t}(sp) &= K \frac{52 \times d_{i,k,t}}{\tilde{Q}_{i,k,t}^{sp}} \\ &+ replen_{sp} \frac{52 \times d_{i,k,t}}{case} \\ &+ pick_{sp} \times \begin{cases} \frac{52 \times d_{i,k,t}}{\tilde{Q}_{i,k,t}^{sp}}, \ if\ sp = each\ or\ inner \\ \frac{52 \times d_{i,k,t}}{case}, \ if\ sp = case \end{cases} \\ &+ ICC_{st} \times c_k \times E(\tfrac{1}{2} \min\left(IP_{i,k}^{1}, IP_{i,k}^{2}\right) + \tfrac{1}{2}\left(IP_{i,k}^{1} + Q_{i,k,t}^{sp}\right)) \\ &+ HC \times 52 \times d_{i,k,t} \\ &+ ExtraHC \times E(ExtraUnits) \times \frac{52 \times d_{i,k,t}}{\tilde{Q}_{i,k,t}^{sp}}, \end{aligned} \qquad (20)$$

Where

$$\tilde{Q}_{i,k,t}^{sp} = \tilde{n}_{i,k,t}^{sp} \times SPQ, \qquad (21)$$

$$\tilde{n}_{i,k,t}^{sp} = E(\left\lceil \frac{OUTL_{i,k,t} - IP_{i,k}}{SPQ} \right\rceil), \qquad (22)$$

and

$$E(ExtraUnits) = E(\max\left(0, IP_{i,k,t} + \tilde{Q}_{i,k,t}^{sp} - MaxShelf_{i,t,k}\right)) \qquad (23)$$

Finally, the annualized total system cost in week t is the summation of the expected cost of all stores plus a DC inventory cost, i.e.,

$$\text{Total Cost} = DC\_InvCost_{k,t}(sp) \; + \sum_{i=1}^{numStore} Cost_{i,k,t}(sp), \tag{24}$$

where *numStore* denotes the total number of stores the DC serves.

With the objective function in place, we can then formulate the following minimization problem.

$$\underset{sp \in \{each, inner, case\}}{MIN} \; Total\ cost$$
$$= DC\_InvCost_{k,t}(sp) \; + \sum_{i=1}^{numStore} Cost_{i,k,t}(sp) \tag{25}$$

The optimal solution to the above problem is valid for only week t. When a multiple-week planning period is in question, we can extend the problem into the following form. Let $numWeek$ be the number of weeks in the planning period. We minimize the average total annual cost.

$$\underset{sp \in \{each, inner, case\}}{MIN} \; \frac{\sum_{t=1}^{numWeek}\left[DC_{InvCost_{k,t}}(sp) + \sum_{i=1}^{numStore} Cost_{i,k,t}\,(sp)\right]}{numWeek} \tag{26}$$

That is, we find the optimal ship-pack that is the most economic decision for the entire planning period from week 1 to week $numWeek$.In the context of Beta, their desire was to determine the best ship-pack quantity for each SKU and for each DC for a planning period of up to 26 weeks; they would then revisit this decision two to four times a year.

Additional complexity arises when we allow multiple ship-pack changes during the planning period. For example, if we allow the warehouse to change its ship-pack each week, then the solution will be a vector of length $numWeek$, whose elements are the optimal ship-pack corresponding to each week. However, due to the physical and practical constraints in the warehouse, it is more appropriate to limit the number of ship-pack changes throughout the planning period. The algorithm that finds such a solution will be discussed in more detail in the results section.

V. Data and Results

Beta provided us with the required cost parameters, including the DC and store handling rates, the fixed order cost and the inventory carrying cost. We also obtained a sample dataset for SKUs from three product families for the set of stores supplied by one DC. Before we show the model output, let us first have a look at the data, which will help us better understand the results presented later.

The sample dataset contains a total of 529 SKUs, three of which have a case quantity of one. Beta terms such circumstance “case of 1”, and no ship-pack analysis is necessary. Thus, we are effectively dealing with 526 SKUs. Moreover, 369 out of the 526 SKUs have an inner quantity of one, meaning there is no inner pack for these SKUs; hence the ship-pack quantity choice is between an each and a case.

The sample data includes 52 weeks of sales forecasts. We have identified three representative annual demand patterns. Figure 3(a) and 3(b) exhibit seasonal demands, with single and multiple peaks respectively, whereas Figure 3(c) shows relatively stable demand throughout the entire year. In each case we plot the total demand forecast for a single SKU, cumulated over all the stores that carry the SKU in the sample. In total, about one fifth of the total SKUs in our data set exhibit the single peak pattern; only 3% exhibits the double peak pattern; and the remaining 77% has relatively stable weekly sales forecasts.

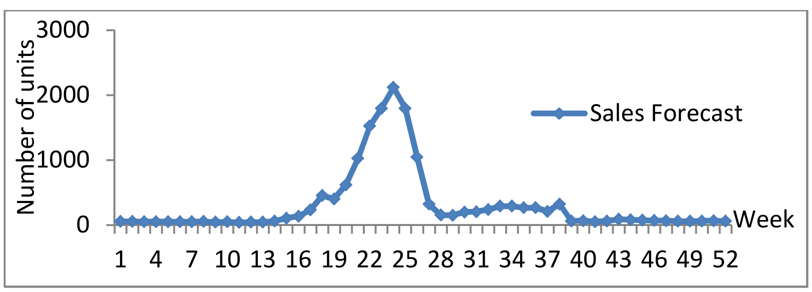


Figure 3(a): Weekly sales forecast for SKU 01

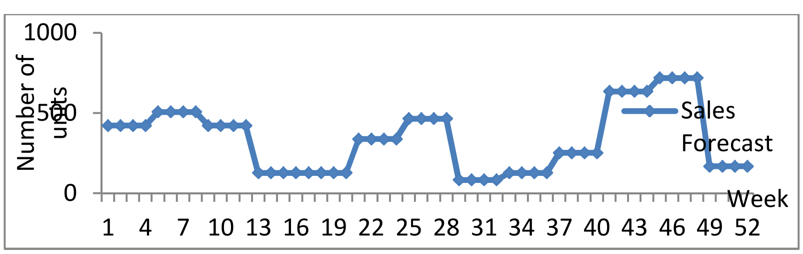


Figure 3(b): Weekly sales forecast for SKU 02

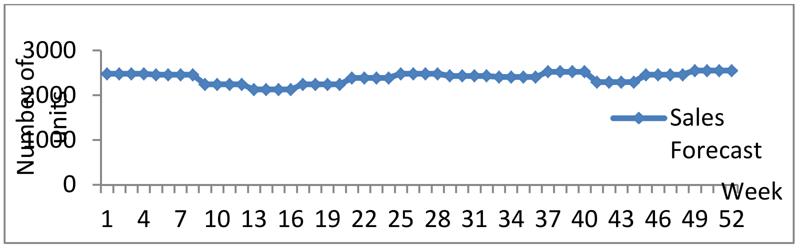
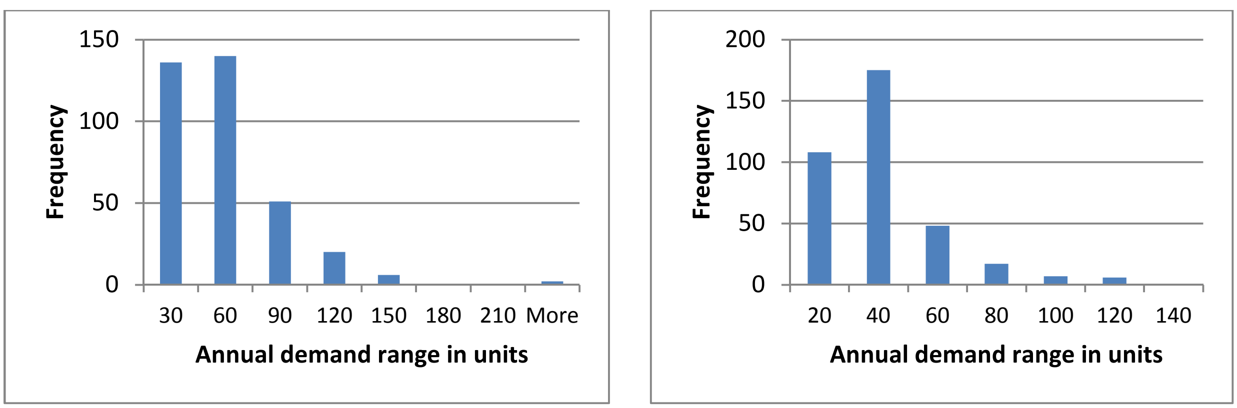


Figure 3(c): Weekly sales forecast for SKU 03

The demand variance across stores is large, but this is not surprising because there are low-, medium- and high-volume stores for the major retailer. A closer look reveals that the coefficient of variation (CV) of sales volume across stores for the given SKUs vary from 0.3 to 3.5. Figures 4(a) and 4(b) show the frequency histograms of the annual demands by stores for two SKUs. Both graphs show that store annual demands vary greatly. In Figure 4(a), the number of low-volume stores (< 30 units per year) is roughly equal to that of the mid-volume (between 30 and 60 units per year) ones. There are two stores having extremely large annual demand such that the bar shows in the "more" column and a handful of stores with annual demand greater than 100 units. In Figure 4(b), about half of the stores are in the mid-volume range, one third are low-volume while the remaining one sixth are high-volume stores.



(a) SKU 1 **(b) SKU 2**

Figure 4: Distribution of annual demand by stores

As mentioned earlier, we consider two types of break-packs (inners and eaches). There is some uncertainty about the DC replenishment cost for inners, relative to that for eaches and for cases. Thus, we perform a sensitivity analysis on this assumed cost. Since the replenishment costs for eaches and cases, given by Beta, are \$0.7789 per case and \$0.1716 per case, respectively, we consider three values for that for inners: the first is exactly in the middle between that for eaches and that for cases ($0.4753 = 0.7789 - \frac{0.7789-0.1716}{2}$), the second cost is closer to that for eaches ($0.6271 = 0.7789 - \frac{0.7789-0.1716}{4}$), and the last cost is equivalent to that for eaches itself, i.e., \$0.7789 per case.

In Table 2 we report the total cost savings compared to the cost from the current ship-pack choices for a 52 week period for roughly 350 stores for 526 SKUs. We present the total cost savings as a percentage in the following three scenarios: (1) restricting one ship-pack change at week 1; (2) restricting to one ship-pack change in any week in the planning horizon; and (3) no restriction at all.

Inner replenishment cost ($ per case)	Percentage Saved		
	One Change at week 1	One change in any week	No Restriction
0.4753	0.32%	0.32%	0.34%
0.6271	0.28%	0.29%	0.30%
0.7789	0.35%	0.35%	0.37%

Table 2: Cost savings in percentage for three inner replenishment costs

From this table, we see that by implementing changes in ship-pack sizes, the retailer can expect a cost reduction of between 0.3% to 0.4%.

Also, the savings percentages are not very sensitive to restrictions on ship-pack change. Basically one can save 0.02% more if no restrictions are enforced on when and how often to change the ship-pack, and almost zero improvement if we relax only the timing constraint. Since coordinating these changes for a single point in time is by far the most practical, we will concentrate our analysis on results for which we limit the ship-pack change to the beginning of the planning period from now on.

Table 3 summarizes the ship-pack change recommendations with one change allowed at week 1, namely the changes that provide the cost savings in column 2 of Table 2.

Inner replenishment cost $ per case	0.4753	0.6271	0.7789
Each to Each	383	387	391
Inner to Inner	102	98	88
Case to Case	10	10	10
subtotal	495 (94%)	495(94%)	489(93%)
Case to Each	3	3	3
Case to Inner	0	0	0
Each to Case	5	6	6
Each to Inner	15	10	6
Inner to Case	0	0	1
Inner to Each	8	12	21
subtotal	31(6%)	31(6%)	37(7%)
Total	526	526	526

Table 3: Summary of ship-pack change recommendations for three inner replenishment costs

In short, it shows that for 94% of the 526 SKUs, Beta is already operating with the optimal ship-pack, whereas only 6% of the SKUs, or slightly more than 30 SKUs, require some sort of action. A closer look at the results reveals that out of the 30-plus changes required, changing just half of those SKUs will provide 80% of the cost savings estimated in Table 2. This means that the cost savings are not only tangible, but also quite doable involving a modest set of changes.

The effects of the DC replenishment costs for inners are also shown from Table 3. As the cost of replenishing inners increases, the number of SKUs recommended for eaches (from each to each, or inner to each) and for cases increases (from each to case, or inner to case), while that recommended for inners (from inner to inner, or each to inner) decreases.

The next question is: Who benefits most from the optimized ship-pack, the retail stores, the DC, or inventory-related operations? In the context of Beta this was a very important question as the implementation of any changes would impact three organizations: Distribution, Store Operations, and Supply Chain, each with their own performance measures. From our model analysis, we can group costs into three major categories (see equations 20-24): distribution costs (the DC replenishment cost and DC picking cost), store costs (store-level handling and extra handling cost), and inventory costs (store and DC inventory costs). We analyze the cost saving from the three major categories of cost in our total cost function, as shown in equation (24). The results are shown in Figure 5, where the vertical bars show the absolute savings for each cost component (exact figures blinded) while the diamond dot shows the percentages. The answer seems clear. The DC saves the most from ship-pack optimization. In our calculation, the DC-level cost (Distribution) is reduced by 2.9%, compared to only 0.08% and 0.03% savings for store costs and inventory-related costs, respectively.

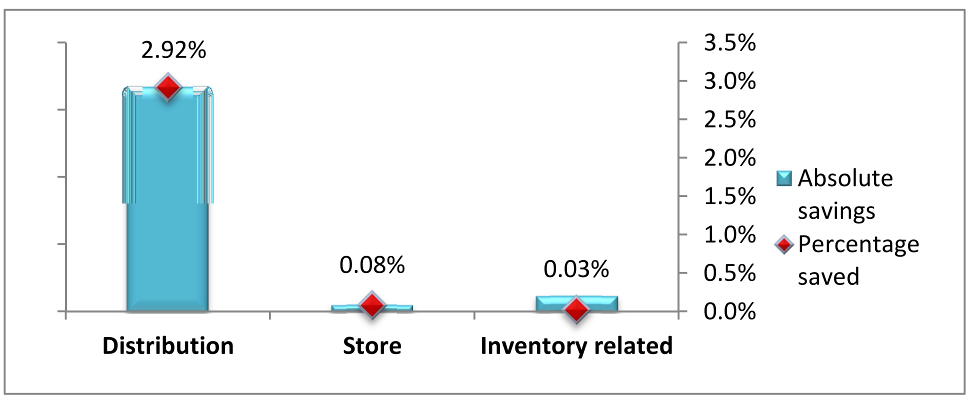


Figure 5: Savings breakdown for inner replenishment cost = $0.4753

VI. Extension: Calculating Optimal Inner-pack Size

Having calculated the potential cost saving from changing ship-pack size, we explore the possibility of further optimizing on the inner-pack sizes. Finding an optimal case configuration is motivated by two factors. First, we observe there is a lack of inner packs for many SKUs and hence maybe an opportunity. Second, Beta might negotiate with vendors to modify their case configuration to include a more economical inner pack, e.g., change the size of an existing inner from (say) 6 units to 8 units. In effect, for each SKU we desire to find the best size for the inner for a given case quantity.

We modify our Warehouse Ship-Pack Cost Model into an Optimal Case Configuration Model. The new model determines the best ship pack size that the DC should replenish its stores. We assume a given case quantity and set it as the upper limit of the pack size, and the lower limit is naturally one, an each. Moreover, we also assume that the inner need be a divisor of the case quantity.

The steps of finding the optimal case configuration are described below (we drop the SKU index for ease of presentation).

1. Given a case quantity SPQ_{case}, we can find all the divisors, e.g. if the case quantity is 12, then the possible inner quantities are 2, 3, 4 and 6, plus an each of size 1.

2. Based on the inner quantity, we approximate the inner replenishment cost by linear extrapolation between each replenishment cost and case replenishment cost. The exact formula used in the model is shown below.

$$replen_{inner} = replen_{each}\left(1 - \frac{SPQ_{inner} - 1}{SPQ_{case} - 1}\right) + replen_{case} * \frac{SPQ_{inner} - 1}{SPQ_{case} - 1} \qquad (27)$$

3. We calculate the annual costs for all possible inner quantities; the smallest one is the preferred inner quantity.

For some SKUs, this analysis can help to identify a preferred inner –size. We show one such example in Figure 6 for a certain SKU.

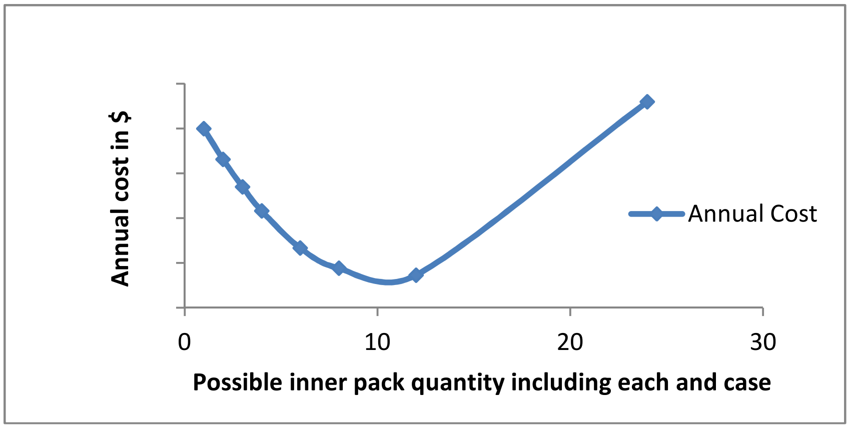


Figure 6: Annual cost over possible inner pack quantities

Here we obtain a nice (discrete) convex curve for this SKU. In this example, an inner pack of 12 is the optimal choice. In fact, it reduces the annual cost by 3.44% from the current ship-pack of 24 units.

However, not every SKU has such a strictly convex shape. Figure 7 shows, for another SKU, a decreasing trend over the feasible range. In this example the best ship-pack is a case of 12 units; actually we might be better off by increasing the case pack size. [3]

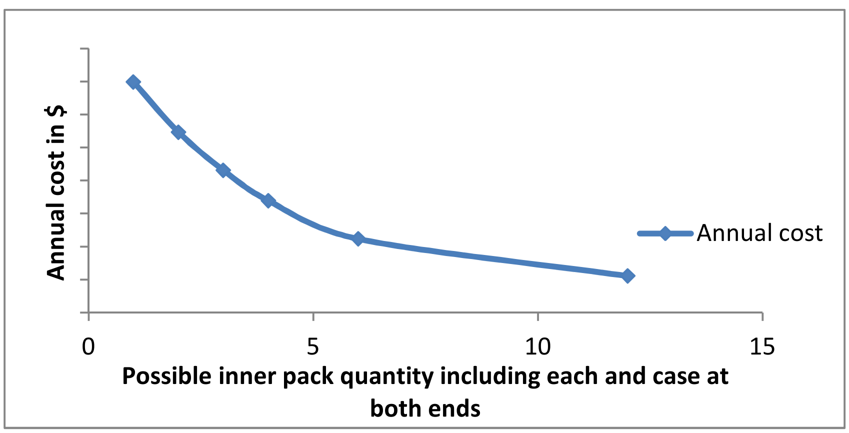


Figure 7: Annual cost over possible inner pack quantities for SKU 05

We run the Optimal Case Configuration Model for all 526 SKUs and find that 171 SKUs would benefit from a more economical inner pack. By changing to the optimal inner packs, the total cost of these 171 SKUs is reduced by 1.30%. For these 171 SKUs, the distribution of the optimal inner pack

[3] One could easily modify the Optimal Case Configuration Model to consider different case-size quantities. We have not done this and leave it for further research. One complication is that changing case quantities can be much more expensive than changing the inner-size, due to the construction and staging of applet loads – which depend critically on the size and shape of a case.

quantity (IPQ) against the current IPQ is shown In Figure 8. Clearly an IPQ of 2 or 4 is the most popular choice according to the results (the two highest bars in Figure 8).

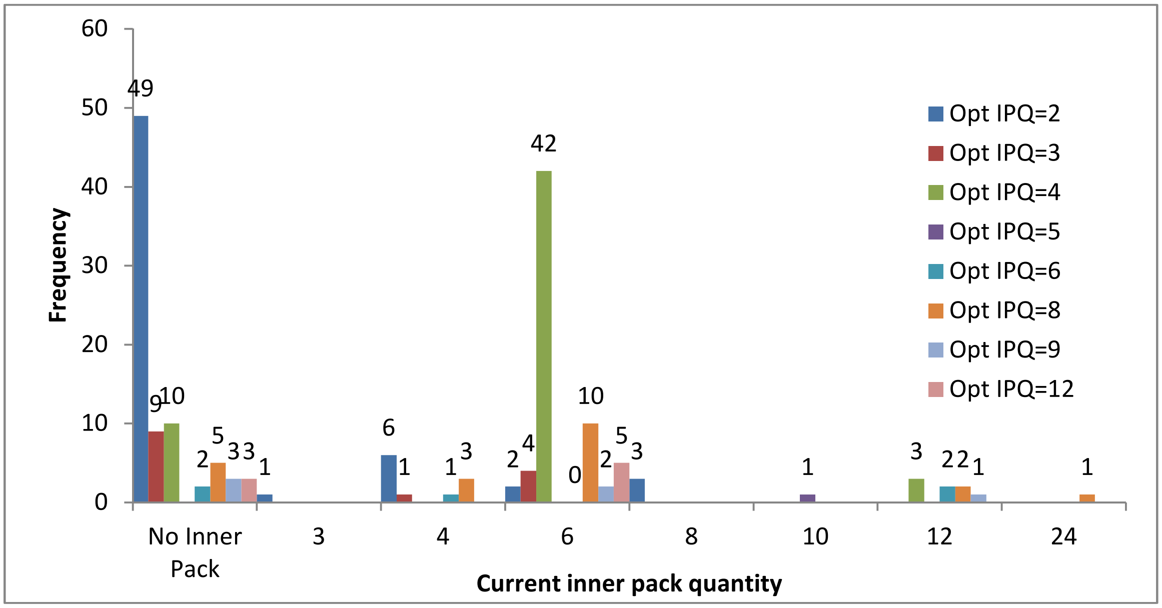


Figure 8: The distribution of optimal IPQ against current IPQ.

VII. Discussion and Conclusion

In this collaborative project with Beta, we establish a cost model that can be used to optimize warehouse ship-pack in the two-echelon distribution system. The three major contributions of this model are the inclusion of store-specific demands, the inclusion of multiple weekly forecasts, and the consideration of extra-handling costs at store level for larger pack size. We determine the optimal ship-pack for the DC given weekly forecasts of store-specific demands over a multi-week planning horizon. Implementation-wise, we have developed for Beta a decision tool that connects to their Microsoft Access database, and can be readily run for any planning period and any DC with its assigned stores.

To combat the large heterogeneity of sales volume across stores, Beta is also considering "dual-slotting" in their warehouses. That is, the DC would set up two picking modules (or slots), instead of the current one picking module, for an SKU; one module would store a larger ship-pack (ie, a case or inner), while the other would store a smaller ship-pack (ie. an inner or each). We therefore have built into our model the capability to determine the best two ship-packs the company should choose for dual-slotting. In fact, with three choices available, there are only three possible combinations. For each of the three combinations, the algorithm will determine for each store and each week the optimal ship-pack and calculate the corresponding total cost. Then the three values are compared and the best combination is selected. Our computation results indicate that compared with optimal single-slotting, optimal dual-slotting can further decrease total costs by another 0.3% - 0.5%.

Two major assumptions in our model are that the demand rate is constant and known, and that the Inventory Position (IP) at stores at the time of order is uniformly distributed. For the latter assumption we have examined the actual distribution of the IP, and found that for many SKUs a geometric or triangular distribution might provide a better empirical fit. Nevertheless, in terms of choosing the optimal *SPQ*, we found the impact from using an alternate distribution was minimal so we adhered to uniform distribution for simplicity.

Some future research questions are to determine the impact of stochastic demand on the ship-pack choice, to incorporate the capital investment of ship-pack changes as well as the capital investment of the dual-slotting, to examine more closely the costs associated with changes to the case configuration, and to incorporate DC replenishment costs and store receipt costs differentiated by SKU.

In summary, in this collaborative project with Beta we establish a cost model that can be used to optimize warehouse ship-pack in a two-echelon distribution system. By exercising the model for one regional DC, we find that Beta can reduce its distribution operating and inventory holding costs by 0.3% to 0.4%. To get some perspective on the potential, consider an example. Suppose the retailer has annual revenues of $10 billion, with inventory turns of ten times a year. Then we estimate the annual inventory holding and distribution handling costs to be in the range of $100 to 200 million. A 0.3% reduction in distribution and inventory holding costs translates to an annual savings of $300,000 to $600,000. This savings can be further enhanced by the possible consideration of dual-slotting and case configuration.

Acknowledgments

The authors acknowledge Beta Corp. for providing us with the research setting, and the support from the Singapore-MIT Alliance, an engineering education and research collaboration among the National University of Singapore, Nanyang Technological University, and MIT.

References:

Broekmeulen, R. A. C. M., J. C. Fransoo, K. H. van Donselaar, T. van Woensel. 2007. Shelf space excesses and shortages in grocery retail stores. Working Paper, Eindhoven University of Technology, The Netherlands.

Erlenkotter, D., 1990. "Ford Whitman Harris and the Economic Order Quantity Model," Operations Research 38(6), 937-946.

Geary, S., Disney, S., Towill, D., 2006. On bullwhip in supply chains—historical review, present practice and expected future impact. International Journal of Production Economics 101, 2–18.

Hopp, W.J. and Spearman, M.L., 1996. Factory Physics, Mc-Graw-Hill, New York, NY.

Lee, H.L., Padmanabhan, P., Whang, S., 1997a. The bullwhip effect in supply chains. Sloan Management Review 38, 93–102.

Silver E.A, D.F. Pyke, R. Peterson, 1998. Inventory management and production planning and scheduling. 3rd ed, Chichester: Wiley.

van Donselaar, K. H., V. Gaur, T. van Woensel, R. A. C. M. Broekmeulen and J. C. Fransoo. 2006. An empirical study of ordering behavior of retail stores. Working paper, Eindhoven University of Technology.

van Zelst, S., K. van Donselaar, T. Van Woensel, R. Broekmeulen and J. Fransoo. 2007. A Model for Shelf Stacking in Grocery Retail Stores: Potential for Efficiency Improvement, International Journal of Production Economics 121, 620-632

Wagner, H.M. 2002. And then there were none. Operations Research 50(1), 217-227.

Yan, X.S., D.J. Robb, E.A. Silver. 2009. Inventory performance under pack size constraints and spatially-correlated demand. International Journal of Production Economics 117, 330–337.

Appendix 1: Derivation of Equation (11)

$$E\left(\frac{1}{2}\min\left(IP_{i,k}^1, IP_{i,k}^2\right) + \frac{1}{2}\left(IP_{i,k}^1 + Q_{i,k,t}^{sp}\right)\right) = E\left(\frac{1}{2}\min\left(IP_{i,k}^1, IP_{i,k}^2\right)\right) + E\left(\frac{1}{2}IP_{i,k}^1\right) + E\left(\frac{1}{2}Q_{i,k,t}^{sp}\right)$$

Clearly, $E\left(IP_{i,k}^1\right) = \frac{ROP_{i,k}}{4}$ and $E\left(\frac{1}{2}Q_{i,k,t}^{sp}\right) = \frac{1}{2}\tilde{Q}_{i,k,t}^{sp}$.

Since $IP_{i,k}^1$ and $IP_{i,k}^2$ both follow discrete uniform distribution in the interval $[0, ROP_{i,k}]$, we can derive a formula which involves only the $ROP_{i,k}$. We assume the two variables are independent, so there are all together $\left(ROP_{i,k} + 1\right) \times (ROP_{i,k} + 1)$ possible pairs for $\left(IP_{i,k}^1, IP_{i,k}^2\right)$. Out of the $\left(ROP_{i,k} + 1\right)^2$ pairs, only one pair will have the minimum value as $ROP_{i,k}$, i.e. when both $IP_{i,k}^1$ and $IP_{i,k}^1$ are equal to $ROP_{i,k}$, and there are three pairs with the minimum value as $(ROP_{i,k} - 1)$, five pairs with the minimum value as $(ROP_{i,k} - 2)$, and so on. In summary, there are $(2i + 1)$ pairs with the minimum value as $(ROP_{i,k} - i)$. Using the definition of the expectation, i.e., $E(x) = x_i Prob(x_i)$, we can calculate the minimum of $\left(IP_{i,k}^1, IP_{i,k}^2\right)$ as follows:

$$E\left(\frac{1}{2}\min\left(IP_{i,k}^1, IP_{i,k}^2\right)\right) = \frac{1}{2}\left(\frac{\sum_{i=0}^{ROP_{i,k}}(2i+1)\left(ROP_{i,k} - i\right)}{\left(ROP_{i,k}+1\right)^2}\right) = \frac{2ROP_{i,k}^3 + 3ROP_{i,k}^2 + ROP_{i,k}}{12\left(ROP_{i,k}+1\right)^2}$$

Thus,

$$E\left(\frac{1}{2}\min\left(IP_{i,k}^1, IP_{i,k}^2\right) + \frac{1}{2}\left(IP_{i,k}^1 + Q_{i,k,t}^{sp}\right)\right)$$

$$= \frac{1}{2}\tilde{Q}_{i,k,t}^{sp} + \frac{1}{4}ROP_{i,k} + \frac{2ROP_{i,k}^3 + 3ROP_{i,k}^2 + ROP_{i,k}}{12\left(ROP_{i,k} + 1\right)^2}$$