Inventory policy for dense retail outlets

Michael Ketzenberg a, Richard Metters b,*, Vicente Vargas c

a Kenan-Flagler Business School, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA
b Cox School of Business, Southern Methodist University, Dallas, TX 75275-0333, USA
c Goizueta Business School, Emory University, Atlanta, GA, USA

Received 1 March 1999; accepted 1 September 1999

Abstract

A potential retail operations strategy is to have a “dense” store. That is, a store that combines high product variety with a small footprint. Retail management desires smaller stores to provide the strategic benefits of convenience and speed to customers, but desires larger stores to provide high product variety. Noting the benefits of smaller, more numerous stores, several retailers well known for their extremely large store size recently have begun experimenting with a small store format. Traditional retail inventory management policies, however, have difficulty combining high variety and small store size. Here, the potential advantages of the dense store type are explored. To facilitate this exploration, inventory policies are developed to help manage small stores by increasing their product density. Results based on grocery industry data indicate that the heuristics compare favorably to optimality and permit the dense store concept to potentially achieve substantial gains compared to current practice. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Service operations; Inventory control

1. Introduction

1.1. Anecdotal evidence for the “dense” store strategy

In recent years, a number of retail firms have made plans to adopt a “denser” store strategy. We use the term “density” to refer to the number of products or categories that a store provides per unit of selling space. This desire has taken two different forms. For some, there has been an increase in the number of product categories or individual stockkeeping units (SKUs) within the same sized store. For others, there has been a desire to reduce store size.

The theory behind increasing category and product variety is that high variety stores are more attractive to customers as they offer a better opportunity for one-stop shopping (Messinger and Narasimhan 1995, 1997). The grocery industry provides a fitting example and will be the main subject of this study. The number of SKUs in the average grocery store increased 96% from 1980 to 1993 (Sansolo and Garry 1994).
Largely, the grocery industry has accommodated increases in SKUs by increasing store size, as average store size has increased nearly 50% from 1985 to 1994 (Sansolo and Garry 1994). However, large stores have several disadvantages well known to retailers: a strategy of a few large stores versus several smaller stores leaves customers with more travel time required to get to a store, usually means more difficulty in parking an automobile close to a store, and requires more search time for customers within a store.

In response to the inherent advantages of keeping stores small, in addition to providing high variety, one of the market leaders in the grocery industry has publicly made “densing up” stores a strategic priority (Safeway annual report 1992, p. 9). Leaders throughout the industry have pursued a strategy of “combination” stores by adding more shelving and product categories to existing stores to incorporate more inventory (Kroger annual report 1997, Safeway annual report 1997). Many supermarkets have become near shopping malls in their own right, adding pharmacies, florists, dine-in restaurants, etc. In one area of expansion, the number of retail banks in US grocery stores has increased from 210 in 1985 to 4400 in 1996 (Williams, 1997).

Examples also abound in other retail industries. Tricon, owner of Pizza Hut, Taco Bell and Kentucky Fried Chicken, has now opened “multi-concept” stores that combine two or even all three of the aforementioned concepts in one physical building (Gibson, 1999). Dollar General, a general merchandise retailer, recently increased their per store SKUs by 20% (Berner, 1998). Southland has revitalized the 7-11 brand and plans to open a new store a day for the next two or three years (Lee, 1998b). A centerpiece of their revitalization has been their new Retail Information System which has allowed them to become more dense, increasing SKUs per selling area an average of 30% (Bennett, 1994, Southland Annual Report 1997).

Prominent examples of firms that are successful with this format include two general merchandise retailers, Dollar General and Family Dollar. Both chains compete for the same target market as Wal-Mart, both are competitive with Wal-Mart on price, and both have over 3000 stores in the US, yet the average store size is 6000 ft² vs. Wal-Mart’s average 92,000 ft². (Berner 1998, Friedman 1998). One strategy of Dollar General is to locate stores next to existing Wal-Marts to highlight their convenience. Wal-Mart has acknowledged these competitors and has begun experimenting with a “Small-Mart” strategy of far smaller stores (Lee, 1998a). Likewise, Home Depot currently has an average store size of 112,000 ft², but in 1999, is experimenting with a small-store strategy (Hagerty, 1999) and Sears is opening “compact” stores (AP, 1998). A key question facing these practitioners is how to combine provide enough variety with the small store convenience to attract customers.

In the grocery industry, the small-store, but non-dense strategy is generally known as a “limited assortment” strategy. Limited assortment stores generally are 20–40% of the size of the average supermarket and stock only a few SKUs in each category. Price competitor chains of over 500 such stores include Aldi, Dia and Save-a-lot (e.g., Lewis, 1997). Other chains, such as Trader Joe’s (Los Angeles) and Harry’s In A Hurry (Atlanta), are upscale competitors. These competitors seek to take advantage of their small size, but have the disadvantage of fewer SKUs.

The purpose of this work is to provide the means to manage a dense store. Current retail inventory practices are not adequate to the task. Consequently, we have devised an inventory replenishment policy for the retail sector that supports a dense store format. We place our research in the context of the grocery industry. Inventory practices are known by the industry to be largely inefficient. Consider dry-groceries. In 1993, the average dry-grocery product was sold to customers 104 days after release by the supplier, with 26 of those days sitting in a retail grocery store (Kurt Salmon Associates 1993, p. 26).

This concern over inventory and replenishment practices has been expressed in an industry-wide movement known as “Efficient Consumer Response” (ECR). This initiative began in the early 1990s and is endorsed by nearly all the major retailers (Kurt Salmon Associates 1993). The tenets of ECR are similar to JIT and time-based competition: the specific goal is “reducing total system costs, inventories and physical assets while improving the customer’s choice.” (Kurt Salmon Associates, 1993, p. 4). A key strategy in ECR is Efficient Store Assortments.
(ESA) which addresses the optimum use of store and shelf space to provide a complete, easy to shop, assortment of products the consumer wants. The idea behind ESA is to enable more space for broader assortments per category or for new categories (Kurt Salmon Associates, 1993, pp. 42–43). Our research supports this and provides a new tool to achieve this result.

1.2. The practical impact of replenishment modeling

There are generally three hierarchical levels of retail inventory policy:

- assortment — deciding which products should be stocked,
- allocation — how much shelf space to give each product in the assortment, and
- replenishment — when and how much to reorder.

Historically, researchers addressing the retail inventory problem have developed models that focus on determining assortment and allocation, rather than on determining inventory stocking and replenishment policies (for example, Anderson, 1979; Corstjens and Doyle, 1983; Borin and Farris 1995; Borin et al., 1994).

The focus here is on finding a simple model for inventory replenishment that enables retail managers to make greater use of their scarce resource: retail space. Methodologically, we construct replenishment heuristics to facilitate category management that are easy to implement. These heuristics are then compared to both optimal policies and current practice. We do so utilizing data from a leading national grocer that is a founder of ECR.

For the data studied, the heuristic reduces inventory levels 24–50% from current practice while maintaining profitability. This inventory reduction has ramifications for the assortment/allocation decisions, as it lessens the need for shelf space assigned to the current assortment. Given a static store size, this enables higher variety through either the addition of more products in a category or the inclusion of more general categories of goods. Alternatively, the reduction in space needs can lead to smaller stores or more aisles per store.

This paper is organized as follows. Section 2 defines the problem and reviews the relevant literature. This is followed by a description of inventory practice at the grocer in question and heuristic development. Heuristics are tested against optimal solutions by dynamic programming in a restricted environment. Then the heuristics are tested under more realistic business conditions by simulation.

2. Problem definition and literature review

The general setting is a retail facility under periodic inventory review that stocks multiple, substitutable products to satisfy stochastic demand. In each time period, the stocking level of any product is restricted because of shelf space allocation decisions made previously. Given a set assortment and allocation scheme, the problem then is to determine the amounts to be ordered in each period so as to maximize profit, where profit is calculated as revenue less the purchasing cost, inventory storage cost, and penalties for substitution and lost sales.

Ordering costs are not included as there is little marginal ordering cost for a given product. As described in detail later, product orders take seconds of employee time to note and transmit electronically, only minutes to kit at a distribution facility, and are loaded onto a common truck filled with hundreds of other orders so there is no marginal transportation cost. As technology improves, ordering costs are likely to decrease from this small amount. One can imagine inventory position being decremented automatically from scanner data captured at the point of sale and the ordering process being computerized to the degree where in-store employees are omitted entirely from the ordering process.

Given a stockout situation in one product, we presume that a known proportion of consumers who originally intended to buy that product will desire to purchase another product as a substitute. As a practical matter, substitutability as a percentage of retail out-of-stock situations has been variously estimated from empirical data at virtually 100% (Motes and Castleberry, 1985), 84% (Walter and Grabner, 1975), 82% (Walter and La Londe, 1975), 73% (Emmelhainz et al., 1981), 58% (Peckam, 1963) and
40% (Nielsen Marketing Research, 1992) depending on the specific retail environment. Product substitutability is important for the replenishment decision as it can act in a similar fashion to the pooling of inventory at a central warehouse; substitutability reduces the need for inventory (McGillvray and Silver, 1978; Moinzadeh and Ingene, 1993).

In Appendix A, we develop a dynamic programming formulation of the replenishment problem. The purpose of developing the formulation is to both link this problem to prior literature and facilitate heuristic comparison to optimal policies for simple cases involving two products. Beyond a few products, optimal solutions for substitutable products are impractical for dynamic programming as the state space expands exponentially (Smith and Agrawal, 2000).

The following notation is useful in mathematically describing the replenishment problem and constructing heuristics.

For products \( j = 1, 2, \ldots, p \) in the assortment, the decision and state variables, are \( Y = (Y_1, Y_2, \ldots, Y_p) \) and \( i = (i_1, i_2, \ldots, i_p) \) where \( Y = (Y_1, Y_2, \ldots, Y_p) \) is the target inventory position on hand after ordering but before demand (a target that will often not be reached due to case–pack sizes) and \( i = (i_1, i_2, \ldots, i_p) \) is the vector of inventory positions at the beginning of the order cycle. Due to the nature of retail demand, we assume that excess demand results in either lost sales or sales of another product, so \( i_j \geq 0 \).

We define the following parameters for each product \( j \). Let \( s_{jk} \) = proportion of unfilled demand that may be filled by substituting product \( k (s_{ij} = 0) \), \( S_j = \sum_{k=1}^{p} s_{jk} \) = the overall substitutability factor, \( 0 \leq S_j \leq 1 \), \( v_j = \) revenue per unit, \( h_j = \) holding cost per unit per period, \( p_j = \) penalty cost per unit of stockout, \( c_j = \) purchase cost per unit, \( R = \) maximum units of product that space allocation allows on the shelf, and \( \phi(x) = \) the joint probability mass function of demand, where \( x = (x_1, x_2, \ldots, x_p) \) denotes a particular realization of demand.

Finally, \( K_j = \) case–pack size. Products cannot be ordered as individual units. They must be ordered in manufacturer prepared “case–packs” of several units each.

Simpler versions of this problem can be solved readily. If case–pack considerations are eliminated and substitutability is zero, traditional marginal analysis is known to be optimal (Karr and Geisler 1956).

If the case–pack size is one and there is no substitutability, we have a multi-period news vendor problem where the form of the solution is a single critical order-up-to number for each product.

Given a case–pack size greater than one, the optimal solution is known to be a single critical reorder point with a target order-up-to point for each product. Note that the \( Y_j \) inventory target can only be attained when \( Y_j - i_j \) is a precise multiple of the case–pack size. Setting \( Y_j \) as a maximum target is consistent with prior literature on batch ordering (Veinott, 1965). Given the complexity of the problem, the form of the solution is not known and the behavior and calculation of optimal solutions and approximating heuristics have not been explored. Consequently, an additional contribution that this manuscript makes is in investigating the behavior of such systems and developing heuristics. These heuristics are relatively simple, lead to good solutions and provide substantial improvement over current practice.

3. Current practice

3.1. An empirical example: existing logistics and inventory practices

This study explores retail inventory management by means of a specific example in the grocery industry. We utilize data from what is known in the trade as the “oil and shortening” category of a store in a leading national grocery chain.

Transportation and logistics have benefited from JIT-based enhancements. Stores can receive product shipments in the oil and shortening category five times weekly as opposed to once a week a decade ago. Delivery lead time is two days. Orders are processed and transmitted electronically between stores and the regional warehouses that resupply them.

Within-store inventory management is less technically advanced. Stores employ a group of night stockers that take the delivered product and place it on the appropriate shelf. Each SKU is visually reviewed to determine if a replenishment order is required. This visual review involves inspecting each
SKU to determine whether a case-pack of the item would fit in the space left vacant by item sales. If so, an order is placed.

As suggested above, inventory policy calls for each product to fill the entire depth of the shelf. The reason given for this policy was to reduce the cost of inventory at the warehouse. However, it is clear that while this may reduce the inventory cost at one location, overall inventory investment is not reduced. This inventory policy does simplify ordering, however, as a visual estimation of the amount of empty shelf space is the only decision required.

The company in question does not have a “state of the art” replenishment system. At the current time scanner based inventory control is still considered too inaccurate to pursue. For a description of a more technologically advanced operating system, we refer the reader to McKenney et al. (1994) and Garry (1996). It is anticipated that this system will become more automated in the future.

3.2. Analyzing an existing shelf set

The oil and shortening category occupies 86 linear feet of shelf space on five separate shelves. The information requirements for analyzing shelf set performance were difficult to obtain. Data were obtained from multiple, disparate, and isolated computer systems. Some data, such as product cost, were only available on manually filed historical order records. Product physical dimensions had to be measured by hand. This is typical of the industry.

A total of 93 distinct SKUs were on display. For a number of products insufficient information was available from store records. As a result, 77 of the 93 possible SKUs, comprising 239 product facings and 70 linear shelf feet, are used in this study. (A product “facing” refers to the front row on a shelf, the only row of product “facing” the customer). Summary statistics on these products are in Table 1.

There is substantial variance throughout the category on each of the measures noted in Table 1. The average case-pack size is 9, but 15 items can be resupplied individually while 34 items have case-pack sizes of 12 to 24.

The average gross margin reported here is 43%, which is consistent throughout the industry in dry goods. In the mind of the public the profit margin for grocery stores is about 1% (e.g., Coleman, 1997). However, this “profit” margin is income netted against all expenses — including the proverbial president’s jet.

Demand differs strongly between products. Sixteen products sell two or less units per week while 43 products average daily unit movement. This is somewhat more homogeneous than the industry average, where 56% of dry grocery items average less than one unit sales per week and 22% of items average at least one unit per day (Kurt Salmon Associates, 1993, p. 27).

Considering the frequent delivery schedule, inventory levels appear high. This excess is emblematic of retailing in general. An industry average of 26 days of inventory of dry grocery items is on display (Kurt Salmon Associates, 1993, p. 26). Inventories of cosmetics at department stores often turn over twice yearly (Seshadri, 1996) and retail clothing stores set inventory levels at nearly 100% service levels by stocking several weeks of demand (Agrawal and Smith, 1996).

Due to relative profitability levels and the influence of case-pack sizes, these inventory levels are not as unreasonable as they appear.

Throughout the retail sector, SKUs that have low sales volume are included in the assortment. Agrawal and Smith (1996) found that two-thirds of the SKUs in retail clothing sell less than one per week. But low-demand SKUs are still profitable. Given the profitability data on Table 1 and assuming holding costs of 25% annually, the profits from a one-unit sale are nearly equivalent to the costs of holding a unit for two years. For these items, case-pack size creates large inventory values relative to demand. Given a case-pack size of 12 and an average demand of one per month inventory turns will not be high.

Safety stock can be large for high demand items as well. Assume inventory decisions are based on a newsvendor inventory model. The cost of understocking, $C_o$, could be construed as the average profit margin above plus, say, 50% of the margin as a penalty cost of lost sales. Assume $C_o$, the cost of overstocking, to be represented by an annual holding cost of 25% of the item cost. The service level that should be attained would be $C_o/(C_o + C_u) = 0.9990$, or stocking at the 99.90 percentile of daily demand.
Table 1

<table>
<thead>
<tr>
<th>Category summary</th>
<th>Case-pack Price (US$)</th>
<th>Cost (US$)</th>
<th>Percent margin (%)</th>
<th>Weekly sales</th>
<th>Maximum inventory¹</th>
<th>Days supply of inventory²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>9</td>
<td>2.92</td>
<td>2.02</td>
<td>43</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>6</td>
<td>2.34</td>
<td>1.27</td>
<td>32</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>20th percentile</td>
<td>1</td>
<td>1.79</td>
<td>1.17</td>
<td>21</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>40th percentile</td>
<td>8</td>
<td>2.10</td>
<td>1.60</td>
<td>31</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>60th percentile</td>
<td>12</td>
<td>2.59</td>
<td>1.81</td>
<td>43</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>80th percentile</td>
<td>12</td>
<td>3.09</td>
<td>2.24</td>
<td>54</td>
<td>15</td>
<td>30</td>
</tr>
</tbody>
</table>

¹Inventory level when all facings allocated to an SKU are full of product.
²Days supply of product at the maximum inventory level. This measure does not equal the mean maximum inventory divided by mean weekly sales. Mean days supply is the numerical average of the days supply of each product individually.

Traditionally there has been little incentive to keep inventories low due to the asymmetric penalties between erring on the low versus high side. Fig. 1 depicts profitability levels of various reorder points assuming the costs on Table 1. Extremely low stocking levels cause substantial stockouts and provide a negative return. Adding inventory increases profitability sharply and profit is maximized at the point shown. Because of the high ratio between profits and holding costs adding large amounts of inventory in excess of the profit maximizing point produces only negligible profitability deterioration.

This traditional view of inventory management, however, neglects the issue of space utilization. Given finite space for a category, including more inventory of one product means excluding another.

Fig. 1. Expected profit vs. order-up-to point.
product entirely, thereby reducing the assortment and making the store less attractive to customers. Or, if all items in a category are allowed to stock to nearly 100% customer service levels, fewer categories can be included in a store.

4. Heuristic development

The central ordering question that must be answered is the number of case-packs, \( N_j \), of product \( j \) to order in each time period. We compare two simple heuristics to both optimal solutions and the current practice utilized by the business studied.

We adapt the literature of base-stock policies (generally, zero ordering cost policies) to the case-pack environment and derive the “base-stock” heuristic \( \text{BAS} \). From the literature on case-packs, we take the essential insights from Veinott (1965), who worked on the single product, back-order problem, and generate the “Modified Veinott” heuristic \( \text{MOD} \) to the lost sales/substitution case with multiple items and lead times. In keeping with the notation of the literature on batch ordering, for each heuristic we define a target inventory level \( Y_j \). The number of cases ordered for a given inventory position is \( N_j = \max(0, [(Y_j - i_j)/(K_j)]) \) where \([x]\) indicates truncation to an integer value.

As a formal heuristic, current practice \( \text{CP} \) seeks to replicate the managerial emphasis on filling all the available space given by the allocation scheme with product.

4.1. Current practice heuristic \( \text{CP} \)

For each product \( j = 1, \ldots, p \), the target inventory level is the maximum stock level,

\[
Y_j = R_j.
\]

Or, equivalently, \( R_j - K_j \) functions as a reorder point. This heuristic has the advantage of operational simplicity — a visual check to see if there is room on the shelf for another case-pack is all that is required.

Some additional notation is useful for the subsequent heuristics. Let \( \phi_j(x) = \sum_{x_j=0}^{\infty} \cdots \sum_{x_{j-1}=0}^{\infty} \sum_{x_{j-2}=0}^{\infty} \cdots \sum_{x_1=0}^{\infty} \phi(x_1, \ldots, x_p) \), the marginal probability mass function for product \( j \), and \( \Phi_j(x) = \sum_{u=0}^{x} \phi_j(u) \), the cumulative marginal demand distribution for product \( j \).

4.2. Base stock heuristic \( \text{BAS} \)

For each product \( j \), determine the target inventory level as follows:

\[
\text{candidate}_j = K_j
\]

\[
+ \min \left\{ U \mid \Phi_j(u) \geq \frac{\pi_j + \nu_j - c_j}{h_j + \pi_j + \nu_j - c_j} \right\}.
\]

(3)

\[
Y_j = \min\{\text{candidate}_j, R_j\}.
\]

(4)

The BAS heuristic calculates a traditional base stock inventory position plus a case-pack \( \text{(3)} \) so that inventory after ordering is above the traditional base stock reorder point. The target inventory value is subject to a maximum of the shelf capacity \( \text{(4)} \). The reason for including the BAS heuristic in this study is due to its benchmark status among researchers, its simplicity in calculation and that it takes into consideration the relative profitability of items.

4.3. Modified Veinott heuristic \( \text{MOD} \)

The MOD heuristic attempts to consider both case-pack size and substitutability in addition to differential product margins. At the core of the MOD heuristic is the form of the optimal policy for ordering by case-packs (Veinott, 1965): if expected profitability, \( E(\text{Profit}) \), is a unimodal function of inventory where the slope changes sign once from positive to negative as inventory positions move from zero to infinity, then the optimal case-pack ordering policy is to order up to, but not over, a point \( B \) where \( E(\text{profit} \mid y = B) = E(\text{profit} \mid y = B - \text{-case-pack}) \) (Fig. 1). Point \( B \) less the case-pack size leaves point \( A \) as the reorder point. We adapt this fundamental insight to the partial lost sales/substitutability case with positive lead times.

We define two functions for each product \( j \). The first is the expected sales for a given stocking level.
The second approximates the profitability of a given stocking level using the expected sales.

\[ x_{ \text{sales}}(u) = \sum_{i=1}^{u} \phi(i) + u(1 - \Phi(u)), \]

\[ x_{ \text{profit}}(u) = (v_j - c_j) \frac{x_{ \text{sales}}(u)}{\text{gross margin}} \]

\[ - h_j(u - x_{ \text{sales}}(u)) \]

\[ - \pi_j(\mu_j - x_{ \text{sales}}(u)) \]

\[ - S_j(v_j - c_j) x_{ \text{sales}}(u), \text{ substitutability} \]

where

\[ \mu_j = \sum_{i=0}^{\infty} i \phi(i). \]

is the expected value of demand.

In the formulation of the \( x_{ \text{profit}} \) function, we assume that a penalty cost applies when a customer is forced to substitute an item. Given product substitutability, however, a sale of one item can be viewed as a decline in sales of another item. Because a consumer usually substitutes within a restricted set, say a premium product for another premium product, the substitutability term in \( x_{ \text{profit}} \) assumes the same revenue and cost functions of the product at hand. The reverse situation, the additional demand seen due to stockouts of other items, is not considered. This serves to reduce the safety stock held and models the pooling effect of product substitutability discussed earlier, while keeping the model simple to use and understand.

For each product \( j \) determine the target inventory level as follows:

\[
\text{difference} = 0,  \\
\text{Target} = K_j,  \\
\text{Reorder} = 1,  \\
\text{While difference} \geq 0 \text{ and Target} < R_j \text{ do begin},  \\
\text{Target} = \text{Reorder} + K_j,  \\
\text{Target}_j_{\text{profit}} = x_{ \text{profit}}_j (\text{Target}),  \\
\text{Reorder}_j_{\text{profit}} = x_{ \text{profit}}_j (\text{Reorder}),
\]

if difference < 0 then Target = Target - 1,
else Reorder = Reorder + 1,
end while,
\( Y_j = \text{Target} \).

The MOD heuristic is a single pass algorithm for determining case–pack orders. For each product, the default order-up-to target is the minimum possible value that forces the product to be included in the assortment, the case–pack size \( K_j \). The profits associated with the order-up-to target and reorder point are compared in the value ‘‘difference.’’ As long as the difference is positive the reorder point and target inventory levels increase. They are increased until either the difference becomes negative or the target inventory level reaches shelf capacity.

It should be noted that there are practical limitations to the heuristics described above – they cannot reasonably be applied to every product in a store. ‘‘Loss leaders’’ represent a common marketing technique where products are sold below cost. ‘‘Slotting fees’’ are often given to grocers for a minimum shelf requirement. This will not usually present a barrier, since slotting fees are typically given just for being present in a store (i.e., receiving a single facing).

5. Experimental design and results: two-product case

5.1. Experimental design

The heuristics are compared to optimality by dynamic programming in a two-product scenario. Later, in Section 6, we compare heuristics by simulation for the full 77-product category. Dynamic programming is not a practical solution method in this case due to the well-known ‘‘curse of dimensionality.’’ The only reason for using dynamic programming is to give a sense of how distant from optimality each heuristic may be in a limited setting. System parameters for both analyses, such as the product revenues, costs, physical capacities, case–pack sizes, etc., are from actual store data. Actual sales data is used here as the ‘‘natural’’ demand (as opposed to demand stemming from substitution) for each product. Normally, sales data excludes lost sales information, and must be
converted to a demand function (e.g., Hill, 1992). In this case, as also occurred with Agrawal and Smith (1996), the actual inventory levels are so large for all but a few products that this step is not necessary.

For the two-product problem, an optimal solution is found and costs are calculated utilizing the dynamic program in Appendix A. As dynamic programming is used to find exact expected values, statistical analysis is not required for this portion of the experimental design.

Parameters are varied from a base case to explore the effects of parameter changes. The base case assumes two identical products, which have retail prices of US$2.92, cost US$2.02, sell an average of 10 units/week, and have a case–pack size of 9, corresponding to the averages on Table 1. (For the CP heuristic, the industry average stocking policies cited earlier are used.) The substitutability percentage is assumed to be 70%, holding costs 25% annually of product cost, the lost sales penalty cost is assumed to be 50% of product margin and the penalty for substitution is 10% of the lost sales penalty. A cycle of length 3 days is modeled, roughly corresponding to the case where 3 days of demand can occur before a replenishment is received. In accordance with both analysis of the data and similar work, Poisson demand distributions are assumed for the base case (e.g., Karr and Geisler, 1956; Hill, 1992).

Ten scenario variations are explored to this base case. The lost sales penalty cost is set to 10% and 200% of product margin. Corresponding to research cited earlier, substitution takes on values of 40% and 100%, and a substitution rate of 0% is also used for comparison. Due to the presence of negative binomial distributions noted in retail practice by Agrawal and Smith (1996), the negative binomial demand distribution is explored. Three values of the variance-to-mean ratio of 1.1, 1.5 and 2, were used for the negative binomial distribution.

5.2. Results

The detailed results from the base case experimental cell (Table 2) mirror those of all other cells. Gross profits are nearly identical for all policies because each heuristic has a fill rate beyond 99%. The reasons for extreme fill rates for all policies have been identified earlier: large differences in holding versus lost sales costs and the effect of case–pack sizes. Net profits differ between the heuristics, but the differences are small, amounting to only 1% in the base case.

The results of interest are the average inventory investment and the return on inventory investment (ROII), a common industry barometer of effectiveness (e.g., Thayer, 1991), and a more useful barometer when implementing a dense store strategy. Both of these results follow a clear pattern: in every experimental cell, MOD is the superior heuristic (Table 3). The reason we focus on ROII rather than net income is because of the relative scarcity of retail space. The business implications of policies which achieve higher ROII values are discussed in detail in Section 6, but include the ability to include more products in an assortment, more product categories in a store or allow a reduction in store size.

Perhaps the most striking aspect of Table 3 is the lack of numerical change of ROII results for the BAS and CP heuristics as parameters change. Decreases in substitutability and increases in penalty costs and demand variance have the effect of increasing the probability of product shortages. However, due to reasons discussed earlier concerning industry cost structure, inventory levels are so high for these heuristics that the increased lost sales due to these factors is insufficient to change the ROII for these heuristics to the level of one-hundredth of a point.

Substitutability does not have a large effect on optimal or MOD heuristic results from 0% to 40%
parameter values. This is due to the large stockout to penalty cost ratio. Changing substitution rates from 0% to 40% may change the optimal service level from, say, 99.9% to 99.5%, which is insufficient to cause an integral change in the optimal inventory reorder point.

6. Experimental design and results — 77-product simulation

6.1. Experimental design

Here, we simulate the heuristics in the more realistic environment of the 77 products described earlier. Varied quantities include the penalty cost for lost sales, $\pi = 10\%$, $50\%$ and $200\%$ of net margin. We assume that the quantities $S_j = a$ constant factor among all products that represents the general substitution percentage and employ values $S = 40\%$, $70\%$ and $100\%$. The individual product substitutability factors, $s_j$, are uniformly distributed based on product type: for example, a premium olive oil would only be substituted for other premium olive oils, and all other premium olive oils are considered equally likely to benefit from a stockout. Due to the absence of a strong demand distribution effect in the dynamic program, only the Poisson distribution is used.

Detailed logistics practices are modeled. Deliveries occur every day except Saturday and Sunday. The demand pattern varies by day of the week, with Monday through Friday accounting for 13% of weekly demand and Saturday and Sunday for 17% and 18% of weekly demand, respectively, corresponding to store sales. There is a lead time of 2 days for ordered product.

Ten replications of 6 years is run for each heuristic and experimental cell. Statistics are collected after the first simulated year to avoid initialization bias. Total CPU time for all experimental cells is approximately 24 h.

6.2. Results and implications for dense stores

The mean results of a typical experimental cell (Table 4) mirror those seen earlier — gross and net profits are similar but ROII differs considerably between heuristics. Standard errors are small for all entities reported due to the extreme length of time simulated.

Evaluating the effect on ROII as parameters change, again MOD responds to changes in the substitutability level and is more effective as substi-

<table>
<thead>
<tr>
<th>Substitutability</th>
<th>Optimal solution</th>
<th>MOD</th>
<th>BAS</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>35.43</td>
<td>29.86</td>
<td>18.01</td>
<td>7.14</td>
</tr>
<tr>
<td>70% (base case)</td>
<td>29.83</td>
<td>26.42</td>
<td>18.01</td>
<td>7.14</td>
</tr>
<tr>
<td>40%</td>
<td>26.41</td>
<td>23.64</td>
<td>18.01</td>
<td>7.14</td>
</tr>
<tr>
<td>0%</td>
<td>26.39</td>
<td>23.63</td>
<td>18.01</td>
<td>7.14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Penalty cost (% of margin)</th>
<th>Optimal solution</th>
<th>MOD</th>
<th>BAS</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>26.41</td>
<td>23.63</td>
<td>18.01</td>
<td>7.14</td>
</tr>
<tr>
<td>50 (base case)</td>
<td>29.83</td>
<td>26.42</td>
<td>18.01</td>
<td>7.14</td>
</tr>
<tr>
<td>10</td>
<td>29.83</td>
<td>26.42</td>
<td>18.01</td>
<td>7.14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Demand distribution</th>
<th>Optimal solution</th>
<th>MOD</th>
<th>BAS</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson (base case)</td>
<td>29.83</td>
<td>26.42</td>
<td>18.01</td>
<td>7.14</td>
</tr>
<tr>
<td>Negative binomial: var/mean = 1.1</td>
<td>26.27</td>
<td>21.38</td>
<td>14.51</td>
<td>7.14</td>
</tr>
<tr>
<td>Negative binomial: var/mean = 2</td>
<td>22.33</td>
<td>17.98</td>
<td>12.84</td>
<td>7.14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Results from a typical experimental cell: substitutability 100%, lost sales cost 50% of margin (annual basis)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MOD</td>
</tr>
<tr>
<td>Gross profit mean</td>
<td>28,985 (23)</td>
</tr>
<tr>
<td>Holding cost mean</td>
<td>241 (0.1)</td>
</tr>
<tr>
<td>Penalty cost mean</td>
<td>28 (0.3)</td>
</tr>
<tr>
<td>Net profit mean</td>
<td>28,716 (22)</td>
</tr>
<tr>
<td>Average inventory mean</td>
<td>964 (0.4)</td>
</tr>
<tr>
<td>ROII mean</td>
<td>29.79 (0.03)</td>
</tr>
</tbody>
</table>
tutability increases (Table 5). Further experimentation (not shown) indicates that the sensitivity of MOD to parameter values outside the range of the experimental design is not large.

Adoption of the MOD heuristic as a replenishment tool assists in achieving the goals of a dense store: either more categories can be offered in the same space, or the same number of categories can be offered in a smaller space. Thus far, we have assumed that the allocation of facings to SKUs is given. That is, each heuristic accords the oil and shortening category 239 facings and 70 linear shelf feet. Given the lower inventory levels, however, the MOD heuristic could reduce the number of facings from 239 to 179 (Table 6). This reduction in facings would reduce the number of linear feet required for the 77 products analyzed from 70 to 53 ft.

Given this space reduction and assuming other products have similar characteristics, 103 products could fit in the same space where the 77 original products now reside. Assuming equivalent net profits to the shelf set analyzed, which is optimistic, this could increase annual profitability of those 70 linear ft 38% by adding more SKUs to the category or by adding more categories to the store.

It could be argued that cutting the shelf space of the category may depress category sales. There is ample evidence that shrinking the space of one product leads to sales shifting to other products in the category. However, we are not aware of any research that indicates that shrinking the shelf space of all products in all categories simultaneously results in category-wide or store-wide revenue reductions (a review of related research is in Mahajan and Van Ryzin, 1999).

Given the assumption that category-wide sales declines may occur, the MOD heuristic can still assist in increasing variety. Using this heuristic, the same number of facings could be achieved by changing shelf depth from the current 20 to 16 in. Given the extra space, and assuming six feet between aisles, an additional aisle can be included for every 13 aisles.

As another alternative to utilizing the decrease in inventory required we can look to the overall reduction in store size. To accommodate the increased number of SKUs and categories, grocers and other retailers have moved to a “superstore” policy of building much larger new stores and increasing the size of current stores. Given an equivalent reduction in inventory in all categories as has been demonstrated here, the volume of SKUs and categories now present in a superstore under the currently practiced replenishment system may fit within a conventional store, eliminating the need for retrofitting older stores.

### 7. Conclusions

We have generated a specific heuristic that performs well in achieving a dense store compared to both actual practice and naïve heuristics such as a news vendor heuristic. The practical importance of this work, however, does not lay in the specific heuristic devised. Rather, it is in the demonstration that excessive inventory levels are impeding profitability and the provision to management of points on a trade-off curve between inventory of current

---

#### Table 5
Return on inventory investment (ROII) cross-tabulated by substitutability and lost sales penalty cost

<table>
<thead>
<tr>
<th>Substitutability</th>
<th>MOD</th>
<th>BAS</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>26.15</td>
<td>15.37</td>
<td>14.47</td>
</tr>
<tr>
<td>70%</td>
<td>21.48</td>
<td>15.40</td>
<td>13.96</td>
</tr>
<tr>
<td>40%</td>
<td>19.93</td>
<td>15.39</td>
<td>13.58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Penalty cost</th>
<th>MOD</th>
<th>BAS</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>200% of margin</td>
<td>20.79</td>
<td>14.96</td>
<td>13.77</td>
</tr>
<tr>
<td>100% of margin</td>
<td>22.45</td>
<td>15.41</td>
<td>14.05</td>
</tr>
<tr>
<td>50% of margin</td>
<td>24.32</td>
<td>15.79</td>
<td>14.20</td>
</tr>
</tbody>
</table>

#### Table 6
Allocation ramifications of replenishment heuristics (substitutability 100%, lost sales cost 50% of margin)

<table>
<thead>
<tr>
<th>MOD</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facings</td>
<td>179</td>
</tr>
<tr>
<td>Linear shelf feet</td>
<td>53</td>
</tr>
<tr>
<td>Number of products in 239 facings</td>
<td>103</td>
</tr>
<tr>
<td>Required shelf depth for 239 facings</td>
<td>16 in.</td>
</tr>
<tr>
<td>Potential profit from 70 linear shelf feet</td>
<td>US$39,130</td>
</tr>
</tbody>
</table>
goods and inclusion of other categories. The excessive inventory levels seen in the industry have a historical and reasonable basis — logistics practices, supply chain relationships and information systems of a decade ago necessitated larger inventories. The implementation of ECR practices will negate the reasons for those inventories, but inventory policy must change or those inventories will persist.

The trade-off offered here is not the classic inventory versus service trade-off: traditionally, lower inventories mean lower costs and lower service. Quite the contrary, as shown in Table 4, the MOD heuristic actually has fewer lost sales than current practice, although the lost sales of all the heuristics are minimal. Instead, the excessive inventory levels held affect profitability through a less obvious route - by crowding out other categories of goods. A retailer with limited shelf space must face the trade-off of putting fewer categories out for sale against holding inventories of current products. But without the context of an opposing inventory policy and the associated space required, that trade-off is not defined.

Implementation of this, or any other, heuristic at virtually any large U.S. grocery chain requires appropriate information systems. At the store in question, incompatible systems hold sales data, product costs, product physical dimensions and store shelving information. The direction of both this company and the industry, however, is toward further systems integration and the automation of inventory analysis by using check-out scanner data to feed computerized ordering. The implementation of scanning systems in grocery stores was cost justified based on check-out labor savings alone. The financial justification of continued ECR systems improvements, however, depend on how store management can use these systems to improve business practice. So the promise of better space utilization and the consequent potential of revenue gains or cost decreases is important in moving the industry forward. The trade-off faced by the entire grocery industry is that of the cost of improving information systems and inventory practices versus providing customers with variety and convenience. This work helps to quantify that trade-off.

Future research is needed on many fronts. The analytic determination of optimal policies in the complex retail environment could guide future heuristic work. On a practical level, combining the currently disparate questions of assortment, allocation and replenishment is also a worthwhile area of pursuit.

The heuristics were chosen due to their relative simplicity, rather than their proximity to optimality. Heuristics of other types may also have potential. Given the similarity of this problem to other work, heuristics based on reparable inventory theory or marginal analysis could be useful. Such methods have experience difficulty with case-pack restrictions, and are better suited for situations that call for overall service level constraints, but adaptation may be possible.

Appendix A. Dynamic programming formulation

A dynamic programming representation used to calculate optimal policies for the two-product problem is formed here. In the interest of notational compactness, we define the following auxiliary variables, determined from the state and decision variables and demand realization. For each product \( j = 1,2 \), let \( Y_j = I_j + N_j K_j \), inventory position after ordering but before demand, and \( s_j = \sum a_j \) round \( \max(0, x_j - y_j) s_{ij} \), substitution demand.

Demand is independent and identically distributed between periods; however, demand need not be independent nor identically distributed among products. There is no disposal or deterioration of inventory. Primary demand is satisfied before any substitution demand is met. There is no secondary substitution (as is shown in the text, due to the service levels utilized in a time-based logistics environment, this assumption does not materially affect the results). For ease of exposition, we assume that there is no delivery lag, but this assumption is relaxed for the simulation.

We now discuss the additional notation associated with a cost of substitution.

Let \( \pi'_{kj} \) for \( k \neq j \) denote the penalty for meeting a demand for product \( k \) with product \( j \). We presume that \( \pi'_{kj} < \pi k \), i.e., it is less costly to meet a demand with a substitute than to leave it unfilled. In the interest of notational compactness, we define the following auxiliary variables for each product \( j = 1,2 \)
and \( m = 1.2 \). Let \( I_j = \max(0, y_j - x_j) \), leftover product \( j \) after primary demand \( d_{jm} = \text{round}(\max(0, x_j - y_j) s_j) \), secondary demand generated by product \( j \) for product \( m \), \( a_{jm} \) secondary demand generated by product \( j \) and met by product \( m \): if \( \sum_k d_{km} \leq l_m \), then \( a_{jm} = d_{jm} \), for all \( j \) (all secondary demand for product \( m \) satisfied), else \( a_{jm} = (d_{jm})/\left(\sum_k d_{km}\right) l_m \), for all \( j \) (secondary demand causes a stockout in product \( m \)).

The optimization can be characterized as an infinite horizon dynamic programming model. The value \( \tilde{c} \) is the equivalent average return per period when an optimal policy is used. The extremal equations are:

\[
\begin{align*}
&f(i) + \tilde{c} = \\
&\max_{Y_1, Y_2, \ldots, Y_p, R_j \geq Y_j \geq K_j} \left\{ \prod_{j=1}^p \left( 1 - \sum_{i=0}^{\infty} \frac{\prod_{j=1}^p \prod_{k=1}^q \prod_{l=1}^m \prod_{s=0}^p a_{jkl}}{\prod_{j=1}^p \prod_{k=1}^q \prod_{l=1}^m \prod_{s=0}^p a_{jkl}} \right) \right\}.
\end{align*}
\]

Since the state space is finite, there exists an optimal stationary policy that does not randomize.

References


Nielsen Marketing Research, 1992. Category Management: Posi-
tioning Your Organization to Win. NTC Business Books, Chicago ILL.


