CHAPTER 3
LITERATURE SURVEY

3.1 BACKGROUND
The retail industry has emerged as a fascinating choice for researchers in the field of supply chain management. It presents a vast array of stimulating challenges that have long provided the context of much of the research in the area of operations research and inventory management. However, in recent years, advances in computing capabilities and information technologies, hyper-competition in the retail industry, emergence of multiple retail formats and distribution channels, an ever increasing trend towards a globally dispersed retail network, and a better understanding of the importance of collaboration in the extended supply chain have led to a surge in academic research on topics in supply chain management. Many supply chain innovations (e.g., vendor managed inventory) were first conceived and successfully validated in this industry, and have since been adopted in others. Conversely, many retailers have been quick to adopt cutting edge practices that first originated in other industries.

However, for every example of leading edge progressive thinking among retailers, there are examples of archaic systems and planning processes. More over, there continue to be a host of open problems facing practitioners and academics. All of this is, of course, good news for academics engaged in research in retail supply chain management. The recent past has witnessed exciting new research- theoretical as well as applied –aimed at addressing some of the retail industry’s many pressing challenges. This chapter is an attempt to summarize some of the research and present a perspective on
what new application may lie ahead. They are:

1) Assortment planning
2) Retail operation

**3.1.1 ASSORTMENT PLANNING:**

A retailer’s assortment is defined by set of products carried in each store at each point in time. The goal of assortment planning is to specify an assortment that maximizes sales or gross margin subject to various constraints, such as a limited budget for purchase of products, limited shelf space for displaying products, and a variety of miscellaneous constraints such as a desire to have at least two vendors for each type of product. Given fixed store space and financial resources, assortment planning requires a tradeoff between three elements: how many different categories does the retailer carry (called a retailer breadth), how many SKUs do they carry in each category (called depths), and how much inventory do they stock of each SKU, which obviously affects their in-stock rate. The breadth vs. depth tradeoff is a fundamental strategic choice faced by all retailers. We have all had the experience of going into a store looking for a particular product, not finding it, and settling for another similar product instead. This is called substitution, and the willingness of customers to substitute within a particular category is an important parameter in assortment planning. We can delineate three patterns with respect to customer substitution:

1) The customer shops a store repeatedly for a daily consumable and one day she finds it stocked out so she buys another. This is called stock-out based substitution.
2) A customer identifies a favorite product based on ads or what she has seen in other stores, but when she tries to find it in a particular store,
she can’t because they don’t carry it, so she buys another product. This is called assortment based substitution.

3) The consumer chooses her favorite product from the ones she sees on the shelf in a store when she is shopping and buys it if it has higher utility than her no purchase option. In this case, there may be other products she would have preferred and in this sense we can say she substituted, although she may not be aware that these other products exist and hence doesn’t herself think of her purchase decision as involving substitution.

The first two patterns are common with daily consumables like food and the later with consumable durables like apparel or consumer electronics. Assortment planning is a relatively new but quickly growing field of academic study. The academic approach to the assortment planning problem rests on the formulation of an optimization problem with which to choose the optimal set of products to be carried and the inventory level of each product. Decisions for each product are interdependent because products are linked in considerations such as shelf space availability, substitutability between products, common vendors (brands), joint replenishment policies. Most of the literature focuses on a single category or subcategory of products at a given point in time. While a retailer might have a different assortment at each store, the academic literature has focused on determining a single assortment for a retailer, which could be viewed as either a common assortment to be carried at all stores or the solution to the assortment planning problem for a single store. In this section we will be reviewing four streams of literature that assortment planning models are build on: product variety and product line, shelf space allocation, multi product inventory systems, and a consumer’s perception of variety.
3.1.1.1 PRODUCT VARIETY AND PRODUCT LINE DESIGN

Product selection and the availability of products has a high impact on the retailer’s sales, and as a result gross profits and assortment planning has been the focus of numerous studies, mostly concerned with whether assortments were too broad or narrow. Retailers have increased product selection in all merchandise categories for a number of reasons, including heterogeneous customer preferences, consumers seeking variety and competition between brands: Quelch and Kenny [61] report that the number of products in the market place increased by 16% per year between 1985 and 1992 while shelf space expanded only by 1.5% per year during the same period. This has raised question as to whether rapid growth in variety is excessive. For example, many retailers are adopting an “efficient assortment” strategy, which primarily seeks to find the profit maximizing level of variety by eliminating low-selling products and “category management”, which attempts to maximize profits within category. There is empirical evidence that variety levels become so excessive that reducing variety does not decrease sales [8, 12, and 23]. And from the perspective of operations within store and across the supply chain, it is clear that variety is costly: a broader assortment implies less demand and inventory per product, which can lead to slow selling inventory, poor product availability, higher handling costs and greater markdown costs.

The literature that studies the economics of product variety is vast. The main model in this field is the oligopoly competition between single product firms based on Hotelling [36]. In the Hotelling model, consumers are distributed uniformly on a line segment and firms choose their positions on the line segment and their prices to maximize profits. Consumers utility from each firm is decreasing in the firm’s price and their physical distance to the firm. Each consumer chooses the firms that provide her the maximum utility. The
objective is to find the number of firms, their location and their prices in equilibrium and the resulting consumer welfare. Extension of this model is used to study product differentiation. There are two type of product differentiation. In a horizontally differentiated market, products are different in features that can’t be ordered. In that case, each of the products is ranked first for some of the consumers. A typical example is shirts of different color. In vertically differentiated markets, products can be ordered according to their objective quality from the highest to the lowest. A higher quality product is more desirable than a lower quality product for any consumer. Anderson et al. [3] and Lancaster [41] provide excellent reviews of this literature.

One of the outgrowths of the literature on the economics of product variety is the product line design problem pioneered by Mussa and Rosen [54] and Moorthy [52]. A monopolist chooses a subset of products from a continuum of vertically differentiated products and their prices to be sold in a market to a variegated set of customer classes in order to maximize total profit. Consider cars as a product with a single attribute, say engine size. The monopolist’s problem is to choose what size engines to put in the cars and how to price the final product. These papers assume convex production costs and do not consider operational issues such as fixed costs, changeover costs, and inventories. Joint consideration of marketing and production decisions in product line design is reviewed by Eliashberg and Steinberg [24]. Dobson and Kalish [21] proposes a mathematical programming solution for this problem in the presence of fixed costs for each product included in the assortment. Desai et al. [20] study the product line design problem with component commonality. De Groote [19] also considers concave production costs and analyzes the product line design problem in a horizontally differentiated market. He shows that the firm chooses a product
line to cover the whole market and the product locations are equally spaced. Alptekinoglu [2] extends his work to two competing firms, one offering infinite variety through mass customization and the other limited variety under mass production. He shows that the mass producer needs to reduce variety in order to mitigate the price competition. Chen et al. [15] is the only paper that considers product positioning and pricing with inventory considerations. They show that the optimal solution for this model under stochastic demand can be constructed using dynamic programming.

These models were early treatments of assortment planning from the manufacturer’s view that were precursors of similar models developed for retailing. The manufacturer’s problem is one of product positioning in an attribute space and pricing. The retailer’s problem is to select products from the product lines of several manufacturers. A more careful consideration of inventories at product level is needed in retail assortment planning, since inventories have a direct impact on both sales and costs for the retailer.

3.1.1.2 MULTI-PRODUCT INVENTORY MODELS

Multi-product inventory problems are also highly relevant to the assortment planning problem. The inventory management of multiple products under a single shelf space or budget constraint is studied extensively in the operations literature and solutions using Lagrangian multipliers is presented in various text books, e.g., Hadley and Whitin [32]. Downs et al. [22] describe a heuristic approximation to the multi-period version of this problem with lost sales. In these models, the demand of products is not dependent on others’ inventory levels.

The other group of inventory models with multiple products considers stock-out based substitution, focusing on the stocking decisions given a selection, but not the selection of the products. These models are based on
an exogenous model of demand. McGillivray and Silver [50] first introduced the problem with two products. Parlar and Goyal [59] study the decentralized version of this problem. Rajaram and Tang [64] present heuristic algorithms for the solution of the case with n products. Netessine and Rudi [57] investigate the case with n product under centralized and decentralized management regimes. The complexity of the problem is prohibitive and it is not possible to obtain an explicit solution to the problem. Netessine and Rudi [57] find that a decentralized regime carries more inventory than the centralized regime because of the competition effects. Mahajan and van Ryzin [47] establish similar results under dynamic customer substitution with the multinomial logit choice model. Parlar [60] study the infinite horizon version of this problem under centralized and competitive scenarios respectively. Lippman and McCardle [44] consider a single period model under decentralized management, where aggregate demand is a random variable and demand for each firm is a result of different rules of initial allocation and reallocation of excess demand. Bassok et al. [7] consider an alternative substitution model, in which the retailer observes the entire demand before allocating the inventory to products. In this retailer controlled substitution model, the retailer may upgrade a customer to a higher quality product. The reallocation solution is obtained by solving a transportation problem.

The literature on assemble-to-order systems is also related. The demands for individual components are linked through the demand for finished goods. An online retailer’s order fulfillment problem when customer can order multiple products can be viewed as assemble-to-order systems. Song [69] estimates the order fill rate in such systems and discusses other examples from retailing.
3.1.1.3 SHELF SPACE ALLOCATION MODELS

In some products segment such as grocery and pharmaceuticals, how much shelf space is allocated to a given product category is an important component of the assortment planning process. This view seems especially relevant for fast moving products whose demand is sufficiently high that a significant amount of inventory is carried on the shelf. This contrast with other categories e.g., shoes, music, books where only one or two units are carried for most SKUs, hence amount of inventory and shelf space are not critical decisions at product level.

In an influential paper Corstjens and Doyle [18] suggest a method for allocating shelf space to categories. They perform store experiments to estimate sales of product. The problem of profit maximization with a shelf space constraint is solved within a geometric programming framework. Their results are significantly better than commercial algorithms that allocate space proportional to sales or to gross profit by ignoring interdependencies between product groups. The estimation and optimization procedures can not be applied to large problems; hence they elect to work with product groups rather than SKUs. Bultez and Naert [13] apply the paper Corstjens and Doyle [18] model at the brand level assuming symmetric cross elasticities within product groups. Their models are tested at four different Belgian supermarket chains, leading to encouraging results. An interesting paper by Borin and Farris [9] reports the sensitivity of the shelf space allocation models to forecast accuracy. They compare the solution with correct parameters to that with incorrect parameter estimates. Even when the error in parameter estimates are 24%, the net loss in category return on inventory is just over 5% compared to the optimal allocation based on true estimates. This proves the robustness of these models to estimate errors. Similar to these shelf space allocation papers, but using an inventory
theoretic perspective, Urban [74] models the own and cross product effects of displayed inventory on demand rate in a mathematical program and solves for shelf space allocation and optimal order-up-to quantities. He reports that on average a greedy heuristic yields solutions that are within 1% of a solution obtained by genetic programming.

Yang [77] paper formulates a model and proposes an approach which is similar to the algorithm used for solving a Knapsack problem. Subject to given constraints, the proposed heuristic allocates shelf space item by item according to a descending order of sales profit for each item per display area or length. Through the use of simulations, the performances of objective value and the computational efficiency of this method are evaluated. Three options are also proposed for improving the heuristics. Compared to an optimal method, the improved heuristic is shown to be a very efficient algorithm which allocates shelf space at near optimal levels.

Reyes and Frazier [65] propose a shelf space allocation model that considers the tradeoff between profitability and customer service level. Their formulation is a nonlinear integer weighted goal program, which allows the manager to evaluate the tradeoff between profitability and customer service level to select the most preferred solution. An experiment is presented that assesses the usefulness of this tool over different parameter values. An alternate approach is also proposed, instead of using goal programming.

Martin et al. [49] proposes a differential game to study retailer’s allocation strategy of shelf space shares between the manufacturers of two competing brands. Each manufacturer can influence the allocation decision by her advertising spending to improve her brand’s goodwill which in turn affects the demand for her product. The game is played a Stackelberg with the
manufacturers as leaders and the retailer as follower. Stackelberg open-loop equilibrium is characterized and shown to be time-consistent.

A retail chain manager must draw on experience based on data available from his points of sale to diagnose space misallocation in stores and to make recommendations. Hwang et al. [38] addresses a problem of retailer who sells various brands of items through displaying on multi-level shelves. It is assumed that the level of shelf on which the product is displayed has a significant effect on sales. They developed an integrated mathematical model for the shelf space allocation problem and inventory-control problem with the objective of maximizing the retailer’s profit. Then, a gradient search heuristic and a genetic algorithm have been proposed for the solution to the model. The validity of the model has been illustrated with example problems.

Hwaang et al. [39] papers dealt with the shelf space design and allocation problem simultaneously considering location effect and space elasticity on demand. The demand rate was assumed as a function of the displayed inventory level and the location on which each brand of items has been displayed. Then the mathematical model was developed with the objective of maximizing retailer’s profit. Recognizing the difficulty in finding an optimal solution of the model, two solution procedures were proposed based on genetic algorithm. To examine the validity of the proposed solution procedures, thirty test problems were solved.

Recently, the progress of information technology makes retailer easily collect daily transaction data at very low cost. Through the point of sale (POS) system, a retailer store can collect a large volume of transaction data. From the huge transaction database, a great quantity of useful information can be extracted to support the retail management. Data mining is frequently
adopted to discover the valuable information from the huge database. In data mining, association rule mining is widely applied to market basket analysis or transactional data analysis. Chen et al. [53] proposes a data mining approach to make decision about which product to stock, how much shelf space allocated to the stocked products and where to display them. Association rules are generated by directly analyzing the transactional database, and these rules can be used to effectively resolve the product assortment and shelf space allocation problems. Their study applies the association instead of the space elasticity to formulate the mathematical model for product assortment. In their paper, multi-level association rules are generated to express the relationships between products and product categories to allocate the products selected in the assortment stage.

Most inventory problems in the real world involve multiple products. A very relevant example is retail stores, it is multi-billion industry and inventory costs form a significant part of it. Even in the multi-product scenario the single product case is able to capture the essential elements of the problem and is more intuitive. Abbot and Palekar [33] analyze the effect of interaction between products and respective shelf space allocations on their replenishment decisions in retail stores. They started their analysis of one product case to understand the dynamics of system and then extend it to multi-product case to include the cross effects of other products on the inventory decisions.

In their paper, Moncer et al. [51] propose an optimization model to determine the product assortment, inventory replenishment, display area and shelf space allocation decisions that jointly maximize the retailer’s profit under shelf space and backroom storage constraints. The variety of products to be displayed in the retail store, their display locations within the
store, their ordering quantities, and the allocated shelf space in each display area are considered as decision variables to be determined by the proposed integrated model. In the model formulation, they included the inventory costs, which are proportional to the average inventory, and storage and display costs as components of the inventory costs and make a clear distinction between showroom and backroom inventories. They also considered the effect of the display area location on the item demand. The developed model was a mixed integer non-linear program that was solved using LINGO software.

Nogales and Suarez [25] study is concerned with the space that national and private labels occupy in the shelves of stores. They explained this issue through a case study using a primary source of information: direct shelf observation. Their exploratory research concludes that the space allocated to private labels is clearly larger than that assigned to whole of brands on average. Other aspects analyzed are assortment, prices and promotions in order to establish their relationship with shelf space management.

Nawel and Zaccour [56] propose a game theoretic model in which one national-brand manufacturer, acting as a leader, maximizes her own profit and one retailer , selling the national brand and her private label and acting as a follower , maximizes her category profit. They characterize the resulting Stackelberg equilibrium in terms of the amount of shelf space allocated to these brands as well as their prices. The results suggest that the allocation of the shelf space depends on the quality of the private label. In their framework, quality is measured by the baseline sales, the degree of brand substitution and the price positioning.
3.1.1.4 PERCEPTION OF VARIETY
Consumer choice models often assume that customers are perfectly knowledgeable about their preferences and the product offerings. Therefore, consumers are always better off when they choose from a broader set of products. However, empirical studies show that consumer choice is affected by their perception of the variety level rather than the real variety level. This perception can be influenced by the space devoted to a category, the presence or absence of a favorite item (Broniarczyk et al. [12]), or the arrangement of the assortment (Simonson [68], Hoch et al. [35]) define a measure of the dissimilarity between product pairs as the count of attributes on which a product pair differs. They show that this measure is critical to the perception of variety of an assortment and that consumers are more satisfied with stores carrying those assortments perceived as offering high variety. Van Herpen and Pieters [75] find the impact of two attribute-based measures that significantly impact the perception of variety. These measures are entropy (whether all products have the same color or different colors) and dissociation between attributes (whether color and fabric choice across products are uncorrelated). The perception of variety at a store is especially important for variety-seeking consumers. Variety seeking consumers tend to switch away from the product consumed on the last occasion. On a final note, variety can even negatively affect consumer’s experience: confusion or complexity due to higher variety may cause dissatisfaction of consumers and decrease sales (Huffman and Kahn [37]).

3.1.2 RETAIL OPERATIONS
Assortment planning and shelf space allocation are important issues in retail, which we would expect to be based on a trade-off between transportation, shelf space, inventory costs and handling costs.
Currently, models that assess the overall operational costs in retail stores on these multiple dimension are not available. Existing research in retail operations mainly concentrates on inventory or marketing decisions separately. (e.g. Corstjens and Doyle [18]; Dreze et al [23]; Urban [74]; Cachon [14]; Hoare and Beasley [34]). Typically, in these models the handling time and its related costs are not considered explicitly. While warehouse handling operations received considerable attention in the literature (Rouwenhorst et al [66]; Tompkins et al [73]), there is still much opportunity for research in the field of store handling operations. The review by Rouwenhorst et al. [66] indicates that most of the research in warehousing is related to automated storage systems (AS/RS systems) and little research has been done discussing conventional warehouse (e.g. with manual picking). Only few researchers have reported on the difference in work pace in a warehouse environment (Bartholdi et al. [6]). Time study approaches are sometimes reported in the warehouse operations research for estimating order-picking times. Gray [29] uses basic multiple regression to derive estimations of the necessary time to pick all items from a pick list for a customer order, and applies it for establishing labor productivity standards. Gray et al [28] consider the general problem of warehouse design and operation, and propose a model in which order-picking time includes three components: walking, stopping and grabbing. However, in retailing literature time-studies are rarely reported. Recently, Van Zelst et al [76] showed that significant efficiency in terms of shelf stacking time could be gained once the impact of most important drivers is well understood. For most retailers, store handling operations are not only labour-intensive, but also very costly. Empirical study (Saghir and Jonson [67]) suggest that the handling costs in the retail chain represent the largest share of operational costs with high shares in the retail stores.