Shelf Space Allocation: A Critical Review and a Model with Price Changes and Adjustable Shelf Heights

SHELF SPACE ALLOCATION: A CRITICAL REVIEW AND A MODEL WITH PRICE CHANGES AND ADJUSTABLE SHELF HEIGHTS

BY

MEHMET ERDEM COSKUN, B.Sc.

A THESIS
SUBMITTED TO THE SCHOOL OF COMPUTATIONAL ENGINEERING & SCIENCE
AND THE SCHOOL OF GRADUATE STUDIES
OF MCMASTER UNIVERSITY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

Master of Applied Science (2012) McMaster University

(Computational Engineering & Science) Hamilton, Ontario, Canada

TITLE: Shelf Space Allocation: A Critical Review and a Model with Price Changes and Adjustable Shelf Heights

AUTHOR: Mehmet Erdem Coskun

B.Sc., (Industrial Engineering)

Yildiz Technical University, Istanbul, TURKEY

SUPERVISOR: Dr. Elkafi Hassini

NUMBER OF PAGES: xxiv, 137

This thesis is dedicated to my parents, who supported me all the way through my entire life with their love and patience and my future wife Derya Kacan, whose continuously growing love encourages me to achieve any task that I start.

Abstract

In today's retail environment, there are many consumer packaged goods (CPG) in the same category with various brands and differential products under the same brand. These differential products appear in different dimension sizes, display facing areas, purchasing costs and selling prices which are competing for a limited space in retail store shelves. Product assortment and space allocation of the chosen products to a limited shelf space is becoming more and more important for retailers. This is especially true for supermarkets that are struggling to make ends meet in a market with razor-thin profit margins. In this thesis we critically review the existing literature of shelf space allocation optimization models and solution techniques. We then propose a comprehensive model for shelf space allocation for a product category. Products are allocated to a two-dimensional area of a shelf section where a shelf section consists of multilevel vertical shelves. We account for adjustable shelf heights and product and brand integrity in a shelf section. Unlike the existing optimization models in the literature, we model our demand not only as a function of the space allocated to a product, in terms of the number of display facings, but also as a function of vertical product location in a shelf section and price sensitivity. Stackability of the products is also considered and products can be stacked depending on their package. Our objective is to maximize the retailer's daily gross profits. We numerically show that incorporating price changes and adjustable shelf spaces can have major impacts on the retailers profits. Finally, we provide directions and suggestions for future research in this growing area of research.

Acknowledgements

The existence of this thesis would not be possible without the help and guidance of my professor, supervisor, my mentor Dr. Elkafi Hassini. I would like to thank him for giving me the opportunity to be his student, offering me to work on this project and supporting my every step with patience and tranquility throughout my research. I have always felt his continuous support and guidance on my back and never doubted about my next step in my personal life as well as in my academic life. Therefore, I take this opportunity to show my sincere appreciation to thank him for encouraging me to accomplish my goals.

I also thank my parents, brother, friends and my future wife for their ongoing love and support.

Notation and abbreviations

Indices:

I	the set of all products in the product category
I^+	the set of products that are included into the product assortment
I^-	the set of products that are not included into the product assortment
i (j)	item index of the products that are included in the assortment ($i, j \in I^+$, note that for models that do not consider product selection $i, j \in I$ is used in order to simplify the notation in the model)
k (k')	item index of the products that are not included in the assortment ($k, k' \in I^-$)
m	shelf index of a shelf in the showroom inventory ($m =$ 1,2,....,M)
n	shelf index of a shelf in the backroom inventory ($n =$ 1,2,....,N)
e	shelf part index of a shelf m in the showroom inventory ($e =$ 1,2,....,E)
z	marketing variable index of the non-space marketing variables ($z =$ 1,2,....,Z)
b	brand index of different brands ($b =$ 1,2,....,B)
o	orientation index of a product i ($o = 1, 2, 3$)

t	time index of the period
t'	period index

Demand and Cost Parameters:

D_i (D_{im})	Demand function of product i (on shelf m) (unit/day)
Π	Total available potential demand of the product category (unit/day)
RD	Random Demand (sum of all RD_i's, $\sum_{i=1}^{I} RD_i$)
PD	Preference Demand ($PD = \Pi - RD$)
LD	Loyal Customer Demand
GD	Gained Demand (demand resulting from product switching within the assortment)
AD	Acquired Demand (demand resulting from product switching from unstocked products to stocked products)
U	Unmodified Preference ($U_i = \pi_i$ for product i) (unit/day)
UMD_i	Unmodified Demand of product i ($UMD_i = RD_i + LD_i$ in additive form, $RD_i\, LD_i$ in multiplicative form, $UMD_i = U_i = \pi_i$) (unit/day)
MD_i	Modified Demand of product i ($MD_i = UMD_i + GD_i = RD_i + LD_i + GD_i$ in additive form, $MD_i = UMD_i\, GD_i$ in multiplicative form) (unit/day)
AD_i	Acquired Demand of product i (unit/day)
$SOBD_i$	Stockout Benefit Demand of product i (unit/day)
$SOLD_i$	Stockout Loss Demand of product i (unit/day)
$MinD_i$	Minimum Daily Demand of product i (Lower Bound-Limit Demand) in units of sales ($= LD_i$) (unit/day)

UD_i	Maximum daily demand of product i (Upper Bound-Limit Demand) in units of sales (unit/day)
ND_i	Natural Demand of product i (demand before in-store support, very similar to π_i, however, π_i reflects the demand when all marketing mix variables are fixed, e.g. when product i has a fixed price and a base level shelf allocation) (unit/day)
ψ_i	parameter for product i's portion of preference demand that is switching preference demand
τ_i	parameter for product i's portion of preference demand
θ_k (θ_j)	parameter that reflects the resistance to compromise customer's original purchase decision from an unstocked (temporary stockout) product $k(j)$ to stocked products in the assortment
Δ_{ij}	a probability ratio of switching from a stocked product i to another stocked product j ($\theta_i + \sum_{\substack{j \in I^+ \\ j \neq i}} \Delta_{ij} = 1$)
δ_{iom}	shelf location-orientation quality-adjustment weight corresponding to display facing area allocated to product i in orientation o on shelf m
ξ	parameter that reflects the range flexibility from the base level shelf space allocation allowed by the decision maker (ranges between $0-1$)
$\zeta_{\bullet i}$	parameters of the quadratic function for product i where $\bullet = 1, 2, 3, ...$
Ω_{ib}	parameter that transforms the purchase unit of product i under brand b into its display unit
Ω'_{ib}	parameter that transforms the purchase unit of product i under brand b into its sales units ($\Omega_{ib} \geq \Omega'_{ib}$)

adv	total advertising level for the product category
di	total distribution level for the product category
$c_i(c_{ib})$	unit cost (purchasing price for the retailer) of product i (under brand b)
C_i	Cost function for product i (\$/day)
P_i	Gross profit function for product i (\$/day)
NP (NP_i)	Net Profit function (for product i) (\$/day)
$MAD_{solution}$	Mean Absolute Deviation for the solution
MAD_{best} (MAD_{worst})	Best (Worst) MAD value achieved when $P_{solution} = P_{worst}$ ($P_{solution} = P_{best}$)
$MAD\%_{solution}$	Percent of MAD value achieved by the solution
$P_{solution}$	Gross Profit achieved by the solution
P_{best} (P_{worst})	Best (Worst) Gross Profit achieved when $MAD = MAD_{worst}$ ($MAD = MAD_{best}$)
$P\%_{solution}$	Percent of Gross Profit achieved by the solution

Elasticity Parameters:

α_i	direct-space elasticity of product i, $0 \leqslant \alpha_i \leqslant 1$
β_{ij}	cross-space elasticity of product i to product j, $-1 \leqslant \beta_{ij} \leqslant 1$ for all $i \neq j$, $\beta_{ii} = \alpha_i$
σ_i	direct-price elasticity of product i, $\sigma_i \leqslant 0$
μ_{ij}	cross-price elasticity of product i to product j , $0 \leqslant \mu_{ij}$ for all $i \neq j$, $\mu_{ii} = \sigma_i$

σ'_{ib}	direct-price-change elasticity of product i under brand b as compared to the previous period, $\sigma'_{ib} \leqslant 0$
$\mu'_{ib,j}$	cross-price-change elasticity of product i under brand b as compared to the previous period, $-1 \leqslant \mu'_{ib,j} \leqslant 1$ for all $i \neq j$, $\mu'_{ib,i} = \sigma'_{ib}$
γ_{zi}	the elasticity of product i relative to z-th marketing variable
λ_i	operating cost elasticity associated with increased sales of product i
ϱ_i	own-advertising elasticity of product i
ϱ'_{ij}	cross-advertising elasticity of product i to product j
ς_i	own-distribution elasticity of product i
ς'_{ij}	cross-distribution elasticity of product i to product j
ϵ	elasticity parameter that reflects the marketing carryover effect

Scale Parameters:

π'_i	scale parameter of product i ($= LD_i$) (unit/day)
π''_j	scale parameter of product j ($\prod_{j=1, j\neq i}^{I} \pi''_j = RD_i$) (unit/day)
π_i	scale parameter of product i (potential demand of product i, market share-strength of product i without in-store support, base level demand rate of product i when base level shelf space allocation is used) (unit/day)
ϕ_i	cost scale parameter of product i (DPC_i is used by Bookbinder and Zarour, 2001)
φ_m	location scale parameter that reflects the increase of the demand rate with respect to the shelf level when products are displayed on shelf m

φ_i — weighted average of location scale parameter of product i when product i is displayed on more than one shelf

Product Parameters:

w_i — width of product i in showroom orientation

h_i — height of product i in showroom orientation

d_i — depth of product i in showroom orientation

wb_i — width of product i in backroom orientation

hb_i — height of product i in backroom orientation

db_i — depth of product i in backroom orientation

$\bar{y}_{ib}$ — parameter representing whether product i under brand b is stackable

p_i — unit sales price of product i

p^{avg} — average sales price of products in the product category

${p_i}^{suggested}$ — suggested sales price of product i by the manufacturer or supplier

g_i — gross margin of product i ($p_i - c_i$, contribution to profit generated by a unit of product i; $DPPR_i$ is used by Bookbinder and Zarour 2001)

$block_i$ — number of product i in a horizontal-vertical block (stack) if the products come in multi-packed cases such as twin-packed paper towels. In Zufryden (1986), the $\bar{o}_i$ will be in incremental block units of $block_i$

u_i — the amount of space required per unit of product i (cm, mm, facing, package, case, cm^3)

v_i (v_{im}, v_{ime}) — per facing profit (value) of product i (on shelf m, on part e of shelf m, respectively)

$deviance$	a limit distance from one shelf to its adjacent shelf
f_i	variable(s) that may affect the demand of product i (space, retail price, advertising, promotions, store characteristics)
f_{zi}	variables (where $z = 1, 2, ..., Z$) that may affect the demand of product i other than space (retail price, advertising, promotions, store characteristics)
$weight$	weight of profit objective in the objective function given by the decision maker (range between $0 - 1$)
$1 - weight$	weight of under base level allocation objective in the objective function given by the decision maker
P_i^{min} (P_i^{max})	minimum (maximum) selling price that can be assigned to product i which is decided by the retailer or determined by the market
$PC_{ib}^{negative}$ ($PC_{ib}^{positive}$)	maximum allowable percentage of price increase (decrease) of product i compared to the previous period which is decided by the retailer

Shelf Parameters:

L (L_m, L_{me})	total width of the available shelf space (shelf m, part e of shelf m, respectively) in showroom inventory
H (H_m)	height of the available shelf space (shelf m) in showroom inventory
DE (DE_m)	depth of available shelf space (shelf m) in showroom inventory
BL	total width of available shelf space in backroom inventory
BH	total height of available shelf space in backroom inventory

BDE	total depth of available shelf space in backroom inventory
TH	Total Height of the shelf section of the product category ($TH = \sum_{m=1}^{M} H_m$)
H^{min} (H^{max})	minimum (maximum) height of a shelf which is decided by the retailer
L_i^{min} (L_i^{max})	minimum (maximum) amount of shelf space to be allocated to product i, in terms of total width-length, which is decided by the retailer.
X_i^{min} (X_i^{max})	minimum (maximum) quantity of product i to be allocated in showroom inventory which is decided by the retailer.
X (X_m)	available shelf space in showroom inventory, in terms of total number of products, which can be allocated to the shelf (shelf m).
BX	available shelf space in backroom inventory, in terms of total number of products, which can be allocated to the shelf.
O	total number of slots to be allocated to the products in showroom inventory
O_i^{min} (O_i^{max})	minimum (maximum) number of slots to be allocated to product i in showroom inventory

Inventory Parameters:

Q^{min} (Q^{max})	minimum (maximum) quantity of product i to be ordered from the suppliers due to production, order size or procurement constraints
LT_i	Lead Time of product i (days)
II_i	Initial Inventory of product i (unit)
$IN_{im}(t)$	displayed Inventory of product i on shelf m at time t

AI_i (AI_{im})	Average Inventory of product i (on shelf m) (unit)
IR	Investment Rate of the products (%)
CT_i (CT_{im}, CT_b)	Cycle Time of product i (on shelf m, joint replenishment of all products under brand b, respectively) (days)
CT_i^{max}	Maximum Cycle Time of product i (expiration date of the product) (days)
OC	Order (Procurement) Cost - fixed ($/order)
OC_b	Order Cost of joint replenishment of all products under brand b ($/order)
HC_i	Holding (Inventory Investment) Cost of product i based on the inventory level (Average inventory level of the showroom and backroom inventories will be used to simplify some of the models, e.g., see Urban (1998) and Hariga et al. (2007), in which demand is a function of display inventory level, not the instantaneous inventory level) ($/day)
SC_i	Storage Cost of product i based on the allocated space in backroom inventory ($/day)
DC_i (DC_{im})	Display Cost of product i (on shelf m) based on the allocated space in showroom inventory (= storage cost for showroom inventory) ($/day)
GC_i	Guarantee Cost of product i (costs incurred by insurance, deterioration and returning) ($/day)
IC_i	Insertion Cost of product i to include it into the assortment - fixed ($)

RC_i	Replacement (Rearranging, Restacking) Cost of product i each time product i is rearranged on the shelves at the beginning of a new period (\$/order)
ET_{ib}	Expiration Time (holding time limitation) of product i (in days)
WD	number of Working Days in a period
BU	available Budget for purchasing products in a period
NR	maximum allowable Number of Replenishments for all different types of products in a period which is decided by the retailer
NR_b	maximum allowable Number of Replenishments for brand b from the manufacturer or supplier
NF_i (NF_{ib}, NB_{ib})	total number of product i (under brand b in showroom inventory, under brand b in backroom inventory, respectively) that can be placed in one facing in case of a fixed shelf height (in vertical and depth column)
DL_i	Desired target level of product i placed on the width-length of the showroom inventory
nDL_i (pDL_i)	negative (positive) deviations from DL_i
DB_i	Desired target level of product i placed on the width-length of the backroom inventory
nDB_i (pDB_i)	negative (positive) deviations from DB_i
DPM	Desired target level placed on Profit Margin by the retailer
$nDPM$ ($pDPM$)	negative (positive) deviations from DPM

p_s	a penalty for nonproductive use of space
p_m	a penalty for negative deviation from DPM
p_{si}	a penalty for negative and positive deviations from DS_i
p_{bi}	a penalty for negative and positive deviations from DB_i
SVS_{ib}	slack variable
TCP	Total Costs of Purchasing per period
TCO	Total Cost of Joint Replenishment (Order Costs for all brands) per period
TCH	Total average Cost of Holding (inventory investment) per period
$TCSP$	Total penalty in a period due to nonproductive use of space
TC	Total Cost per period
TR	Total Revenue per period

Decision and Consequence Variables:

x_i (x_{im}, x_{ime}, x_{iom}, x_{ib}, $x_{ib,m}$)	number of display facings of product i (on shelf m, on part e of shelf m, in orientatation o on shelf m, under family b or brand b, under family b on shelf m, respectively) to be allocated to the width-length of the shelf in showroom inventory
x_i^{base}	number of display facings of product i when product i is in base level allocation ($x_i^{base} = \frac{\pi_i}{\Pi} X$)
$\bar{x}_{ib,m}$	number of display facings of product i from brand b to be allocated to the height-length ($\lfloor \frac{H_m}{h_i} \rfloor$) of shelf m in showroom inventory
x'_i (x'_{ib})	number of product i (under brand b) to be allocated to the width-length of the shelf in backroom inventory

$\bar{x}'_i$	number of product i to be allocated to the height-length $(\lfloor \frac{H_n}{h_i} \rfloor)$ in backroom inventory
$\bar{o}_i$	number of slots to be allocated to product i (area allocation)
l_i (l_{ib})	shelf space allocation of product i (under brand b), in terms of total width-length, associated with the number of display facings in showroom inventory
$\bar{l}_i$ (l_{ib})	shelf space allocation of product i (under brand b), in terms of total height-length $(\lfloor \frac{H_m}{h_i} \rfloor h_i)$, in showroom inventory
l'_i (l'_{ib})	shelf space allocation of product i (under brand b), in terms of total width-length, in backroom inventory
$\bar{l}'_i$ ($\bar{l}'_{ib}$)	shelf space allocation of product i (under brand b), in terms of total height-length $(\lfloor \frac{H_n}{h_i} \rfloor h_i)$, in backroom inventory
a_i (a_{ib}, a_{iom})	shelf space allocation of product i (under brand b, in orientation o on shelf m, respectively), in terms of total area allocation, associated with the number of display facings in showroom inventory
$\bar{c}_i$ ($\bar{c}_b$)	center of the display facings of product i (under family-brand b) on a shelf in horizontal dimension (e.g. the distance from the total facing area center to the left end of the shelf)
PC_{ib}	percentage of price change for product i under brand b as compared with the previous period
$PC_{ib,j}$	percentage of price change for product i under brand b's substitute products (as compared with the previous period or average-suggested price by the market)

q_i (q_{im}, q_{ib})	order quantity of product i (on shelf m, under brand b, respectively)
r_i	reorder point of product i (quantity of products in inventory)
y_i	binary variable representing whether product i is in the assortment
y_{im}	binary variable representing whether product i is located on shelf m ($y_i = 1$, when the product is already in the assortment)
y_{ime}	binary variable representing whether product i is located on part e of shelf m
y_{bm}	binary variable representing whether family b is located on shelf m
$y_{ib,m}$	binary variable representing whether product i under family-brand b is located on shelf m
y'_{ij}	binary variable representing whether product i is located to left of product j
$left_{bm}$	continuous variable representing the left end (coordinate) of family-brand b on shelf m
$right_{bm}$	continuous variable representing the right end (coordinate) of family-brand b on shelf m
top_i (top_{ib})	integer variable representing the top shelf on which product i (under family-brand b) is located
$bottom_i$ ($bottom_{ib}$)	integer variable representing the bottom shelf on which product i (under family-brand b) is located
uv_i	under-achievement goal variable for product i allocation
ov_i	over-achievement goal variable for product i allocation

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

Shelf space is one of the most important scarce resources that a retailer has to work with. The importance stems from the fact that the shelf space provides the display facing area of the products to the public. Therefore, the wise management of this very limited space is crucial.

The average number of new products, or so called ”store keeping units” (SKUs), in a supermarket have increased by %20 in the 1970s and %75 in 1980s (Greenhouse 2005). This increase put a lot of pressure on the limited, and often fixed, amount of space in the retailers facilities. Therefore, the questions (1) which product to include in the assortment, (2) how much space to allocate and (3) at what price to sell, become the most important challenges facing a retailer. Applying cost minimization techniques to increase profitability may not take the retailer one step ahead against its competitors, where the average net profit after taxes is around 1%, although the average gross margin can be up to 28% (Reyes and Fraizer 2007). Attracting customers to switch stores through promotions, better product category placements within the store and better product assortment and shelf space allocation might be the only strategy in order to increase overall demand and profitability.

In todays retail environment there are many consumer packaged goods (CPG) in the same category with various brands and differential products under the same brand. These differential products appear in different dimension sizes, display facing areas, purchasing costs and selling prices which are competing for a limited space in retail store shelves. Similar to how products' sales prices influence the customers' purchasing decisions in the stores; product assortment, product display facing areas, number of facings of the product and location of the products on the shelves also influence the customers' in-store buying decision (Dreze et al. 1994). While consumers would prefer to purchase a product by paying less rather than paying more, they would not necessarily purchase the cheapest product from a retail store assortment because of the other marketing variables mentioned above and factors such as loyalty to a particular brand, desire to try different things, perceived quality of the product etc. (Reyes and Frazier 2007).

Many shoppers decide on what to purchase inside the store; this rate could be as high as nine out of ten (e.g., see Silvera et al. 2008). Thus, it is expected that different shelf space allocations and displays will lead to different sales. Especially when shoppers do not receive aid in their shelf "browsing" except perhaps for high-value items such as electronics and beauty products. Given the importance of packaging and the information it conveys (color, text, image, texture, etc...) and that it is perhaps the only information that is readily accessible to all self-serve shoppers (as opposed to other advertising media), it is crucial that such a function is used properly by retailers in the stores. In effect packaging serves as the "silent salesman" (Dichter 1957) in the absence of a real salesperson in the world of self-service retail.

Through the utilization of information technologies, the collection of data in retail stores became very efficient. With the help of point of sale (POS) systems, retail stores can now easily collect transactions data from consumers to know what kind of brand, which exact item, when and from which store (among their retail store chains) they purchased. Using this transactions data, they are able to analyze and forecast future demand, inter-relation between brands and product or products with product categories and estimate the parameters for space allocation. This data availability makes shelf space allocation decision making more reliable and provides an incentive for management to use the data as a way of getting a return on their information technology infrastructure investments.

1.2 Thesis Focus

In general, the decision of shelf space allocation can be divided into two levels: the retailer should decide (1) how much shelf space to allocate to a product category and (2) how much shelf space to allocate to a product within each product category. In the shelf space allocation literature, we generally work on the optimization of a shelf space allocation problem within a product category with a limited shelf space or multiple shelves allocated to that category. The main concern of the problem is to determine which products to display on shelves (it is referred to as "product selection" or "product assortment" in the literature) and how much shelf space will be allocated to each selected product. In shelf space allocation literature, product assortment and shelf space allocation decisions are generally solved simultaneously for a certain time period. As Borin et al. (1994) noted, this is a routine decision-making process.

Brown and Tucker (1961) list three product classes: "Unresponsive Products," such as salt and spices, which are considered as inelastic, "General Use Products," such as canned foods and cereals, which are consumed on a daily basis and "Occasional Purchase Products," such as sardines and nuts, which are not the generally purchased products and not looked for at first when shopping. This last class of products is slow in responding to shelf space changes until the display facing area is large enough to get the attention of the consumers. At that point, the sales curve rises steeply. In the literature this phenomenon is called the "threshold effect". Although there is no strict distinction, unresponsive and general use products are considered as staple products and occasional purchase products are considered as impulse products (Curhan 1973).

Cox (1970) found that there is no relation between the amount of shelf space given to a staple product and the total unit of sales of that product. Therefore, we do not worry about the staple product assortments and allocation. As long as we allocate the minimum amount of shelf space needed to a staple product based upon past sales history (Anonymous 1960), the amount of space allocated to that product will not impact consumers demand for that product. Thus in shelf space allocation studies, our focus is on impulse products in which the consumers make their decisions in a retail store in front of the display shelves.

Another important conclusion from the literature is that the changes in shelf space effects are more in large retail stores than in small retail stores (Frank and Massy 1970). Thus, in our thesis we focus on retailers with large retail surfaces, such as supermarkets and department stores.

1.3 Shelf Space Allocation Approaches

The early studies implied that the shelf space should be allocated in proportion to sales (e.g. see Progressive Grocer publications in 1951, 1955, 1958, 1960). However, to allocate shelf space according to the market share of the product (or in proportion to product's market strength, which implicitly assumes that all products' prices are fixed within the same category) is often considered as a trivial solution to overcome the shelf space allocation problem (Buttle 1984b; Yang 2001), as such an approach neglects space and price elasticity.

Another approach to shelf space allocation is direct product profitability (DPP). The aim of DPP was to calculate standard costs continuously and use them to determine the end products. Thus through DPP, we can calculate the unit contribution of products to overhead and profits and retailers could accordingly reallocate shelf space to maximize their return on investment. Many results supported the fact that profits can be increased by reallocating products and shifting their spaces from low profit products to high profit products (Curhan 1973). Of course, these increases were dependent on the location of the space allocated to that product and space elasticities of the products in the assortment. But DPP can not show the effects due to space changes or the gains because of the increased sales from replacing a product with another product (Bookbinder and Zarour 2001). Since it has been assumed that unit handling costs are stable against the reallocation changes, total profit would increase only by the increase of the products sold because of the different space elasticities of the new products included in the assortment (McKinsey-General Foods Study 1963). However, a more practical model should include these cost implications (Corstjens and Doyle 1981).

Cosmos 1 package used a rule of thumb, which has been proposed by Buzzell et al. (1965), to reallocate the space by taking the space from the least profitable products and assigning them to the most profitable products (Hansen and Heinsbroek 1979). But there were no product selection in the package.

All of these early studies have the unrealistic assumptions that products have uniform space elasticities on demand and zero cross elasticities between products (Corstjen and Doyle 1981).

1.4 Computer Packages

A planogram is a traditional shelf space management tool which uses heuristics for shelf space display facings allocation. Since the shelf space allocation problem is complex and planograms often do not consider the demand parameters of retailer stores and only simple heuristics rules have been developed, planograms are not sufficient for globally optimizing shelf space (Zufryden 1986; Desmet and Renaudin 1998; Yang 2001; Lim et al. 2004). PC-based tools such as Apollo (IRI) and Spaceman (Nielsen) use historical product sales, gross profit margins and turnover to allocate shelf space and use handling and inventory costs for the constraints of the problem (Lim et al. 2004) in order to reduce the time spent to change space allocation manually, however, they are still generally used for planogram accounting (Dreze et al. 1994; Yang 2001). Although they help retailers to visualize alternative shelf space designs and show the shelf location of the products in the assortment, they are not used for optimization purposes. Galaxi (Space Solutions) is another planogram, which is used by Tesco in U.K., does not optimize the shelf space allocation but uses manual drag-and-drop procedures (Bai 2003; Lim et al. 2004).

1.5 Suppliers Perspective

In a market which generally consists of loyal customer demand, the changes of shelf space or any other marketing variables would hardly affect products' demands. In contrast, in a market which generally consists of random customer demand, the changes of shelf space or other marketing variables for a product would affect products' demand easily. Although retailers' purpose is to manipulate customers' purchasing preference using in-store support such as space allocation and change his purchasing decision to more profitable products within their assortment in order to maximize their profit, the manufacturers do not want to adhere to the retailers' space allocations or any other in-store marketing variables. Manufacturers want to create strong relationship between their products and their customers through product differentiation. The manufacturer's aim is to maximize his own profit using his product diversity and their shelf space allocation. Therefore, in this reciprocal relationship between the retailer and the manufacturer, the manufacturer wants to create customer brand-loyalty in order to have more shelf space in retail shelves while the retailer wants to lessen these brand loyalties using in-store support such as space allocation, pricing and in-store promotions so as to maximize his own profit. Furthermore, retailers are also increasingly using the same support system to built up customer loyalties for their own store brands.

To guarantee a shelf space suppliers buy (rent) space from the retailer. The retailer would want to sell the shelf space to the supplier that can offer a price for a unit space which exceeds the opportunity cost of this space. This opportunity cost is the most gross profit a retailer can get by allocating the products into the shelves the best way he can to maximize his profit. Otherwise, the retailer would instead proceed with his own product assortment. Suppliers' interest in renting shelf space is driven by their desire to expose their products to potential customers in the store so as to increase their profits. For suppliers, the location and the amount of the shelf space is very important (Cairns 1962). The sales of a product on the shelves depend not only on the price, but also on determinants such as the number of facings, total display facing areas and the location of the shelf space allocated to that product.

If the manufacturer's product has a high demand on the market, then the retailer may not have the choice of not carrying the product even if its profit margin is so low. Since otherwise the absence of this product may cause loyal customers to switch to other retail stores. Furthermore, if the demand is high enough, the retailer might consider a priority shelf space allocation policy for those products that reserves for them a better location and more amount of space. The manufacturer may still rent the retailer's space, in spite of a zero gross profit, due to the presence of the manufacturer's competitor products in the assortment (Cairns 1962, 1963).

1.6 Private Brands

In today's retailer environment, most of the retailers have their own "private brand" (store brand) along with a "national brands". Since these private brands have a tendency to have a large amount of shelf space in their own retail stores, it is likely that they will have their own promotions and best location on the shelves. As a result, the retailer should balance the gross profit per unit of space coming from their private brand with the gross profit per unit of space coming from the national brand (Cairns 1962). Curhan (1973) also found that private brands have a higher elasticity than national brands. Nogales and Suarez (2005) investigated the effects of private brands and promotions on shelf space allocation in a case study.

1.7 Shelf Level and Store Location

Although Frank and Massy (1970) found that there is no significant change in sales variation with shelf levels, retailers believe that shelves closer to the eye level would have a greater effect than the shelves above and below eye level (Curhan 1973). Dreze et al. (1994)'s experiments showed that the vertical location of the products (the shelf level on which the product is displayed in a multi-level shelf environment) had a great impact on products' sales even more important than changes in the number of facings allocated to those products as long as a threshold was maintained. On average, the sales increased 39% from the worst vertical shelf to the best vertical shelf and additionally 15% from worst horizontal location to best horizontal location. The sales changed almost an average of 59% from the worst positioning to the best positioning with the combined effects of both. Furthermore, Underhilll (1999) noted that the zone between slightly above eye level and knee level is a reliable zone for products to be noticed. The type of consumer to whom the retailer is selling the products also plays an important role in the shelf space allocation. If the target audience are children then the products, such as toys, sweets, chocolates etc., should be located on the lower levels of the shelves which are more visible and accessible to children in order to attract their purchasing preference (Silva et al. 2009).

Underhill (1999) found that products at the end of aisles are more visible and noticeable to customers because they do not tend to walk the whole aisle completely in order to make a choice. He called this phenomenon the "boomerang effect". Dreze et al. (1994) found that half of the category sales at the end of the aisles increased.

Larson et al. (2005)'s study had similar results where customers tend to choose the aisles for shopping instead of entering all the aisles in the store and have a tendency to enter and exit aisles instead of walking the whole aisle from one end to another end.

1.8 How Much to Stock on the Displaying Shelves?

Larson and DeMarais (1990) described the stimulus effect of the display facings on customers as "psychic stock" where they found, through experimental studies, that a retailer can achieve better sales by keeping the display shelves fully stocked. They suggested a policy that keeps the shelves full all the time and called it "full-shelf merchandising". However, the effects of out-of-stock (temporary) or decreasing number of display facings have not been studied in the literature. Although in some product categories decreasing the number of display facings might decrease the demand of that product, we can observe the reverse effect in other product categories.

1.9 Life Cycles

Most of the products in today's environment have a life cycle in the market and because of the consumers' changing desires, manufacturers introduce new products or a modified version of the existing products with stylish and colorful display facings. At the end of the product's life cycle, the product demand decreases. Thus retailers should also take these dynamics into account in their product assortments and shelf space allocations (Cortsjen and Doyle 1983). Fresh products, such as produce, dairy and meat, have a short shelf life that continuously deteriorates through time where the freshness of the product decays (Bai et al. 2008). In contrast to existing models in the literature, the demand rate of such products should not be assumed constant between replenishments. Some of the inventory models in the literature (Nahmias 1982; Raafat 1991; Goyal and Giri 2001) modeled the demand for these kind of products using a fixed deteriorating rate, but with non-decaying utilities between replenishments (before the expiration date of the product). The assumption of non-decaying utilities is unrealistic because the freshness is a quality factor that affects the product's demand (Bai et al. 2008). Besides, the cost of shelf space allocation of such fresh products is significantly higher than that of durable goods due to low temperature requirements to increase their freshness time. Additionally, due to the lack of durability of fresh products, the inventory replenishments of such products will be more frequent compared to durable products. Moreover, Kar et al. (2001) observed that in the developing countries fresh products are divided into different categories depending on their freshness and generally exhibited in different store types.

1.10 Space Elasticity

Space elasticity, α_i, can be defined as "the ratio of relative change in unit sales to relative change in shelf space" (Curhan 1973). It means that the amount of shelf space assigned to a product affects its sales per unit of space which is referred to as "unit sales". Therefore, we say the unit sales of the products are space elastic. When the unit sales increase as shelf space increases and vice versa, space elasticity should be positive. However, researchers such as Bultez and Naert (1988) and Dreze et al. (1994) found that as space allocated to a product increases, marginal returns will first increase then decrease in a S-shaped curve for $\alpha_i > 1$ (the increasing case would occur only for a very small number of display facings, such as one facing, but it would be very difficult to identify such an S-shape) or strictly concave for $0 < \alpha_i < 1$. However, it is generally assumed that an increase in the number of display facings will only show diminishing returns on the demand where $0 \leqslant \alpha_i \leqslant 1$. Lee (1959) and Bates (1970) hypothesized that as shelf space increases, unit sales will increase at a decreasing rate until maximum sales level is reached. For example, assigning two display facings of a product instead of one display facing, might double the sales. However, an increase in the number of display facings from two to three would generally result in less increase than from one to two. This especially true for impulse products (Brown and Tucker 1961), the subject of our research.

Generally, with a finite product quantity, as the number of display facings increases, the demand is an increasing concave function which reaches to a maximum beyond which increasing the number of facings will not increase the sales anymore. For inelastic products, the quantity display of staple products such as salt and pepper does not affect the product's demand as long as a threshold is maintained. In practice, however, it has been assumed that there is a linear relationship between shelf space and unit sales for all products (Cox 1964).

In the literature there have been different estimates for space elasticities, depending on whether it is used at a product or store level. In Curhan (1973), the space elasticity for all products are averaged at 0.212 and unit sales showed a 8% positive change in the same direction of shelf space changes. Hansen and Heinsbroek (1979) used a value of 0.15 on average at the item level and Thurik (1988) used a value of 0.6 on average at the store level. Corsjen and Doyle (1981) used 0.086. Desmet and Renaudin (1998) found that the space elasticity ranged between negative values and

0.8 for different product categories. They model the demand function using the market share (π) of a product and suggested that if the buying rate of a product category increases (an increase in the overall category demand) then the space elasticities of the products in that product category would increase.

1.11 Product Interdependency

As well as the sales of one product affects the sales of another product in the same category, the sales of one product may affect the sales of another product in another category (such as a camera and batteries). Therefore, retail store managers should consider another key concept of the relationship between two products which is called interdependency. There are two kinds of interdependencies: substitution and complementarity. When a customer can not find the product he is looking for on the shelves, he would substitute his choice with another product from the assortment under certain conditions. This is called customer-based stock-out substitution. However, if one product is sold and another product is purchased along with the main product, this is a complementary effect. These are defined as cross elasticities between products and we may see these effects in inter-groups and intra-groups. For example, in a shaving category, a sale increase of a particular safety razor brand may increase the sales of shaving lotions but it may decrease the sales of another brand of a safety razor and electric razors, and vice versa (Urban 1969).

Most of the studies show that in many cases, customers are not willing to purchase any product if their preferred product is not within the displayed showroom. In other words, these loyal customers do not want to substitute their preferred products. In some cases up to 40% choose to not substitute (Borin et al. 1994). It is important to note here that, in fact, premium products have much more resistance to compromise customer preference than low-end products. Therefore, it is more likely to substitute for low-end products than for premium products since premium products have a higher brand loyalty (Fadiloglu et al. 2010). In our research we use the notation θ_k to indicate the ratio of such loyal customer demand for an unstocked product k. The ratio θ_k is higher for premium products than for low-end products. Furthermore, as this ratio decreases for a product according with its low profit margin, the product is likely to be eliminated from the product assortment since a significant amount of that product's demand can be directed to the other products within the assortment (Fadiloglu et al. 2010).

We should note here that it has been implicitly assumed, in the literature and retail practice, that a customer who prefers a product which is not in the store's product assortment will substitute by a product from the assortment depending on his loyalty. However, if the product assortment of the retail store does not satisfy the customers' product variety perception, and thereby not trigger their purchasing instinct, then not only loyal customers but also a portion of the customers who are willing to switch would be lost. This kind of (lost) demand has been neglected in the literature. Given the focus on satisfying the consumer demand on highly demanded products and include them into the product assortment, a constant overall category demand has been assumed for all periods and lost customers are neglected in the long run. This effect can be neglected in short run decisions, however, it should be considered in the long run in a dynamic model.

1.12 Cross Effects in Products' Demand

Starting from the 1970's, academic research has focused on how to model the demand function in shelf space allocation literature. Given the marketing findings that the amount of space dedicated to a product in a retail store (display area in the show-room inventory) affects the demand of the product, space elastic functions were used to model demand. Although most of studies took into account product space elasticities, only some of them took into account inter-elasticities between products. As we mentioned above, the own (direct) space elasticity of a product measures the effect of increasing or decreasing the number of display facings on demand for a single product. An increase of a display facing of a product affects the demand of that product positively. On the other hand, an increase of that product's demand will affect most of the other products' demand in the same category negatively, such as the same products under different brands. However, the increase in the product's demand can affect some of the products' demand in the same category positively, such as shaving razors and lotions. Similarly, an increase for a product's demand may also affect another product's demand positively in some other product category such as cameras and batteries, an intra-group effect.

Cross-space elasticity of a product measures these substitute or complementary effects between two products. We note that the cross elasticity of a product to another product (product i to product j) is not equal to the cross elasticity of the direct opposites (product j to product i). For example, an increase in camera sales may affect the battery sales positively but not vice versa ($\beta_{ij} \neq \beta_{ji}$). Furthermore, since the relation between products can be positive (supplementary), negative (substitute) or neutral, the cross elasticity values can take positive values, negative values or 0, respectively.

Bultez and Naert (1988) stated that the cross elasticities between products would not occur if there was enough shelf space. The existence of a limited shelf space causes interdependencies between products to occur because an increasing space of a product causes a decreasing space for the other products in the assortment. If there was no limitation on the available shelf space area or if there was enough shelf space corresponding to the maximum demand that could be reached by the space increments of each product in the assortment, these interdependencies would not occur, since increasing space of a product would not increase its demand after a threshold has been reached. Therefore, if there is enough space to fulfill the threshold demand for each product and to allocate the available space corresponding to this maximal space of each product, then there would not be any interdependencies (space) between products within the assortment. An important conclusion from Anderson (1979)'s work is that, if the products are perfect substitutes, which means that they satisfy the customers equally with no financial incentives, then the retailer can encourage the customers to purchase the high-profit-margin product by allocating more space to that product, while discouraging the low-profit-margin products.

1.13 Knapsack Model in Shelf Space Allocation Models

Due to the highly non-linear nature of the space elasticities, all the models that included these effects are complex and generally hard to solve. Furthermore, since the estimation of own-space elasticity and cross-space elasticity parameters are difficult, many studies chose to approximate the space allocation problem with a special version of the simple knapsack problem: maximizing the benefit of including items in a knapsack while not exceeding the knapsack's capacity. The space allocation problem is more general than the knapsack problem in that it may have additional constraints to the capacity constraint and the latter may be represented by multiple constraints each for a different section of the shelf.

Consequently, shelf space allocation problem can also be considered as a modified version of a multi-dimensional knapsack problem where the goal is to choose the most profitable products and allocate them into a scarce space to gain the most profit from the customers. In this formulation, the cross elasticity between two products are neglected due to the difficulty of estimating these elasticity parameters. Furthermore, the nature of own-space elasticity is assumed linear for an interval of a certain number of display facings (Yang and Chen, 1999; Yang, 2001; Lim et al. , 2004). We know that the products which are less elastic to space changes such as staple products will result in a diminishing return in demand earlier than the products which are more elastic to space changes, however, as Yang and Chen (1999) mentioned, the concave nature of this space elasticity effect can be controlled in a certain interval of display facing quantities and the effect of the space elasticity could be assumed as a linear in this interval of display facings.

We emphasize that the effects of cross elasticities had to be ignored when modeling the problem as a knapsack problem because the nature of the knapsack problem is unable to define and observe the opportunity cost of a substitute or a complementary product when a product is allocated to a shelf. Cross elasticity factors in shelf space allocation problems using the nonlinear demand function models take all these inter-relations between two products into account and measures the substitute or complimentary effects one by one for the entire product set.

1.14 Multi Display Locations Within the Store

During promotional activities in retail stores, some products are displayed in more than one location within the store such as store entrances, shelf endings, checkout points, and particular places in the store where store management can stack products. This type of marketing technique stimulates a threshold effect on many consumers who do not have any preferences on that product category. Also, some products may be displayed in more than one shelf within the store because the products could be a part of different categories at the same time and to satisfy all consumers various shopping behavior, store management should premeditate this phenomenon. For example, 100% fresh juices can be categorized under soft drinks as well as they can be categorized under natural products. Some retail stores even place these kind of soft drinks right behind the fresh fruits category to emphasize the naturality of these products. Another good example is alcoholic drinks which can be grouped under manufacturers, distributors or price sets (Russell and Urban 2010). In our thesis we focus on products that are likely to be displayed in one single location within the store.

1.15 Thesis Structure

The remainder of this thesis is as structured. In this chapter we discuss the most related issues to shelf space allocation problem. We define the basic concepts and components of the problem. In the Literature Review Chapter, we present the optimization models that have been proposed in a critical perspective and investigate each component of the models in a profound structure. First we examine how each study formed their demand model, second we examine each optimization model's problem definition, assumptions and components in detail, and third we criticize each model's contributions and weaknesses. We use detailed figures and table structures to define and compare all the models at the same time. In the New Model chapter, we propose our comprehensive optimization model of a shelf space allocation problem for a product category which considers almost all needs and necessities that a retailer is facing in todays retail environment. We then present the results from our experiments and compare the results of our model with the early proposed models. Finally, in the Conclusion chapter, we summarize our study and provide directions and suggestions for future research.

Chapter 2

Critical Review of Optimization Models

In this chapter, we critically discuss the demand functions and optimization models that have been proposed in the shelf space allocation literature and describe the main differences between these models. In each optimization model, we will discuss the model's main assumptions and define its objective function and its constraints. Afterwards, we examine the gaps in the literature in lights of the main needs of the current real world retail environment and describe the contributions of our thesis.

2.1 Demand Function

In this section, we define various demand functions that have been proposed in the shelf space allocation literature. Most of the demand functions are based on models that have been proposed in the 1980's. To have a unified presentation we use a common notation (see the Notation and Abbreviations on pages vii–xix).

2.1.1 Urban (1969)

Urban (1969) suggested a mathematical model to represent the interaction between products in a product category. He determined the three most basic and important effects on sales as price, advertising and distribution and combined the effects of these variables in an aggregate demand function. He mentioned the disadvantage of the simplest (linear) form of the combined equation. Since the linear form would not allow the mix effects of the marketing variables and the decreasing returns, because of the linear response to these variables, he suggested a linear log function, which can represent the mix effects of the marketing variables. The demand function of a product can then be expressed as

$$D_i = \pi_i' p_i^{\sigma_i} (adv)_i^{\varrho_i} (di)_i^{\varsigma_i} \tag{2.1.1.1}$$

This form allows for the nonlinearity of the demand function for a product which is reflected by the own-elasticity parameters σ_i, ϱ_i and ς_i. Urban defined the elasticity parameters as "the proportionate changes in the product sales resulting from a proportionate change in one variable". For example, when the value of the elasticity parameter of a product price is between 0 and 1, the response of the demand function will be decreasing as the variable increases. When the value of the elasticity parameter is less than 0, the response of the demand function will be increasing as the variable increases.

Urban then introduced the interdependency effects to the demand function by considering the interdependencies among products and emphasized the complementary and substitution effects between products. The complementary effect induces consumers to purchase a product along with another product and the substitution effect causes disloyal consumers to change their purchase decisions from one product to another product. The interdependency effects for the original demand formation of a product can be expressed as

$$\pi''_j p_j^{\mu_{ij}} (adv)_j^{\varrho'_{ij}} (di)_j^{\varsigma'_{ij}} \tag{2.1.1.2}$$

where μ_{ij}, ϱ'_{ij} and ς'_{ij} are the cross-product elasticities of product i to product j. For example, if the cross (price) elasticity parameter is positive then this means that there is a negative relationship between the two products and the products are substitutes. If the cross (price) elasticity parameter is negative then the products are complements. The marketing mix variables and the interdependency effects between products are then combined to form the following model:

$$D_i = \pi_i p_i^{\sigma_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} p_j^{\mu_{ij}} (adv)_i^{\varrho_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} (adv)_j^{\varrho'_{ij}} (di)_i^{\varsigma_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} (di)_j^{\varsigma'_{ij}} \tag{2.1.1.3}$$

Urban's demand model contains the basic variables such as price, advertising and distribution where the interdependencies among products is considered by using direct and cross elasticities of the variables.

2.1.2 Anderson and Amato (1973)

Anderson and Amato are the first researchers who modeled the potential demand and its components in the shelf space literature and used this demand function to simultaneously select the products that should be displayed on shelves and allocate shelf space for the chosen products. We think that one of the most important models which defined demand is Anderson and Amato's. They decomposed the expected potential demand Π, which is the product quantity limit that would be sold if all the products were in the assortment and displayed, into two important components and subcomponents:

1. Random Demand: Demand that originates from consumers that do not select the product on purpose or deliberately. We define RD_i as the demand for product i from potential consumers who do not have an incentive to purchase that particular product, but nevertheless purchase it without any selection preference. Therefore, the sum of all random demand for each product generates the total random demand, $RD = \sum_{i=1}^{I} RD_i$.

2. Preference Demand: If consumers choose a product within an assortment deliberately and purposely, then the demand is a preference demand. Therefore $PD = \Pi - RD$. Potential preference demand for a product has two important disjoint components: The demand originating from consumers which are loyal to a product (LD) and consumers who have a calculated product preference among all products ($GD + AD$). It is obvious that both the loyal consumer demand (LD) and the calculated choice demand ($GD + AD$) decisions contain a consumer preference. The second category of consumers make their choices based on their calculated valuations among the products. Therefore, consumers may or may not switch from one product to another product. We describe these components of the preference demand below in more details. The demand that arises from consumers who prefer or choose to purchase a product within the assortment but purchase product i through calculated valuations, can not be categorized under LD_i since the consumers who generate this demand are not

loyal to product i. Preference demand for a product i in an assortment constitutes of three parts, $PD_i = LD_i + GD_i + AD_i$. The notations are defined in the switching and non-switching preference demand subsections below.

(a) Switching preference demand: The demand originates from consumers who make their choices on their calculated valuations among products. There are two important components of switching preference demand:

   i. Switching preference demand arises from consumers who are willing to prefer a product within the assortment but through their calculated valuations, choose a product i within the assortment. We call this gained demand for a product i (GD_i) from the products within the assortment.

   ii. Switching preference demand arises from consumers who are not willing to prefer any product within the assortment but through their calculated valuations, choose a product i within the assortment. Borin (1994, 1995) refer to this as acquired demand for a product i (AD_i).

   We link the definition of the switching preference demand by Anderson and Amato (1973) for a product i in the assortment to $GD_i + AD_i$, since consumers switch to product i regardless of whether they prefer a product within the assortment or not. Furthermore, AD_i can be further partitioned into two components:

   - consumers who are willing to purchase or prefer a product which is not in the assortment and are willing to switch to a product i in the assortment. This is referred to as acquired demand (AD_i) in the literature.

- consumers who are willing to purchase or prefer a product which is not in the assortment and are not willing to switch to a product in the assortment. But, Anderson and Amato hypothesized that if there is any customer who prefers a product which is not in the assortment, then he/she will substitute his/her choice from a product in the assortment. This means that they neglected the lost demand-sales (LD of the products that are not in the assortment) due to permanent stock-out and did not consider it in their demand model.

(b) Nonswitching preference demand arises from consumers who are loyal to a product i and do not consider to switch to another product under any circumstances. We call this demand for a product i in the assortment, loyal customer demand (LD_i).

Since potential demand Π has been defined in two components, it is obvious that

$$RD + PD = \Pi \tag{2.1.2.1}$$

where

$$RD = \sum_{i=1}^{I} RD_i. \tag{2.1.2.2}$$

It is then posited that a consumer would purchase a particular product i within the assortment with probability

$$Pr(i) = x_i(\frac{w_i}{L}), \tag{2.1.2.3}$$

i.e., the probability is proportional to the number of facings, x_i, and the ratio of the horizontal space (w_i) reserved for the product to that available in the shelf (L). The random demand for product i becomes

$$RD_i = Pr(i)RD = x_i(\frac{w_i}{L})RD \tag{2.1.2.4}$$

A product's preference demand coming from a nonswitching preference is

$$LD_i = (\Pi - RD)\tau_i(1 - \psi_i) \tag{2.1.2.5}$$

and the part of this product's preference demand coming from switching preference is as follows

$$GD_i = (\Pi - RD)\tau_i\psi_i. \tag{2.1.2.6}$$

This demand can be categorized under the consumer demand who chooses to purchase a product or any product with a calculated valuation within the assortment and then chooses to purchase product i. Not to forget, consumers' first choice is not to purchase product i within the assortment. The consumers who prefer any product within the assortment but choose to purchase product i constitute this demand.

The other part of the product's preference demand comes from switching preference and is as follows:

$$AD = (\Pi - RD)\left(\sum_{i \in I^-} \tau_i\psi_i\right) \tag{2.1.2.7}$$

This demand can be categorized under the consumer demand who chooses to purchase the product with a calculated valuation from unstocked products and then chooses to purchase product i within the assortment with a probability. And not to forget again, consumers first choice is not to purchase any product within the assortment and there is a probability for product i to be picked among the available product assortment. Since there is a probability for product i to be chosen along all the products within the assortment, the probability is posited by

$$Pr\left(i|i \in I^+\right) = \begin{cases} \frac{\tau_i}{\sum_{i \in I^+} \tau_i}, & \text{if } \tau_i > 0 \\ \frac{1}{|I^+|}, & \text{if } \tau_i = 0 \end{cases}, \tag{2.1.2.8}$$

where $|I^+|$ denotes the cardinality, or number of items, of I^+. Anderson and Amato (1973) assume that $\tau_i > 0$. Therefore the acquired demand for product i becomes

$$AD_i = \left(\frac{\tau_i}{\sum_{i \in I^+} \tau_i}\right) (\Pi - RD) \left(\sum_{i \in I^-} \tau_i \psi_i\right) \tag{2.1.2.9}$$

Thus, we may find the total preference demand coming from the stocked products' preferences and unstocked products' preferences by adding the following demands

$$GD_i + AD_i = (\Pi - RD) \left[\tau_i \psi_i + \left(\frac{\tau_i}{\sum_{i \in I^+} \tau_i}\right) \left(\sum_{i \in I^-} \tau_i \psi_i\right)\right] \tag{2.1.2.10}$$

To find the total demand for product i we add up both the random demand and preference demands for product i:

$$\begin{aligned} D_i &= RD_i + LD_i + (GD_i + AD_i) \\ &= \left[x_i(\frac{w_i}{L})RD + (\Pi - RD)\tau_i(1 - \psi_i) + \left((\Pi - RD)\left[\tau_i\psi_i + \left(\frac{\tau_i}{\sum_{i \in I^+} \tau_i}\right)\left(\sum_{i \in I^-} \tau_i \psi_i\right)\right]\right)\right] \\ &= \left(x_i(\frac{w_i}{L})RD + (\Pi - RD)\left[\tau_i + \left(\frac{\tau_i}{\sum_{i \in I^+} \tau_i}\right)\left(\sum_{i \in I^-} \tau_i \psi_i\right)\right]\right) \end{aligned} \tag{2.1.2.11}$$

To better understand the different demand components we present in Figure 2.1 the consumers' decision making process and related demands for all the customers who enter the store. In Figure 2.2, we show the customer's decision process for purchasing a product i and how the total demand for a product i is structured. It is important to note here that both figures consider all possible purchase options of a customer and therefore the assortment may or may not be satisfactory for customers, unlike the assumptions in Anderson and Amato's.

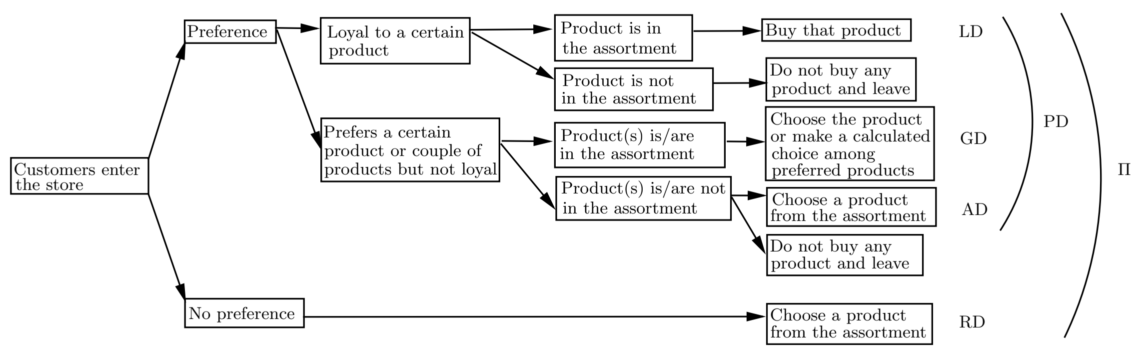


Figure 2.1: Consumers' Decision Making

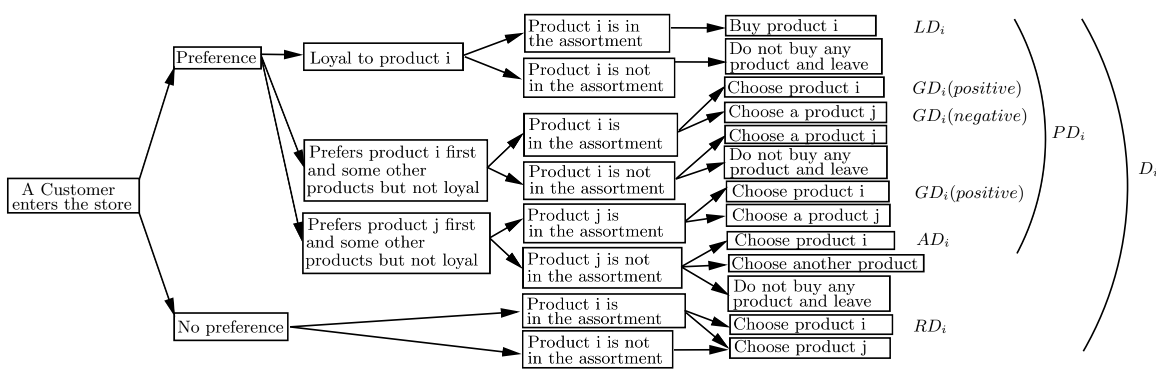


Figure 2.2: One customer's decision making and the role it plays in the overall demand of that product

2.1.3 Hansen and Heinsbroek (1979)

Hansen and Heinsbroek modeled the demand function of their shelf space allocation model using the main shelf space effect and ignored the cross effects between products. They used a multiplicative form, instead of using an additive form, to shape the demand function. They assumed that the demand only depends on the shelf space allocated to that product and this means that sales will be proportional to the space allocated to the products. They included the space elasticity factor into their demand function and assumed that it is constant for all products in the assortment. Thus, the demand model is represented by

$$D_i = \pi_i l_i^{\alpha_i}. \tag{2.1.3.1}$$

2.1.4 Corstjen and Doyle (1981, 1983)

Corstjens amd Doyle (1981) are the first researchers who incorporated both the product-space elasticity effect and inter-cross elasticity effect between products within the store in a shelf space optimization model. But, the optimization framework is actually presented in Urban (1969)'s study where the cross effects of the products are modeled and estimated. Their contribution to shelf space allocation literature has been used in many studies directly or with few extensions to the original demand model. They mentioned that there are many variables that effect the demand besides the space factors, such as marketing, promotions and price, and they assumed that these factors will be constant in their demand model. The model is represented by

$$D_i = \pi_i x_i^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ij}}. \tag{2.1.4.1}$$

Corstjen and Doyle's (1983) second demand model involved a dynamic case of the model presented above where they account for the growth of the market. Their model introduces the time dimension and assumes that past demand (and thus consumer preferences) influences current period demand. The model for demand is represented by a multiplicative form as

$$D_{it'} = \pi_i x_{it'}^{\alpha_{it'}} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_{jt'}^{\beta_{ijt'}} D_{i(t'-1)}^{\epsilon}. \tag{2.1.4.2}$$

2.1.5 Zufryden (1986)

Zufryden assumed that demand for a product is a function of a vector which includes all variables that may affect product demand such as advertising, promotions, shelf space and retail price. The unit demand for a product is defined as

$$D_i(f_i) = D_i(\bar{o}_i, f_{1i}, f_{2i}, ..., f_{zi}), \tag{2.1.5.1}$$

where $\bar{o}_i$ is the number of slots to be allocated to product i and the f_i, f_{1i} denote the function and factors, respectively, that affect demand other than space.

For tractability, Zufryden considered the space effect as the main effect, assumed the other non-space demand variables as fixed values, and ignored the consideration of cross-elasticities between products in the store. The resulting demand model was the first model that did not consider the shelf space in one dimension (only the width of the shelf). His study introduced the stacking process into the literature, a factor that was largely ignored in the literature. The unit demand for a product becomes

$$D_i(f_i) = D_i(\bar{o}_i) = D_i(\bar{o}_i, f'_{1i}, f'_{2i}, ..., f'_{zi}) = \pi_i(\bar{o}_i)^{\alpha_i} \prod_{z=1}^{Z} f'^{\gamma_{zi}}_{zi} \qquad (2.1.5.2)$$

where f'_{zi}'s are fixed values of non-space demand variables.

2.1.6 Borin et al. (1994) and Borin and Farris (1995)

Borin et al. (1994) defined the demand in four key concepts: unmodified demand (UMD), modified demand (MD), acquired demand (AD) and stockout demand ($SOBD$ and $SOLD$) where UMD, MD and AD influence the product demand positively and stockout demand influences the product demand negatively. UMD represents the intrinsic preference for the product. It is similar to π_i, as used in Shugan (1989) and Corstjen and Doyle (1981), where it represents the potential market share of the product in the store without in-store support such as advertisement, promotions, space and specific location in store. UMD_i reflects the demand of product i when all marketing variables are fixed. Thus, UMD of a product can be represented by

$$UMD_i(x_i, f'_{1i}, f'_{2i}, ..., f'_{zi}) = U_i = \pi_i \qquad (2.1.6.1)$$

where space (x_i) and f'_{zi}'s are fixed values of demand variables.

MD represents the demand for the product from its UMD with in-store merchandising support such as advertisement, promotions, space and specific location in the store. They assumed that the modified demand is solely a function of space while the other marketing variables are held constant and therefore modified demand is the "modified" version of the unmodified demand with its differential space allocation. If some of the other marketing variables were not fixed, they would be a part of MD. Since marketing variables other than space are fixed in the unmodified demand, space is the only in-store attractiveness variable that affects sales. Thus, MD of a product can be represented by

$$MD_i = \pi_i \prod_{j \in I^+} x_j^{\beta_{ij}} = \pi_i x_i^{\alpha_i} \prod_{\substack{j \in I^+ \\ j \neq i}} x_j^{\beta_{ij}} \tag{2.1.6.2}$$

AD represents the demand that unstocked products capture. Some of the consumers may change their preferences and purchase a product within the assortment and some of them may stay loyal to their first choice and decide not to purchase the product. A product's AD would consist of three basic parts:

$$AD_i = \sum_{k \in I^-} \left[\left(\pi_k x_k^{\alpha_k} \prod_{k' \in I^-} x_{k'}^{\beta_{kk'}} \right) (1 - \theta_k) \left(\frac{\beta_{ik} \pi_i x_i^{\alpha_i} \prod_{j \in I^+} x_j^{\beta_{ij}}}{\sum_{i \in I^+} \beta_{ik} \pi_i x_i^{\alpha_i} \prod_{j \in I^+} x_j^{\beta_{ij}}} \right) \right] \tag{2.1.6.3}$$

where the first part, in the product expression, is the potential MD that the unstocked products would have received if they were in the assortment. The second part is the fraction of consumers which would change their opinion and purchase a product within the assortment. The third part represents the distribution of the demand amongst the stocked products in the assortment. If the sum of these demands exceeds the product's quantity in the stock, this causes a temporary stockout and there exists a potential demand (the difference between the sum of the demands and product stock) for products in the stock. A temporary stockout demand for the other products can then be determined by using the resistance to compromise factors (θ_j) and this potential (temporary stockout) demand can be allocated to products in the assortment as the same proportion which is used in AD_i. The resultant amount allocated to a product is a stockout gain ($SOBD_i$). They assumed that temporary out-of-stock situations and permanent stockout situations have the same resistance to compromise factors which might not be realistic in the real life. With all of these differentially categorized demands which have positive and negative effects, they formed the unit demand for a products in an additive form as

$$
\begin{aligned}
D_i &= MD_i + AD_i + SOBD_i + SOLD_i \\
&= \pi_i x_i^{\alpha_i} \prod_{\substack{j \in I^+ \\ j \neq i}} x_j^{\beta_{ij}} + \sum_{k \in I^-} \left[\left(\pi_k x_k^{\alpha_k} \prod_{k' \in I^-} x_{k'}^{\beta_{kk'}} \right) (1-\theta_k) \left(\frac{\beta_{ik} \pi_i x_i^{\alpha_i} \prod_{j \in I^+} x_j^{\beta_{ij}}}{\sum_{i \in I^+} \beta_{ik} \pi_i x_i^{\alpha_i} \prod_{j \in I^+} x_j^{\beta_{ij}}} \right) \right] \\
&\quad + SOBD_i + SOLD_i
\end{aligned}
\tag{2.1.6.4}
$$

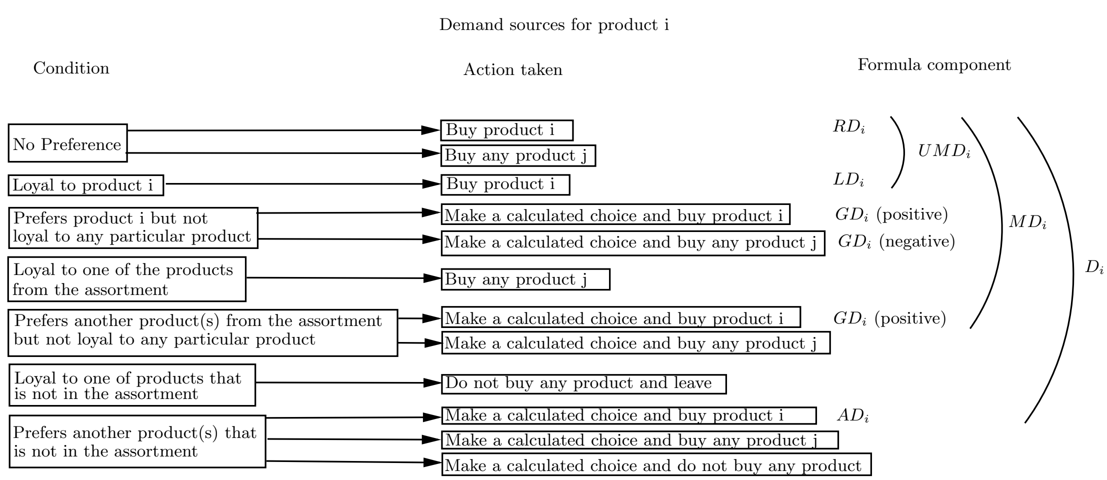


Figure 2.3: Demand Sources for a Product i

In Figure 2.3 we show the demand sources for a product i and each related formula component. In Figure 2.4, we show a general demand distribution of the overall

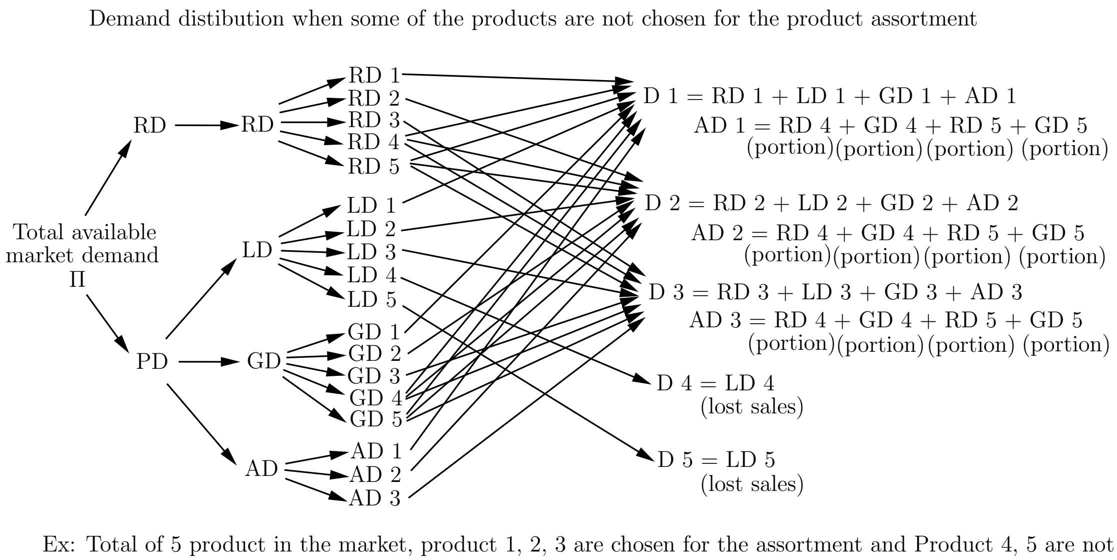


Figure 2.4: Demand Distrubution

category demand and how each product's demand is being structured. Figure 2.4 is useful for understanding how a product's demand in the assortment is shaped and how a retailer loses sales. Figure 2.4, unlike Figure 2.2 and 2.3, assumes a satisfactory product assortment, since lost sales are only generated by loyal demand for the non-stocked products. If the assortment is not satisfactory to the customer, a portion of each product's gained demand (a portion of $\sum_{i \in I^+ GD_i} = GD$) will also be lost. In general, all the optimization models in the shelf space allocation literature, implicitly, assume a satisfactory product assortment for the product category.

The figures we used to reflect Anderson and Amato (1973) and Borin et al. (1994) demand functions might look similar, however, the Anderson and Amato function does not consider the effects of marketing variables. They construct their demand function using probabilistic relations. Although they did use MD_i in their demand structure, the GD_i component in MD_i is assumed to be fixed by using probabilistic values and the RD_i component is assumed to be the active component that reflects the shelf space allocation of product i. But since Borin et al. considers the effects of marketing variables, such as space, they modify the structure as $UMD_iGD_i = MD_i$. Therefore, MD_i in Anderson and Amato (1973) would not reflect the same results as MD_i in Borin et al. (1994).

2.1.7 Urban (1998)

Urban (1998) was the first model which incorporated the conventional inventory-control decisions. All existing models by that time implicitly assumed that the shelves are always kept fully stocked with the products in the assortment, since they used number of facings to calculate the demand. Decreasing number of display facings and its effect on demand was not considered before in the demand formulations.

To account for changing display facings we need to introduce the concepts of backroom inventory and shelf space inventory (backroom space). Backroom inventory can be defined as the inventory where the products are being delivered to the retailer and being stocked in a warehouse. Backroom space is the space in the shelf behind the display facings where the products are also being stocked in order to satisfy the consumer demand right on the shelf. Backroom inventory depletion, backroom space depletion and decreasing the number of display facings therefore should be considered in order to represent the demand.

Urban (1998) formulated the demand as a function of the display inventory level and assumed a deterministic and constant demand rate and that the inventory is replenished instantaneously with a constant lead time. Hence, as long as the inventory level of the product's display facing is constant, the demand rate of the product will be kept constant. This means that the products in the backroom inventory or backrooom space are the ones that are being depleted since the facing products are dedicated for display. After all the backroom inventory and the products in the backroom space are depleted and it is the turn for the facing products to be depleted, then the number of facings will start to decrease and therefore the demand rate decreases too, since the demand rate depends on the displayed inventory level. In addition, the demand rate of a product will decrease at a decreasing rate after the backroom inventory is depleted because the demand rate of a product depends on the instantaneous display inventory level. Since the demand rate of a product will always change for every instance of depletion after the backroom inventories are depleted, they simplified the demand rate of a product by making the demand rate as a function of space allocation of all the other products and not the instantaneous inventory level. Thus, the unit demand for a product is represented by

$$D_i = \pi_i x_i^{\alpha_i} \prod_{\substack{j \in I^+ \\ j \neq i}} x_j^{\beta_{ij}} \left[1 + \sum_{k \in I^-} (1 - \theta_k) f(\pi_k, \beta_{ki}) \right] \tag{2.1.7.1}$$

The expression in the brackets shows the demand coming from the consumers which are willing to purchase a product within the assortment if their preferred product is not included in the assortment. This function depends on potential demand (π_k) and the degree of substitutability (β_{ki}) with the products that are not in the assortment. This part is analogous to the expression for AD in Borin (1994).

2.1.8 Yang and Chen (1999)

Yang and Chen (1999) additionally considered the location effect of a product in different shelves. Their model also contained cross-elasticity between products and included other marketing variables besides space, but assumed that the other marketing variables are fixed. They modeled the demand of a product as a function of the display facing of the shelf on which the product is displayed. Thus, the unit demand for a product on shelf m in the assortment is defined as

$$D_{im}(x_{im}) = D_{im}(x_{im}, f'_{1i}, f'_{2i}, ..., f'_{zi}) = \pi_i x_{im}^{\alpha_{im}} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ij}} \prod_{z=1}^{Z} {f'}_{zi}^{\gamma_{zi}} \tag{2.1.8.1}$$

where f'_{zi}'s are fixed values of non-space demand variables and $x_j = \sum_{m=1}^{M} x_{jm}$ is the total amount of display facing of product j. Their main contribution to the literature was to propose an alternative optimization model that is applicable to the retail practice. Since we focus on the demand models in this section, we will discuss their alternate model formulation in the optimization models section.

2.1.9 Irion et al. 2004

Unlike the other polynomial demand models, Irion et al. (2004) model the demand using not only the number of facings of the products but also the width-lengths of the products. They describe the width-length diversity of the products in assortments and the interrelation between the product and its dimensions (width-length in this model) and explicitly implied that the width-length of a display facing would affect the demand since the product and its width-length are interrelated. Therefore, they model the demand not using the number of facings but using the total width-length area (since total width-length area is in one dimensional space, we may just call it total width-length). This demand model is similar to that of Hansen and Heinsbroek (1979) in terms of modeling the demand using the total width-length of a product, not the number of products. However, Hansen and Heinsbroek did not consider the cross-space effect between products within the assortment. Irion et al. defined demand as

$$D_i = \pi_i \left(w_i x_i\right)^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} \left(w_j x_j\right)^{\beta_{ij}} . \qquad (2.1.9.1)$$

They model the demand using one-dimensional area of the product to allocate space and then stacking the products on top of each other for every product in the assortment until there is no space left to stack one more product. If we look at the demand model carefully, although it seems like the demand model can be expressed using a two-dimensional space, because of the stacking process, the authors argue that "as long as the products are stacked until there isn't enough space left to stack one more product on top of the other, one-dimensional area allocation (total width-length of the product allocation) determines the demand because the area remained between the stacked products and the height of the shelf does not affect the demand". Hence they did not consider the total area of the product, a_i, and used the one-dimensional area of the product, $(w_i x_i)$, to form their demand function. We believe that this may be a reasonable assumption in a product category, if the height of the different products are very close to each others. The similar product heights would result in the same number of display facings to be stacked in the shelf height and therefore the total height of the different products would be very close. Since the percentage cover of the stacked products to the height of the shelf display will be very similar and consumers' space allocation perception depends more on the allocated space of the total width-length of the products (consistent with the shelf space allocation literature), therefore the demand might not be affected by the deviations of the uncovered area on top of the shelf caused by different (but close) product heights.

2.1.10 Reyes and Frazier (2005)

Reyes and Frazier (2005) model their demand function in order to determine the initial shelf space allocation of the retail store and therefore their demand model has very similar characteristics to Anderson and Amato (1973)'s demand model. The main difference is that Reyes and Frazier incorporated space elasticity with probabilistic values when calculating the different demand components. In order to define the cross space elasticity between two products, a ratio of the relative difference in shelf space allocation to the two products, $1 + \frac{x_i - x_j}{X}$ was used. Thus, they model their demand in an additive form as

$$\begin{aligned} D_i &= RD_i + LD_i + (GD_i + AD_i) \\ &= RD_i + LD_i + (GD_i(negative) + GD_i(positive) + AD_i) \\ &= RD\frac{x_i}{X} + ND_i - \sum_{\substack{j \in I^+ \\ j \neq i}} \frac{X + x_j - x_i}{X} ND_i \, \Delta ij + \sum_{\substack{j \in I^+ \\ j \neq i}} \frac{X + x_i - x_j}{X} ND_j \, \Delta ji \\ &+ \sum_{k \in I^-} \frac{x_i}{X} ND_k(1 - \theta_k) \end{aligned} \tag{2.1.10.1}$$

The two parts of GD_i, $GD_i(negative)$ and $GD_i(positive)$, reflect the losses and gains from the other products in the assortment. In particular, $GD_i(negative)$ represents the lost demand resulting from switching product i to a stocked product j and $GD_i(positive)$ represents the gained demand resulting from switching a stocked product j to product i.

2.1.11 Hwang et al. (2005)

Hwang et al. (2005) model is inspired by the study of Dreze et al. (1994) in which they concluded the importance of shelf level on which products are displayed within the shelves and emphasized that eye-level position has the most of effect on sales. They modified Corstjen and Doyle (1981) model and incorporated a location effect multiplier ($\varphi_m \geqslant 1$) in the demand function. Each shelf has a weight for the position relevant to the eye-level and the shelf which is on eye level has the highest weight. The bottom or the top shelf have lower weights depending on the characteristics of a store and the lowest weight of 1 is assigned to the worst shelf location. If the product is displayed on different shelves at the same time, they assumed an average value of φ_i, depending on the amount of products on the different shelves. Hence, the unit demand for a product is defined as

$$D_i = \pi_i x_i^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ij}} \varphi_i \tag{2.1.11.1}$$

where $x_i = \sum_{m=1}^{M} x_{im}$ and $\varphi_i = \frac{\sum_{m=1}^{M} x_{im}\varphi_m}{x_i}$ is the average value of a φ_i if the product is displayed on more than one shelf.

2.1.12 Reyes and Frazier (2007)

Reyes and Frazier (2007) mentioned the various factors which impact the customer's choice among different products within a product category such as price, space, in-store promotion, customer loyalty, brand reputation, quality of the product, and the desire to try something different. They incorporated two key factors in their demand model: price and space. Most consumers prefer to pay less for a product, however not all consumers prefer to purchase the cheapest products because of the other factors that alters consumers' preference. That is why highly price-sensitive consumers are affected more by price changes. Thus, reducing the price affects the demand in a positive way and increasing the price affects the demand in a negative way. The difference between Reyes and Frazier (2007) demand model and existing demand models is that they represent the impact of the price sensitivity of demand by dividing the price of a particular product by the average price of all the products in the assortment or product category ($\left(\frac{p_{avg}}{p_i}\right)^{\sigma_i}$, where they used $\sigma_i = 0.5$ for the examples in their study). When $p_i = p_{avg}$ then the price sensitivity factor of that product is 1, implying that product i is not sensitive to price changes. They also represent the space elasticity factor in a different form from the existing demand models, which used a multiplicative form. They form the base level shelf space allocation for a product, which is determined by the ratio of base level demand of that product to total demand for the product category, and multiply this ratio by the total available shelf space in terms of number of product units ($x_i^{base} = \left(\frac{\pi_i}{\Pi}\right) X$). Then using this knowledge, they represent the space sensitivity factor of a product by dividing the actual shelf space allocation of that product by the base shelf space allocation ($\left(\frac{x_i}{x_i^{base}}\right)^{\alpha_i}$ where they used $\alpha_i = 0.5$ for the examples in their study). When $x_i = x_i^{base}$ then the space sensitivity factor of that product is 1, implying that product i is not sensitive to space changes. Therefore, they defined the unit demand for a product in the assortment as

$$D_i = \pi_i \left(\frac{x_i}{x_i^{base}}\right)^{\alpha_i} \left(\frac{p_{avg}}{p_i}\right)^{\sigma_i}. \tag{2.1.12.1}$$

2.1.13 Hariga et al. (2007)

Hariga et al. (2007) model incorporates both the main-space effect and the cross effects between products with location effects. But, in their work they consider the location effect as being displayed in more than one location within the store. Therefore, they model the unit demand of a product on shelf m (or m-th display location) in the assortment as

$$D_{im}(t) = \pi_i IN_{im}^{\alpha_{im}}(t) \prod_{\substack{j \in I^+ \\ j \neq i}} IN_{jm}^{\beta_{ij}}(t) \tag{2.1.13.1}$$

and to simplify the model, they assumed that the demand will depend on total space allocation of the other products in the assortment and not the instantaneous inventory level of display facings. They have done this by replacing the time dependent $IN_j(t)$ by x_j. Thus, the unit demand for a product on m-th display location in the assortment becomes

$$D_{im}(t) = \pi_i x_{im}^{\alpha_{im}}(t) \prod_{\substack{j \in I^+ \\ j \neq i}} (1 - y_j + x_j y_j)^{\beta_{ij}} \tag{2.1.13.2}$$

where y_i is a binary variable in order to determine whether the product is included in the assortment. The authors also mentioned that this model can be extended by using price as a decision variable in the demand function using the main effect and the cross effects between the products in the assortment under the circumstances of no advertising and in-store promotions.

2.1.14 Murray et al. (2010)

Most of the shelf space allocation models consider the number of facings as the decision variables in the demand model. Unlike the common demand models, Murray et al.

(2010) consider the product facing areas and three display orientations, assuming three dimensional quadrangular sizes for product packaging. They refer to the study of Dreze et al. (1994) about the importance of product facing areas and aesthetic determinants of product packaging such as size, color coordinations related to display orientation and model the demand using display facing areas rather than the number of facings. They argue that "a full facing packaging orientation on an eye-level shelf is likely to be much more better in terms of quality of a given amount of display facing area allocated to the product than a side facing packaging orientation on a bottom shelf." motivated by the findings of Bimolt et al. (2005) and Bucklin et al. (1998), the incorporate the price decisions and cross-product price interactions. In addition they consider display facing areas (two dimensional space) for stackable products. They modeled the unit demand for a product by

$$D_i(p, x_i) = \pi_i \left(\sum_{o=1}^{3} \sum_{m=1}^{M} \delta_{iom} w_{io} h_{io} \bar{x}_{iom} x_{iom} \right)^{\alpha_i} p_i^{\sigma_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} p_j^{\mu_{ij}} \tag{2.1.14.1}$$

where $w_{io}h_{io}$ represents the display surface area of product i and $\bar{x}_{iom}x_{iom}$ represents the quantity of product i that can be placed in shelf m (two-dimensional facing area of shelf m). Thus, the product $w_{io}h_{io}\bar{x}_{iom}x_{iom}$ represents the shelf space allocation of the product in display orientation o on shelf m (a_{iom}). In addition, δ_{iom} is a parameter that indicates a shelf location-orientation quality-adjustment weight corresponding to the display facing area of product i in display orientation o on shelf m.

2.1.15 Russell and Urban (2010)

Russell and Urban (2010) based their model on the model by Dreze et al. (1994), which noted that the sales tend to act as a quadratic function in the horizontal dimension in a shelf and as a cubic function in a vertical dimension. They also noted that the model results would not differ much if sales are assumed to be quadratic in the vertical dimension. Additionally, they also used a quadratic formulation for the demand function for the effects of the display facing numbers of products on sales, because it can reflect diminishing returns and provide a good estimate of concave functions' diversity, especially for a small number of products' display facings. The demand formulation for a product in the assortment is defined as

$$\begin{aligned} D_i &= \zeta_{0i} + \zeta_{1i}\bar{c}_i + \zeta_{2i}(\bar{c}_i)^2 \\ &\quad + \sum_{m=1}^{M} \left[\zeta_{3i}(H_m y_{im}) + \zeta_{4i}(H_m y_{im})^2 + \zeta_{5i}(H_m y_{im})^3 + \zeta_{6i}(x_{im}) + \zeta_{7i}(x_{im})^2\right] \end{aligned} \tag{2.1.15.1}$$

where $\zeta_{\bullet}i$ are appropriate coefficients for particular products. Since (y_{im}) is a binary variable (implying that that $y_{im} = (y_{im})^2 = (y_{im})^3$) then D_i can be simplified to

$$D_i = \zeta_{0i} + \zeta_{1i}\bar{c}_i + \zeta_{2i}(\bar{c}_i)^2 + \sum_{m=1}^{M} \left[(\zeta_{3i}H_m + \zeta_{4i}H_m^2 + \zeta_{5i}H_m^3)y_{im} + \zeta_{6i}x_{im} + \zeta_{7i}x_{im}^2\right] \tag{2.1.15.2}$$

2.1.16 Lotfi et al. (2011)

Lotfi et al. (2011) model the demand using a minimum-maximum approach where the maximum demand is defined by the price change of the product and space allocation.

They have looked at the problem in a different way and claimed that any price change in a product's price affects the demand of the product as well as the other products' demand in the assortment. For example, if a product's price increases and its substitutes price stay the same, increase less or decrease, then the consumers who are not willing to purchase any particular product, or are not totally loyal to any brand or product, tend to alter their choice in terms of their benefit and gain from the products. They assume a minimum daily demand (in units of sales) which is also assumed to be constant for a product in a previously determined time period and use this information as an input in the demand model. They calculate the maximum daily demand for that product which it can meet due to the price change and space allocation. Eventually, they form the maximum demand for a product in the assortment as

$$UD_{ib} = MinD_{ib}\frac{(1 - PC_{ib}\sigma'_{ib})}{(1 - PC_{ib,j}\mu'_{ib,j})}x_{ib}^{\alpha_i} \tag{2.1.16.1}$$

Since they applied a minimum-maximum approach to demand, this information will be used in according with $MinD_{ib}$ as a constraint in the main optimization model. The constraint will guarantee that the order quantity, which is a decision variable, will be between UD_{ib} and $MinD_{ib}$ of the product. The quantity $\frac{1-PC_{ib}\sigma'_{ib}}{1-PC_{ib,j}\mu'_{ij}}$ represents how price changes affect the customer demand, since the numerator represents the price change effects of the product and the denominator represents the price change effects of the product's substitutes. If μ_{ij} is positive then the demand of product i increases, while a positive μ_{ji} decreases the demand and vice versa.

2.2 Optimization Models

2.2.1 Anderson and Amato (1973)

Anderson and Amato (1973) demand model is a touchstone among the demand models that have been proposed until now. It almost includes all the aspects of a potential demand Π. However, the other demand models tried to model the demand as a function other factors such as the space allocated to the product and its price. Anderson and Amato (1973) demand model works only for short-run periods and can not include the effects of any advertising or promotion activities in the demand model (since they partitioned the potential demand of a short-run period). Thus, it may not be suitable for real time shelf space allocation. The optimization side of the model has basic assumptions such as:

1. The total display facing area for a product (l_i) has a homogeneous quality in fixed physical size and is able to contain at least one display facing of the product.

2. There is enough stock of products in the assortment to satisfy the minimum demand for each displayed product.

3. All the products' operating costs and profit margins are fixed and given in the period.

4. All products is the assortment have the same physical size (since we allocate products using the width of a product w_i) and require the same amount of display facing per placement.

5. Total display area for a product is a multiple area of one product and total display area (l_i) is a multiple of a product's display facing area (or width-length w_i). i.e., $l_i/w_i = K$ where K is an integer.

Based on these assumptions, we see that they assume one, horizontal, homogeneous in quality shelf space. Since they consider simultaneously determining the product selection within a product category and allocation of the chosen products within this product category, they used binary variables (y_i) to decide whether the product is chosen for the assortment or not in addition to the decision variables which represents the number of products to be allocated (x_i's). Their aim was to maximize the gross profit of the product category and therefore they proposed their optimization problem as

$$\max \quad P = \sum_{i=1}^{I} y_i g_i D_i$$

$$\text{s. t.} \quad \sum_{i=1}^{I} x_i \leqslant X, \tag{2.2.1.1}$$

$$X^{min} \leqslant x_i y_i, \qquad \forall_i \tag{2.2.1.2}$$

where $g_i D_i = P_i$ is the gross profit of a product in the assortment. Their model resembles a classical "knapsack" problem where the gross profit multipliers ($g_i D_i$) imply the values in the objective function of the knapsack problem and the width-length of each product (w_i) imply the weights and total width-length of the shelf (L) imply the weight capacity of the knapsack. Since they assume the same product sizes (width-lengths) for all products, they are translated to x_i and X values, respectively.

It can easily been seen that they allocated the space among products depending on the products' potential demand, i.e., the market share of the product in a short-run period. Even though they ignored the cross-space elasticities between products, which generally uses the absolute space of a product to be allocated, the main-space effects were accounted for in the share of a product (RD_i) in the assortment. In addition, the lost sales due to stockouts were neglected and they assumed that all the stockout demand will be split up to the products in the assortment.

2.2.2 Hansen and Heinsbroek (1979)

Hansen and Heinsbroek (1979) optimization model incorporates the product selection (y_i) among a set of products. Their model also takes into account the minimum shelf space needed for a product to be allocated in a shelf. Furthermore, they require an integer solutions to allocate the space (l_i) for integer number of display facings. Hansen and Heinsbroek proposed an optimization model to maximize the total profit by selecting the products from a given set and optimally allocating the chosen products. The model has basic assumptions such as

1. If a product is chosen to be in the assortment and to be allocated to the shelves, a minimum quantity must be given to that product.
2. Total shelf space (L) is limited to products to be allocated.
3. Space elasticity of the products is constant for all products.

Based on these assumptions, we see that they assume one, horizontal, homogeneous in quality shelf space. Since the purpose is to maximize the total profit from each product that can be chosen to the assortment, the problem can be defined as

$$\max \quad \sum_{i=1}^{I}(g_i \pi_i l_i^{\alpha_i} - DC_i l_i) - f(NR, LT)$$

$$\text{s. t.} \quad \sum_{i=1}^{I} l_i \leqslant L, \tag{2.2.2.1}$$

$$L_i^{min} y_i \leqslant l_i \leqslant L y_i, \qquad \forall_i \tag{2.2.2.2}$$

$$y_i \in \{0, 1\}, \qquad \forall_i \tag{2.2.2.3}$$

$$\frac{l_i}{w_i} \in \mathbb{N}^+, \qquad \forall_i \tag{2.2.2.4}$$

where $g_i \pi_i l_i^{\alpha_i} - DC_i l_i = P_i$ is the profit (not the gross profit) of a product in the assortment and $f(NR, LT)$ is the replenishment cost of the shelf stock function which depends on the number of replenishment (times per week) and lead time of the orders (in days). Constraint (1) ensures that the total amount of shelf space to be allocated to all the products in the assortment does not exceed the total shelf space. Constraint set (2) ensures that if product i is chosen to be in the assortment then the space allocated to the product should be between the minimum and maximum shelf space (the total length of the shelf) which is under the control of the retailer. Constraints (3) represent the binary variable which determines whether the product will be in the assortment or not, and constraints (4) ensures that the number of product i to be allocated to the shelf is an integer value.

2.2.3 Corstjen and Doyle (1981)

Corstjen and Doyle (1981)'s demand model is the most well known space allocation demand model in the literature. As we indicated in the previous section, they modeled the demand using both own product space elasticities and cross (inter-product) elasticities between products in a multiplicative form (polynomial form). Therefore, the amount of shelf space allocated to a product determines the product's demand. Since retailers seek an allocation which maximizes their profit from the products in the assortment, they mentioned that the operating cost side also plays an important role within the model. For each product in the assortment, procurement costs, carrying and out-of-stock costs are different. They included two cost components in their model: the different gross profit margins (caused by different purchase costs and different sales prices of the products) and the operating costs such as procurement, carrying and out-of-stock costs associated with the incremented sales of various product assortments caused by alternative allocation of products. They modeled the operating cost function of the model as

$$C_i = \phi_i \left(\pi_i x_i^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ij}} \right)^{\lambda_i}$$

and the gross profit of a product as

$$P_i = g_i \pi_i x_i^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ij}}.$$

Corstjen and Doyle modeled the problem of maximizing the net profit subject to the constraints of total available shelf space, limiting sales of the products due to production or carrying limit of the products and lower (essential for retailer's image in the market) and upper (life cycle of a product may be on a later stage) bounds of the shelf space allocation of a product where the number of products to be allocated is the decision variable (x_i) for their problem. Hence, their proposed their optimization model is

$$\max \quad \sum_{i=1}^{I} \left[g_i \left(\pi_i x_i^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ij}} \right) - \phi_i \left(\pi_i x_i^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ij}} \right)^{\lambda_i} \right]$$

$$\text{s. t.} \quad \sum_{i=1}^{I} x_i \leqslant X, \tag{2.2.3.1}$$

$$X_i^{min} \leqslant x_i \leqslant X_i^{max}, \qquad \forall_i \tag{2.2.3.2}$$

$$D_i \leqslant Q_i^{max}, \qquad \forall_i \tag{2.2.3.3}$$

$$x_i \geqslant 0, \qquad \forall_i \tag{2.2.3.4}$$

where constraint (1) guarantees that the total number of products to be located does not exceed the available shelf space. Constraints (2) ensures that the number of product i which will be located on the shelf should be between the minimum and maximum number of display facings which the retailer decided. Constraints (3) represents the limiting sales (upper bound of sales) because of the production or carrying restrictions.

Based on these assumptions, we see that they assume one, horizontal, homogeneous in quality shelf space. They also assumed that the physical size of the products in the assortment are the same in terms of dimensions. Although, their model is the most well known model in the literature and almost all the following models were inspired by their work, one of the weaknesses of their model is that the solution of their model may not be integer.

Corstjen and Doyle (1983)'s dynamic model is more of a theoretical model where they eliminated most of the constraints and only included the shelf space availability constraints (constraints 1). The practical use this model is shallow since the constraints are pretty much simplified and most importantly the cost side of the model is ignored.

2.2.4 Zufryden (1986)

Zufryden (1986)'s aim was to develop a tractable model that might be used efficiently by computers. He modeled the demand using the own-space elasticity of the product but ignored the cross-elasticity between products in the store. He has considered both the demand side and the cost side in the objective function to maximize the profit. The cost function of the product is defined by

$$C_i = \phi_i \left(\pi_i \bar{o}_i^{\alpha_i} \prod_{z=1}^{Z} f'_{zi} \right)^{\lambda_i}$$

where f'_{zi}'s are fixed values of non-space demand variables. The optimization problem that maximizes net profit is

$$\begin{aligned} \max \quad & \sum_{i=1}^{I} \left[g_i \left(\pi_i \bar{o}_i^{\alpha_i} \prod_{z=1}^{Z} f'_{zi} \right) - \phi_i \left(\pi_i \bar{o}_i^{\alpha_i} \prod_{z=1}^{Z} f'_{zi} \right)^{\lambda_i} \right] \\ \text{s. t.} \quad & \sum_{i=1}^{I} \bar{o}_i \leqslant O, \end{aligned} \tag{2.2.4.1}$$

$$O_i^{min} \leqslant \bar{o}_i \leqslant O_i^{max}, \qquad \forall_i \tag{2.2.4.2}$$

$$D_i \leqslant Q_i^{max}, \qquad \forall_i \tag{2.2.4.3}$$

$$\bar{o}_i = 0, block_i, 2block_i, 3block_i...., \qquad \forall_i \tag{2.2.4.4}$$

$$block_i \geqslant 1, \qquad \forall_i \tag{2.2.4.5}$$

$$block_i \in \mathbb{N}^+, \qquad \forall_i \tag{2.2.4.6}$$

Based on their assumptions, they assumed one, horizontal, homogeneous in quality shelf space and the whole shelf space is divided into rectangular slots. A product can fit and can be allocated to one slot since all the products in the assortment have the same physical size (width and height). The space allocated to a product then can be the multiples of slots (slot areas). Therefore, display facing area (number of slots, $\bar{o}_i$ considered as the decision variable) is considered rather than the space volume allocation, and thus they guarantee integer solutions. This is the first model that considers product stacking (vertically) on a shelf inline with retailers desire to allocate the shelf space as uniform and complete columns (Dreze et al. 1994).

Constraint (1) ensure that the total number of slots to be allocated to all the products in the assortment does not exceed the available slot numbers in the showroom inventory. Constraint set (2) guarantees that the number of slots which will be allocated to product i will be between the minimum and maximum number of slots, as desired by the retailer. Constraint set (3) sets the limiting sales (upper bound of sales) because of the production or carrying restrictions. Constraint set (4) permits the consideration of space allocation in block units such as twin-packed paper towels and six-packed soft drink cans (display facing of the six-packed can be allocated into the shelf in blocks of three or two).

2.2.5 Borin et al. (1994), Borin and Farris (1995)

Borin et al. (1994)'s demand model is one the most comprehensive model in the literature. Their style of categorizing the demand for a product into different components gives a clear idea of how to identify each type of demand. They formulated an optimization problem where the retailer is to maximize the category's return on inventory investment. Their model also has basic assumptions such as:

1. The space variable in the model is assumed to be the number of display facings of a product.

2. Retailer's inventory investment represents the retailer's purchase cost of a full shelf for all the products in the assortment.

3. Products which will be allocated to the shelves are given. It means that there is no product selection in the optimization model.

4. There is adequate backroom space (shelf depth, not backroom inventory) for extra product stocking. The depth of the shelf is being considered in the model and it is fixed.

5. Minimum shelf space needed (L_i^{min} for this model) for a product to be allocated is the minimum number of display facings of the product and the maximum shelf space for a product ($L_i^{max} = L$ for this model) to be allocated is the number of display facings that available total shelf space can take.

6. The loyalty factors for both cases of temporary out-of-stocks and permanent stockouts are the same.

Based on the assumptions, they assume one, horizontal, homogeneous in quality shelf space. Their goal is to find l_i in order to:

$$\max \quad \frac{\sum_{i=1}^{I} g_i[p_i(MD_i + AD_i + SOBD_i + SOLD_i)]}{\sum_{i=1}^{I}(1-g_i)p_i II_i}$$

$$\text{s. t.} \quad \sum_{i=1}^{I} l_i = L \tag{2.2.5.1}$$

$$L_i^{min} \leqslant l_i \leqslant L, \quad \forall_i \tag{2.2.5.2}$$

where constraint (1) stipulate that the number of products to be allocated to all products into the shelf space should corresponds to the total number of products which can be allocated to the shelf. Constraint set (2) sets the minimum (one facing) and the maximum (total shelf space) for the number of number of product i to be allocated.

2.2.6 Urban (1998)

As we mentioned in the previous section, Urban was the first researcher who incorporated the inventory level dependent decisions to explicitly model the demand function as a displayed inventory level. All of the previous demand models implicitly assumed the shelves are kept fully stocked at all times. He was also the first researcher who made an explicit distinction between the backroom inventory and the showroom inventory, since the backroom inventory has no effect on the sales. He incorporated the effect of the showroom (display) inventory on demand when the showroom shelves are not fully stocked. To do so, he made the following assumptions:

1. The products are being depleted from the showroom (displayed) inventory but

the replacement from the backroom inventory is instantaneous. This means that the showroom shelves are being fully stocked and the backroom inventory is being depleted until there is no product in the backroom inventory left.

2. Since the backroom inventory is being depleted until all the products are gone, the showroom inventory is fully stocked and therefore the demand rate is constant.

3. Once the backroom inventory is depleted then the showroom inventory will be depleted and the demand rate of the product will start to decrease as the showroom inventory level decreases.

4. The total inventory replenishment is also instantaneous with a known and constant lead time. It means that the entire order from the suppliers is received instantaneously as soon as the inventory is depleted. Replenishments are independent for each product in the assortment. It means that there is no joint replenishments.

5. The total order is being received in the backroom inventory.

6. There is a limited dedicated display area (shelf space) for the product and when other products are being depleted from the showroom inventory (in case of their backroom is depleted already), the product can not be allocated to the other products dedicated display area. It means that the limited shelf space for the product can not be exceeded.

7. All the of the prices and costs (selling price, unit purchase cost of the products, holding costs, carrying costs, display costs) are known and constant.

Based on the assumptions, we conclude that Urban considered one, horizontal, homogeneous in quality shelf space for the showroom inventory and one backroom inventory with a limited capacity. Inventory decisions are incorporated through two variables: the order quantity (q_i) and the reorder point (r_i). In addition, a binary variable (y_i) is used to determine which products to be chosen for the assortment. The objective of the model is to maximize the net profit from all the products in the assortment. The net profit per cycle from a product in the showroom inventory is defined as

$$NP_i = \frac{(p_i - c_i)q_i}{CT_i} - \frac{OC_i}{CT_i} - HC_i AI_i - DC_i x_i, \tag{2.2.6.1}$$

where AI_i is assumed to be the average inventory, not the instantaneous inventory level, to simplify the model. In the above expression the net profit is obtained by subtracting the purchase, ordering, inventory and display costs from the revenue.

There are two components of the total inventory level (see Urban (1998) p. 20 and 21). The first part is when the backroom is being depleted and the showroom inventory is kept fully stocked. The second part is when there is no product left in the backroom inventory and the showroom inventory is being depleted. In the first part the total inventory level will decrease linearly since the demand rate is constant, but the showroom inventory level will stay the same and be fully stocked. Therefore, the amount of displayed inventory will be equal to the space allocated to that product. In the second part the showroom inventory will be depleting and so the demand rate will also decrease. It has been show by Silver and Peterson (1985, p. 176) that if there is a deterministic, constant demand rate and instantaneous replenishment, it is optimal to reorder when the inventory level reaches to zero. This means that it is optimal to reorder when all the inventory is depleted. On the other hand, the retailer does not want to take the risk for very low demand rates. So the best time to reorder is sometime in the second part of the period. Therefore, the net profit of a product is calculated as

$$NP_i = \frac{[(p_i - c_i)q_i - OC]\pi_i(1-\alpha_i)x_i^{\alpha_i} \prod_{\substack{j \in I^+ \\ j \neq i}} x_j^{\beta_{ij}} \left[1 + \sum_{k \in I^-} (1-\theta_k) f(a_k, \beta_{ki})\right]}{(1-\alpha_i)(q_i + r_i) + \alpha_i x_i - r_i^{1-\alpha_i} x_i^{\alpha_i}} - \frac{\frac{HC_i(1-\alpha_i)}{2(2-\alpha_i)}[(2-\alpha_i)(q_i+r_i)^2 + \alpha_i x_i^2 - 2r_i^{2-\alpha_i} x_i^{\alpha_i}]}{(1-\alpha_i)(q_i + r_i) + \alpha_i x_i - r_i^{1-\alpha_i} x_i^{\alpha_i}} - DC_i x_i$$

where the first part takes care of the revenue and purchase and ordering costs, the second part is the average inventory cost (see Urban (1998) for more details) and the last part is the display cost. Since the demand is a function of the displayed inventory level of the products in the assortment, the demand rate of a product will change every instance when the level of display inventory falls below the space allocation of that product. Therefore, they simplified the evaluation of the demand function and approximate it as a total shelf space allocation of all the other products in the assortment and neglected the instantaneous inventory level case. Hence, HC_i depends on this average inventory (of the total inventory which contains the backroom and the showroom inventory) level. The objective of the model is to maximize the net profit from the products, which have been selected in the assortment, and can be formulated as

$$\begin{aligned} \max \quad & \sum_{i=1}^{I} y_i NP_i \\ \text{s. t.} \quad & \sum_{i=1}^{I} x_i u_i \leqslant X, \end{aligned} \tag{2.2.6.2}$$

$$\sum_{i=1}^{I}(q_i + r_i)u_i \leqslant BX, \tag{2.2.6.3}$$

$$x_i \leqslant Xy_i, \forall_i \tag{2.2.6.4}$$

$$r_i \leqslant BXy_i, \qquad \forall_i \tag{2.2.6.5}$$

$$q_i \leqslant BXy_i, \qquad \forall_i \tag{2.2.6.6}$$

$$Q_i^{min}y_i \leqslant q_i \leqslant Q_i^{max}, \qquad \forall_i \tag{2.2.6.7}$$

$$r_i \leqslant x_i \leqslant q_i + r_i', \qquad \forall_i \tag{2.2.6.8}$$

$$X_i^{min} \leqslant x_i \leqslant X_i^{max}, \qquad \forall_i \tag{2.2.6.9}$$

$$y_i = \{0, 1\}, \qquad \forall_i \tag{2.2.6.10}$$

$$q_i, r_i, x_i \geqslant 0, \qquad \forall_i. \tag{2.2.6.11}$$

Constraint (1) and (2) represent the showroom and backroom space capacity limit, respectively. The parameter u_i allows for different measure units such as facings, packages, cases. Constraint sets (3), (4) and (5) make sure that if a product is not selected ($y_i = 0$) then the number of products, reorder and order quantities are set to zero, respectively. Constraints (6) represent maximum and minimum bounds on the order quantity, usually imposed by a supplier. Note including binary variable on the left hand side of this constraint is necessary (Urban did not include it) otherwise the problem will be infeasible because of the inconsistency between constraints (5) and (6) when $y_i = 0$. Constraint set (7) ensures that the products allocated to the shelf space exceed the reorder point ($r_i \leqslant x_i$) and the replenishment provides at least enough quantity to cover the allocated space ($q_i + r_i' \geqslant x_i$, where r_i' is certain inventory level). Constraint set (8) represents bounds on the number of product i to be allocated to the showroom inventory, usually imposed by the retailer.

2.2.7 Yang and Chen (1999)

The first demand model which considered the location of the product on different shelves in a product category and its effect on sales was Yang and Chen (1999). Their optimization model also introduced different shelf space sizes (different width-length of showroom-display shelves) for each shelf which can be very useful in retail practice. They used the same assumptions as those of Corstjen and Doyle (1981). Their optimization model's objective is also to maximize the total profit from the products in the assortment, where the number of products to be allocated to each shelf (x_{im}) is the decision variable. As in the basic model of Corstjen and Doyle (1981), the cost function can be represented as

$$C_i(x_i) = \phi_i \left(\sum_{m=1}^{M} D_{im}(x_{im}) \right)^{\lambda_i} = \phi_i \left(\sum_{m=1}^{M} \pi_i x_{im}^{\alpha_{im}} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ij}} \prod_{z=1}^{Z} f_{zi}^{\gamma_{zi}} \right)^{\lambda_i}$$

The optimization model is to

$$\max \quad \sum_{i=1}^{I} \left[g_i \left(\sum_{m=1}^{M} \pi_i x_{im}^{\alpha_{im}} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ij}} \prod_{z=1}^{Z} f_{zi}^{\gamma_{zi}} \right) - \phi_i \left(\sum_{m=1}^{M} \pi_i x_{im}^{\alpha_{im}} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ij}} \prod_{z=1}^{Z} f_{zi}^{\gamma_{zi}} \right)^{\lambda_i} \right]$$

$$\text{s. t.} \quad \sum_{i=1}^{I} w_i x_{im} \leqslant L_m, \quad \forall_m \qquad (2.2.7.1)$$

$$X_i^{min} \leqslant \sum_{m=1}^{M} x_{im} \leqslant X_i^{max}, \qquad \forall_i \qquad (2.2.7.2)$$

$$\sum_{m=1}^{M} \pi_i x_{im}^{\alpha_{im}} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ij}} \prod_{z=1}^{Z} f_{zi}^{\gamma_{zi}} \leqslant Q_i^{max}, \qquad \forall_i \qquad (2.2.7.3)$$

$$x_{im} \in \mathbb{N}^{+}, \qquad \forall_{i,m} \tag{2.2.7.4}$$

Constraint set (1) enforces the total shelf space limit in the width-length measure for each shelf m. Constraint set (2) sets bounds on the total number of display facing of product i as per the retailer requirements. Constraint set (3) makes sure that the demand for product i does not exceed the supply limit of that product. Finally, constraint set (4) ensures integrality of the solution (number of display facing).

They also proposed an alternative model that can be considered as an extension of a knapsack problem. Within a small range of the display facing amount (number of facings) for a product, they assume that the profit is linear which can be controlled using X_i^{min} and X_i^{max} sizes. Thus, using per facing profit for a product (v_i) on a shelf, their objective is still to make the maximum profit from the products in the assortment. They transformed the shelf space allocation problem to an extension of a knapsack problem where the objective is to

$$\max \sum_{i=1}^{I} \sum_{m=1}^{M} v_{im} x_{im}$$

subject to the same constraints except constraint set (3) since they assumed that the profit is linear within a small range of the display facing amount of a product.

The main problem of the alternate model is that it ignores the cross effects between products since the process of picking the highest gross profit margin products neglects the complementary and substitute effects among products within the assortment or in the store. Murray et al. (2010) have also disputed the validity of this assumption arguing that assuming a minimum number of display facings contradicts the minimum required display facings or the product inventory availability assumption.

2.2.8 Bookbinder and Zarour (2001)

Bookbinder and Zarour incorporated direct the product profitability (DPP) approach within the optimization models in order to enhance the utility of the models. Since DPP is a method to calculate an item's net profit margin, which goes beyond the gross margin of a product and accounts for all the discounts, deals and promotions per product from different supply chain parties, it is a more reliable way to estimate the net profit of a product within the assortment. DPP accounts for all the revenue and cost elements of a product in a unified analysis (As Stern and El-Ansary 1992). However, because DPP is static, it does not account properly for costs that vary depending on the order and inventory level such as procurement and transportation.

Bookbinder and Zarour (2001) incorporated the DPP approach in order to provide a better way to estimate how the space allocation of a product, and its compliment and substitute products, may affect a SKU's profitability. They use the Corstjen and Doyle (1981) demand model which incorporates the own-space and cross-space effects. The DPP approach is included in the objective function of the problem in order to calculate the net profit from a product in the assortment. Thus, the new adjusted gross profit margin is computed by multiplying the DPP margin rate ($DPPR_i = \frac{DPP_i}{DPC_i}$, where DPP_i is the direct product profit and DPC_i is the direct product cost) of a product with the unit demand of a product. As Corstjen and Doyle (1981) mentioned in their study, the cost of each product should be computed for a better optimization model and thus DPC_i is used instead of ϕ in the DPP approach to refer to the total costs of a product. The optimization model is to

$$\max \quad \sum_{i=1}^{I} \left[DPPR_i \left(\pi_i l_i^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} l_j^{\beta_{ij}} \right) - DPC_i \left(\pi_i l_i^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I} l_j^{\beta_{ij}} \right)^{\lambda_i} \right]$$

$$\text{s. t.} \quad \sum_{i=1}^{I} l_i \leqslant L \tag{2.2.8.1}$$

$$L_i^{min} \leqslant l_i \leqslant L_i^{max}, \qquad \forall_i \tag{2.2.8.2}$$

$$D_i \leqslant Q_i^{max}, \qquad \forall_i \tag{2.2.8.3}$$

where constraint (1) represents the shelf space limit. Constraint set (2) sets the retailer bounds on the number of display facings for each product. Constraint set (3) represents the limit on sales because of the production or carrying space restrictions.

2.2.9 Irion et al. (2004)

Irion et al. (2004) optimization model is the second model, after Urban (1998), that considers the diminishing display inventory level and product selection process. But unlike Urban's model, they let the display inventory level reach zero and replenish when there is no product left on the shelf. Furthermore, they implicitly assume a constant demand rate in the time period of a decreasing inventory level. This is not a reasonable assumption since shelf space allocated to a product is the main marketing variable that affects demand. As such one would expect the demand rate to change if the shelf space is changing. The following is a summary of their major assumptions:

1. All shelves are owned by the retailer.

2. There is no backroom inventory to stock products in the warehouse.

3. A product can be stacked on top of the other until there is no space left to stack one more.

4. There is backroom space to stock products behind display facings.

5. The products are being depleted from the showroom (displayed) inventory but the replacement from the backroom space is instantaneous. This means that the showroom shelves are being fully stocked and the backroom space is being depleted until there is no product in the backroom space left. The replacement cost is negligible.

6. When the backroom space is being depleted until all the products are gone, the demand rate of the product is kept constant.

7. There is a limited dedicated display area (shelf space) for the product and when other products are being depleted from the showroom inventory (in case their backroom is depleted already), the product can not be allocated to the other products dedicated display area. It means that the limited shelf space for the product can not be exceeded.

8. Unit cost of each product contains all procurement costs.

9. Products are replenished individually from the manufacturer or the supplier. This means that the order cost for each product is different since procurement (order placement and transportation) costs are different.

10. As soon as the number of products on the shelves reaches zero, the product on the shelf will be replenished by the amount that the shelf would take (considering both display facing and the backroom space).

11. There exist related costs such as, holding costs, unit replenishment costs (covers product insurance, deterioration and processing costs of sending products back to the supplier if there are complications on the product or they are no longer needed), fixed replacement (re-arraging, re-stacking) costs to restack the products to the shelves and fixed costs to include a product into the assortment.

Based on the above assumptions, they assume one, horizontal, homogeneous in quality shelf space for the showroom inventory. Since they do not consider a backroom inventory, there is no space to stock the product except behind the display facing. Therefore, their model does not need to determine the order quantity or the reorder point. The only decision variables are the binary selection variables (y_i) and variables corresponding to the number of products i's to be allocated (x_i). The objective is to maximize the net profit from all products in the assortment. They model the net profit for a product per day as

$$\begin{aligned} NP_i &= (p_i - c_i)D_i - GC_i D_i - HC_i AI_i - RC_i \left(\frac{D_i}{x_i NF_i} \right) \\ &= (p_i - c_i)D_i - GC_i D_i - (c_i \, IR) \left(\frac{x_i NF_i}{2} \right) - RC_i \left(\frac{D_i}{x_i NF_i} \right) \\ &= \left(p_i - c_i - GC_i - \frac{RC_i}{x_i NF_i} \right) D_i - (c_i \, IR) \left(\frac{x_i NF_i}{2} \right). \end{aligned}$$

The optimization model is

$$\max \quad \sum_{i=1}^{I} y_i \left[\left(p_i - c_i - GC_i - \frac{RC_i}{x_i NF_i} \right) D_i - (c_i \, IR) \left(\frac{x_i NF_i}{2} \right) - IC_i \right]$$

$$\text{s. t.} \quad \sum_{i=1}^{I} y_i w_i x_i \leqslant L \tag{2.2.9.1}$$

$$(w_i x_i - w_i)(w_i y_i - w_i) \geqslant 0, \qquad \forall_i \tag{2.2.9.2}$$

$$X_i^{min} \leqslant x_i \leqslant X_i^{max}, \qquad \forall_i \tag{2.2.9.3}$$

$$x_i \in \{1, 2, 3...\}, \qquad \forall_i \tag{2.2.9.4}$$

$$y_i \in \{0, 1\}, \qquad \forall_i \tag{2.2.9.5}$$

where IC_i is an insertion cost that is subtracted from the revenue. Constraint (1) models the space availability. Constraint set (2) ensures that there would be at least one product i in the assortment if product i is not chosen for the assortment. This is introduced to avoid the case where x_i takes a value of zero resulting in negative demand. Constraint set (3) represents the retailers bounds on the space to be allocated for each product. Constraint sets (4) and (5) represent integrality of the space variable and the binary nature of the selection variable, respectively.

Recently, Irion et al. (2011) extended this model by including the price effect without the cross-price effect.

2.2.10 Reyes and Frazier (2005)

Reyes and Frazier (2005) consider inventory ordering and holding costs in a basic optimization model:

$$\max \quad \sum_{i=1}^{I} (p_i - c_i) D_i - \frac{OC_i}{x_i} D_i - HC_i x_i$$

$$\text{s. t.} \quad \sum_{i=1}^{I} x_i = X \tag{2.2.10.1}$$

$$x_i^{base}(1 - \xi) \leqslant x_i \leqslant x_i^{base}(1 + \xi), \qquad \forall_i \tag{2.2.10.2}$$

$$x_i > 0, \qquad \forall_i \tag{2.2.10.3}$$

$$x_i \in \{1, 2, 3, ...\}, \qquad \forall_i \tag{2.2.10.4}$$

Constraint (1) represents the space limit and constraint set (2) defines the bounds for the space variable. We can see from their optimization model that they assume one, horizontal, homogeneous in quality shelf space and same product sizes for all the products in the assortment.

2.2.11 Hwang et al. (2005)

Hwang et al. consider the location effect on multi-level shelves and incorporate the inventory level effects on the optimization model, by including the backroom inventory. But in their model, this backroom inventory is not interpreted as a replenishment stock as in Urban (1998) model, but as the “backroom space” where products are placed in the back side of the display facing products. Their model also has some other assumptions:

1. The display area is always kept fully stocked.

2. The products are being depleted from the showroom (displayed) inventory but the replacement from the backroom inventory (from now on backroom space) is instantaneous. This means that the showroom shelves are being fully stocked and the backroom space is being depleted until there is no product in the backroom space left. The replacement cost is negligible.

3. Since the backroom space is being depleted until all the products are gone in the backroom space, the showroom inventory is fully stocked and therefore the demand rate is constant.

4. The products are ordered with an (r, Q) policy and the total order is being received in the backroom space.

5. The total inventory replenishment is also instantaneous with a known and constant lead time. It means that the entire order from the suppliers is received instantaneously as soon as the inventory is depleted. Replenishments are independent for each product in the assortment. It means that there is no joint replenishments.

6. All the of the prices and costs (selling price, unit purchase cost of the products, holding costs, carrying costs, display costs) are known and constant.

7. The shelves in the category have different shelf space sizes (different width-length).

Since the demand rate is constant, the inventory level will decrease linearly. Because the assumption was to keep display facings fully stocked, the retailer should order before all products in the backroom space are depleted. They use a zero inventory order policy: order only when the inventory level reaches to zero. This means that the reorder point is x_i for all products where x_i is the inventory level when all the products stock in the backroom space is depleted and only the fully stocked display facing inventory is left. The decision variables are q_i, x_i and x_{im}. Since the cycle time for a product is $CT_i = \frac{q_i}{D_i}$ and the average inventory for a product is $AI_i = \frac{2x_i+q_i}{2}$ (see Hwang et al. 2005 p. 188) then the net profit for a product during a unit period can be defined as

$$
\begin{aligned}
NP_i &= \frac{(p_i - c_i)q_i}{CT_i} - \frac{OC_i}{CT_i} - HC_i AI_i - DC_i x_i \\
&= \frac{(p_i - c_i)q_i - OC_i}{\frac{q_i}{D_i}} - HC_i \frac{2x_i + q_i}{2} - DC_i x_i \\
&= \left[(p_i - c_i) - \frac{OC_i}{q_i}\right] D_i - HC_i \frac{2x_i + q_i}{2} - DC_i x_i
\end{aligned}
$$

The optimization model can be expressed as

$$\max \quad \sum_{i=1}^{I} \left[\left[(p_i - c_i) - \frac{OC_i}{q_i} \right] D_i - HC_i \frac{2x_i + q_i}{2} - DC_i x_i \right]$$

$$\text{s. t.} \quad \sum_{i=1}^{I} w_i x_{im} \leqslant L_m, \qquad \forall_m \tag{2.2.11.1}$$

$$X_i^{min} \leqslant \sum_{m=1}^{M} x_{im} \leqslant X_i^{max}, \qquad \forall_i \tag{2.2.11.2}$$

$$x_i = \sum_{m=1}^{M} x_{im}, \qquad \forall_i \tag{2.2.11.3}$$

$$x_{im} \geqslant 0, \qquad \forall_{i,m} \tag{2.2.11.4}$$

$$q_i > 0, \qquad \forall_i \tag{2.2.11.5}$$

Constraint set (1) ensures that the space allocated to a shelf does not exceed the shelf space size (width-length). Constraint set (2) represents the bounds on the total amount of display facing for each product. Constraint set (3) defines the shelf space variable for a product as the sum of all its allocations in all shelves. Finally, constraint sets (4) and (5) represent non-negativity for the shelf storage variable and strict positiveness for the order quantity variable.

2.2.12 Reyes and Frazier (2007)

Reyes and Frazier (2007) consider two important performance measure: profitability in terms of improving return on inventory investment through incremental sales, and thereby increasing profit margins, and customer service in terms of reducing the out-of-stock situations. This is achieved by having enough amount of shelf space dedicated to every product within the assortment relative to the proportion of each product's demand to the total available demand (Yang and Chen 1999; Yang 2001). They model the problem as a nonlinear integer weighted goal program for which the solution can be evaluated through weighting the trade-off between profitability and customer service level. Therefore, while maximizing the first objective, which is the profitability objective, will try to allocate more space to products which have higher profit margins, minimizing the second objective (since the second term has a negative sign in the objective function, the model attempts to minimize that objective), which is related to customer service, will try to allocate a base level to each product to satisfy all possible consumer demand for each product and reduce the risk of stock-outs. Accordingly, the number of products to be allocated (x_i), under-achievement and over-achievement goal variables (uv_i and ov_i, respectively) are defined as the decision variables.

Their model is

$$\max \quad weight \sum_{i=1}^{I} \left[(p_i - c_i)\pi_i \left(\frac{x_i}{x_i^{base}} \right)^{\alpha_i} \left(\frac{p_{avg}}{p_i} \right)^{\sigma_i} \right] - (1 - weight) \sum_{i=1}^{I} uv_i$$

$$\text{s. t.} \quad \sum_{i=1}^{I} x_i = X \qquad (2.2.12.1)$$

$$x_i + uv_i - ov_i = x_i^{base}, \qquad \forall_i \qquad (2.2.12.2)$$

$$x_i \geqslant (1 - \xi) x_i^{base}, \qquad \forall_i \qquad (2.2.12.3)$$

$$x_i, uv_i, ov_i \geqslant 0, \qquad \forall_i \qquad (2.2.12.4)$$

$$x_i \in \{1, 2, ...\}, \qquad \forall_i. \qquad (2.2.12.5)$$

Constraint (1) represents the space limit. Constraint set (2) represents the goal constraints for the base level allocations. Constraint set (3) places a lower bound on the space variable. Constraint sets (4) and (5) represent the non-negativity and integrality requirements, respectively.

They also proposed an alternate approach, which was a nonlinear integer program, however not a goal program, to avoid the decision maker to make the choice of desired weights for the objectives with goal programming. They find the best composition of the two objectives, maximizing the profitability and minimizing the mean absolute deviation (MAD) between the actual shelf space allocation and base level allocation. The purpose is to maximize the combined average percentage achievement of the two objectives. To do so, new decision variables are defined such as, total profit for a solution ($P_{solution}$), best profit (P_{best}), worst profit (P_{worst}) and percent of optimal profit achieved with a solution ($P\%_{solution}$) and MAD for a solution ($MAD_{solution}$), best value of MAD (MAD_{best}), worst value of MAD (MAD_{worst}) and percent of optimal MAD value achieved with a solution ($MAD\%_{solution}$). The aim of their alternate approach model is to maximize the average percentage achievement of the two objectives, profitability and MAD. The alternate model is defined as

$$\max \quad 100\frac{P\%_{solution} + MAD\%_{solution}}{2}$$

$$\text{s. t.} \quad \sum_{i=1}^{I} x_i = X \tag{2.2.12.6}$$

$$x_i \geqslant (1-\xi)x_i^{base}, \qquad \forall_i \tag{2.2.12.7}$$

$$P_{solution} = \left[(p_i - c_i)\pi_i \left(\frac{x_i}{x_i^{base}}\right)^{\alpha_i} \left(\frac{p_{avg}}{p_i}\right)^{\sigma_i}\right] \tag{2.2.12.8}$$

$$P\%_{solution} = \frac{P_{solution} - P_{worst}}{P_{best} - P_{worst}} \tag{2.2.12.9}$$

$$MAD_{solution} = \frac{1}{I}\sum_{i=1}^{I} |x_i - x_i^{base}| \tag{2.2.12.10}$$

$$MAD\%_{solution} = \frac{MAD_{worst} - MAD_{solution}}{MAD_{worst} - MAD_{best}} \tag{2.2.12.11}$$

$$\text{All variables} \geqslant 0 \tag{2.2.12.12}$$

$$x_i \in \{1, 2, ...\}, \qquad \forall_i \tag{2.2.12.13}$$

2.2.13 Hariga et al. (2007)

Hariga et al. (2007) considered many decision variables: the variety of the products (product selection, y_i), the various locations a product is being displayed within a store (y_{im}), the total order quantity of the product (q_i) and their order quantity to each location (q_{im}), number of products to be allocated (x_i) and number of products to be allocated in each location (x_{im}), cycle time of the product (CT_i) and cycle time of the location (CT_{im}) and average inventory of the product in a particular location (AI_{im}). They also made a clear distinction between the showroom inventory and the backroom inventory. They made several assumptions that are similar to those made in Urban (1998) and defined the new assumptions as:

1. The products are being depleted from the showroom (displayed) inventory but the replenishment from the backroom inventory is instantaneous. This means that the showroom shelves are being fully stocked and the backroom inventory is being depleted.

2. While the backroom inventory is being depleted, the showroom inventory is fully stocked and therefore the demand rate is constant for that time interval.

3. Once the backroom inventory is depleted for a product in a shelf (or a display location) then the showroom inventory will be depleted and the demand rate

of the product will start to decrease as the showroom inventory level decreases (see Hariga et al. 2007, p. 243 and 244).

4. The total order is being received in the backroom inventory.

5. The total inventory replenishment is instantaneous with a known and constant lead time. It means that the entire order from the suppliers is received instantaneously as soon as the inventory is depleted.

6. There is a limited dedicated display area (shelf space) for the product and when other products are being depleted from the showroom inventory, the product can not be allocated to the other products' dedicated display area. It means that the limited shelf space for the product can not be exceeded.

7. All the of the prices and costs (selling price, unit purchase cost of the products, holding costs, carrying costs, display costs) are known and constant.

In Urban (1998) the retailer did not want to take the risk for very low demand rates and tried to find the optimal reorder point, which was in the second period of the cycle time and did not let the inventory level of the product reach to zero. In contrast, Hariga et al. (2007) assume that products will not be replenished from the backroom inventory until all products stock allocated to the shelf is totally depleted. While Urban's model does not let the showroom inventory reach zero, because of the possible decrease of the demand, Hariga et al. let the showroom inventory reach zero for all the products in the same shelf (or location). Hariga et al. do not allow an unstocked product i in shelf m (location) to be replenished as soon as its inventory is depleted, rather they wait until all the other stock of products in shelf m is depleted. In their model the new inventory cycle begins at that time when the shelf is replenished, since inventory replenishment is instantaneous. For a detailed explanation of the model see Hariga et al. p. 245–247. The net profit per unit of time for a product can be expressed as

$$NP_i(x_{im}, q_{im}) = \sum_{m=1}^{M} \left[\frac{(p_i - c_i)q_{im}}{CT_i} - \frac{OC_i}{CT_i} - HC_iAI_i - SC_i(q_{im} - x_{im}) - DC_ix_{im} \right]$$

and therefore the profit per a unit of time for a product whether it is included in the assortment or not can be expressed as

$$NP_i(x_{im}, q_{im}) = \sum_{m=1}^{M} \left[\frac{(p_i - c_i)q_{im} - OC_iy_i - HC_iAI_{im} - SC_i(q_{im} - x_{im})CT_{im} - DC_ix_{im}CT_{im}}{(1 - y_i + CT_iy_i)} \right] \tag{2.2.13.1}$$

and the optimization model is defined as

$$\max \quad \sum_{i=1}^{I} \sum_{m=1}^{M} \left[\frac{(p_i - c_i)q_{im} - OC_iy_i - HC_iAI_{im} - SC_i(q_{im} - x_{im})CT_{im} - DC_ix_{im}CT_{im}}{(1 - y_i - CT_iy_i)} \right]$$

$$\text{s. t.} \quad \sum_{i=1}^{I} u_ix_{im} \leqslant X_m, \qquad \forall_m, \tag{2.2.13.2}$$

$$\sum_{i=1}^{I} u_iq_i \leqslant BX \tag{2.2.13.3}$$

$$x_i = \sum_{m=1}^{M} y_{im}x_{im}, \qquad \forall_i \tag{2.2.13.4}$$

$$q_i = \sum_{m=1}^{M} y_{im}q_{im}, \qquad \forall_i \tag{2.2.13.5}$$

$$x_i \leqslant X_i^{max}, \qquad \forall_i \tag{2.2.13.6}$$

$$q_i \leqslant Q_i^{max}, \qquad \forall_i \tag{2.2.13.7}$$

$$y_{im}X_i^{min} \leqslant x_{im} \leqslant y_{im}X_i^{max}, \qquad \forall_i \tag{2.2.13.8}$$

$$q_{im} \leqslant y_{im} Q_i^{max}, \qquad \forall_{i,m} \tag{2.2.13.9}$$

$$x_{im} \leqslant q_{im}, \qquad \forall_{i,m} \tag{2.2.13.10}$$

$$D_{im}(0) = \pi_i x_{im}^{\alpha_{im}} \prod_{\substack{j \in I^+ \\ j \neq i}} (1 - y_j + x_j y_j)^{\beta_{ij}}, \qquad \forall_{i,m} \tag{2.2.13.11}$$

$$CT_{im} = \frac{q_{im} - x_{im}\left(\frac{x_{im}}{1-\alpha_{im}}\right)}{1 - y_{im} + y_{im} D_{im}(0)}, \qquad \forall_{i,m} \tag{2.2.13.12}$$

$$0 \leqslant CT_{im} \leqslant y_{im} CT_i, \qquad \forall_{i,m} \tag{2.2.13.13}$$

$$CT_i \leqslant y_i CT_i^{max}, \qquad \forall_i \tag{2.2.13.14}$$

$$AI_{im} = \frac{1/2(q_{im} - x_{im})^2 + x_{im}\left(q_{im} - \left[\frac{1-\alpha_{im}}{2-\alpha_{im}}\right] x_{im}\right)}{1 - y_{im} - D_{im}(0) y_{im}}, \quad \forall_{i,m} \tag{2.2.13.15}$$

$$y_{im} \leqslant y_i, \qquad \forall_{i,m} \tag{2.2.13.16}$$

$$y_i \leqslant \sum_{m=1}^{M} y_{im}, \qquad \forall_i \tag{2.2.13.17}$$

$$y_{im} = \{0, 1\}, \qquad \forall_{i,m} \tag{2.2.13.18}$$

$$y_i \in \{0, 1\}, \qquad \forall_i. \tag{2.2.13.19}$$

Constraint set (1) represents the storage limit for each shelf. Constraint (2) takes care of the backroom inventory space limit. Constraint sets (3) and (4) defines the logical relationship between a product's total allocated number of units and order quantity and the allocations and orders sizes to individual shelves, respectively. Constraint set (5)–(9) set the appropriate limits on shelf allocations and order quantities. Constraint set (10) defines the demand rate at time 0 for product i in shelf m, $D_{im}(0)$, for the selected products, which accounts for the own-space elasticity of the product and cross-elasticities among products. Constraint set (11) defines the cycle time for selected products, which depends on the order quantity, space allocation, space elasticity and the demand function of the product on shelf m. Constraint set (12) ensures that if product i is in the assortment of shelf m then the cycle time of product i on shelf m does not exceed the cycle time of product i. Constraint set (13) defines the bound on a product cycle time, CT_i, as the maximum value among all its cycle times on the shelves. Constraint set (14) defines the average inventory of product i on shelf m per unit of time as a function of the average backroom and showroom inventories. Constraint sets (15) and (16) define the logical relationships between the binary variables for product selection and shelf placement. Finally constraint sets (17) and (18) state the binary variables.

2.2.14 Silva et al. (2009)

Silva et al. (2009) use Yang and Chen (1999)'s idea of modeling the shelf space allocation model as a modified version of a multidimensional knapsack problem. They consider the shelf levels (vertically) and shelf parts (horizontally). They model this problem for convenience stores and accordingly made the following assumptions:

1. One shelf block consists of three shelves in a horizontal dimension: the top level, eye level and bottom level.

2. Each shelf consists of three parts in a vertical dimension: the right part, middle part and left part.

3. The shelves which are on the eye level and middle parts of each shelf are assumed to have the highest profitable parts.

4. Each shelf and each part has a priority based on their position since each section contributes to profits differentially.

5. There is no stacking process and therefore the height of the products are ignored in the model. However, the shelf is tall enough to accommodate each products' height.

6. The depth of the products are also ignored.

7. Each products' display facing should be located on the same shelf. This assumption ignores the situation of having multiple display facings on different shelves.

Their decision variables are the number of products to be allocated on a part of a shelf (x_{ime}) and a binary variable to determine whether the product is located on a part of shelf or not (y_{ime}). The modified version of Yang and Chen (1999) optimization model is then defined as

$$\max \quad \sum_{i=1}^{I}\sum_{m=1}^{M}\sum_{e=1}^{E} v_{ime}x_{ime}$$

$$\text{s. t.} \quad \sum_{i=1}^{I} w_i x_{ime} \leqslant L_{me}, \qquad \forall_{m,e} \tag{2.2.14.1}$$

$$\sum_{m=1}^{M}\sum_{e=1}^{E} y_{ime} = 1, \qquad \forall_i \tag{2.2.14.2}$$

$$X_i^{min} \leqslant \sum_{m=1}^{M}\sum_{e=1}^{E} x_{ime} \leqslant X_i^{max}, \qquad \forall_i \tag{2.2.14.3}$$

$$y_{ime} \leqslant x_{ime} \leqslant X_i^{max} y_{ime}, \qquad \forall_{i,m,e} \tag{2.2.14.4}$$

$$x_{ime} \in \mathbb{N}^+, \qquad \forall_{i,m,e} \tag{2.2.14.5}$$

$$y_{ime} \in \{0, 1\}, \qquad \forall_{i,m,e} \tag{2.2.14.6}$$

Constraint set (1) takes care of space limitations. Constraint set (2) ensures that product i can only be located to a particular part of a particular shelf. Constraint sets (3) and 4 models the bounds on the total number of display facing of product i, depending on whether the product is located on a part of a shelf or not. Finally, constraint sets (5) and (6) represents the positive number of display facings for each product and the binary variable requirements.

The authors mention the potential weakness of the model when shelf parts doe not correspond to the space where products are allocated. In such cases, there would be an unused space and potentially lower profitability.

2.2.15 Murray et al. (2010)

Murray et al. (2010) consider display facing areas of a product, their various display orientations on a shelf and the location of the product within a product category. Their model is also the first optimization model that considers the product price (p_i) with shelf space allocation (display facing areas, shelf locations and display orientations, x_{iom}) as decision variables. Since they consider the display facing area of a product, unlike existing models which considers the width of the product , they take into account the height of the product too. However, since they allow the products to be placed on a shelf in diverse orientations, for each orientation of the product, different dimension lengths are used as the width and the height of the product. The height of the shelf also became an important parameter because of product stacking. They made the following assumptions:

1. The products to be allocated have already been chosen.

2. Each display shelf has different dimension sizes for each dimension.

3. The model does not consider the backroom inventory (space) which is behind the product display facings.

4. The use of shelf depth dimension and product depth dimension is solely for the purpose of ensuring the product's depth does not exceed the shelf's depth.

5. Product i can only be stacked on top of itself and it must be in the same orientation.

The gross profit of a product can be defined as

$$P_i = (p_i - c_i)\ D_i(p, x_i)$$

The aim of their optimization model is to maximize the gross profit from the products:

$$\begin{aligned}
\max \quad & \sum_{i=1}^{I} (p_i - c_i)\ D_i(p, x_i) \\
\text{s. t.} \quad & \sum_{i=1}^{I} \sum_{o=1}^{3} w_{io} x_{iom} \leqslant L_m, \qquad \forall_m && (2.2.15.1) \\
& x_{iom} = 0 \quad \text{for all } i,\, o,\, m \text{ such that} \quad \bar{x}_{iom} = 0 && (2.2.15.2) \\
& x_{iom} = 0 \quad \text{for all } i,\, o,\, m \text{ such that} \quad \lfloor \frac{D_m}{d_{io}} \rfloor = 0 && (2.2.15.3) \\
& X_i^{min} \leqslant \sum_{o=1}^{3} \sum_{m=1}^{M} \bar{x}_{iom} x_{iom} \leqslant X_i^{max}, \qquad \forall_i && (2.2.15.4) \\
& P_i^{min} \leqslant p_i \leqslant P_i^{max}, \qquad \forall_i && (2.2.15.5) \\
& x_{iom} \in \{0, 1, 2, ...\}, \qquad \forall_i && (2.2.15.6)
\end{aligned}$$

Constraint set (1) ensures that the space allocated to all products (width of all products) on shelf m does not exceed the space (width) of shelf m. Constraint sets (2) and (3) ensure that if the product i is placed on shelf m in orientation o and product's height or depth exceed the shelf's height or depth, respectively, then the product can not be placed on shelf m in orientation o. Constraint sets (4) and (5) put bounds on the products allocated space and price. Constraint (6) ensures the number of display facings is integer.

2.2.16 Russell and Urban (2010)

Underhill (1999) noted the positive effect on sales of the region from slightly above eye level to knee level (horizontally) and the region from one end to another end of the aisles (vertically). To model this effect, and inspired by Dreze et. al. (1994), Russel and Urban (2010) assumed sales to be quadratic in the horizontal and vertical dimensions. The objective function of their model is to maximize the net profit from products in the assortment. As Dreze et al. (1994) noted the importance of allocating space to products in a uniform column, they took into account the location of the product family within the shelf category and used a limit deviance (*deviance*) to control the distance from one shelf to its adjacent. Their decision variables are the location of product i (center of the products' display facings in a shelf, $\bar{c}_i$), the number of product i (x_{im}) to be allocated on shelf m, a binary variable which represents whether the product i will be placed on shelf m (y_{im}), a binary variable which represents whether a family will be placed on shelf m (y_{bm}), and (when the family is placed on shelf m) a continuous variables which represent the location of the family b from left end to the right end ($left_{bm}$ and $right_{bm}$) and top shelf to the bottom shelf

(top_b and $bottom_b$). The optimization model is

$$\max \quad \sum_{i=1}^{I} g_i D_i$$

$$\text{s. t.} \quad \bar{c}_i \geqslant \frac{w_i}{2} x_{im}, \qquad \forall_{i,m} \tag{2.2.16.1}$$

$$\bar{c}_i \leqslant L_m - \frac{w_i}{2} x_{im}, \qquad \forall_{i,m} \tag{2.2.16.2}$$

$$\sum_{m=1}^{M} y_{im} = 1, \qquad \forall_i \tag{2.2.16.3}$$

$$\sum_{m=1}^{M} x_{im} \geqslant X_i^{min}, \qquad \forall_i \tag{2.2.16.4}$$

$$x_{im} \leqslant X_i^{max} y_{im}, \qquad \forall_{i,m} \tag{2.2.16.5}$$

$$\bar{c}_i - \frac{w_i}{2} x_{im} \geqslant \bar{c}_j + \frac{w_i}{2} x_{jm} - L_m\, y'_{ij} - L_m(2 - y_{im} - y_{jm}), \quad \forall_i \tag{2.2.16.6}$$

$$y'_{ij} + y'_{ji} = 1, \qquad \forall_{i,j}(i < j) \tag{2.2.16.7}$$

$$left_{bm} \leqslant \bar{c}_{ib} - \frac{w_{ib}}{2} x_{ib,m} + L_m(1 - y_{im}), \qquad \forall_{i,b,m} \tag{2.2.16.8}$$

$$right_{bm} \geqslant \bar{c}_{ib} - \frac{w_{ib}}{2} x_{ib,m} - L_m(1 - y_{im}), \qquad \forall_{i,b,m} \tag{2.2.16.9}$$

$$right_{bm} - left_{bm} = \sum_{ib=1}^{I} w_{ib} x_{ib,m}, \qquad \forall_{b,m} \tag{2.2.16.10}$$

$$y_{bm} \leqslant \sum_{ib=1}^{I} y_{ib,m}, \qquad \forall_{b,m} \tag{2.2.16.11}$$

$$y_{bm} \geqslant y_{ib,m}, \qquad \forall_{i,b,m} \tag{2.2.16.12}$$

$$top_b \geqslant m\, y_{bm}, \qquad \forall_{b,m} \tag{2.2.16.13}$$

$$bottom_b \leqslant M - (M - m) y_{bm}, \qquad \forall_{b,m} \tag{2.2.16.14}$$

$$top_b - bottom_b = \sum_{m=1}^{M} y_{bm} - 1, \qquad \forall_b \tag{2.2.16.15}$$

$$top_b \geqslant bottom_b, \qquad \forall_b \tag{2.2.16.16}$$

$$left_{bm} - left_{b,m+1} \leqslant deviance + L_m(2 - y_{bm} - y_{b,m+1}), \quad \forall_{b,m}(m \leqslant M-1) \tag{2.2.16.17}$$

$$left_{b,m+1} - left_{bm} \leqslant deviance + L_m(2 - y_{bm} - y_{b,m+1}), \quad \forall_{b,m}(m \leqslant M-1) \tag{2.2.16.18}$$

$$right_{bm} - right_{b,m+1} \leqslant deviance + L_m(2 - y_{bm} - y_{b,m+1}), \quad \forall_{b,m}(m \leqslant M-1) \tag{2.2.16.19}$$

$$right_{b,m+1} - right_{bm} \leqslant deviance + L_m(2 - y_{bm} - y_{b,m+1}), \quad \forall_{b,m}(m \leqslant M-1) \tag{2.2.16.20}$$

Constraint sets (1) and (2) ensures that the center of product i sequence does not go beyond the shelf m's both ends. Constraint set (3) stipulates that each product can only be allocated to one shelf. Constraint sets (4) and (5) put bounds on the total number of product i to be allocated to the all shelves and the number items of product i placed on shelf m, respectively. Constraint set (6) states that the left end of the product i sequence is to the right of product j sequence's right end, if product i is not placed to the left of product j ($-L_m\, y'_{ij}$) or is not on a different shelf ($-L_m(2 - y_{im} - y_{jm})$). Constraint set (7) ensures that product i should be placed either to the left of product j or to the right of product j. Constraint sets (8) and (9) model the locations of both ends of family (brand) b in each shelf. Constraint set (10) ensures that all family (brand) members will be adjacent to each other. Constraint sets (11) and (12) represents logical relations for for the definition of the binary variable y_{bm} indicating whether family b is on shelf m or not. Constraint sets (13) and (14) model the location (top shelf and bottom shelf in a vertical dimension, respectively) of family b using the binary variables and shelf numbers. Constraint

(15) ensures that all the family members will be adjacent to each other in terms of shelves. Constraint set (16) secures logical conditions of the top and bottom shelf. Constraint sets (17–20) model the limits on adjacent shelves from their left (17–18) and right (19–20) sides.

2.2.17 Lotfi et al. (2011)

In Lotfi et al. (2011) each product is represented by the form of multiple brands. They considered costs related to nonproductive use of space along with the inventory investment cost, replenishment cost and inventory holding costs. They assumed an estimated minimum daily demand for a product which is used as an input for the model and tried to calculate the maximum (potential) demand due to the price change and space allocation. Their decision variables are the maximum daily demand of product i (UD_{ib}), order quantity of product i (q_{ib}), number of display facings and total width-length of product i to be allocated in showroom inventory (x_{ib} and s_{ib}), number of product i and total width-length of product i to be allocated in backroom inventory (x'_{ib} and s'_{ib}) and cycle time of the joint replenishment of all products under brand b (CT_b). The objective is to minimize the total weight penalty due to the deviations from different profit margins and the space allocated to a product (variables such as $nDPM$, $pDPM$, nDL_i, pDL_i, nDB_i, pDB_i are the deviations from desired target levels and SVS_{ib} is the slack variable). They considered a planning horizon (say three months) and assumed that:

1. Each product is represented by the form of multiple brands.
2. Data such as demand (daily demand), prices, and costs for the products are constant during the time period.

3. Seasonal variations and periodic trends are noticed in the estimation of the parameters.

4. If the products are under the same brand, then they are replenished jointly.

5. There is a limited number of replenishments during a period (especially for perishable products).

6. Products are being held in backroom inventory and showroom inventory.

7. A full shelf strategy is being applied.

8. The products are being depleted from the showroom (displayed) inventory but the replenishment from the backroom inventory is instantaneous. This means that the showroom shelves are being fully stocked and the backroom inventory is being depleted. This also means that backroom inventory decreasing rate is equal to the constant demand rate.

9. The replenishment lead-time for a brand is zero.

10. The best display facing of a product has already been chosen by the retailer.

11. The dimension sizes of the showroom inventory (display shelves) and backroom inventory (holding shelves) are predetermined by the retailer.

12. If the consumer faces a shortage of product i, then he will either purchase another product or does not purchase any product within the assortment. If the consumer does not purchase any product in case of a shortage, then the shortage is identified as a lost sale.

13. A penalty is introduced for the nonproductive use of shelf space, because to place a different kind of product on or behind of a product in showroom inventory (shelves) is not possible due to aesthetic reasons and accessibility limitations.

The optimization model is

$$\min \quad p_m\, nDPM + \frac{p_s}{DL}\sum_{i=1}^{I}(pDL_i + nDL_i) + \frac{p_{bi}}{DB}\sum_{i=1}^{I}(pDB_i + nDB_i)$$

$$\text{s. t.} \quad UD_{ib} = MinD_{ib}\frac{(1-\sigma'_{ib}PC_{ib})}{(1-\mu'_{ij}PC_{ib,j})}x_{ib}^{\alpha_i}, \qquad \forall_{i,b} \tag{2.2.17.1}$$

$$MinD_{ib}CT_b \leqslant q_{ib}\Omega'_{ib} \leqslant UD_{ib}CT_b, \qquad \forall_{i,b} \tag{2.2.17.2}$$

$$CT_b \leqslant ET_{ib}, \qquad \forall_{i,b} \tag{2.2.17.3}$$

$$\sum_{i=1}^{I}\sum_{b=1}^{B} l_{ib} \leqslant L \tag{2.2.17.4}$$

$$\sum_{i=1}^{I}\sum_{b=1}^{B} l'_{ib} \leqslant BL \tag{2.2.17.5}$$

$$x_{ib} = \frac{l_{ib}}{w_{ib}}, \qquad \forall_{i,b} \tag{2.2.17.6}$$

$$x'_{ib} = \frac{l'_{ib}}{w_{ib}}, \qquad \forall_{i,b} \tag{2.2.17.7}$$

$$q_{ib} + SVS_{ib} = \frac{x_{ib}NF_{ib}}{\Omega_{ib}} + x'_{ib}NB_{ib}, \qquad \forall_{i,b} \tag{2.2.17.8}$$

$$WD\sum_{i=1}^{I}\sum_{b=1}^{B}\frac{c_{ib}q_{ib}}{CT_b} \leqslant BU \tag{2.2.17.9}$$

$$WD\sum_{b=1}^{B}\frac{1}{CT_b} \leqslant NR \tag{2.2.17.10}$$

$$TCP = WD\sum_{i=1}^{I}\sum_{b=1}^{B}\frac{c_{ib}q_{ib}}{CT_b} \tag{2.2.17.11}$$

$$TCO = WD\sum_{b=1}^{B}\frac{OC_b}{CT_b} \tag{2.2.17.12}$$

$$TCH = \sum_{i=1}^{I} HC_i \sum_{b=1}^{B} \frac{c_{ib} q_{ib} \Omega'_{ib}}{2} \tag{2.2.17.13}$$

$$TCSP = p_s \sum_{i=1}^{I} \sum_{b=1}^{B} \frac{SVS_{ib} w_{ib}}{NB_{ib}} \tag{2.2.17.14}$$

$$TC = TCP + TCO + TCH + TCSP \tag{2.2.17.15}$$

$$TR = \sum_{i=1}^{I} \sum_{b=1}^{B} \frac{p_{ib} q_{ib} \Omega'_{ib}}{CT_b} WD \tag{2.2.17.16}$$

$$NP = TR - TC \tag{2.2.17.17}$$

$$\frac{NP}{TC} + (nDPM - pDPM) = DPM \tag{2.2.17.18}$$

$$\sum_{b=1}^{B} l_{ib} + nDL_i - pDL_i = DL_i, \qquad \forall_i \tag{2.2.17.19}$$

$$\sum_{b=1}^{B} l'_{ib} + nDB_i - pDB_i = DB_i, \qquad \forall_i \tag{2.2.17.20}$$

$$nDPM, pDPM, nPL_i, pDL_i, nDB_i, pDB_i, SVS_{ib} \geqslant 0, \quad \forall_{i,b} \tag{2.2.17.21}$$

$$x_{ib}, x'_{ib}, q_{ib} \in \{1, 2, ..\}, \qquad \forall_{i,b} \tag{2.2.17.22}$$

Constraint set (1) states that the maximum daily demand of product i depends on the minimum daily demand of product i, number of display facing and the price change effects for product i and its substitute (since positive PC_{ib} increases the demand for product i while positive $PC_{ib,j}$ decreases the demand for product i). Constraint set (2) puts limits on the order quantity of product i. Constraint set (3) states that the cycle time for a $brand_b$ should be less than the holding time limit of all the products under $brand_b$. Constraint sets (4) and (5) state showroom inventory and backroom inventory space limits, respectively. Constraint sets (6) and (7) ensure that the space allocated to product i is a multiple length of the product's width dimension so that the quantity of product i to place on shelves can be an integer value for showroom inventory and backroom inventory, respectively. Constraint set (8) represents the relationship between the order quantity of product i and its space allocation. Constraint set (9) sets a limit on the total cost of ordering for all the products in the assortment . Constraint set (10) puts a limit on the total number of replenishments for a $brand_b$. Constraint sets (11-14) define the cost components of the model. Constraints (15), (16) and (17) define the total cost, revenues, and net profit for the period, respectively. Constraint sets (18-20) represent the goal constraints of the model. Finally constraint sets (21-22) represent non-negativity and integrality requirements.

2.3 A Critical Discussion about the Literature

From the reviewed literature on shelf space allocation optimization models we infer that there are two main modeling approaches: models that incorporate space elasticity in the demand function and models that assume a linear relation between shelf space and sales. Although a large majority of the models considered space elasticity functions, some of the models such as Yang and Chen (1999), Yang (2001), Lim et al. (2004), Silva et al. (2009) and Hansen et al. (2010) assumed a linear relation in a knapsack-like models and provided solution algorithms in order to solve the linear version much faster. Yang (2001) proposed the first heuristic to the linear formulation of the shelf space allocation problem. The solution heuristic is composed of four phases:

(1) A preparatory phase where enough shelf space is being checked to allocate the minimum required number of facings for all the products and assigning a priority

index to each product based on their "profit weight" (g_i/w_i).

(2) An allocation phase where an initial solution is designed by assigning shelf space to each product following the ordered priority indexes with minimum required number of facings.

(3) An adjustment phase which consists of three adjustment methods. The first adjustment swaps a display facing between two products allocated on the same shelf. The second adjustment swaps a facing between two products allocated on different shelves. The last adjustment, an extension of the second adjustment, looks for additional shelf space to allocate more products before swapping products on different shelves again.

(4) A termination phase where the total gross profit of the shelf space allocation is being calculated.

Lim et al. (2004) and Silva et al. (2009) improved the efficiency of this heuristic by using different elements in the above phases. Hansen et al. (2010) proposed a genetic algorithm as a method to improve the outcome of the shelf space allocation problem in their comparative analysis study.

The justification for the linearity relationship is that obtaining the direct-space elasticity parameters for each product is difficult. Furthermore, assuming a linear relation between shelf space allocation and sales in a limited range of display facing for each product is reasonable. It is important here to also note that knapsack-like models do not consider the inter-relation between products and therefore ignores factors like cross-space and -price elasticities. Furthermore, these models assume all marketing mix variables are fixed. Thus, we think that knapsack-like models, although appealing from a computational tractability perspective, they are restricted in their ability to represent real word retail conditions. As Corstjen and Doyle (1981) showed, including cross-elasticity and inter-product factors leads to significantly different shelf space allocations than when they are not included. In the sequel we focus on models that use space elasticity and consider inter-product marketing factors interactions. For convenience, we summarize the features of the optimization models and solutions approaches in Tables 1 and 2, respectively.

Marketing research (Dreze et al. 1994) shows that products should be allocated on shelfs as uniform and complete columns. Although models in the literature considered using complete columns in a single shelf, in a real retail environment most of the products can be located vertically on multi-shelves. Yang and Chen (1999), Yang (2001), Lim et al. (2004), Hwang et al. (2005) and Murray (2010) considered the allocation of a product vertically on multilevel shelves. However, Hwang et al. (2005) model can result in space allocations on discreet shelves in a vertical column. For example, if there are 7 shelves in a product category, their model can make space allocation on the second and fifth shelves vertically, an unrealistic arrangement. The same problem occurs with Murray et al. (2010) model which considers product display orientation, location on vertically different shelves and stacking. For example, if there are six shelves in the category, their model can make a shelf space allocation on the second shelf with the display orientation 1 (say main facing of the product) and the sixth shelf with the display orientation 3 (say side facing of the product). Besides, both of these models did not enforce uniform and complete allocation in a shelf section. In our thesis, we enforce both uniformness and completeness in the case of allocations on multi-shelves by having the same number of products on vertically neighboring shelves.

TABLE 1: Optimization Models Major Features.

	PAPER	FACTORS: Product display facing decision: Number of display facings	Total Width-length	Area	Product Selection	Objective Function: Max. Net Profit	Max. Gross Profit	Max. Return on Inventory Investment	Min. total weighted penalty due to the deviations from targets and non-productive use of space	Min. under base level allocation	Max. Average Optimal achievement	Cost Components in the objective function
1	Anderson and Amato 1973	√			√		√					
2	Hansen and Heinsbroek 1979		√		√	√						
3	Corstjen and Doyle 1981	√				√						
3	Corstjen and Doyle 1983	√					√					
5	Zufryden 1986			√		√						
4	Bultez and Naert 1988	√				√						
6	Borin et al. 1994(a)	√			√(2)			√				
7	Urban 1998	√			√	√						OC, HC, DC
8	Yang and Chen 1999	√				√						
9	Yang and Chen 1999(b)	√					√					
10	Bookbinder and Zarour 2001		√			√						
11	Irion et al. 2004		√		√	√						OC(*), HC, GC, IC, RC
11	Irion et al. 2004(c)		√		√	√						
12	Reyes and Frazier 2005	√				√						OC, HC
13	Hwang et al. 2005	√				√						OC, HC, DC
14	Reyes and Frazier 2007	√					√			√		
15	Reyes and Frazier 2007(d)	√									√	
16	Hariga et al. 2007	√			√	√						OC, HC, DC(**), SC
17	Silva et al. 2009	√					√					
18	Murray et al. 2010			√			√					
19	Russell and Urban 2010	√			√(6)		√					
20	Lotfi et al. 2011			√					√			OC(***), HC

	PAPER	FACTORS: Demand Model Formation: Quadratic	Additive	Multiplicative	Demand Function Dependents: Space allocation: Own-space elasticity	Cross-space elasticity	Price: Own-price elasticity	Cross-price elasticity	Own-price change elasticity	Cross-price change elasticity	Other marketing variables (fixed)	Operating Cost elasticity associated with increased sales
1	Anderson and Amato 1973		√									
2	Hansen and Heinsbroek 1979			√	√							
3	Corstjen and Doyle 1981			√	√	√						√
3	Corstjen and Doyle 1983			√	√	√						
5	Zufryden 1986			√	√						√	√
4	Bultez and Naert 1988			√	√	√					√	√
6	Borin et al. 1994(a)		√	√	√	√						
7	Urban 1998		√	√	√	√						
8	Yang and Chen 1999			√	√	√					√	√
9	Yang and Chen 1999(b)		√									
10	Bookbinder and Zarour 2001			√	√	√						√
11	Irion et al. 2004			√	√	√						
11	Irion et al. 2004(c)			√	√	√	√				√	
12	Reyes and Frazier 2005		√		√	√						
13	Hwang et al. 2005			√	√	√						
14	Reyes and Frazier 2007			√	√				√			
15	Reyes and Frazier 2007(d)			√	√				√			
16	Hariga et al. 2007			√	√	√						
17	Silva et al. 2009		√									
18	Murray et al. 2010			√	√		√	√				
19	Russell and Urban 2010	√										
20	Lotfi et al. 2011			√	√				√	√		

TABLE 1: Optimization Models Major Features.

	PAPER	FACTORS													
		Demand Function Considerations							Considering different product sizes for showroom and backroom inventory	Product sizes for all products					
		Display inventory level	Vertical Location effect on different shelves	Horizontal Location effect on a shelf	Permenant Stockout demand (Acquired Demand): Does not consider loyal customer preference	Permenant Stockout demand (Acquired Demand): Considers loyal customer preference	Temporary Stockout demand (Stockout Demand): Does not consider loyal customer preference	Temporary Stockout demand (Stockout Demand): Considers loyal customer preference		Width: Same	Width: Different	Height: Same	Height: Different	Depth: Same	Depth: Different
1	Anderson and Amato 1973					√					√	not specified	not specified	not specified	not specified
2	Hansen and Heinsbroek 1979										√	not specified	not specified	not specified	not specified
3	Corstjen and Doyle 1981									√		not specified	not specified	not specified	not specified
3	Corstjen and Doyle 1983									√		not specified	not specified	not specified	not specified
5	Zufryden 1986										√(1)	√		not specified	not specified
4	Bultez and Naert 1988									√		not specified	not specified	not specified	not specified
6	Borin et al. 1994(a)					√		√			√(1)	not specified	not specified	not specified	not specified
7	Urban 1998	√				√					√	not specified	not specified	not specified	not specified
8	Yang and Chen 1999		√								√	not specified	not specified	not specified	not specified
9	Yang and Chen 1999(b)		√								√	not specified	not specified	not specified	not specified
10	Bookbinder and Zarour 2001									√		not specified	not specified	not specified	not specified
11	Irion et al. 2004	√(4)									√		√		√
11	Irion et al. 2004(c)	√(4)				√		√			√		√		√
12	Reyes and Frazier 2005					√				√		not specified	not specified	not specified	not specified
13	Hwang et al. 2005	√	√								√	not specified	not specified	not specified	not specified
14	Reyes and Frazier 2007									√		not specified	not specified	not specified	not specified
15	Reyes and Frazier 2007(d)									√		not specified	not specified	not specified	not specified
16	Hariga et al. 2007	√	√(5)								√	not specified	not specified	not specified	not specified
17	Silva et al. 2009		√	√							√	not specified	not specified	not specified	not specified
18	Murray et al. 2010		√								√		√		√
19	Russell and Urban 2010		√	√							√	not specified	not specified	not specified	not specified
20	Lotfi et al. 2011								√		√		√		√

	PAPER	FACTORS											
		Shelf sizes for showroom inventory								Shelf sizes for backroom inventory			Stacking
		One, horizontal, homogeneous in quality shelf	Width: Same	Width: Different	Height: Same	Height: Different	Depth: Same	Depth: Different	Considering shelf depth for stocking products	Backroom Inventory	One, horizontal, homogeneous in quality shelf space	One, whole space capacity (volume)	
1	Anderson and Amato 1973	√(1D)											
2	Hansen and Heinsbroek 1979	√(1D)											
3	Corstjen and Doyle 1981	√(1D)											
3	Corstjen and Doyle 1983	√(1D)											
5	Zufryden 1986	√(2D)											√
4	Bultez and Naert 1988	√(1D)											
6	Borin et al. 1994(a)	√(3D)							√				
7	Urban 1998	√(3D)							√	√		√	
8	Yang and Chen 1999			√	not specified	not specified	not specified	not specified					
9	Yang and Chen 1999(b)			√	not specified	not specified	not specified	not specified					
10	Bookbinder and Zarour 2001	√(1D)											
11	Irion et al. 2004	√(3D)			not specified	not specified	not specified	not specified	√				√
11	Irion et al. 2004(c)	√(3D)			not specified	not specified	not specified	not specified	√	√		√	√
12	Reyes and Frazier 2005	√(1D)											
13	Hwang et al. 2005			√	not specified	not specified	not specified	not specified	√				
14	Reyes and Frazier 2007	√(1D)											
15	Reyes and Frazier 2007(d)	√(1D)											
16	Hariga et al. 2007			√	not specified	not specified	not specified	not specified		√		√	
17	Silva et al. 2009			√	not specified	not specified	not specified	not specified					
18	Murray et al. 2010			√		√		√					√
19	Russell and Urban 2010			√	not specified	not specified	not specified	not specified					
20	Lotfi et al. 2011	√(3D)			not specified	not specified	not specified	not specified	√	√	√(3D)		√

TABLE 1: Optimization Models Major Features.

	PAPER	FACTORS															
		Display Inventory Level Decisions						Display orientation	Considering space allocation of a product on multi level shelves	Product integrity on display shelves when products are displayed on multi level shelves	Considering Brand or Family Classification	Brand or Family integrity on display shelves when products under same brand are displayed on multi level	Considering minimum and maximum order quantity from manufacturer	Inventory Replenishment		Cycle Time	
		Display shelves kept fully stocked		Let display inventory level reach zero		Let display inventory for a stockout											
		Yes	No	Yes	No	Yes	No							Independent	Joint	Independent	Joint
1	Anderson and Amato 1973	√															
2	Hansen and Heinsbroek 1979	√												√		√	
3	Corstjen and Doyle 1981	√											√(only maximum)				
3	Corstjen and Doyle 1983	√															
5	Zufryden 1986	√											√(only maximum)				
4	Bultez and Naert 1988	√															
6	Borin et al. 1994(a)		√(3)	√(3)		√(3)											
7	Urban 1998		√		√		√							√		√	
8	Yang and Chen 1999	√							√	No integrity			√(only maximum)				
9	Yang and Chen 1999(b)	√							√	No integrity							
10	Bookbinder and Zarour 2001	√															
11	Irion et al. 2004		√(3)	√(3)			√							√		√	
11	Irion et al. 2004(c)		√(3)	√(3)			√							√		√	
12	Reyes and Frazier 2005	√															
13	Hwang et al. 2005	√							√	No integrity				√		√	
14	Reyes and Frazier 2007	√															
15	Reyes and Frazier 2007(d)	√															
16	Hariga et al. 2007		√	√		√							√(only maximum)	√(CT(i))	√(CT(i,m))	√(CT(i))	√(CT(i,m))
17	Silva et al. 2009	√															
18	Murray et al. 2010	√						√	√	No integrity							
19	Russell and Urban 2010	√									√	√					
20	Lotfi et al. 2011	√									√				√		√

(a) and Borin and Farris 1995
(b) and Yang and Chen 1999 (alternate approach); Yang 2001; Lim et al. 2004; Hansen et al. (2010)
(c) extended version
(d) alternate approach

(1) muptiples of slots/products e.g. twin-packed
(2) by letting x(i) reach zero
(3) however, the effect of decreasing number of facings on demand have not been considered
(4) display level changes but demand does not get affected
(5) the shelves are not vertically located, the shelves are located on different part of the store
(6) extended version

(*) in the extentions
(**) different for all products also on different locations
(***) same under same brands, different for same products

TABLE 2: Solution Methodologies

	PAPER	Solution					Decision Variables			Experimental Data Size (Number of products in a product category)
		Solution Method	Modeling Language	Processor and RAM	Computation Time (CPU Time)	Comparison of the solution	Number of Display Facings (x)	Order Quantity (q)	Cycle Time (CT)	
1	**Anderson and Amato 1973**	Specialized Heuristic					Integer			4
2	**Hansen and Heinsbroek 1979**	Lagrange Multiplier, Specialized Heuristic		IBM 370	√	√	Integer			6443
3	**Corstjen and Doyle 1981**	Geometric Programming				√	Continuous			5
4	**Bultez and Naert 1988**	Specialized Heuristic				√	Continuous			6
5	**Zufryden 1986**	Dynamic Programming	BASICA	IBM	√		Integer			9, 10, 11, 20, 30, 40
6	**Borin et al. 1994**	Simulated Annealing	C++			√	Integer			6,18
7	**Urban 1998**	Generalized Reduced Gradient (Greedy Heuristic) and Genetic Algorithm				√	Integer	Integer		6,18,54
8	**Yang and Chen 1999 (alternate approach)**	QSB		IBM Pentium 75	√	√	Integer			6
10	**Yang 2001**	Specialized Heuristic		IBM Pertium 225	√	√	Integer			4, 6, 8, 10
12	**Bookbinder and Zarour 2001**						Continuous			2
	Lim et al. 2002	Adjusted Network Flow		Pentium 450		√	Integer			10, 50, 100
11	**Lim et al. 2004**	Specialized Heuristic, Tabu Search, Squeaky-Wheel Optimization		2.5 GHz Pentium 4 with 512 KB of RAM	√	√	Integer			10, 30, 50, 100
13	**Irion et al. 2004**	LINGO 8.0 (CPLEX 8.1) , Piecewise Linearization		0.93 GHz Pentium 3 with 384 KB of RAM	√	√	Integer			6,20
14	**Reyes and Frazier 2005**	MINLP (LINGO), Branch and Bound					Integer			6
15	**Hwang et al. 2005**	Gradient Search and Genetic Algorithm	Visual Basic	2.4 GHz Pentium	√	√	Continuous	Continuous		4
16	**Reyes and Frazier 2007**					√	Integer			4
17	**Hariga et al. 2007**	MINLP (LINGO - CPLEX), Branch and Bound					x(i) Integer, x(i,k) Continuous	q(i) and q(i,k) Continuous	CT(i) and CT(i,m) Continuous	4
18	**Silva et al. 2009**	Specialized Heuristic			√	√	Integer			135, 907
19	**Murray et al. 2010**	MINLP (BONMIN), Branch and Bound, Outer-apprixmation (B-OA), Hybrid (B-Hyb)		1.83 GHz Intel Core Duo with 2 GB of RAM	√	√	Integer			3, 5, 7, 10, 25, 50, 75, 100
20	**Russell and Urban 2010**	MIQP (CPLEX), Branch and Bound	ILOG OPL	3.4 GHz Pentium 4 with 2 GB of RAM	√	√	Integer			6, 8, 10, 12, 14, 17, 20, 34, 103
21	**Lotfi et al. 2011**	MILNP (LINGO), Branch and Bound				√	Integer	Integer	Continuous and Integer	8

Hariga et al. (2007) consider the joint inventory ordering and allocation of shelf space to a product on different locations within a retail store. These situation arise in practice when a product can be a member of different product categories (e.g. fresh juice) or should be located under a certain product category (e.g. batteries under use-and-throw camera category) so as to increase its sales. Hariga et al. (2007) assume that the replenishment of a shelf occurs when all its products are depleted. This would mean that the first product that has been depleted can only be replenished when all the other products on the same shelf are depleted. They also divided the order quantity from the manufacturer into order quantities for each different shelf location. Therefore, when one dedicated inventory of a shelf for a product in the backroom inventory reaches zero, there would still be products of the same kind in the backroom inventory, but they would not be used to replenish the depleted items because they are not dedicated to that particular shelf. This assumption of joint replenishment of the different products in a shelf may cause the shelves to be partially full, mostly empty, for a significant period of the time.

There is still some ambiguity in the literature over whether demand for a product should be a function of the total display space or display inventory. Studies such as Urban (1998) and Hariga et al. (2007) assume that if the display facings of a product are decreasing, then the demand should be decreasing in a decreasing rate. Borin et al. (1994) and Irion et al. (2004) modeled their problem in terms of space allotted to a product but they let the inventory level reach zero in the showroom inventory. Both of these studies implicitly assumed constant demand rates in the time interval of decreasing number of display facings, since they modeled the demand in terms of space allocation, not the instantaneous inventory level (or average inventory level). Urban

(2005) demonstrated when the model costs are not dependent on the inventory level throughout the period then models that assume demand depends on the instantaneous inventory level can be shown to be equivalent to models that assume that demand depends only on the initial inventory level.

The only studies that considered the effects of product's horizontal location in a shelf were Silva et al. (2009) and Russell and Urban (2010). However, in order to simplify the problem structure, they artificially partitioned one horizontal shelf into smaller parts and assigned priority weights based on the horizontal location in a shelf. Partitioning one shelf causes smaller widths and could affect feasibility in terms of space allocation. When a product allocation can not fit into the remaining space in a shelf part, it would be allocated to another part of this shelf and it would leave unused productive space. In general, any kind of fragmentation of space in a shelf limits the use of scarce shelf space. An alternative is to consider the center of facings as a continuous decision variable as modeled in the main model of Russell and Urban (2010).

Lotfi et al. (2011) used a minimum maximum approach in order to model demand. They assumed a minimum (loyal demand in a day) as a constant (input) and model the maximum demand as a decision variable. They developed a rich model by including different practical aspects such as products under brands, inventory decisions, cycle times of a brand, different sizes of products for backroom and showroom inventory, expiration time of the products, budget of the product category, and maximum number of replenishments for a product or a brand. The downside is that their model resulted in a complex problem where they were able to solve small problem sizes of 8 products under 4 brands. Their model's purpose was to minimize the usage of nonproductive use of space behind display facing.

To conclude our critical review of the literature, we present all the studies in a chronological and clustered structure on Figure 2.5 in order to visualize each study's path in the literature.

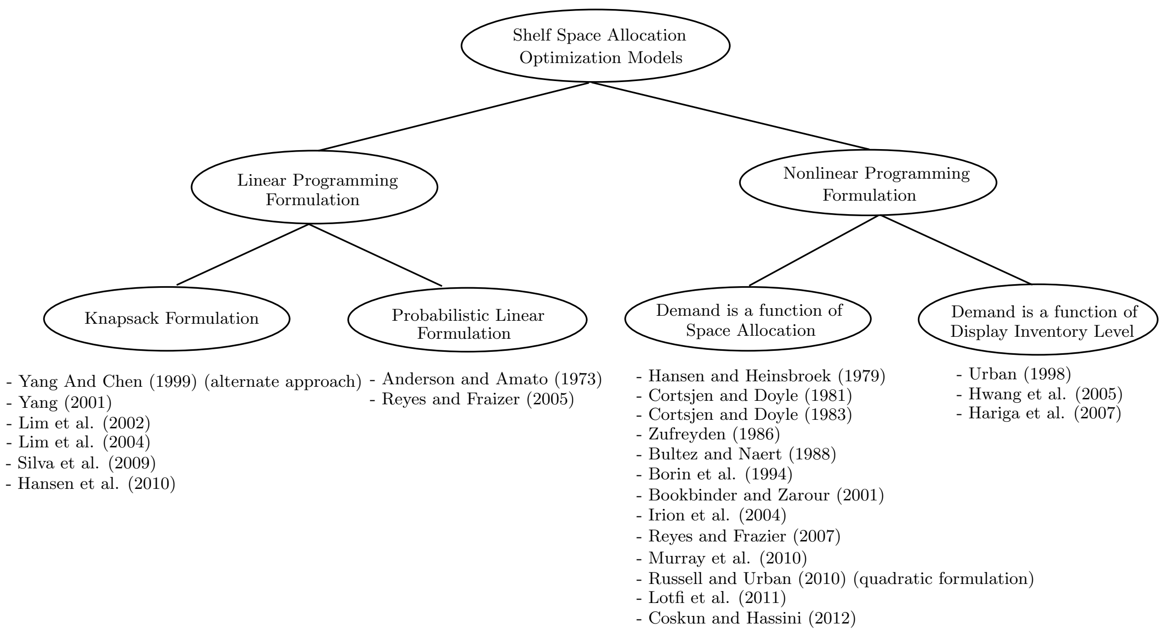


Figure 2.5: Chronological Schema of the Optimization Models

2.4 Contributions

Our major contributions are

1. *Critical review.* We critically review the major optimization and demand models for shelf space allocation. We presented a uniform notation to help us analyze the similarities, differences and limitations of the models.

2. *A Comprehensive and Practical Model.* Motivated by the practical issues facing retailers and the existing literature, we developed a comprehensive model that incorporated:

   (a) *Adjustable Shelf Heights.* We allow for the flexibility of adjusting shelf heights so as to maximize shelf space utilization. Our numerical results show that including this feature leads to significant profit gains.

   (b) *Joint minimization of non-productive space and maximization of net profits.* Several objective functions have been considered in the literature: maximizing gross or net profit, return on investment and minimizing the usage of nonproductive space behind display facing. A more promising model would be one that can jointly minimize nonproductive use of space in the showroom inventory (2 dimensional area) and backroom space behind display facings (3 dimensional space) while maximizing the gross or net profit of the product category. To achieve such a goal, in our study we consider adjustable shelf heights as a decision variable to use more space (display facing area) in the shelf section, increase the visibility of the products. This would also allow us to respond to space elastic demand more wisely than models that use fixed shelf heights and hereby minimizing the nonproductive use of space. As we show in our experiments, considering adjustable shelf heights significantly improves the usage of productive space and increase profitability.

   (c) *Product integrity.* We insure product and brand integrity in a shelf section (the location and numbers in a vertical column). Motivated by the real world needs for uniform space allocation and stackability, in our thesis we enforce both uniformness and completeness, for product allocations on multi-shelves, by having the same number of products on vertically neighboring shelf allocations.

(d) *Price Changes* We allow for prices to be controllable by the retailer by considering the option to increase or decrease the current price by a certain percentage. Given the different cross-elasticities we include in our model, changing prices provides the retailer with a level to optimize revenues depending on the product marketing mix. Our numerical results show that this ability significantly increases the retailer profits.

2.5 Open Research Problems

Based on our analysis of the existing literature and the needs arising in the retail industry we suggest the following topics as possible research topics:

1. Providing a decision support system with alternative demand functions and models. One can build on the spreadsheet-based decision support system provided by Ramaseshan (2008), which is based on his work that integrates shelf space and inventory management (Ramaseshan, 2009). Such a decision support system would be interactive with the user to identify the retail environment and then selects the right demand functions and constraints for optimization.

2. Incorporate inter-temporal effects such as promotions and reference prices. For example, when a retailer carries a time-limited promotion how should the shelf space allocation change? Should more space be given to the product in the same (old) shelf or should additional (new) display area be provided elsewhere?

3. Jointly optimizing shelf space with with inventory management such as:

    - Minimizing the inter-replenishment time as it disrupts the customer shipping experience. The model developed by Hassini (2008) for backroom inventory storage management can be used as a starting point.
    - Inventory coordination and transshipment between stores. Retails are using virtual integrations, through sharing real time inventory data, to achieve the benefits of inventory aggregation. This, in a way, allows the retailer to extend their product display areas. How should retailers that use virtual integration design their space allocations?
    - How should the demand be modeled in the event of a stockout occurs, for example due to promotions and/or supply scarcity? Most of the literature of shelf space allocation employs models that assume that as the shelf display area decreases the demand will also decrease. But, wouldn't the reverse happen in situations where there is a promotion or scarce capacity?

4. Many retails have opted for e-tailing channels: selling through the internet to the consumers. How can we define a demand function in such environments? Which factors would be more important in generating demand. For example, is it plausible to think of the web site layout as a proxy for store shelf space?

Chapter 3

New Model

We consider the problem of allocating shelf space for a product category in a shelf section that is composed of multi-level vertical shelves. Each shelf in this shelf section has a scarce space and has a different quality for allocating products. Our goal is to maximize the daily gross profit of the product category. As a basic assumption, the assigned shelf section's location in the retail store, product display facing and total space limit have already been determined by the management.

Although a product may be located on a shelf in different display orientations, in practice most products have one practical orientation. Different display orientations are possible in some product categories, such as books and CDs, however, are often limited to two orientation (e.g., book front cover and back cover) that have to be displayed simultaneously side by side. Retailers often count on the shopper to examine the product before purchase and so one display orientation is often enough. Therefore, in our model we consider that the product will be displayed in only one orientation.

Our main contributions to the literature is the considerations of adjustable shelf heights, product and brand integrity in a shelf section (the location and numbers in a vertical column) in a shelf section while considering space (in terms of number of display facings), price change and product location as the main determinants which affect product's demand.

We should note here that we assume the selling prices are predetermined by the manufacturer, the supplier or the market. We can think of these prices as the "suggested price" for a product by the manufacturer or "the average price" on the market. Except for space, all the other marketing mix variables, such as advertising and promotions, were fixed in the market share-strength of the product (π). Since the market share-strength of the product (π) has been determined by the marketing mix variables such as "suggested price" or "average price" of the product, we assume that only percentage changes from that suggested price affect its sales. We introduce this new structure of the demand function to the literature where the demand of a product is a function of price changes from its suggested price rather than the product's price itself. In this context, our demand model is similar to those of Reyes and Frazier (2007) and Lotfi et al. (2011). Reyes and Frazier (2007) use a "price sensitivity factor", $\left(\frac{p_{avg}}{p_i}\right)^{\sigma_i}$, that represents the deviation of the product selling price from the average price of the category. We model our price sensitivity factor as the deviation of the selling price of the product from its suggested price ($\left(\frac{p_i^{new}}{p_i^{suggested}}\right)^{\sigma_i} = (1 - PC_i)^{\sigma_i}$). Lotfi et al. (2011) used a similar price change idea, however , they modeled their demand using a minimum-maximum approach.

3.1 Model Assumptions

We consider products, stackable or non-stackable, that are grouped into brands that have to be allocated to one shelf section. This section consists of vertical multi-level adjustable shelves for each product category. In each of these shelf levels, there is a parameter assigned to each shelf to reflect the location effect on demand. We note here that although we use adjustable shelves we assume fixed vertical shelf effects on demand. This is plausible as the number of the vertical shelves is fixed in a given shelf section and we assume given minimum and maximum shelf heights.

Our objective is to maximize the gross profits generated from the displayed products. An alternative objective could be to maximize the display area used, since we are using adjustable shelf heights, by penalizing the unused display space. However, it is not easy to estimate these penalties. Our numerical results indicate that our profit maximization model leads to significantly higher use of display space than models that do not use flexible shelf heights. Although we do not account for operating costs, such as inventory costs, we note that our model leads to a more efficient use of the shelf space, include the space behind the display facings, and in turn it releases the pressure on the backroom inventory space and leads to longer cycle times for the brands, thus, decreasing inventory costs.

Motivated from the realities of the retail world and the literature, we make the following additional assumptions:

1. The shelves are kept fully stocked with the display facings at all times. Thus, we avoid temporary stockout situations and the possible resulting change in demand rates. This assumption is realistic in many retail setting, where often employees will be going around shelves to replenish them or to reorganize products from the back of the shelf into the front facings. In effect we are assuming that the replenishment from the backroom space and inventory to showroom space is instantaneous. In addition, we assume that the cost of this

replenishment is negligible.

2. There is adequate backroom space (to the shelf depth, not backroom inventory) for extra product stocking and it is fixed.

3. The shelf section which is dedicated to a product category has a fixed length of width, height and depth. However, the height of each shelf in the shelf section can be adjustable as long as total height of the shelves will be equal to the height of the shelf section. The model also considers the policy of the retail management about the minimum and maximum shelf height conditions. Considering shelf height and depth in the showroom inventory is not just for the purpose of ensuring the products' depth or height doesn't exceed the shelfs' height or depth, but also to stack products vertically (to the height) on top of the other and stock products horizontally (to the depth).

4. Products come in rectangular size packages. For non-rectangular physical contours of a product, we consider the least area rectangular contour dimensions that the product can fit into (Murray et al. 2010).

5. The products under the same brand will be adjacent to each other on the display shelves. This means that each product under a brand will be located to the left, to the right, to the top or to the bottom of a product under the same brand. A brand will be visible to the public on display shelves as a cluster.

6. Unlike the vertical location, we assume that the horizontal location of a product does not affect its demand.

7. Stackable products should be stacked in a uniform and complete column (Dreze

et al. 1994). It means that the left end and the right end of each horizontal stack and the top end and the bottom end of each vertical stack should be on the same level .

8. Products may be located on adjacent shelves. In such cases they should be located in uniform and complete columns (Dreze et al. 1994). It means that the number of products on one shelf (in a horizontal stack, $x_{ib,m}$) should be the same as the number of products on the adjacent shelf (in a horizontal stack $x_{ib,m+1}$ or $x_{ib,m-1}$). Since the horizontal location of product does not effect its demand, the left end coordinate and the right end coordinate of each horizontal stack on adjacent shelves can be arranged as long as there are same number of products in each horizontal stack.

9. The best display facing of a product has already been chosen by the retailer. This means that the width, height and depth length of the product for showroom inventory is already known.

10. For each product, minimum and maximum number of display facings, the purchasing cost and suggested selling price are known and given.

3.2 Demand Function

We define the unmodified demand (UMD) as the intrinsic preference for the products by

$$UMD_i(x_i, PC_i, f'_{1i}, f'_{2i}, ..., f'_{zi}) = U_i = \pi_i.$$

where all the marketing mix variables are fixed. The parameter π_i represents the potential market share-strength of a product i in the store without in-store support, such as advertisement, promotions, specific location in store, that are represented by f_{zi}. In our study, the demand is a function of space and price change (instead of price) while other marketing variables are fixed.

Switching preference demand (PD) arises from (1) consumers who are willing to prefer a product within the assortment and (2) those who are not willing to prefer any product within the assortment, but through their calculated valuations, chose a product within the assortment (although there may be consumers who are not willing to prefer a product within the assortment and still prefer another product which is also not in the assortment, such scenario has always been neglected in the literature, please see Chapter 1 for detailed explanation of such demand). We call the demand gained by product i from the switching preference of the first consumer group the "Gained Demand" (GD_i) and the demand gained by product i by switching preference of the second consumer group the "Acquired Demand" (AD_i).

Incorporating UMD_i with GD_i results in "Modified Demand", MD_i. MD_i represents the demand for the product that results from its UMD_i through in-store merchandising support such as advertisement, promotions, space and specific location in the store. In our model, modified demand is the "modified" version of the unmodified demand with its differential "space allocation", "price change" decisions in a multi-level shelf concept (Dreze et al. 1994). Since marketing variables, other than space and price change, are fixed in the unmodified demand (UMD), we expect that space, display location of the product within the shelves and price changes to be the in-store attractiveness that affects sales. Therefore, we may define the modified demand for a product i as

$$\begin{aligned} MD_i &= \pi_i \left(\frac{\sum_{m=1}^{M} \varphi_m x_{im} \bar{x}_{im}}{x_i} \right) \prod_{j=1}^{I_1} x_j^{\beta_{ij}} \prod_{j=1}^{I_1} (1 - PC_j)^{\mu_{ij}} \\ &= \pi_i \left(\frac{\sum_{m=1}^{M} \varphi_m x_{im} \bar{x}_{im}}{x_i} \right) x_i^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I_1} x_j^{\beta_{ij}} (1 - PC_i)^{\sigma_i} \prod_{\substack{j=1 \\ j \neq i}}^{I_1} (1 - PC_j)^{\mu_{ij}} \end{aligned}$$

where $\left(\frac{\sum_{m=1}^{M} \varphi_m x_{im} \bar{x}_{im}}{x_i} \right)$ is the display location factor (weighted average value of φ_m when product i is displayed on more than one shelf), $x_i^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I_1} x_j^{\beta_{ij}}$ is the space allocation factor in terms of number of display facings and $(1 - PC_i)^{\sigma_i} \prod_{\substack{j=1 \\ j \neq i}}^{I_1} (1 - PC_j)^{\mu_{ij}}$ is the price change factor. As in Hwang et al. (2005), the shelf that is in the worst position in the shelf section, generally the bottom shelf, has the smallest value of 1 for the display factor.

AD represents the demand that unstocked products capture, where some consumers may change their preferences and purchase a product within the assortment and some of them stay loyal to their first choice and decide not to purchase any product. For tractability purposes we leave out this type of demand from our model. As a result, our unit demand for a product i in the assortment can be represented in a multiplicative form as

$$\begin{aligned} D_i &= UMD_i \times GD_i = MD_i \\ &= \pi_i \left(\frac{\sum_{m=1}^{M} \varphi_m x_{im} \bar{x}_{im}}{x_i} \right) x_i^{\alpha_i} \prod_{\substack{j=1 \\ j \neq i}}^{I_1} x_j^{\beta_{ij}} (1 - PC_i)^{\sigma_i} \prod_{\substack{j=1 \\ j \neq i}}^{I_1} (1 - PC_j)^{\mu_{ij}} \end{aligned}$$

3.3 Optimization Model

The gross profit for a product i under brand b is defined as

$$P_{ib} = (p_{ib} - p_{ib}PC_{ib} - c_{ib}) \left(\pi_{ib} \left(\frac{\sum_{m=1}^{M} \varphi_m x_{ib,m} \bar{x}_{ib,m}}{x_{ib}} \right) \cdot x_{ib}^{\alpha_{ib}} \prod_{\substack{j=1 \\ j \neq i}}^{I} x_j^{\beta_{ib,j}} (1 - PC_{ib})^{\sigma_{ib}} \prod_{\substack{j=1 \\ j \neq i}}^{I} (1 - PC_j)^{\mu_{ib,j}} \right) \tag{3.3.0.23}$$

and the shelf space allocation problem is formulated as

$$\max \quad \sum_{i=1}^{I} \sum_{b=1}^{B} P_{ib}$$

$$\text{s. t.} \quad 1 \leqslant \sum_{m=1}^{M} y_{ib,m}, \qquad \forall_{i,b} \tag{3.3.1}$$

$$\sum_{m=1}^{M} y_{ib,m} \leqslant A, \qquad \forall_{i,b} \tag{3.3.2}$$

$$y_{ib,m} \leqslant y_{b,m}, \qquad \forall_{i,b,m} \tag{3.3.3}$$

$$y_{b,m} \leqslant \sum_{i=1}^{I} y_{ib,m}, \qquad \forall_{b,m} \tag{3.3.4}$$

$$top_b - \frac{\sum_{m=1}^{M} m y_{b,m}}{\sum_{m=1}^{M} y_{b,m}} - \frac{\sum_{m=1}^{M} y_{b,m} - 1}{2} = 0, \qquad \forall_b \tag{3.3.5}$$

$$bottom_b - \frac{\sum_{m=1}^{M} m y_{b,m}}{\sum_{m=1}^{M} y_{b,m}} + \frac{\sum_{m=1}^{M} y_{b,m} - 1}{2} = 0, \qquad \forall_b \tag{3.3.6}$$

$$top_b \geqslant m\, y_{bm}, \qquad \forall_{b,m} \tag{3.3.7}$$

$$bottom_b \leqslant M - (M - m) y_{bm}, \qquad \forall_{b,m} \tag{3.3.8}$$

$$top_{ib} - \frac{\sum_{m=1}^{M} m y_{ib,m}}{\sum_{m=1}^{M} y_{ib,m}} - \frac{\sum_{m=1}^{M} y_{ib,m} - 1}{2} = 0, \quad \forall_{i,b} \tag{3.3.9}$$

$$bottom_{ib} - \frac{\sum_{m=1}^{M} m y_{ib,m}}{\sum_{m=1}^{M} y_{ib,m}} + \frac{\sum_{m=1}^{M} y_{ib,m} - 1}{2} = 0, \quad \forall_{i,b} \tag{3.3.10}$$

$$top_{ib} \geqslant m\, y_{ib,m}, \qquad \forall_{i,b,m} \tag{3.3.11}$$

$$bottom_{ib} \leqslant M - (M-m) y_{ib,m}, \qquad \forall_{i,b,m} \tag{3.3.12}$$

$$x_{ib,m} \leqslant x_{ib}, \qquad \forall_{i,b,m} \tag{3.3.13}$$

$$y_{ib,m} \leqslant x_{ib,m}, \qquad \forall_{i,b} \tag{3.3.14}$$

$$x_{ib,m} \leqslant y_{ib,m} R, \qquad \forall_{i,b} \quad \text{where } R \text{ is a very large number} \tag{3.3.15}$$

$$y_{ib,m}\, x_{ib,m+1} = x_{ib,m}\, y_{ib,m+1}, \qquad \forall_{i,b,m} \tag{3.3.16}$$

$$\sum_{i=1}^{I} \sum_{b=1}^{B} w_{ib} x_{ib,m} \leqslant L, \qquad \forall_{i,b,m} \tag{3.3.17}$$

$$H^{min} \leqslant H_m \leqslant H^{max}, \qquad \forall_m \tag{3.3.18}$$

$$\sum_{m=1}^{M} H_m = TH \tag{3.3.19}$$

$$\bar{x}_{ib,m} = y_{ib,m} \quad \text{for all } i, b,\, m, \text{ such that } \bar{y}_{ib} = 0 \tag{3.3.20}$$

$$\bar{x}_{ib,m} h_{ib} \leqslant H_m y_{ib,m} \quad \text{for all } i,\, b,\, m, \text{ such that } \bar{y}_{ib} = 1 \tag{3.3.21}$$

$$x_{ib} = \sum_{i=m}^{M} x_{ib,m}\, \bar{x}_{ib,m}, \qquad \forall_{i,b} \tag{3.3.22}$$

$$X_{ib}^{min} \leqslant x_{ib} \leqslant X_{ib}^{max}, \qquad \forall_{i,b} \tag{3.3.23}$$

$$PC_{ib}^{negative} \leqslant PC_{ib} \leqslant PC_{ib}^{positive}, \qquad \forall_{i,b} \tag{3.3.24}$$

$$y_{ib,m}, y_{b,m} \text{ are binary} \qquad \forall_{i,b,m} \tag{3.3.25}$$

$$x_{ib,m}, \bar{x}_{ib,m}, x_{ib}, top_b, bottom_b, top_{ib}, bottom_{ib} \text{ are integer} \qquad \forall_{i,b,m} \tag{3.3.26}$$

where constraint set (1) ensures that product i under brand b will be located at least on one shelf. Constraint set (2) represent the policy constraint of the retail management where product i can be located on a limited number (where A is a constant decided by the retailer) of shelves. Constraint set (3) ensures that if a brand is not chosen for a shelf m, then a product i under that brand should not be chosen to be located on that shelf. Constraint set (4) ensure that if a brand is chosen for a shelf m then at least one product i under that brand should be chosen to be located on that shelf. Constraint sets (5–12) define the top and bottom shelves of brand b (any product under brand b) and product i. The constraints ensure that items for a certain product, or brand, should be adjacent to each other if they are located on multiple shelves. We provide more explanation on the derivation of these constraints in Appendix A.1. Constraint sets (13–15) represent the interdependencies between the allocations to shelf sections (x_{ib}), shelf levels (x_{ibm} and the decision of allocating to a shelf (y_{ibm}). Constraint set (16) guarantees that the number of items from product i under brand b located on a shelf m will be the same as that on the adjacent shelves, if the product is located on multiple shelves. Constraint set (17) ensures that the total amount of shelf space allocation to the width-length of a shelf m does not exceed that shelf m's width-length. Constraint set (18) defines the bounds on the height-length of a shelf m. Constraint (19) guarantees that the total height-length of the shelves should be equal to the height-length of the shelf section. Constraint sets (20–22) take care of the product stackability requirements. Constraint sets (23-24) define the bounds on the number of facings and price changes.

3.4 Numerical Analysis

In our numerical analysis, we solve the optimization problem starting from small size test problems to relatively large size problems. To benchmark the performance of our model, we have solved the following varieties of our model: fixed shelf heights shelf space allocation model (FSHSSA), adjustable shelf heights shelf space allocation model (ASHSSA), price change and fixed shelf heights shelf space allocation model (PCFSHSSA), and price change and adjustable shelf heights shelf space allocation model (PCASHSSA). For simplicity, all the test samples were solved without considering product brands.

3.4.1 Parameter Estimation

For each test problem parameters, we chose different lower and upper bounds between the ranges that has been given in Table 3 and generate the data using a uniform distribution between the chosen these bounds.

Table 3: Test problems parameters.

Parameter	Range
width-length (w)	$[4, 9]$
height-length (h)	$[7, 13]$
purchasing cost (c)	$[5, 8]$
suggested selling price (p)	$[9, 13]$
market share-strength (π)	$[40, 100]$
own-space elasticity (α)	$[0.3, 0.5]$
cross-space elasticity (β)	$[0.005, 0.03]$
own-price change elasticity (σ')	$[-1, -4]$
cross-price elasticity (μ')	$[0, 0.2]$

The highest value of the location scale parameter is assigned to the a shelf that is closer to the eye level (e.g. second shelf from the top is closer to the eye-level for

3 shelves, 4 shelves and 5 shelves problems). The next highest value is assigned to the top shelf and the lowest value, 1, is assigned to the bottom shelf as suggested by Hwang et al. (2005). The dimensions of the shelf section/shelves are determined in such a way that sufficient space was available to allow a feasible solution while limiting the maximum number of display facings. For all test problems, the bounds for minimum and maximum number of display facings were 2 and 16, respectively and allowable (percentage of) price change was %40.

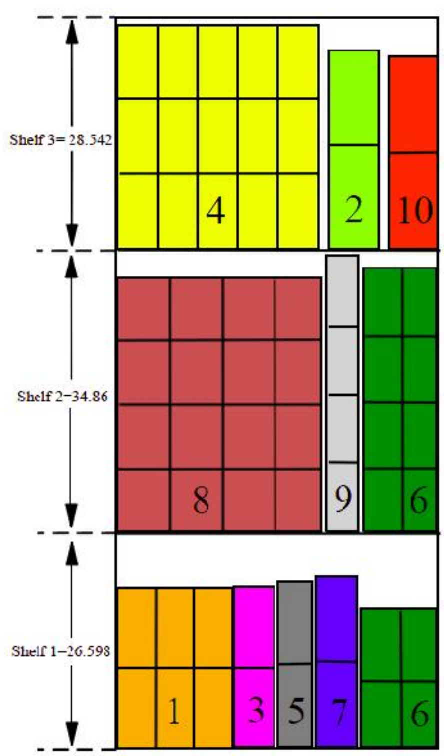
As per Talluri and van Ryzin (2005, p. 325) a price increase of a product does not increase the total demand (Π) of the product category. However, such assumption allows that a price increase (decrease) for all products in the category can decrease (increase) the overall demand. In this context, the overall demand can be controlled by the cross effect parameters of the products corresponding with the own effect parameters of the products. In all of our test samples, the associated cross-price change elasticity parameters are generated consistent with this aforesaid assumption. If any solution is observed by the price increase or decrease of all products in the assortment, this would mean that total product category demand would be decreasing or increasing, respectively. Such scenarios are allowed in our model, but in our test samples we randomly generated such price change elasticity parameters in a certain range that does not reflect such situations. As we can easily observe from the results of all our experiments, when some of the products' demand rises with better shelf space allocation and price reduction, then some of the products' demand decreases with worse shelf space allocation and price increase. In general, the overall demand increases or decreases are neglected.

The ranges for the cross-space and -price changes elasticities were taken from the literature. Space elasticity parameters for a product i, α_i and β_i, can be estimated based on experimental field data (Dreze et al. 1994, Frank and Massy 1970). Yang and Chen (1999) state that store location is one of the factors that influence the allocation of space in a store and for chain stores they should take into account this factor when estimating elasticity parameters. Price change elasticity parameters can be estimated through scanning actual transaction data (Mace and Neslin 2004, Russell and Peterson 2000).

We present an example shelf space allocation of a 10 products-3 shelves problem to visualize the results of our solutions in Table 4. As expected, the shelf with the higher values of location scale parameter has more space than the shelf with lower values of the location scale parameter. Products with many display facings are allocated closer to the eye level and have discounts on their prices, on the other hand, products with few display facings are allocated to the bottom and top shelves with price increases.

Table 4 Solution of the 10 product - 3 Shelves test problem

Item	$x_i = x_{im} \times \bar{x}_{im}$	top$_i$	bottom$_i$	PC$_i$		$p_i^{suggested}$	p_i^{new}
1	6 = 3×2	1	1	-0.4	raise	9.442	13.2188
2	2 = 1×2	3	3	-0.4	raise	11.010	15.414
3	2 = 1×2	1	1	-0.4	raise	9.641	13.4974
4	5 = 5×3	3	3	0.4	discount	12.490	7.494
5	2 = 1×2	1	1	-0.4	raise	10.060	14.84
6	12 = 2×2 + 2×4	2	1	0.236	discount	10.143	7.749
7	2 = 1×2	1	1	-0.4	raise	11.376	15.9264
8	16 = 4×4	2	2	0.273	discount	11.891	8.6447
9	4 = 1×4	2	2	-0.4	raise	11.513	16.1182
10	2 = 1×2	3	3	-0.4	raise	10.855	15.197



3.4.2 Computer Packages

Our models are nonconvex Mixed Integer Nonlinear Programs (MINLP). In general, MINLPs are very difficult to solve and there is no method that can guarantee finding a global optimal solutions for such problems (Bussieck and Pruessner, 2003). We have used the optimization modeling language GAMS and tried three solvers:

- BONMIN (Basic Open Source Nonlinear Mixed Integer programming) which is designed for solving MINLPs.
- BARON, a solver that is designed to find global optimal solutions for nonconvex optimization problems.
- KNITRO solver package, a NLP solver.

Through experimental trials we found that BARON was very slow and often gave solutions with lower profits than KNITRO and BONMIN. Therefore we decided to focus our numerical efforts on the latter two solvers. Although KNITRO is designed for nonlinear problems, it can be also be efficient for solving discrete optimization problems with integer or binary variables and therefore has been suggested for solving moderate size MINLPs. KNITRO implements both state-of-the-art interior-point methods, where the nonlinear problem is replaced by a series of subproblems, and active set methods for solving nonlinear optimization problems. It provides two procedures for the computing steps within the interior point approach: Interior/CG where each step is computed using a projected conjugate gradient iteration and Interior/Direct which attempts to compute a new iterate by solving the primal-dual KKT matrix using direct linear algebra. For all of our models, Interior/Direct procedure is automatically chosen by KNITRO. KNITRO found slightly better solutions much faster than BONMIN for small size problems. Therefore, we kept solving all the problems with KNITRO. Unfortunately, the integer decision variables, such as number of display facings on a shelf (x_{im}) and total number of display facings for a product (x_i), were unable to converge for relatively large size problems. As we show our results in Table 5, although we found solutions for models using the fixed shelf heights strategy (FSHSSA and PCFSHSSA) in all test cases, KNITRO was unable to converge for models using the adjustable shelf heights strategy (ASHSSA and PCASHSSA) for 8 products 3-shelves test problems and larger size problems .

The BONMIN solver provides two algorithms for solving MINLPs: (1) B-OA (outer approximation) that features outer approximation based decomposition in which the objective function and constraints are linearized at various points and (2) B-Hyb (Hybrid) which is a hybrid algorithm composition of outer approximation and branch-and-cut algorithms rather than the default algorithm, B-BB. The latter is a basic branch-and-bound algorithm in which a continuous nonlinear relaxation is solved at each node of the search tree. Because we use binary and integer variables to implement the product integrity and uniformity constraints, we were not able to make use of the BONMIN algorithms, unlike Murray et al. (2010). Moreover, Bonami and Lee (2007) suggested the use of B-BB for nonconvex MINLPs since outer-approximation algorithms have not been tailored to treat nonconvex problems. They also noted that B-BB is only a heuristic for such problems.

Table 5 Results of the 4 models under each test case

Problem size	MODEL	Solvers			
		KNITRO		BONMIN	
		Optimal Solution	Solution time (in seconds)	Optimal Solution	Solution time (in seconds)
3 Products-2 Shelves	FSHSSA	1996	0.749	1981	53.618
	PCFSHSSA	2492	0.358	2392	134.41
	ASHSSA	2092	11.716	2092	111.915
	PCASHSSA	2599	6.692	2486	84.444
4 Products-2 Shelves	FSHSSA	2432	0.53	2400	74.319
	PCFSHSSA	3518	0.172	3518	17.41
	ASHSSA	2589	28.8	2560	165.77
	PCASHSSA	3788	23.44	3711	81.792
5 Products-2 Shelves	FSHSSA	1491	2.262	1491	22.573
	PCFSHSSA	1694	1.451	1673	32.713
	ASHSSA	1563	110.027	1563	326.244
	PCASHSSA	1763	97.079	1763	149,402
6 Products-2 Shelves	FSHSSA	2391	4.633	2367	27.456
	PCFSHSSA	2525	14.258	2480	180.384
	ASHSSA	2500	262.316	2503	605.128
	PCASHSSA	2677	252.659	2677	366.368
4 Products-3 Shelves	FSHSSA	3322	35.163	3263	515.303
	PCFSHSSA	4889	13.931	4869	364.793
	ASHSSA	3731	511	3638	2781.638
	PCASHSSA	5543	632	5417	659.529
5 Products-3 Shelves	FSHSSA	2321	45.864	2291	695.78
	PCFSHSSA	2432	72.899	2429	121.571
	ASHSSA	2515	655.938	2503	2198.475
	PCASHSSA	2647	1064.938	2594	2892.024
6 Products-3 Shelves	FSHSSA	1981	15.912	1937	589.918
	PCFSHSSA	2257	49.843	2155	464.478
	ASHSSA	2040	3627.195	1978	7013.805
	PCASHSSA	2363	3679.533	2340	3039.149
7 Products-3 Shelves	FSHSSA	4540	10.576	4512	2813.634
	PCFSHSSA	4744	103.429	4647	999.624
	ASHSSA	4905	7830.923	4838	3348.234
	PCASHSSA	5102	52803.4775	5095	21503.739
	FSHSSA	5094	16.146	5018	971.465

8 Products-3 Shelves	PCFSHSSA	5863	602.024	5830	508.72
	ASHSSA	5315	50000 (time limit)	5263	3422.366
	PCASHSSA	NA	50000 (time limit)	5993	10431.413
9 Products-3 Shelves	FSHSSA	5549	2801.731	5471	909.017
	PCFSHSSA	7933	8350.422	7909	1142.037
	ASHSSA	NA	50000 (time limit)	5629	10051.269
	PCASHSSA	NA	50000 (time limit)	8159	25301.272
10 Products-3 Shelves	FSHSSA	6968	887.022	6944	1602.333
	PCFSHSSA	9504	13365.338	9450	1440.638
	ASHSSA	NA	50000 (time limit)	6946	50830.57
	PCASHSSA	NA	50000 (time limit)	9579	22993.362
6 Products-4 Shelves	FSHSSA	2573	471.528	2502	932.231
	PCFSHSSA	2878	803.749	2817	3336.221
	ASHSSA	2892	100000 (time limit)	2862	100000 (time limit)
	PCASHSSA	NA	100000 (time limit)	3145	50765.268
8 Products-4 Shelves	FSHSSA	5725	5.928	5560	6494.275
	PCFSHSSA	6565	2155.59	6441	4045.917
	ASHSSA	NA	100000 (time limit)	5868	33810.549
	PCASHSSA	NA	100000 (time limit)	6788	15502.007
10 Products-4 Shelves	FSHSSA	7884	100000 (time limit)	7844	24800.321
	PCFSHSSA	10535	100000 (time limit)	10459	8571.819
	ASHSSA	NA	100000 (time limit)	8203	150000 (time limit)
	PCASHSSA	NA	100000 (time limit)	11937	50929.927
13 Products-5 Shelves	PCASHSSA	NA	NA	11828	149946
14 Products-5 Shelves	PCASHSSA	NA	NA	NA	150000 (time limit)

15 Products-5 Shelves	PCASHSSA	NA	NA		NA	150000 (time limit)
20 Products-6 Shelves	PCASHSSA	NA	NA		NA	150000 (time limit)

3.4.3 Test Samples

The results of the test problems are displayed in Table 5. In total 56 problems were solved by KNITRO and BONMIN solvers. Moreover, we tried to solve the price change and adjustable shelf heights shelf space allocation model (PCASHSSA) for relatively large size problems such as 13 product-5 shelves, 14 product-5 shelves, 15 products-5 shelves and 20 products-6 shelves. However, we could only find efficient solutions for the 13 product-5 shelves problem. We tried to provide “good” initial solutions to the solvers, but this did not help. The stopping times for the solvers were between 40,000 seconds and 100,000 seconds depending on the initial solutions that we have given to the solver.

3.5 Results and Insights

3.5.1 Adjustable shelves effect

One benefit of our general model should be that we would use more space, and thus make more profit, due to the flexibility we get from adjustable shelves. The concept suggests that if we were able to adjust shelf heights in a shelf section, we would be able to use the space more efficiently and respond to space elastic parameters more delicately and accurately. As we can examine in all of our experiments, the shelf space allocations of the four different models are different for fixed shelf heights and adjustable shelf heights (FSHSSA versus ASHSSA and PCFSHSSA versus PCASHSSA) in terms of the number of display facings. A small percentage of profit increase has been observed, as expected, in all of our test cases. We note that ASHSSA model shows a minimum of %2 to a maximum of %14, with an average of %6 increase on gross profits compared to the conventional model, FSHSSA. Very similar profit increases (a minimum of %3 to a maximum of %14, with an average of %7) are observed when comparing PCASHSSA to PCFSHSSA. It is worth noting that these percentage increases in profits are significant in many retail sectors, such as groceries, where profit margins can be as low as 1 to 2 %.

3.5.2 Price changes effects

We observe that PCFSHSSA model shows a minimum %4 to a maximum %47, with an average of %21 increase on gross profits compared to the conventional model, FSHSSA. Very similar profit increases (a minimum of %4 to a maximum of %48, with an average of %22 on gross profits) have also been observed for PCASHSSA compared to the ASHSSA. The wide range of profit increases depends on the product category price elasticity parameters. In the real world retail environment, each product category's sensitivity to prices changes varies depending on the characteristics of its products. Although categories such as fresh juices, yogurts and cereals can be very sensitive to price changes, categories such as spices and lamps can be quire insensitive to price changes. Therefore, we have used a variety of price change elasticity ranges for different test sample sets. The lower values of the elasticity parameters (more sensitive to price changes) we tested, the higher profit margins we obtained, the higher values (less sensitive to price changes) we tested, the lower profit margins we obtained.

3.5.3 Combined effect of adjustable shelves and prices changes

The most important result of our study is the comparison of the conventional shelf space allocation model of fixed shelf heights, FSHSSA, where the demand is only a function of the space allocated to a product, with our shelf space allocation model of adjustable shelf heights and price change effects, PCASHSSA. This comparison shows that a minimum of %11 and a maximum of %66, with an average of %30 of increase on gross profits, can be achieved through our model. The large increases are achieved when the price changes are large (in our experiment as large as %40). In a realistic competitive environment the retailers may not have the luxury to make such large price changes and so we expect that in practice moderate profits savings would be observed.

3.5.4 Location Effect

In general, the shelves with higher values of location scale parameters (the shelves closer to the eye-level) tend to have more space than the shelves with lower values (shelves closer to the bottom shelf). Products allocated closer to the eye-level shelves tend to have many display facings and have discounts on their prices in order to increase their demand. In contrast, products allocated closer to the bottom shelves tend to have few display facings (e.g. closer to the lower bound) and price increases. However, we have also seen cases where some products with few display facings and price increases have been allocated closer to the eye-level shelves because of the extra shelf space that could not be used by other products.

3.6 Sensitivity Analysis

We have also conducted a sensitivity analysis for own-price change elasticity parameters in order to evaluate the response of the price changes of each product for the case of 10 products and 3 shelves. Each case is solved under a range of own-price change parameters starting from -4 to -1 with increments of 0.25 while keeping other parameters fixed. We find that for larger values of own price-change parameters (less sensitive to price change) such as between -1.25 and -1, all the prices were set up to their upper bounds. But, with each 0.25 reduction on the own-price change elasticity values, the price increases started to decrease starting from the product with the highest profit margin, as expected. As per our sensitivity analysis results in Table 6, we see that as products become more price sensitive, cutoffs start from the highly profitable products to the lower profitable products.

In our early test problems, the price change elasticity parameters were chosen in a range that did not allow overall category demand increases or decreases for any of our problem cases. However in our sensitivity analysis overall demand decreases were allowed (e.g. own price change elasticity parameters are chosen in the range of -1.25 and -1 when cross price change elasticity parameters are fixed between 0 and 0.1). In general, a price decrease for a product provoked a price increase of another product in order to balance the overall effect on the demand of the category. This has been observed in all of our test cases.

Table 5 Sensitivity Analysis of the 10 product - 3 Shelves test problem

Product i	p(i)	c(i)	p(i) - c(i)
i1	9.442	5.827	3.615
i2	11.01	5.235	5.775
i3	9.641	5.628	4.013
i4	12.49	5.093	7.397
i5	10.06	5.677	4.383
i6	10.143	5.364	4.779
i7	11.376	6.291	5.085
i8	11.891	6.121	5.77
i9	11.513	6.54	4.973
i10	10.855	5.596	5.259

PC(i) - percentage of price change of the product i

Price Change elasticity parameters	-1.25 and -1	-1.5 and -1.25	-1.75 and -1.5	-2 and -1.75	-2.25 and -2	-2.5 and -2.25	-2.75 and -2.5	-3 and -2.75	-3.25 and -3	-3.5 and -3.25	-3.75 and -3.5	-4 and -3.75
i1	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4
i2	-0.4	-0.4	-0.4	-0.4	-0.4	**-0.078**	-0.4	0.18	**0.239**	**0.289**	**0.299**	**0.327**
i3	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	**-0.008**	**0.081**	**0.051**	**0.135**
i4	-0.4	-0.4	-0.4	**-0.045**	**0.168**	**0.273**	**0.316**	**0.355**	**0.379**	**0.4**	**0.4**	**0.4**
i5	-0.4	-0.4	-0.4	-0.4	-0.4	**-0.256**	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4
i6	-0.4	-0.4	-0.4	-0.4	-0.4	**-0.124**	-0.4	-0.4	-0.4	-0.4	**0.092**	**0.161**
i7	-0.4	-0.4	-0.4	-0.4	**-0.258**	**-0.081**	-0.4	-0.4	**0.074**	-0.4	**0.085**	**0.063**
i8	-0.4	-0.4	-0.4	**-0.288**	**-0.041**	**0.057**	**0.11**	**0.183**	**0.214**	**0.238**	**0.277**	**0.288**
i9	-0.4	-0.4	-0.4	-0.4	**-0.34**	**-0.285**	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4
i10	-0.4	-0.4	-0.4	-0.4	**-0.205**	**-0.01**	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4

x(i) - number of product to be allocated to the shelf section

Price Change elasticity parameters	-1.25 and -1	-1.5 and -1.25	-1.75 and -1.5	-2 and -1.75	-2.25 and -2	-2.5 and -2.25	-2.75 and -2.5	-3 and -2.75	-3.25 and -3	-3.5 and -3.25	-3.75 and -3.5	-4 and -3.75
i1	6	6	6	3	6	3	3	3	3	3	3	3
i2	3	3	3	9	3	3	6	12	12	15	12	15
i3	3	3	3	3	3	3	3	3	3	3	3	3
i4	6	6	6	6	6	9	12	12	15	15	15	15
i5	4	4	4	4	4	4	4	4	4	4	4	4
i6	3	3	3	6	3	6	3	3	3	3	3	3
i7	16	16	16	16	12	12	4	4	8	4	4	4
i8	12	12	12	12	12	16	16	16	16	16	16	16
i9	6	6	6	6	9	3	3	3	3	3	3	3
i10	15	15	15	9	15	12	12	9	3	3	3	3

*The values for the price change elasticity parameters are randomly generated between the specified values according to a uniform distribution while keeping all the other parameters fixed.

Our sensitivity analysis shows that for the product categories with higher values of price change elasticity parameters (products in the product category are less sensitive to price changes), it is not reasonable to examine the structure of the product category. It is better to increase the prices for all the products to their upper bounds. This can be seen in the real retail environment where for products that are unresponsive to price changes ("unresponsive products" as described by Brown and Tucker, 1961), such as lamps, the retailer does not change the prices significantly from period to period and charges the customer high prices. However for product categories with lower values of price change elasticity parameters (products in the product category are very sensitive to price changes), it is important to define each product's market strength and response to space changes and price changes. Thus, the retailer can adjust his space allocation and decrease the prices of some products in order to increase their demand while increase the prices of some products in order to balance the overall category demand.

Chapter 4

Conclusion

4.1 Summary

In this thesis we have reviewed the shelf space allocation optimization models in the literature. We then developed a model that optimizes retailer's daily gross profit of a product category which makes the decisions of price changes, shelf space allocation and display location in a shelf section with adjustable shelf heights.

In Chapter 2 we carefully examine the literature and classify different demand function forms and components. We also tried to present a unified framework that links the different demand forms together. Our investigation of the major shelf space allocation optimization models in the literature identified some gaps in the literature.

In Chapter 3 we present a new model where we introduce price changes as a factor in the demand function. In addition we consider adjustable shelf heights and enforce product and brand integrity in the shelves. Our numerical results show that considering price changes increases profitability by about 21 to 22 percent compared to the conventional models. While the effect of adjustable shelf heights increases profitability about 6 to 7 percent compared to the conventional models. Finally, the combined effects of price changes and shelf heights increased profits by an average of 30 percent over conventional models.

4.2 Future Work

Through our numerical experiments we found that with todays computing technology we can not solve a comprehensive model of shelf space allocation in a reasonable time (say in a couple of hours). Therefore better solution algorithms should be developed in order to solve these problems in practical times.

Apart from the computational challenges, there are several other future venues for research in this field. One such direction is to add inventory decisions to our model. Including price change effects into the demand function would cause the demand to fluctuate in a wide range and so the related costs such as ordering costs, transportation costs, storage costs and holding costs would also vary. Another direction is to incorporate the effects of uncertain demand in the model. Finally, it would be practical and interesting to extend the model to a supply chain context where decisions related to product assortments and categories inventory management necessarily involve the suppliers.

Finally, extending shelf space allocation models to include inter-temporal effects is also worth investigating. For example, it is plausible to expect the demand to dynamically change depending on previous periods demand as well as prices.

Appendix A

Top Shelf and Bottom Shelf Constraints

In the following we clarify the description of the constraints (3.3.5– 3.3.8) in our model. Since top_i defines the top and $bottom_i$ defines the bottom shelf where product i is located in a shelf section, we can define the top and bottom shelf of each product as

$$top_i = max_m\{m\, y_{i,m}\}$$

and

$$bottom_i = min_m\{m\, y_{i,m}\}.$$

To enforce the adjacency of the products in neighboring shelves we should have

$$top_i - bottom_i + 1 = \sum_{m=1}^{M} y_{i,m}.$$

Finally, we add

$$bottom_i \leqslant top_i$$

to ensure that the top shelf should be on top of the bottom shelf. Using the above four constraints and the idea that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ we get

$$\frac{top_i(top_i + 1)}{2} - \frac{(bottom_i - 1)bottom_i}{2} = \sum_{m=1}^{M} m\, y_{i,m}$$

or

$$top_i + bottom_i = \frac{2\sum_{m=1}^{M} m\, y_{i,m}}{\sum_{m=1}^{M} y_{i,m}}$$

and so we have

$$\begin{aligned} top_i &= \frac{\sum_{m=1}^{M} m\, y_{i,m}}{\sum_{m=1}^{M} y_{i,m}} + \frac{\sum_{m=1}^{M} y_{i,m} - 1}{2} \\ bottom_i &= \frac{\sum_{m=1}^{M} m\, y_{i,m}}{\sum_{m=1}^{M} y_{i,m}} - \frac{\sum_{m=1}^{M} y_{i,m} - 1}{2}. \end{aligned}$$

Bibliography

Anderson, E. E. (1979). An analysis of retail display space: Theory and methods. *Journal of Business*, **52**(1), 103 – 118.

Anderson, E. E. and Amato, H. N. (1974). A mathematical model for simultaneously determining the optimal brand-collection and display-area allocation. *Operations Research*, **22**(1), 13 – 21.

Bai, R. and Kendall, G. (2003). An investigation of automated planograms using a simulated annealing based hyper-heuristic. *In the Proceedings of the 5th Meta-heuristics International Conference (MIC 2003), Kyoto, Japan.*

Bai, R. and Kendall, G. (2008). A model for fresh produce shelf-space allocation and inventory management with freshness-condition-dependent demand. *Journal on Computing*, **20**(1), 78–85.

Bates, A. D. (1970). Estimating the relationship between shelf space and sales for supermarket products. *Dissertation, Indiana University, 1970.*

Bimolt, T. H. A., Van Heerde, H. J., and Pieters, R. G. M. (2005). New emperical generalizations on the determinants of price elasticity. *Journal of Marketing Research*, **38**, 41–56.

Bonami, P. and Lee, J. (2008). *BONMIN Users' Manual*. Available at http://coin-or.org/Bonmin.

Bookbinder, J. H. and Zarour, F. H. (2001). Direct product profitability and retail shelf-space allocation models. *Journal of Business Logistics*, **22**(2), 183–208.

Borin, N. and Farris, P. (1995). A sensitivity analysis of retailer shelf management models. *Journal of Retailing*, **71**(2), 153 – 171.

Borin, N., Farris, P., and Freeland, J. R. (1998). A model for determining retail product category assortment and shelf space allocation. *Decision Sciences*, **25**(3), 359 – 384.

Brown, W. M. and Tucker, W. T. (1961). Vanishing shelf apace. *Atlantic Economic Review*, **9**(10), 9 – 13.

Bucklin, R. E., Russell, G. J., and Srinivasan, V. (1998). A relationship between market share elasticities and brand switching probabilities. *Jourmal of Marketing Research*, **35**, 99–113.

Bultez, A. and Naert, P. (1988). Sh.a.r.p.: Shelf allocation for retailers' profit. *Marketing Science*, **7**(3), 211 – 231.

Buttle, F. (1984). Retail space allocation. *International Journal of Physical Distribution and Material Management*, **14**(4), 3–23.

Buzzle, R. D., Salmon, W. J., and Vancil, R. F. (1965). Product profitability, measurement and merchandising decisions. *Harvard University, Division of Research, Graduate School of Business Administration, Boston.*

Cairns, J. P. (1962). Suppliers, retailers and shelf space. *Journal of Marketing*, **6**, 34 – 36.

Cairns, J. P. (1963). Allocate space for maximum profits. *Journal of Retailing*, **39**, 41–5.

Corstjens, M. and Doyle, P. (1981). A model for optimizing retail space allocations. *Management Science*, **27**(7), 822 – 833.

Corstjens, M. and Doyle, P. (1983). A dynamic model for strategically allocating retail space. *The Journal of the Operational Research Society*, **34**(10), 943 – 951.

Cox, K. K. (1970). The effect of shelf space upon sales of branded products. *Journal of Marketing Research*, **7**(8), 55 – 57.

Curhan, R. C. (1972). The relationship between shelf space and unit sales in supermarkets. *Journal of Marketing Research*, **12**, 406 – 412.

Curhan, R. C. (1973). Shelf space allocation and profit maximization in mass retailing. *Journal of Marketing*, **37**(3), 54 – 60.

Desmet, P. and Renaudin, V. (1998). Estimation of product category sales responsiveness to allocated shelf space. *International Journal of Research in Marketing*, **15**, 443 – 457.

Dichter, E. (1957). The package and the label. *Packaging Carton Research Council, London*.

Dreze, X., Hoch, S. J., and Purk, M. E. (1994). Shelf management and space elasticity. *Journal of Retailing*, **70**(4), 301 – 326.

Fadiloglu, M. M., Karasan, O. E., and Cinar, M. C. (2010). A model and case study for efficient shelf usage and assortment analysis. *Annals of Operations Research*, **180**, 105 – 124.

Frank, P. and Massy, V. (1970). Shelf position and space effects on sales. *Journal of Marketing Research*, **7**(1), 59 – 66.

Gajjar, H. K. and Adil, G. K. (2010). A piecewise linearization for retail shelf space allocation problem and a local search heuristic. *Annals OR*, **179**(1), 149–167.

Gajjar, H. K. and Adil, G. K. (2011). A dynamic programming heuristic for retail shelf space allocation problem. *APJOR*, **28**(2), 183–199.

Goyal, S. and Giri, B. (2001a). Inventory of multi deteriorating items sold from two shops under single management with constraints on space and investment. *Comput Operational Research*, **28**, 1203–1221.

Goyal, S. and Giri, B. (2001b). Recent trends in modelling of deteriorating inventory. *European Journal of Operational Research*, **134**, 116.

Greenhouse, S. (2005). How costco became the anti-wal-mart. *New York Times*.

Grocer, P. (1951a). The foodtown study. *Publications*, **34**.

Grocer, P. (1951b). Sales and margins by commodities. *Publications*, **30**.

Grocer, P. (1958). The super valu study. *Publications*, **37**.

Grocer, P. (1960). The dillon study. *Publications*, **39**.

Grocer, P. (1964). The colonial study. *Publications*, **42-43**.

Hansen, J. M., Raut, S., and Swami, S. (2010). Retail shelf allocation: A comparative analysis of heuristic and meta-heuristic approaches. *Journal of Retailing*, **86**(1), 94 – 105.

Hansen, P. and Heinsbroek, H. (1979). Product selection and space allocation in supermarkets. *European Journal of Operational Research*, **3**(6), 474 – 484.

Hariga, M. A.and Al-Ahmari, A. and Mohamed, A. A. (2007). A joint optimization model for inventory replenishment, product assortment, shelf space and display area allocation decisions. *European Journal of Operational Research*, **181**(1), 239–251.

Hwang, H., Choi, B., and Lee, M. (2005). A model for shelf space allocation and inventory control considering location and inventory level effects on demand. *International Journal of Production Economics*, **97**(2), 185–195.

Irion, J., Al-Khayyal, F., and Lu, J. (2004). A piecewise linearization framework for retail shelf space management models (technical report). *School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta.*

Irion, J., Lu, J.-C., Al-Khayyal, F. A., and Tsao, Y.-C. (2011). A hierarchical decomposition approach to retail shelf space management and assortment decisions. *Journal of the Operational Research Society*, **62**, 1861–1870.

Larson, J. S., Bradlow, E. T., and P.S., F. (2005). An exploratory look at supermarket shopping paths. *International Journal of Research in Marketing*, **22**(4), 395–414.

Larson, P. D. and Robert, A. D. (1990). Psychic stock: An independent variable category of inventory. *International Journal of Physical Distribution and Logistics Management*, **20**(7), 28–34.

Lee, W. (1959). Space management in retail stores and implications to agriculture. *Marketing Keys to Profits in the 1960's, American Marketing Association, Chicago, IL*, pages 523–532.

Lim, A., Rodrigues, B., Xiao, F., and Zhang, X. (2002). Adjusted network flow for the shelf-space allocation problem. In *International Conference on Tools with Artificial Intelligence*, pages 224–.

Lim, A., Rodrigues, B., and Zhang, X. (2004). Metaheuristics with local search techniques for retail shelf-space optimization. *Management Science*, **50**(1), 117–131.

Lotfi, M. M., Rabbani, M., and Ghaderi, S. F. (2011). A weighted goal programming approach for replenishment planning and space allocation in a supermarket. *Journal of the Operational Research Society*, **62**(6), 1128–1137.

Murray, C., Talukdar, D., and Gosavi, A. (2010). Joint optimization of product price, display orientation and shelf-space allocation in retail category management. *Journal of Retailing*, **86**(2), 125–136.

Nahmias, S. (1982). Perishable inventory theory: A review. *Oper. Res.*, **30**, 680708.

Raafat, F. (1991). Survey of literature on continuously deteriorating inventory models. *Journal of Operational Research Society*, **42**, 2737.

Reyes, P. M. and Frazier, G. V. (2005). Initial shelf space considerations at new grocery stores: An allocation problem with product switching and substitution. *International Entrepreneurship and Management Journal*, **1**(2), 183–202.

Reyes, P. M. and Frazier, G. V. (2007). Goal programming model for grocery shelf space allocation. *European Journal of Operational Research*, **181**(2), 634–644.

Russell, R. A. and Urban, T. L. (2010). The location and allocation of products and product families on retail shelves. *Annals of Operations Research*, **179**(1), 131–147.

Shugan, S. M. (1989). Product assortment in a triopoly. *Management Science*, **15**, 304–320.

Silva, D. L., Marikar, F., and Le, K. (2009). Heuristic approach for automated shelf space allocation. In *SAC*, pages 922–928.

Silver, E. A. and Peterson, R. (1985). *Decision Systems for Inventory Management and Production Planning*. New York: Wiley.

Silvera, D. H., Lavack, A. M., and Kropp, F. (2008). Impulse buying: The role of affect social influence, and subjective wellbeing. *Journal of Consumer Marketing*, **25**, 23–33.

Stern, L. W. and El-Ansary, A. I. (1992). *Marketing Channels*. 4th Ed. Englewood Cliffs, NJ: Prentice Hall.

Thurik, R. (1988). Les grandes surfaces en france:etude de la relation ventes/surface du magasin. *Rescherhe et Applications en Marketing*, **3**(3), 21 – 37.

Underhill, P. (1999). Why we buy: The science of shopping. In *New York: Simon and Schuster*.

Urban, G. L. (1969). A mathematical modeling approach to product line decisions. *Journal of Marketing Research*, **6**, 40 – 47.

Urban, T. L. (1998). An inventory-theoretic approach to product assortment and shelf-space allocation. *Journal of Retailing*, **74**(1), 15 – 35.

Urban, T. L. (2002). The interdependence of inventory management and retail shelf management. *International Journal of Physical Distribution and Logistics*, **32**(1), 41 – 58.

White Plains, N. G. F. C. (1963). Mckinsey-general foods study. *The Economics of Food Distributors.*

Yang, M. (2001). An efficient algorithm to allocate shelf space. *European Journal of Operational Research*, **131**(1), 107 – 118.

Yang, M. and Chen, W. (1999). A study on shelf space allocation and management. *International Journal of Production Economics*, **60-61**(1), 309–317.

Zufryden, F. S. (1986). A dynamic programming approach for product selection and supermarket shelf-space allocation. *The Journal of the Operational Research Society*, **37**(4), 413 – 422.